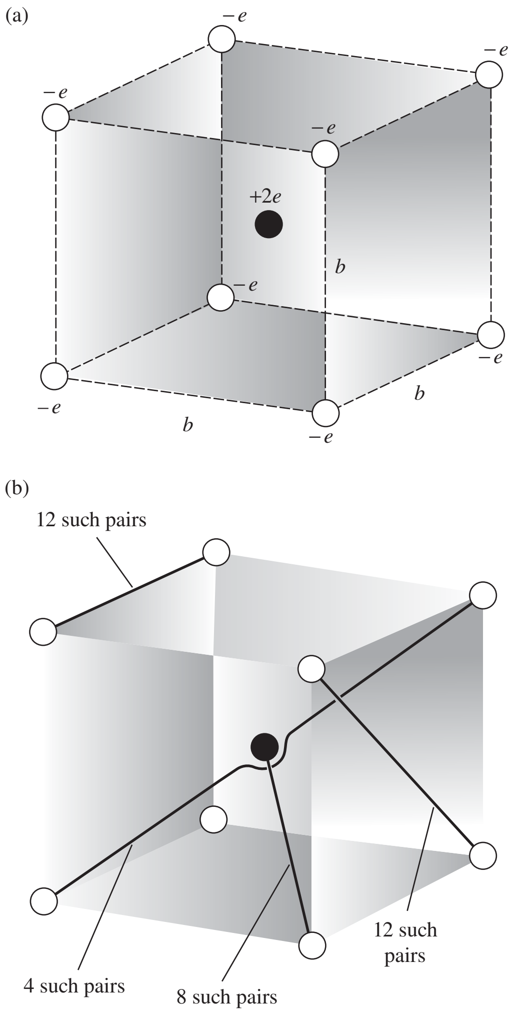


Figure 1.6.
(a) The potential energy of this arrangement of nine point charges is given by Eq. (1.14).
(b) Four types of pairs are involved in the sum.

Example (Charges in a cube) What is the potential energy of an arrangement of eight negative charges on the corners of a cube of side b, with a positive charge in the center of the cube, as in Fig. 1.6(a)? Suppose each negative charge is an electron with charge $-e$, while the central particle carries a double positive charge, $2e$.

Solution Figure 1.6(b) shows that there are four different types of pairs. One type involves the center charge, while the other three involve the various edges and diagonals of the cube. Summing over all pairs yields

$$U = \frac{1}{4\pi\epsilon_0}\left(8\cdot\frac{(-2e^2)}{(\sqrt{3}/2)b} + 12\cdot\frac{e^2}{b} + 12\cdot\frac{e^2}{\sqrt{2}\,b} + 4\cdot\frac{e^2}{\sqrt{3}\,b}\right) \approx \frac{1}{4\pi\epsilon_0}\frac{4.32e^2}{b}. \tag{1.14}$$

The energy is positive, indicating that work had to be done on the system to assemble it. That work could, of course, be recovered if we let the charges move apart, exerting forces on some external body or bodies. Or if the electrons were simply to fly apart from this configuration, the *total kinetic energy* of all the particles would become equal to U. This would be true whether they came apart simultaneously and symmetrically, or were released one at a time in any order. Here we see the power of this simple notion of the total potential energy of the system. Think what the problem would be like if we had to compute the resultant vector force on every particle at every stage of assembly of the configuration! In this example, to be sure, the geometrical symmetry would simplify that task; even so, it would be more complicated than the simple calculation above.

One way of writing the instruction for the sum over pairs is this:

$$U = \frac{1}{2}\sum_{j=1}^{N}\sum_{k\neq j}\frac{1}{4\pi\epsilon_0}\frac{q_j q_k}{r_{jk}}. \tag{1.15}$$

The double-sum notation, $\sum_{j=1}^{N}\sum_{k\neq j}$, says: take $j=1$ and sum over $k=2,3,4,\ldots,N$; then take $j=2$ and sum over $k=1,3,4,\ldots,N$; and so on, through $j=N$. Clearly this includes every pair *twice,* and to correct for that we put in front the factor $1/2$.

1.6 Electrical energy in a crystal lattice

These ideas have an important application in the physics of crystals. We know that an ionic crystal like sodium chloride can be described, to a very good approximation, as an arrangement of positive ions (Na^+) and negative ions (Cl^-) alternating in a regular three-dimensional array or lattice. In sodium chloride the arrangement is that shown in Fig. 1.7(a). Of course the ions are not point charges, but they are nearly spherical distributions of charge and therefore (as we shall prove in Section 1.11) the electrical forces they exert on one another are the same as if each ion were replaced by an equivalent point charge at its center. We show this electrically equivalent system in Fig. 1.7(b). The electrostatic potential energy of the lattice of charges plays an important role in the explanation of the stability and cohesion of the ionic crystal. Let us see if we can estimate its magnitude.

We seem to be faced at once with a sum that is enormous, if not doubly infinite; any macroscopic crystal contains 10^{20} atoms at least. Will the sum converge? Now what we hope to find is the potential energy per unit volume or mass of crystal. We confidently expect this to be independent of the size of the crystal, based on the general argument that one end of a macroscopic crystal can have little influence on the other. Two grams of sodium chloride ought to have twice the potential energy of one gram, and the shape should not be important so long as the surface atoms are a small fraction of the total number of atoms. We would be *wrong* in this expectation if the crystal were made out of ions of one sign only. Then, 1 g of crystal would carry an enormous electric charge, and putting two such crystals together to make a 2 g crystal would take a fantastic amount of energy. (You might estimate how much!) The situation is saved by the fact that the crystal structure is an alternation of equal and opposite charges, so that any macroscopic bit of crystal is very nearly neutral.

To evaluate the potential energy we first observe that every positive ion is in a position equivalent to that of every other positive ion. Furthermore, although it is perhaps not immediately obvious from Fig. 1.7, the arrangement of positive ions around a negative ion is exactly the same as the arrangement of negative ions around a positive ion, and so on. Hence we may take one ion as a center, it matters not which kind, sum over *its* interactions with all the others, and simply multiply by the total number of ions of both kinds. This reduces the double sum in Eq. (1.15) to a single sum and a factor N; we must still apply the factor $1/2$ to compensate for including each pair twice. That is, the energy of a sodium chloride lattice composed of a total of N ions is

$$U = \frac{1}{2} N \sum_{k=2}^{N} \frac{1}{4\pi\epsilon_0} \frac{q_1 q_k}{r_{1k}}. \tag{1.16}$$

Taking the positive ion at the center as in Fig. 1.7(b), our sum runs over all its neighbors near and far. The leading terms start out as follows:

$$U = \frac{1}{2} N \frac{1}{4\pi\epsilon_0} \left(-\frac{6e^2}{a} + \frac{12e^2}{\sqrt{2}\,a} - \frac{8e^2}{\sqrt{3}\,a} + \cdots \right). \tag{1.17}$$

The first term comes from the 6 nearest chlorine ions, at distance a, the second from the 12 sodium ions on the cube edges, and so on. It is clear, incidentally, that this series does not converge *absolutely*; if we were so

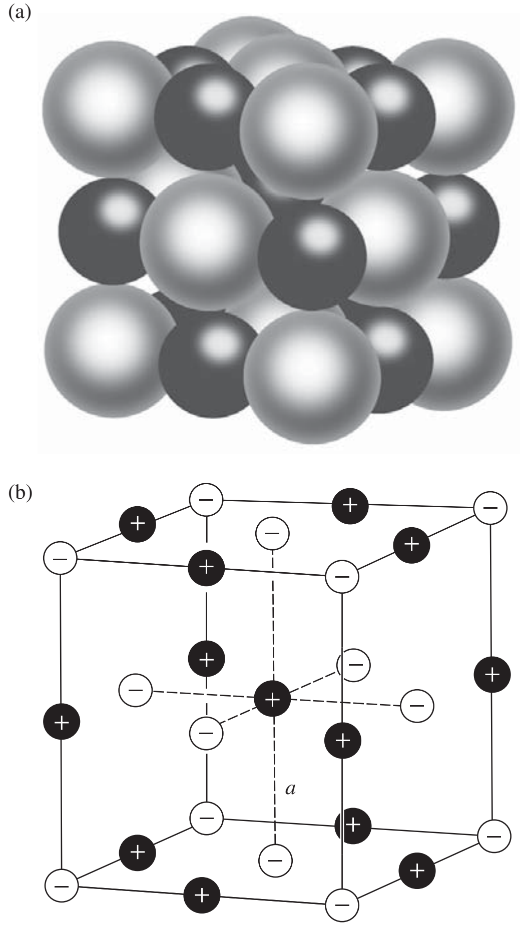


Figure 1.7.
A portion of a sodium chloride crystal, with the ions Na^+ and Cl^- shown in about the right relative proportions (a), and replaced by equivalent point charges (b).

foolish as to try to sum all the positive terms first, that sum would diverge. To evaluate such a sum, we should arrange it so that as we proceed outward, including ever more distant ions, we include them in groups that represent nearly neutral shells of material. Then if the sum is broken off, the more remote ions that have been neglected will be such an even mixture of positive and negative charges that we can be confident their contribution would have been small. This is a crude way to describe what is actually a somewhat more delicate computational problem. The numerical evaluation of such a series is easily accomplished with a computer. The answer in this example happens to be

$$U = \frac{-0.8738Ne^2}{4\pi\epsilon_0 a}. \tag{1.18}$$

Here N, the number of ions, is twice the number of NaCl molecules.

The negative sign shows that work would have to be *done* to take the crystal apart into ions. In other words, the electrical energy helps to explain the cohesion of the crystal. If this were the whole story, however, the crystal would collapse, for the potential energy of the charge distribution is obviously *lowered* by shrinking all the distances. We meet here again the familiar dilemma of classical – that is, nonquantum – physics. No system of stationary particles can be in stable equilibrium, according to classical laws, under the action of electrical forces alone; we will give a proof of this fact in Section 2.12. Does this make our analysis useless? Not at all. Remarkably, and happily, in the quantum physics of crystals the electrical potential energy can still be given meaning, and can be computed very much in the way we have learned here.

1.7 The electric field

Suppose we have some arrangement of charges, $q_1, q_2, \ldots, q_N$, fixed in space, and we are interested not in the forces they exert on one another, but only in their effect on some other charge q_0 that might be brought into their vicinity. We know how to calculate the resultant force on this charge, given its position which we may specify by the coordinates x, y, z. The force on the charge q_0 is

$$\mathbf{F} = \frac{1}{4\pi\epsilon_0}\sum_{j=1}^{N}\frac{q_0 q_j \hat{\mathbf{r}}_{0j}}{r_{0j}^2}, \tag{1.19}$$

where $\mathbf{r}_{0j}$ is the vector from the jth charge in the system to the point (x, y, z). The force is proportional to q_0, so if we divide out q_0 we obtain a vector quantity that depends only on the structure of our original system of charges, $q_1, \ldots, q_N$, and on the position of the point (x, y, z). We call this vector function of x, y, z the *electric field* arising from the $q_1, \ldots, q_N$

and use the symbol $\mathbf{E}$ for it. The charges $q_1, \ldots, q_N$ we call *sources* of the field. We may take as the *definition* of the electric field $\mathbf{E}$ of a charge distribution, at the point (x, y, z),

$$\mathbf{E}(x, y, z) = \frac{1}{4\pi\epsilon_0} \sum_{j=1}^{N} \frac{q_j \hat{\mathbf{r}}_{0j}}{r_{0j}^2}. \tag{1.20}$$

The force on some other charge q at (x, y, z) is then

$$\boxed{\mathbf{F} = q\mathbf{E}} \tag{1.21}$$

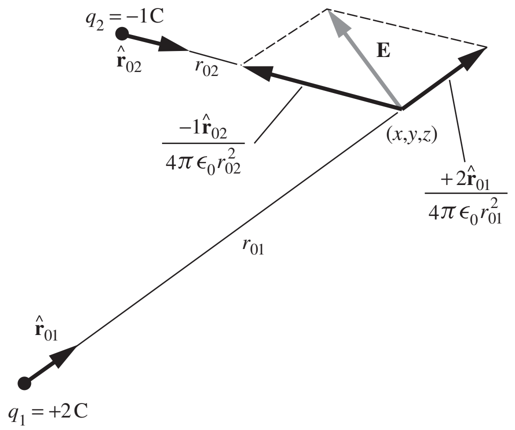


Figure 1.8.
The field at a point is the vector sum of the fields of each of the charges in the system.

Figure 1.8 illustrates the vector addition of the field of a point charge of 2 C to the field of a point charge of −1 C, at a particular point in space. In the SI system of units, electric field strength is expressed in newtons per unit charge, that is, newtons/coulomb. In Gaussian units, with the esu as the unit of charge and the dyne as the unit of force, the electric field strength is expressed in dynes/esu.

After the introduction of the electric potential in Chapter 2, we shall have another, and completely equivalent, way of expressing the unit of electric field strength; namely, volts/meter in SI units and statvolts/centimeter in Gaussian units.

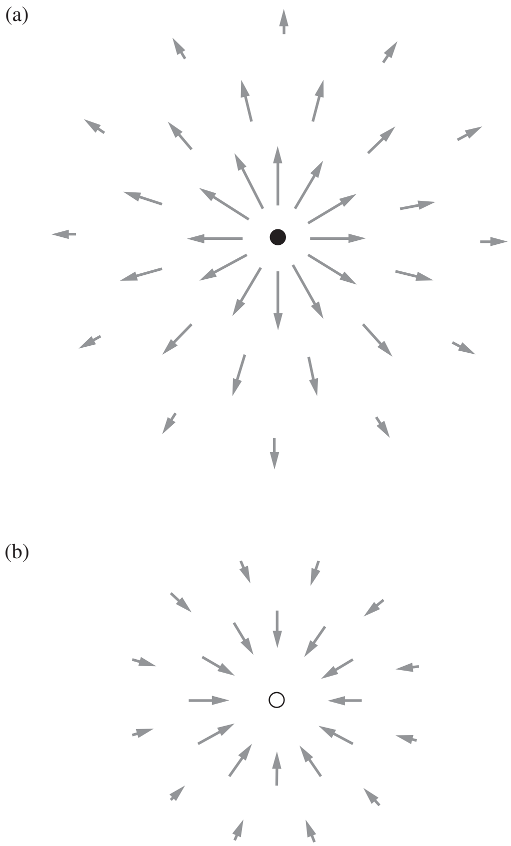
So far we have nothing really new. The electric field is merely another way of describing the system of charges; it does so by giving the force per unit charge, in magnitude and direction, that an exploring charge q_0 would experience at any point. We have to be a little careful with that interpretation. Unless the source charges are really immovable, the introduction of some finite charge q_0 may cause the source charges to shift their positions, so that the field itself, as defined by Eq. (1.20), is different. That is why we assumed fixed charges to begin our discussion. People sometimes define the field by requiring q_0 to be an "infinitesimal" test charge, letting $\mathbf{E}$ be the limit of $\mathbf{F}/q_0$ as $q_0 \to 0$. Any flavor of rigor this may impart is illusory. Remember that in the real world we have never observed a charge smaller than e! Actually, if we take Eq. (1.20) as our *definition* of $\mathbf{E}$, without reference to a test charge, no problem arises and the sources need not be fixed. If the introduction of a new charge causes a shift in the source charges, then it has indeed brought about a change in the electric field, and if we want to predict the force on the new charge, we must use the new electric field in computing it.

Perhaps you still want to ask, what *is* an electric field? Is it something real, or is it merely a name for a factor in an equation that has to be multiplied by something else to give the numerical value of the force we measure in an experiment? Two observations may be useful here. First, since it works, it doesn't make any difference. That is not a frivolous answer, but a serious one. Second, the fact that the electric field vector at a point in space is all we need know to predict the force that will act on *any* charge at that point is by no means trivial. It might have been otherwise! If no experiments had ever been done, we could imagine that, in two different situations in which unit charges experience equal force, test charges of strength 2 units might experience unequal forces, depending on the nature of the other charges in the system. If that were true, the field description wouldn't work. The electric field attaches to every point in a system a *local property,* in this sense: if we know **E** in some small neighborhood, we know, *without further inquiry,* what will happen to any charges in that neighborhood. We do not need to ask what produced the field.

To visualize an electric field, you need to associate a vector, that is, a magnitude and direction, with every point in space. We shall use various schemes in this book, none of them wholly satisfactory, to depict vector fields.

It is hard to draw in two dimensions a picture of a vector function in three-dimensional space. We can indicate the magnitude and direction of **E** at various points by drawing little arrows near those points, making the arrows longer where E is larger.[7] Using this scheme, we show in Fig. 1.9(a) the field of an isolated point charge of 3 units and in Fig. 1.9(b) the field of a point charge of -1 unit. These pictures admittedly add nothing whatsoever to our understanding of the field of an isolated charge; anyone can imagine a simple radial inverse-square field without the help of a picture. We show them in order to combine (side by side) the two fields in Fig. 1.10, which indicates in the same manner the field of two such charges separated by a distance a. All that Fig. 1.10 can show is the field in a plane containing the charges. To get a full three-dimensional representation, one must imagine the figure rotated around the symmetry axis. In Fig. 1.10 there is one point in space where **E** is zero. As an exercise, you can quickly figure out where this point lies. Notice also that toward the edge of the picture the field points more or less radially outward all around. One can see that at a very large distance from the charges the field will look very much like the field from a positive point charge. This is to be expected because the separation of the charges cannot make very much difference for points far away, and a point charge of 2 units is just what we would have left if we superimposed our two sources at one spot.

Another way to depict a vector field is to draw *field lines.* These are simply curves whose tangent, at any point, lies in the direction of the field at that point. Such curves will be smooth and continuous except at singularities such as point charges, or points like the one in the example of Fig. 1.10 where the field is zero. A field line plot does not directly give



● Charge +3
○ Charge −1

Figure 1.9.
(a) Field of a charge $q_1 = 3$. (b) Field of a charge $q_2 = -1$. Both representations are necessarily crude and only roughly quantitative.

[7] Such a representation is rather clumsy at best. It is hard to indicate the point in space to which a particular vector applies, and the range of magnitudes of E is usually so large that it is impracticable to make the lengths of the arrows proportional to E.

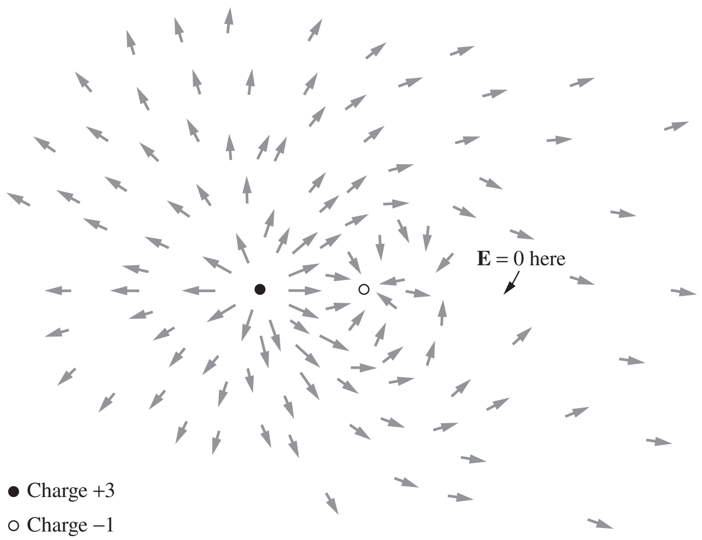


Figure 1.10.
The field in the vicinity of two charges, $q_1 = +3$, $q_2 = -1$, is the superposition of the fields in Figs. 1.9(a) and (b).

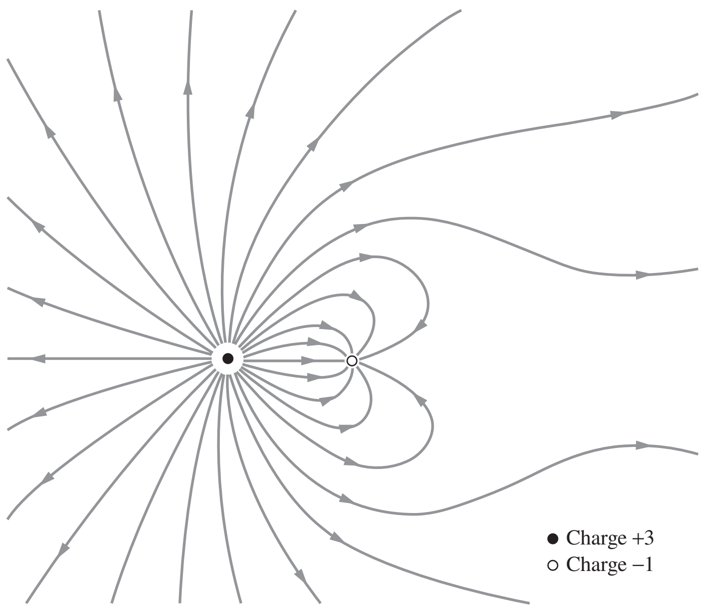


Figure 1.11.
Some field lines in the electric field around two charges, $q_1 = +3$, $q_2 = -1$.

the magnitude of the field, although we shall see that, in a general way, the field lines converge as we approach a region of strong field and spread apart as we approach a region of weak field. In Fig. 1.11 are drawn some field lines for the same arrangement of charges as in Fig. 1.10, a positive charge of 3 units and a negative charge of 1 unit. Again, we are restricted by the nature of paper and ink to a two-dimensional section through a three-dimensional bundle of curves.

1.8 Charge distributions

This is as good a place as any to generalize from *point charges* to *continuous charge distributions*. A volume distribution of charge is described by a scalar charge-density function ρ, which is a function of position, with the dimensions *charge/volume*. That is, ρ times a volume element gives the amount of charge contained in that volume element. The same symbol is often used for mass per unit volume, but in this book we shall always give charge per unit volume first call on the symbol ρ. If we write ρ as a function of the coordinates x, y, z, then $\rho(x, y, z)\, dx\, dy\, dz$ is the charge contained in the little box, of volume $dx\, dy\, dz$, located at the point (x, y, z).

On an atomic scale, of course, the charge density varies enormously from point to point; even so, it proves to be a useful concept in that domain. However, we shall use it mainly when we are dealing with large-scale systems, so large that a volume element $dv = dx\, dy\, dz$ can be quite small relative to the size of our system, although still large enough to contain many atoms or elementary charges. As we have remarked before, we face a similar problem in defining the ordinary mass density of a substance.

If the source of the electric field is to be a continuous charge distribution rather than point charges, we merely replace the sum in Eq. (1.20) with the appropriate integral. The integral gives the electric field at (x, y, z), which is produced by charges at other points (x', y', z'):

$$\mathbf{E}(x, y, z) = \frac{1}{4\pi\epsilon_0} \int \frac{\rho(x', y', z')\hat{\mathbf{r}}\, dx'\, dy'\, dz'}{r^2}. \tag{1.22}$$

This is a volume integral. Holding (x, y, z) fixed, we let the variables of integration, x', y', and z', range over all space containing charge, thus summing up the contributions of all the bits of charge. The unit vector $\hat{\mathbf{r}}$ points from (x', y', z') to (x, y, z) – unless we want to put a minus sign before the integral, in which case we may reverse the direction of $\hat{\mathbf{r}}$. It is always hard to keep signs straight. Let's remember that the electric field points *away* from a positive source (Fig. 1.12).

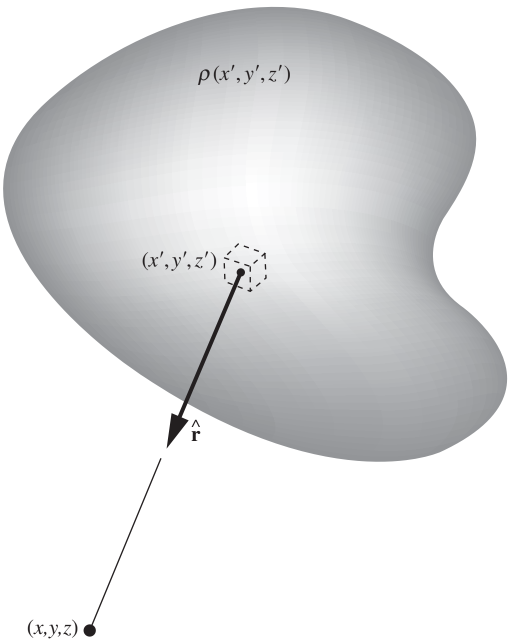


Figure 1.12.
Each element of the charge distribution $\rho(x', y', z')$ makes a contribution to the electric field $\mathbf{E}$ at the point (x, y, z). The total field at this point is the sum of all such contributions; see Eq. (1.22).

Example (Field due to a hemisphere) A solid hemisphere has radius R and uniform charge density ρ. Find the electric field at the center.

Solution Our strategy will be to slice the hemisphere into rings around the symmetry axis. We will find the electric field due to each ring, and then integrate over the rings to obtain the field due to the entire hemisphere. We will work with polar coordinates (or, equivalently, spherical coordinates), which are much more suitable than Cartesian coordinates in this setup.

The cross section of a ring is (essentially) a little rectangle with side lengths dr and $r\,d\theta$, as shown in Fig. 1.13. The cross-sectional area is thus $r\,dr\,d\theta$. The radius of the ring is $r\sin\theta$, so the volume is $(r\,dr\,d\theta)(2\pi r\sin\theta)$. The charge in the ring is therefore $\rho(2\pi r^2\sin\theta\,dr\,d\theta)$. Equivalently, we can obtain this result by using the standard spherical-coordinate volume element, $r^2\sin\theta\,dr\,d\theta\,d\phi$, and then integrating over ϕ to obtain the factor of 2π.

Consider a tiny piece of the ring, with charge dq. This piece creates an electric field at the center of the hemisphere that points diagonally upward (if ρ is positive) with magnitude $dq/4\pi\epsilon_0 r^2$. However, only the vertical component survives, because the horizontal component cancels with the horizontal component from the diametrically opposite charge dq on the ring. The vertical component involves a factor of $\cos\theta$. When we integrate over the whole ring, the dq simply integrates to the total charge we found above. The (vertical) electric field due to a given ring is therefore

$$dE_y = \frac{\rho(2\pi r^2\sin\theta\,dr\,d\theta)}{4\pi\epsilon_0 r^2}\cos\theta = \frac{\rho\sin\theta\cos\theta\,dr\,d\theta}{2\epsilon_0}. \tag{1.23}$$

Integrating over r and θ to obtain the field due to the entire hemisphere gives

$$\begin{aligned} E_y &= \int_0^R\int_0^{\pi/2}\frac{\rho\sin\theta\cos\theta\,dr\,d\theta}{2\epsilon_0} = \frac{\rho}{2\epsilon_0}\left(\int_0^R dr\right)\left(\int_0^{\pi/2}\sin\theta\cos\theta\,d\theta\right) \\ &= \frac{\rho}{2\epsilon_0}\cdot R\cdot\frac{\sin^2\theta}{2}\bigg|_0^{\pi/2} = \frac{\rho R}{4\epsilon_0}. \end{aligned} \tag{1.24}$$

Note that the radius r canceled in Eq. (1.23). For given values of θ, $d\theta$, and dr, the volume of a ring grows like r^2, and this exactly cancels the r^2 in the denominator in Coulomb's law.

REMARK As explained above, the electric field due to the hemisphere is vertical. This fact also follows from considerations of symmetry. We will make many symmetry arguments throughout this book, so let us be explicit here about how the reasoning proceeds. Assume (in search of a contradiction) that the electric field due to the hemisphere is *not* vertical. It must then point off at some angle, as shown in Fig. 1.14(a). Let's say that the **E** vector lies above a given dashed line painted on the hemisphere. If we rotate the system by, say, 180° around the symmetry axis, the field now points in the direction shown in Fig. 1.14(b), because it must still pass over the dashed line. But we have *exactly the same hemisphere* after the rotation, so the field must still point upward to the right. We conclude that the field due to the hemisphere points both upward to the left and upward to the right. This is a contradiction. The only way to avoid this contradiction is for the field to point along the symmetry axis (possibly in the negative direction), because in that case it doesn't change under the rotation.

In the neighborhood of a true point charge the electric field grows infinite like $1/r^2$ as we approach the point. It makes no sense to talk about the field *at* the point charge. As our ultimate physical sources of field are

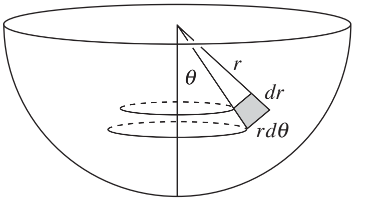
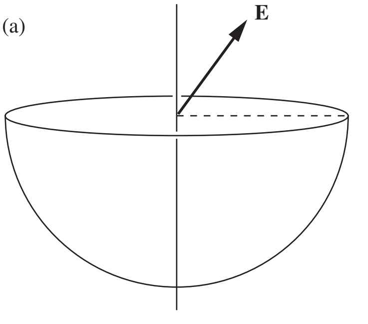


Figure 1.13.
Cross section of a thin ring. The hemisphere may be considered to be built up from rings.

(a)



(b)

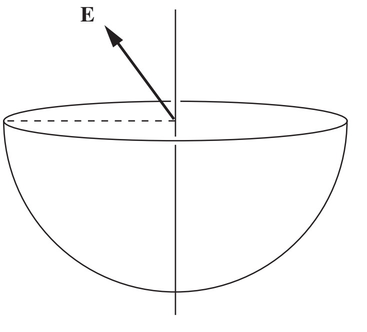


Figure 1.14.
The symmetry argument that explains why **E** must be vertical.

not, we believe, infinite concentrations of charge in zero volume, but instead finite structures, we simply ignore the mathematical singularities implied by our point-charge language and rule out of bounds the interior of our elementary sources. A continuous charge distribution $\rho(x', y', z')$ that is nowhere infinite gives no trouble at all. Equation (1.22) can be used to find the field at any point within the distribution. The integrand doesn't blow up at $r = 0$ because the volume element in the numerator equals $r^2 \sin\phi \, d\phi \, d\theta \, dr$ in spherical coordinates, and the r^2 here cancels the r^2 in the denominator in Eq. (1.22). That is to say, so long as ρ remains finite, the field will remain finite everywhere, even in the interior or on the boundary of a charge distribution.

1.9 Flux

The relation between the electric field and its sources can be expressed in a remarkably simple way, one that we shall find very useful. For this we need to define a quantity called *flux*.

Consider some electric field in space and in this space some arbitrary closed surface, like a balloon of any shape. Figure 1.15 shows such a surface, the field being suggested by a few field lines. Now divide the whole surface into little patches that are so small that over any one patch the surface is practically flat and the vector field does not change appreciably from one part of a patch to another. In other words, don't let the balloon be too crinkly, and don't let its surface pass right through a singularity[8] of the field such as a point charge. The area of a patch has a certain magnitude in square meters, and a patch defines a unique direction – the outward-pointing normal to its surface. (Since the surface is closed, you can tell its inside from its outside; there is no ambiguity.) Let this magnitude and direction be represented by a vector. Then for every patch into which the surface has been divided, such as patch number j, we have a vector $\mathbf{a}_j$ giving its area and orientation. The steps we have just taken are pictured in Figs. 1.15(b) and (c). Note that the vector $\mathbf{a}_j$ does not depend at all on the shape of the patch; it doesn't matter how we have divided up the surface, as long as the patches are small enough.

Let $\mathbf{E}_j$ be the electric field vector at the location of patch number j. The scalar product $\mathbf{E}_j \cdot \mathbf{a}_j$ is a number. We call this number the *flux* through that bit of surface. To understand the origin of the name, imagine a vector function that represents the velocity of motion in a fluid – say in a river, where the velocity varies from one place to another but is constant in time at any one position. Denote this vector field by $\mathbf{v}$, measured in

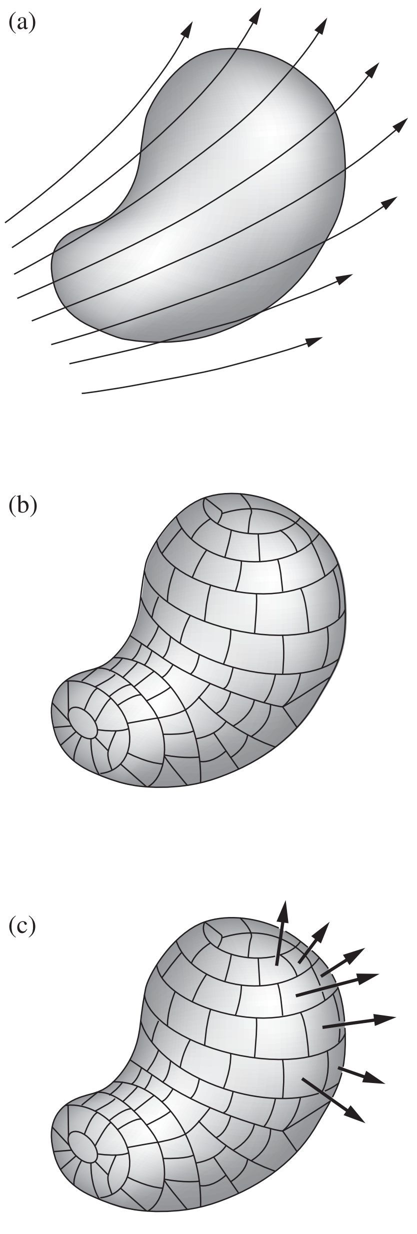


Figure 1.15.
(a) A closed surface in a vector field is divided (b) into small elements of area. (c) Each element of area is represented by an outward vector.

[8] By a singularity of the field we would ordinarily mean not only a point source where the field approaches infinity, but also any place where the field changes magnitude or direction discontinuously, such as an infinitesimally thin layer of concentrated charge. Actually this latter, milder, kind of singularity would cause no difficulty here unless our balloon's surface were to coincide with the surface of discontinuity over some finite area.

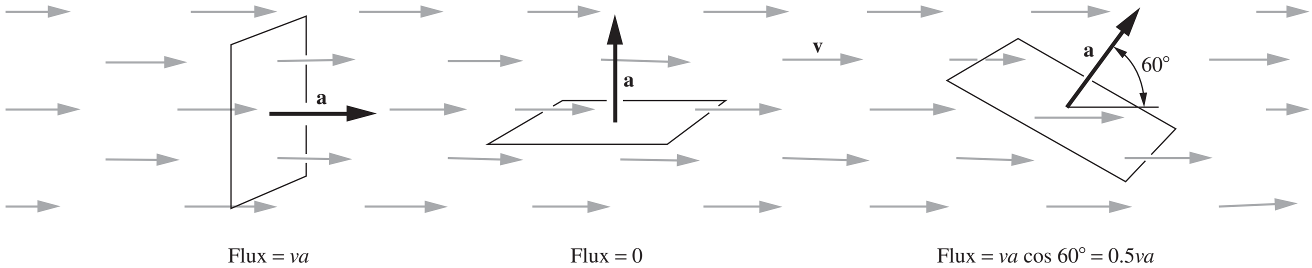


Figure 1.16.
The flux through the frame of area $\mathbf{a}$ is $\mathbf{v} \cdot \mathbf{a}$, where $\mathbf{v}$ is the velocity of the fluid. The flux is the volume of fluid passing through the frame, per unit time.

meters/second. Then, if $\mathbf{a}$ is the oriented area in square meters of a frame lowered into the water, $\mathbf{v} \cdot \mathbf{a}$ is the *rate of flow* of water through the frame in cubic meters per second (Fig. 1.16). The $\cos\theta$ factor in the standard expression for the dot product correctly picks out the component of $\mathbf{v}$ along the direction of $\mathbf{a}$, or equivalently the component of $\mathbf{a}$ along the direction of $\mathbf{v}$. We must emphasize that our definition of flux is applicable to any vector function, whatever physical variable it may represent.

Now let us add up the flux through all the patches to get the flux through the entire surface, a scalar quantity which we shall denote by Φ:

$$\Phi = \sum_{\text{all } j} \mathbf{E}_j \cdot \mathbf{a}_j. \tag{1.25}$$

Letting the patches become smaller and more numerous without limit, we pass from the sum in Eq. (1.25) to a surface integral:

$$\Phi = \int_{\substack{\text{entire}\\ \text{surface}}} \mathbf{E} \cdot d\mathbf{a}. \tag{1.26}$$

A surface integral of any vector function $\mathbf{F}$, over a surface S, means just this: divide S into small patches, each represented by a vector outward, of magnitude equal to the patch area; at every patch, take the scalar product of the patch area vector and the local $\mathbf{F}$; sum all these products, and the limit of this sum, as the patches shrink, is the surface integral. Do not be alarmed by the prospect of having to perform such a calculation for an awkwardly shaped surface like the one in Fig. 1.15. The surprising property we are about to demonstrate makes that unnecessary!

1.10 Gauss's law

Take the simplest case imaginable; suppose the field is that of a single isolated positive point charge q, and the surface is a sphere of radius r centered on the point charge (Fig. 1.17). What is the flux Φ through this surface? The answer is easy because the magnitude of $\mathbf{E}$ at every point on the surface is $q/4\pi\epsilon_0 r^2$ and its direction is the same as that of the outward normal at that point. So we have

$$\Phi = E \cdot (\text{total area}) = \frac{q}{4\pi\epsilon_0 r^2} \cdot 4\pi r^2 = \frac{q}{\epsilon_0}. \tag{1.27}$$

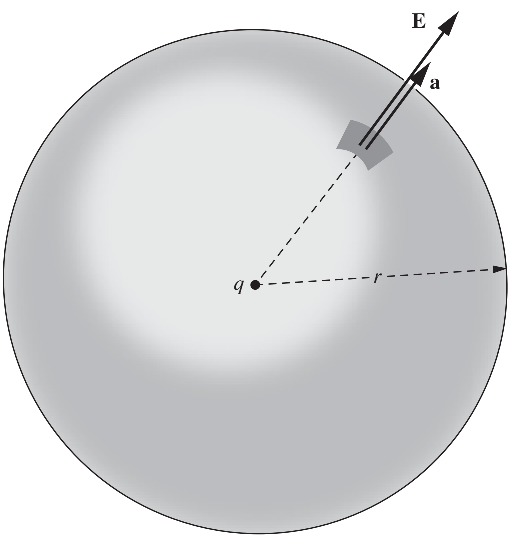


Figure 1.17.
In the field $\mathbf{E}$ of a point charge q, what is the outward flux over a sphere surrounding q?

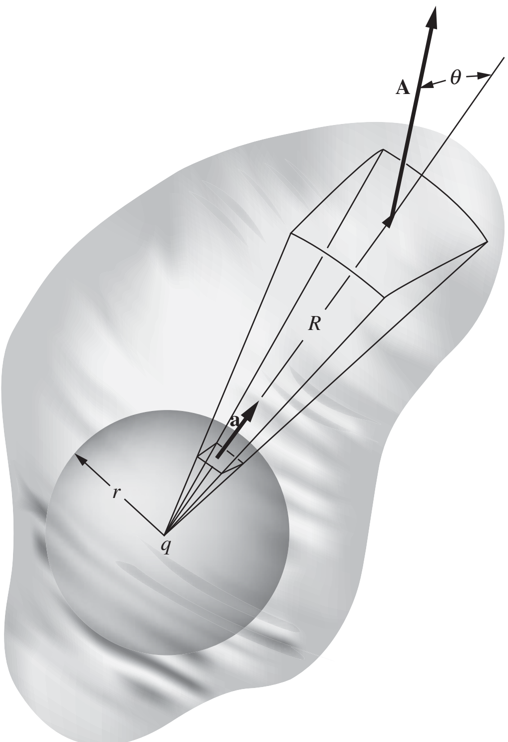


Figure 1.18.
Showing that the flux through any closed surface around q is the same as the flux through the sphere.

The flux is independent of the size of the sphere. Here for the first time we see the benefit of including the factor of $1/4\pi$ in Coulomb's law in Eq. (1.4). Without this factor, we would have an uncanceled factor of 4π in Eq. (1.27) and therefore also, eventually, in one of Maxwell's equations. Indeed, in Gaussian units Eq. (1.27) takes the form of $\Phi = 4\pi q$.

Now imagine a second surface, or balloon, enclosing the first, but *not* spherical, as in Fig. 1.18. We claim that *the total flux through this surface is the same as that through the sphere.* To see this, look at a cone, radiating from q, that cuts a small patch **a** out of the sphere and continues on to the outer surface, where it cuts out a patch **A** at a distance R from the point charge. The area of the patch **A** is larger than that of the patch **a** by two factors: first, by the ratio of the distance squared $(R/r)^2$; and second, owing to its inclination, by the factor $1/\cos\theta$. The angle θ is the angle between the outward normal and the radial direction (see Fig. 1.18). The electric field in that neighborhood is reduced from its magnitude on the sphere by the factor $(r/R)^2$ and is still radially directed. Letting $\mathbf{E}_{(R)}$ be the field at the outer patch and $\mathbf{E}_{(r)}$ be the field at the sphere, we have

$$\begin{aligned}\text{flux through outer patch} &= \mathbf{E}_{(R)} \cdot \mathbf{A} = E_{(R)} A \cos\theta,\\ \text{flux through inner patch} &= \mathbf{E}_{(r)} \cdot \mathbf{a} = E_{(r)} a.\end{aligned} \tag{1.28}$$

Using the above facts concerning the magnitude of $\mathbf{E}_{(R)}$ and the area of **A**, the flux through the outer patch can be written as

$$E_{(R)} A \cos\theta = \left[E_{(r)}\left(\frac{r}{R}\right)^2\right]\left[a\left(\frac{R}{r}\right)^2 \frac{1}{\cos\theta}\right]\cos\theta = E_{(r)} a, \tag{1.29}$$

which equals the flux through the inner patch.

Now every patch on the outer surface can in this way be put into correspondence with part of the spherical surface, so the total flux must be the same through the two surfaces. That is, the flux through the new surface must be just q/ϵ_0. But this was a surface of *arbitrary* shape and size.[9] We conclude: the flux of the electric field through *any* surface enclosing a point charge q is q/ϵ_0. As a corollary we can say that the total flux through a closed surface is *zero* if the charge lies *outside* the surface. We leave the proof of this to the reader, along with Fig. 1.19 as a hint of one possible line of argument.

There is a way of looking at all this that makes the result seem obvious. Imagine at q a source that emits particles – such as bullets or photons – in all directions at a steady rate. Clearly the flux of particles through a window of unit area will fall off with the inverse square of the window's distance from q. Hence we can draw an analogy between the electric field strength E and the intensity of particle flow in bullets per unit area per

[9] To be sure, we had the second surface enclosing the sphere, but it didn't have to, really. Besides, the sphere can be taken as small as we please.

unit time. It is pretty obvious that the flux of bullets through any surface completely surrounding q is independent of the size and shape of that surface, for it is just the total number emitted per unit time. Correspondingly, the flux of E through the closed surface must be independent of size and shape. The common feature responsible for this is the inverse-square behavior of the intensity.

The situation is now ripe for superposition! Any electric field is the sum of the fields of its individual sources. This property was expressed in our statement, Eq. (1.19), of Coulomb's law. Clearly flux is an additive quantity in the same sense, for if we have a number of sources, $q_1, q_2, \ldots, q_N$, the fields of which, if each were present alone, would be $\mathbf{E}_1, \mathbf{E}_2, \ldots, \mathbf{E}_N$, then the flux Φ through some surface S in the actual field can be written

$$\Phi = \int_S \mathbf{E} \cdot d\mathbf{a} = \int_S (\mathbf{E}_1 + \mathbf{E}_2 + \cdots + \mathbf{E}_N) \cdot d\mathbf{a}. \tag{1.30}$$

We have just learned that $\int_S \mathbf{E}_i \cdot d\mathbf{a}$ equals q_i/ϵ_0 if the charge q_i is inside S and equals zero otherwise. So every charge q inside the surface contributes exactly q/ϵ_0 to the surface integral of Eq. (1.30) and all charges outside contribute nothing. We have arrived at Gauss's law.

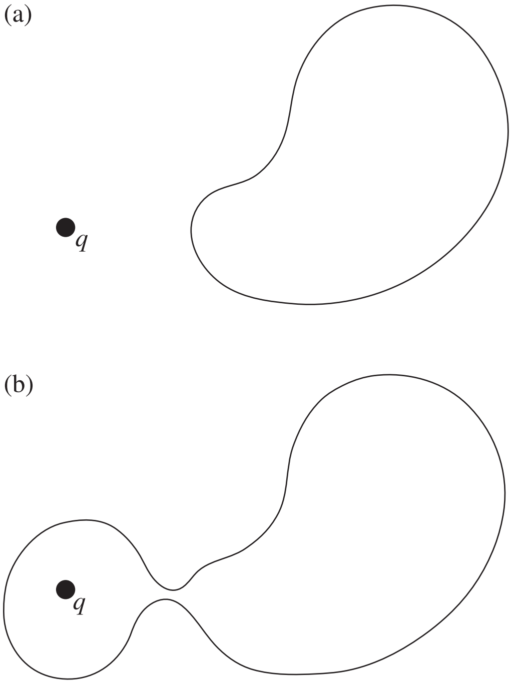


Figure 1.19.
To show that the flux through the closed surface in (a) is zero, you can make use of (b).

> The flux of the electric field $\mathbf{E}$ through any closed surface, that is, the integral $\int \mathbf{E} \cdot d\mathbf{a}$ over the surface, equals $1/\epsilon_0$ times the total charge enclosed by the surface:
>
> $$\int \mathbf{E} \cdot d\mathbf{a} = \frac{1}{\epsilon_0} \sum_i q_i = \frac{1}{\epsilon_0} \int \rho \, dv \qquad \text{(Gauss's law)} \tag{1.31}$$

We call the statement in the box a *law* because it is equivalent to Coulomb's law and it could serve equally well as the basic law of electrostatic interactions, after charge and field have been defined. Gauss's law and Coulomb's law are not two independent physical laws, but the same law expressed in different ways.[10] In Gaussian units, the $1/\epsilon_0$ in Gauss's law is replaced with 4π.

Looking back over our proof, we see that it hinged on the inverse-square nature of the interaction and of course on the additivity of interactions, or superposition. Thus the theorem is applicable to any inverse-square field in physics, for instance to the gravitational field.

[10] There is one difference, inconsequential here, but relevant to our later study of the fields of moving charges. Gauss's law is obeyed by a wider class of fields than those represented by the electrostatic field. In particular, a field that is inverse-square in r but not spherically symmetrical can satisfy Gauss's law. In other words, Gauss's law alone does not imply the symmetry of the field of a point source which is implicit in Coulomb's law.

It is easy to see that Gauss's law would *not* hold if the law of force were, say, inverse-cube. For in that case the flux of electric field from a point charge q through a sphere of radius R centered on the charge would be

$$\Phi = \int \mathbf{E} \cdot d\mathbf{a} = \frac{q}{4\pi\epsilon_0 R^3} \cdot 4\pi R^2 = \frac{q}{\epsilon_0 R}. \tag{1.32}$$

By making the sphere large enough we could make the flux through it as small as we pleased, while the total charge inside remained constant.

This remarkable theorem extends our knowledge in two ways. First, it reveals a connection between the field and its sources that is the converse of Coulomb's law. Coulomb's law tells us how to derive the electric field if the charges are given; with Gauss's law we can determine how much charge is in any region if the field is known. Second, the mathematical relation here demonstrated is a powerful analytic tool; it can make complicated problems easy, as we shall see in the following examples. In Sections 1.11–1.13 we use Gauss's law to calculate the electric field due to various nicely shaped objects. In all of these examples the symmetry of the object will play a critical role.

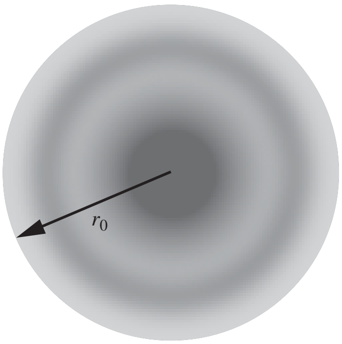


Figure 1.20.
A charge distribution with spherical symmetry.

1.11 Field of a spherical charge distribution

We can use Gauss's law to find the electric field of a spherically symmetrical distribution of charge, that is, a distribution in which the charge density ρ depends only on the radius from a central point. Figure 1.20 depicts a cross section through some such distribution. Here the charge density is high at the center, and is zero beyond r_0. What is the electric field at some point such as P_1 outside the distribution, or P_2 inside it (Fig. 1.21)? If we could proceed only from Coulomb's law, we should have to carry out an integration that would sum the electric field vectors at P_1 arising from each elementary volume in the charge distribution. Let's try a different approach that exploits both the symmetry of the system and Gauss's law.

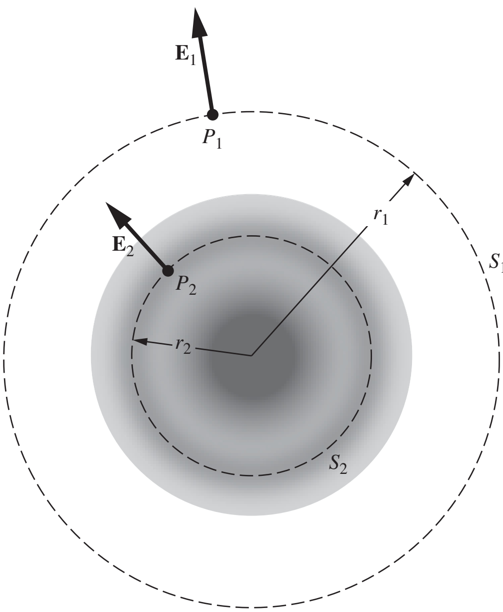


Figure 1.21.
The electric field of a spherical charge distribution.

Because of the spherical symmetry, the electric field at any point must be radially directed – no other direction is unique. Likewise, the field magnitude E must be the same at all points on a spherical surface S_1 of radius r_1, for all such points are equivalent. Call this field magnitude E_1. The flux through this surface S_1 is therefore simply $4\pi r_1^2 E_1$, and by Gauss's law this must be equal to $1/\epsilon_0$ times the charge enclosed by the surface. That is, $4\pi r_1^2 E_1 = (1/\epsilon_0) \cdot (\text{charge inside } S_1)$ or

$$E_1 = \frac{\text{charge inside } S_1}{4\pi\epsilon_0 r_1^2}. \tag{1.33}$$

Comparing this with the field of a point charge, we see that *the field at all points on S_1 is the same as if all the charge within S_1 were concentrated at the center.* The same statement applies to a sphere drawn *inside* the charge distribution. The field at any point on S_2 is the same as if all charge within S_2 were at the center, and all charge *outside* S_2 absent. Evidently the field inside a "hollow" spherical charge distribution is zero (Fig. 1.22). Problem 1.17 gives an alternative derivation of this fact.

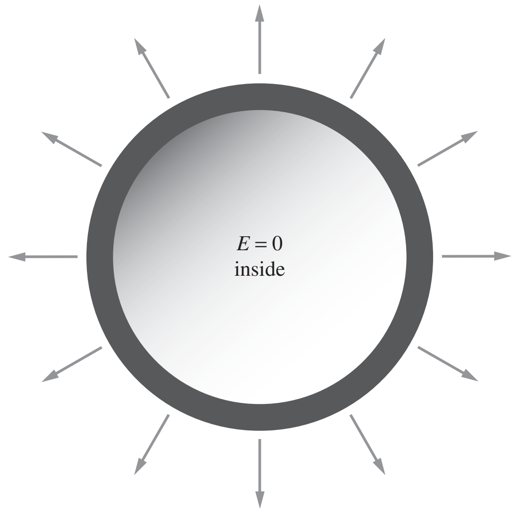


Figure 1.22.
The field is zero inside a spherical shell of charge.

Example (Field inside and outside a uniform sphere) A spherical charge distribution has a density ρ that is constant from $r = 0$ out to $r = R$ and is zero beyond. What is the electric field for all values of r, both less than and greater than R?

Solution For $r \geq R$, the field is the same as if all of the charge were concentrated at the center of the sphere. Since the volume of the sphere is $4\pi R^3/3$, the field is therefore radial and has magnitude

$$E(r) = \frac{(4\pi R^3/3)\rho}{4\pi\epsilon_0 r^2} = \frac{\rho R^3}{3\epsilon_0 r^2} \qquad (r \geq R). \tag{1.34}$$

For $r \leq R$, the charge outside radius r effectively contributes nothing to the field, while the charge inside radius r acts as if it were concentrated at the center. The volume inside radius r is $4\pi r^3/3$, so the field inside the given sphere is radial and has magnitude

$$E(r) = \frac{(4\pi r^3/3)\rho}{4\pi\epsilon_0 r^2} = \frac{\rho r}{3\epsilon_0} \qquad (r \leq R). \tag{1.35}$$

In terms of the total charge $Q = (4\pi R^3/3)\rho$, this can be written as $Qr/4\pi\epsilon_0 R^3$. The field increases linearly with r inside the sphere; the r^3 growth of the effective charge outweighs the $1/r^2$ effect from the increasing distance. And the field decreases like $1/r^2$ outside the sphere. A plot of $E(r)$ is shown in Fig. 1.23. Note that $E(r)$ is continuous at $r = R$, where it takes on the value $\rho R/3\epsilon_0$. As we will see in Section 1.13, field discontinuities are created by surface charge densities, and there are no surface charges in this system. The field goes to zero at the center, so it is continuous there also. How should the density vary with r so that the magnitude $E(r)$ is uniform inside the sphere? That is the subject of Exercise 1.68.

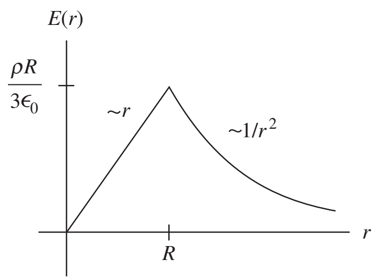


Figure 1.23.
The electric field due to a uniform sphere of charge.

The same argument applied to the gravitational field would tell us that the earth, assuming it is spherically symmetrical in its mass distribution, attracts outside bodies as if its mass were concentrated at the center. That is a rather familiar statement. Anyone who is inclined to think the principle expresses an obvious property of the center of mass must be reminded that the theorem is not even true, in general, for other shapes. A perfect cube of uniform density does *not* attract external bodies as if its mass were concentrated at its geometrical center.

Newton didn't consider the theorem obvious. He needed it as the keystone of his demonstration that the moon in its orbit around the earth and a falling body on the earth are responding to similar forces. The delay of nearly 20 years in the publication of Newton's theory of gravitation

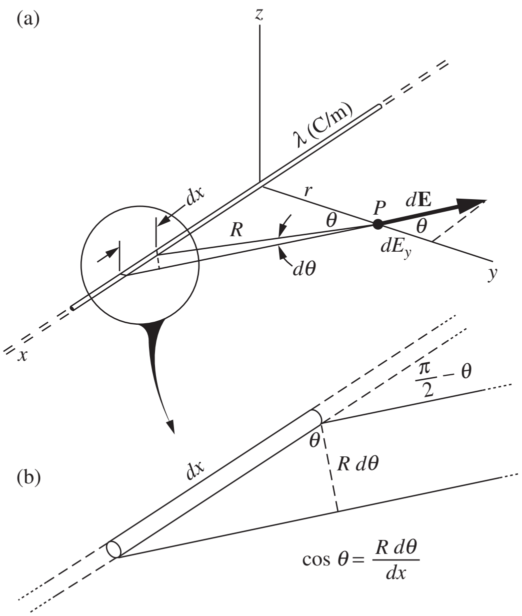


Figure 1.24.
(a) The field at P is the vector sum of contributions from each element of the line charge. (b) Detail of (a).

was apparently due, in part at least, to the trouble he had in proving this theorem to his satisfaction. The proof he eventually devised and published in the *Principia* in 1686 (Book I, Section XII, Theorem XXXI) is a marvel of ingenuity in which, roughly speaking, a tricky volume integration is effected without the aid of the integral calculus as we know it. The proof is a good bit longer than our whole preceding discussion of Gauss's law, and more intricately reasoned. You see, with all his mathematical resourcefulness and originality, Newton lacked Gauss's law – a relation that, once it has been shown to us, seems so obvious as to be almost trivial.

1.12 Field of a line charge

A long, straight, charged wire, if we neglect its thickness, can be characterized by the amount of charge it carries per unit length. Let λ, measured in coulombs/meter, denote this *linear charge density*. What is the electric field of such a line charge, assumed infinitely long and with constant linear charge density λ? We'll do the problem in two ways, first by an integration starting from Coulomb's law, and then by using Gauss's law.

To evaluate the field at the point P, shown in Fig. 1.24, we must add up the contributions from all segments of the line charge, one of which is indicated as a segment of length dx. The charge dq on this element is given by $dq = \lambda\, dx$. Having oriented our x axis along the line charge, we may as well let the y axis pass through P, which is a distance r from the nearest point on the line. It is a good idea to take advantage of symmetry at the outset. Obviously the electric field at P must point in the y direction, so that E_x and E_z are both zero. The contribution of the charge dq to the y component of the electric field at P is

$$dE_y = \frac{dq}{4\pi\epsilon_0 R^2}\cos\theta = \frac{\lambda\, dx}{4\pi\epsilon_0 R^2}\cos\theta, \tag{1.36}$$

where θ is the angle the electric field of dq makes with the y direction. The total y component is then

$$E_y = \int dE_y = \int_{-\infty}^{\infty} \frac{\lambda\cos\theta}{4\pi\epsilon_0 R^2}\, dx. \tag{1.37}$$

It is convenient to use θ as the variable of integration. Since Figs. 1.24(a) and (b) tell us that $R = r/\cos\theta$ and $dx = R\, d\theta/\cos\theta$, we have $dx = r\, d\theta/\cos^2\theta$. (This expression for dx comes up often. It also follows from $x = r\tan\theta \Longrightarrow dx = r\, d(\tan\theta) = r\, d\theta/\cos^2\theta$.) Eliminating dx and R from the integral in Eq. (1.37), in favor of θ, we obtain

$$E_y = \int_{-\pi/2}^{\pi/2} \frac{\lambda\cos\theta\, d\theta}{4\pi\epsilon_0 r} = \frac{\lambda}{4\pi\epsilon_0 r}\int_{-\pi/2}^{\pi/2}\cos\theta\, d\theta = \frac{\lambda}{2\pi\epsilon_0 r}. \tag{1.38}$$

We see that the field of an infinitely long, uniformly dense line charge is proportional to the reciprocal of the distance from the line. Its direction is of course radially outward if the line carries a positive charge, inward if negative.

Gauss's law leads directly to the same result. Surround a segment of the line charge with a closed circular cylinder of length L and radius r, as in Fig. 1.25, and consider the flux through this surface. As we have already noted, symmetry guarantees that the field is radial, so the flux through the ends of the "tin can" is zero. The flux through the cylindrical surface is simply the area, $2\pi rL$, times E_r, the field at the surface. On the other hand, the charge enclosed by the surface is just λL, so Gauss's law gives us $(2\pi rL)E_r = \lambda L/\epsilon_0$ or

$$E_r = \frac{\lambda}{2\pi\epsilon_0 r}, \tag{1.39}$$

in agreement with Eq. (1.38).

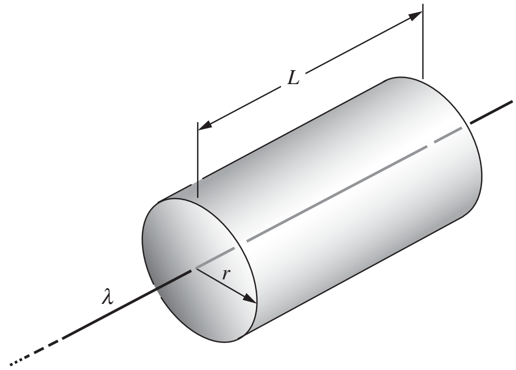


Figure 1.25.
Using Gauss's law to find the field of a line charge.

1.13 Field of an infinite flat sheet of charge

Electric charge distributed smoothly in a thin sheet is called a *surface charge distribution.* Consider a flat sheet, infinite in extent, with the constant surface charge density σ. The electric field on either side of the sheet, whatever its magnitude may turn out to be, must surely point perpendicular to the plane of the sheet; there is no other unique direction in the system. Also, because of symmetry, the field must have the same magnitude and the opposite direction at two points P and P' equidistant from the sheet on opposite sides. With these facts established, Gauss's law gives us at once the field intensity, as follows: draw a cylinder, as in Fig. 1.26 (actually, any shape with uniform cross section will work fine), with P on one side and P' on the other, of cross-sectional area A. The outward flux is found only at the ends, so that if E_P denotes the magnitude of the field at P, and $E_{P'}$ the magnitude at P', the outward flux is $AE_P + AE_{P'} = 2AE_P$. The charge enclosed is σA, so Gauss's law gives $2AE_P = \sigma A/\epsilon_0$, or

$$E_P = \frac{\sigma}{2\epsilon_0}. \tag{1.40}$$

We see that the field strength is independent of r, the distance from the sheet. Equation (1.40) could have been derived more laboriously by calculating the vector sum of the contributions to the field at P from all the little elements of charge in the sheet.

In the more general case where there are other charges in the vicinity, the field need not be perpendicular to the sheet, or symmetric on either side of it. Consider a very squat Gaussian surface, with P and P' infinitesimally close to the sheet, instead of the elongated surface in Fig. 1.26. We can then ignore the negligible flux through the cylindrical "side" of the pillbox, so the above reasoning gives $E_{\perp,P} + E_{\perp,P'} = \sigma/\epsilon_0$, where the "$\perp$" denotes the component perpendicular to the sheet. If you want

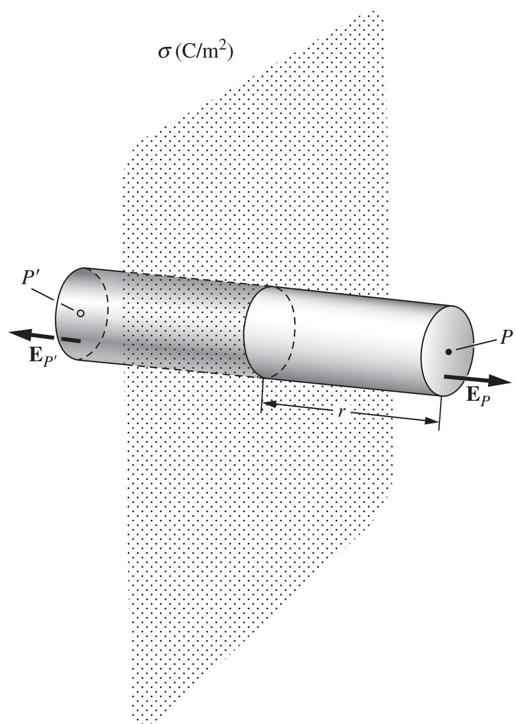


Figure 1.26.
Using Gauss's law to find the field of an infinite flat sheet of charge.

to write this in terms of vectors, it becomes $\mathbf{E}_{\perp,P} - \mathbf{E}_{\perp,P'} = (\sigma/\epsilon_0)\hat{\mathbf{n}}$, where $\hat{\mathbf{n}}$ is the unit vector perpendicular to the sheet, in the direction of P. In other words, the discontinuity in $\mathbf{E}_\perp$ across the sheet is given by

$$\Delta\mathbf{E}_\perp = \frac{\sigma}{\epsilon_0}\hat{\mathbf{n}}. \tag{1.41}$$

Only the normal component is discontinuous; the parallel component is continuous across the sheet. So we can just as well replace the $\Delta\mathbf{E}_\perp$ in Eq. (1.41) with $\Delta\mathbf{E}$. This result is also valid for any finite-sized sheet, because from up close the sheet looks essentially like an infinite plane, at least as far as the normal component is concerned.

The field of an infinitely long line charge, we found, varies inversely as the distance from the line, while the field of an infinite sheet has the same strength at all distances. These are simple consequences of the fact that the field of a point charge varies as the inverse square of the distance. If that doesn't yet seem compellingly obvious, look at it this way: roughly speaking, the part of the line charge that is mainly responsible for the field at P in Fig. 1.24 is the near part – the charge within a distance of order of magnitude r. If we lump all this together and forget the rest, we have a concentrated charge of magnitude $q \approx \lambda r$, which ought to produce a field proportional to q/r^2, or λ/r. In the case of the sheet, the amount of charge that is "effective," in this sense, increases proportionally to r^2 as we go out from the sheet, which just offsets the $1/r^2$ decrease in the field from any given element of charge.

1.14 The force on a layer of charge

The sphere in Fig. 1.27 has a charge distributed over its surface with the uniform density σ, in $\mathrm{C/m^2}$. Inside the sphere, as we have already learned, the electric field of such a charge distribution is zero. Outside the sphere the field is $Q/4\pi\epsilon_0 r^2$, where Q is the total charge on the sphere, equal to $4\pi r_0^2\sigma$. So just outside the surface of the sphere the field strength is

$$E_{\text{just outside}} = \frac{\sigma}{\epsilon_0}. \tag{1.42}$$

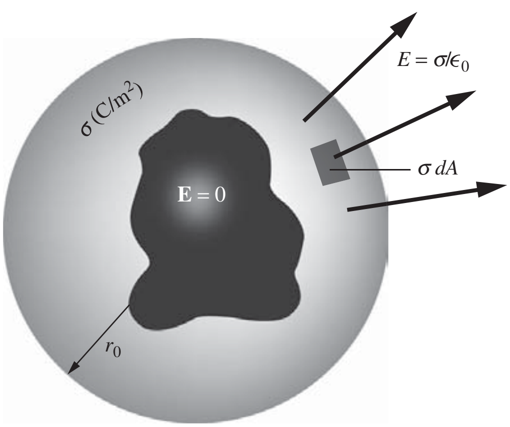


Figure 1.27.
A spherical surface with uniform charge density σ.

Compare this with Eq. (1.40) and Fig. 1.26. In both cases Gauss's law is obeyed: the *change* in the normal component of $\mathbf{E}$, from one side of the layer to the other, is equal to σ/ϵ_0, in accordance with Eq. (1.41).

What is the electrical force experienced by the charges that make up this distribution? The question may seem puzzling at first because the field $\mathbf{E}$ arises from these very charges. What we must think about is the force on some small element of charge dq, such as a small patch of area dA with charge $dq = \sigma\, dA$. Consider, separately, the force on dq due to all the other charges in the distribution, and the force on the patch due to the charges within the patch itself. This latter force is surely zero. Coulomb repulsion between charges within the patch is just another example of