Chapter 2

Electricity and Magnetism

2.1 Coulomb's Law, Electric Fields

Starting from this section, we dive into another important branch of physics—**electromagnetism** (EM)—which, roughly speaking, studies interactions between particles with electric charges via electromagnetic fields. It is the second fundamental interaction studied in this course and is the dominant force in the interactions of electrons, atoms, and molecules. The modern science of electromagnetism was developed by scientists in many countries. One of the most important was Michael Faraday, a gifted experimenter with a talent for physical intuition and visualization. In the mid-nineteenth century, James Clerk Maxwell put Faraday's ideas into mathematical form, introduced many new ideas of his own, and put electromagnetism on a sound theoretical basis. We will present Maxwell's unified theory of electromagnetism in Section 2.7, in the form of the famous *Maxwell's equations*.

In this section, we study the simplest form of electromagnetism, called *electrostatics*, which is a branch of EM theory that studies electric charges at rest, with a main focus on electric fields due to static electrically charged objects. (However, we will also consider moving charged particles in electric fields—it is just the electric field that is "static".)

2.1.1 Electric charges

There are two types of electric charge, named by the American scientist Benjamin Franklin as *positive charge* and *negative charge*. He could have called them anything, but using algebraic signs as names comes in handy when we add up charges to find the net charge. In most everyday objects, there are about equal numbers of negatively charged particles and positively charged particles, and so the net charge is zero, in which case the object is said to be *electrically neutral*. But sometimes, the positive and negative charges are unbalanced, then the net charge of the object is called *excess charge*.

Generally, we denote the charge of a point particle or a charged object by q. q is positive (resp. negative) if the excess charge of the object is positive (resp. negative). The SI unit of charges is C (coulomb). For practical reasons having to do with the accuracy of measurements, the coulomb unit is derived from the SI unit A (ampere) for electric current i. We shall discuss current in detail in Section 2.3. For now, we only note that current i is the rate at which charge moves past a point:

$$i = \frac{\mathrm{d}q}{\mathrm{d}t}.$$

This gives that $\mathrm{A} = \mathrm{C/s}$, or $\mathrm{C} = \mathrm{A} \cdot \mathrm{s}$.

Charge is quantized. Charges of objects are all due to charged particles in atoms, which consist of positively charged protons, negatively charged electrons, and electrically neutral neutrons. The charge of a single electron and that of a single proton have the same magnitude but are opposite in sign. Hence, an electrically neutral atom contains equal numbers of electrons and protons. Electrons are held near the nucleus because they have the electrical sign opposite that of the protons in the nucleus and thus are attracted to the nucleus. The measured value of the charge of a proton is approximately

$$e = 1.602 \times 10^{-19}\ \mathrm{C},$$

and the charge of an electron is given by $-e$. Thus, any positive or negative charge q of an object can be written as

$$q = ne, \quad n \in \mathbb{Z},$$

where n is equal to the number of protons minus the number of electrons in the object. This gives a very important feature of charges that is very different from mass:

Charges are not continuous; they are **quantized** *and appear as integer multiples of e.*

For this reason, people call e as the *elementary charge*. The 2019 redefinition of the SI base units fixed the numerical value of e as $1.602176634 \times 10^{-19}$ when expressed in coulombs, i.e, the ampere (A) is defined as the electrical current equivalent to 10^{19} elementary charges moving every 1.602176634 seconds. But for historical reasons, people still use A instead of C as a base unit.

Quarks, the constituent particles of protons and neutrons, have charges of $\pm e/3$ or $\pm 2e/3$, but they cannot be detected individually. For this and for historical reasons, people do not take their charges to be the elementary charge. The elementary charge e is one of the important constants of nature, like the speed of light c. In modern physics, it is believed that the quantization of electric charges is related to some topological properties of the gauge structure of the EM theory.

Charge is conserved. An important feature of electric charge is that it is always *conserved*, i.e., the net charge of a closed system of bodies is unchanged. In other words, in all processes, a positive or negative charge is not created but only transferred from one body to another, changing the net charge of each body. This hypothesis of conservation of charge, first put forward by Benjamin Franklin, has stood up under close examination, both for large-scale charged bodies and for atoms, nuclei, and elementary particles. No exceptions have ever been found. A simple example of charge conservation occurs when an electron e^- and its antiparticle, the positron e^+, undergo an annihilation process, transforming into two gamma rays:

$$\mathrm{e}^- + \mathrm{e}^+ \to \gamma + \gamma.$$

Here, e^- has charge $-e$ and e^+ has charge e, while γ is neutral. The converse of annihilation also occurs: in pair production, a gamma ray transforms into an electron and a positron:

$$\gamma \to \mathrm{e}^- + \mathrm{e}^+.$$

So far, we have seen four conservation laws: energy, momentum, angular momentum, and electric charge. Noether's theorem, a central result in theoretical physics, asserts that each conservation law is associated with a symmetry of the underlying physics. The conservation of energy is associated with time translation symmetry. The conservation of momentum and the conservation of angular momentum are associated with (space) translation symmetry and rotation symmetry, respectively. The symmetry that is associated with charge conservation is the *global gauge invariance of the electromagnetic field.*

2.1.2 Coulomb's law

With experiments, scientists discovered that: particles with the same sign of electrical charge repel each other, and particles with opposite signs attract each other. Quantitatively, the strength of the attraction or repulsion is described by the famous **Coulomb's law** of electrostatic force (or electric force) between charged particles.

Physics law 9 (Coulomb's law)**.** *Consider two charged particles, where particle 1 has charge q_1 and locates at $\vec{r}_1$ and particle 2 has charge q_2 and locates at $\vec{r}_2$. Then, the electric force acting on particle 2 due to particle 1 is given by*

$$\vec{F}_{12} = \frac{kq_1q_2}{r^2}\hat{r}, \tag{2.1.1}$$

where $r = |\vec{r}_2 - \vec{r}_1|$ is the separation between the particles, $\hat{r} = (\vec{r}_2 - \vec{r}_1)/r$ is the unit vector along the direction pointing from particle 1 to particle 2, and k is a positive constant called the Coulomb constant *and has value*

$$k = 8.99 \times 10^9\ \mathrm{N \cdot m^2/C^2}.$$

Remark. The Coulomb constant is often written as

$$k = \frac{1}{4\pi\varepsilon_0},$$

where ε_0 is called the *permittivity constant* and has value

$$\varepsilon_0 = 8.85 \times 10^{-12}\ \mathrm{C^2/(N \cdot m^2)}.$$

In some sense, it is a more fundamental quantity than k, and is one of the two physical constants that will appear in Maxwell's equations (another one is the Vacuum permeability μ_0).

Let's first check the direction of the force on particle 2 as given by (2.1.1). If q_1 and q_2 have the same sign (i.e., they are both positive or both negative), $q_1 q_2$ is positive and (2.1.1) tells us that the force on particle 2 is in the direction of $\hat{r}$, pointing from particle 1 to particle 2. That is, particle 2 is being repelled from particle 1. Conversely, if q_1 and q_2 have opposite signs, then (2.1.1) tells us that the force on particle 2 is in the direction $-\hat{r}$. That is, particle 2 is being attracted toward particle 1.

You may notice something that is very curious. Although the two types of forces are wildly different, the form of Coulomb's law is the same as that of Newton's law of gravitation (1.6.1): they both satisfy the inverse square law (i.e., the $1/r^2$ dependence) that involve a product of a property of the interacting particles—the charge in one case and the mass in the other. The main difference between these two laws is that gravitational forces are always attractive but electrostatic forces may be either attractive or repulsive, depending on the signs of the charges. This difference arises from the fact that there is only one type of mass but two types of charge.

Due to the similarity with Newton's law, some results developed in Section 1.6 also apply to Coulomb's law. Let's list them here for your convenience.

1. Superposition of forces. As with all forces in this book, the electrostatic force obeys the principle of superposition. In general, consider an object with charge density at $\vec{r} = (x, y, z)$ given by $\rho(\vec{r})$. Then, similar to (1.6.2), the total electric force on a particle of charge q at position $\vec{r}'$ is given by a triple integral

$$\iiint_D kq\rho(\vec{r})\frac{\vec{r}' - \vec{r}}{|\vec{r}' - \vec{r}|^3}\mathrm{d}x\mathrm{d}y\mathrm{d}z, \quad \vec{r} = (x, y, z), \tag{2.1.2}$$

where $D \subset \mathbb{R}^3$ denotes the domain occupied by an object.

2. Shell theories. Analogous to the shell theories (Proposition 1.6.2 and Proposition 1.6.3) for the gravitational force, we have two shell theories for the electrostatic force.

Proposition 2.1.1. *(1) A charged particle outside a shell with a radially symmetrical charge density is attracted or repelled as if the shell's charge were concentrated as a particle at its center.*

(2) A charged particle inside a shell with a radially symmetrical charge density has no net force acting on it due to the shell.

3. Electric potential energy. Similar to Proposition 1.6.4, the electrostatic force is a conservative force and is associated with some potential energy.

Proposition 2.1.2 (Electric potential energy)**.** *Coulomb (electrostatic) force is conservative. Furthermore, the potential energy of a system of two particles with charges q_1 and q_2 is*

$$V = \frac{kq_1q_2}{r} + C, \tag{2.1.3}$$

where r is the distance between the two particles. In particular, we can choose $C = 0$ if the reference potential energy at ∞ is taken to be 0.

If our system contains more than two particles, we consider each pair of particles in turn, calculate the gravitational potential energy of that pair as if the other particles were not there, and then algebraically sum the results. In general, the potential energy between an object occupying $D \subset \mathbb{R}^3$ and a point particle of mass q at position $\vec{r}'$ is calculated as an integral:

$$V = \iiint_D \frac{kq\rho(\vec{r})}{|\vec{r} - \vec{r}'|}\mathrm{d}x\mathrm{d}y\mathrm{d}z, \quad \vec{r} = (x, y, z). \tag{2.1.4}$$

Finally, following Section 1.4, we have the celebrated energy conservation for an isolated system interacting only through Coulomb forces: *the mechanical energy of the system, i.e., the kinetic energy plus the electric potential energy, does not change.* The electric potential energy is associated with a so-called *electric potential* of the charged object, which will be discussed in detail in Section 2.2.

4. Gauss's law. Similar to gravitation, the electrostatic force also satisfies Gauss's law as discussed in Section 1.6.5. We will discuss Gauss's law and its consequence in detail in Section 2.2.

2.1.3 Electric fields

Consider two positively charged particles. We know that an electrostatic force acts on particle 2 due to the presence of particle 1. We can also calculate the force direction and the force magnitude using Coulomb's law. There is another convenient (and actually deeper) way of looking at the interaction between the two particles: particle 1 sets up an *electric field* at all points in the surrounding space, even if the space is a vacuum. If we place particle 2 at any point in that space, it is affected by the electric field particle 1 has already set up at that point.

In modern physics, fields are more fundamental concepts than forces. In particular, for EM theory, the electric and magnetic fields are the two basic subjects of study. An electric field is a *vector field* that consists of a distribution of electric field vectors $\vec{E}$, one for each point in the space around a charged object. In other words, an electric field is a vector-valued function of the space points, whose value $\vec{E}(\vec{r})$ at a point $\vec{r}$, is defined as follows. At point $\vec{r}$, we place a particle with a small positive charge q_0, called a test charge. (We can think of the charge to be small so that it does not disturb the object's charge distribution.) We then measure the electrostatic force $\vec{F}$ that acts on the test charge. The electric field at that point is then defined as

$$\vec{E}(\vec{r}) = \frac{\vec{F}}{q_0}. \tag{2.1.5}$$

From (2.1.5), we see that the SI unit for the electric field is N/C. Since the test charge is positive, the two vectors in (2.1.5) are in the same direction. We can shift the test charge around to various other points, to measure the electric fields there, so that we can figure out the distribution of the electric field set up by the charged object. Note that *the electric field exists independent of the test charge.* It is something that a charged object sets up in the surrounding space (even the vacuum), independent of whether we happen to come along to measure it. (Be sure to distinguish between force and field: force is a push or pull between two charged objects, while the electric field is an abstract property set up by one given charged object.)

By Coulomb's law (2.1.1), the electric field at $\vec{r}'$ due to a point charge q located at $\vec{r}$ is given by

$$\vec{E}(\vec{r}') = \frac{kq}{|\vec{r}' - \vec{r}|^2}\hat{r}, \quad \hat{r} = \frac{\vec{r}' - \vec{r}}{|\vec{r}' - \vec{r}|}. \tag{2.1.6}$$

Similar to forces, the electric fields also obey the principle of superposition. That is, if several electric fields are set up at a given point by several charged particles, we can find the net field by adding them as vectors. Hence, to calculate the net electric field at a given point due to several particles, find the electric field due to each particle and then sum the fields as vectors. The electric field set up by a general charged object can be calculated as in (2.1.2). As a special example, the field due to a shell with a radially symmetrical charge density can be calculated easily with the help of Proposition 2.1.1. We will see more examples in Section 2.1.4 below.

The idea of electric fields was introduced by Michael Faraday, who also introduced a useful way to visualize an electric field in space. He envisioned lines, called *electric field lines*, in the space around any given charged particle or object The electric field lines are drawn according to the following rules: (1) at any point, the electric field vector must be tangent to the electric field line through that point and in the same direction; (2) in a plane perpendicular to the field lines, the relative density of the lines represents the relative magnitude of the field there, with greater density for greater magnitude. Mathematically, it is not hard to show that such a family of electric field lines exists and is unique.

Figure 2.1 gives an example of electric field lines near a sphere uniformly covered with negative charges. At every point around the sphere, an electric field vector points radially inward toward the sphere, and we can represent this electric field with electric field lines as in Figure 2.1.

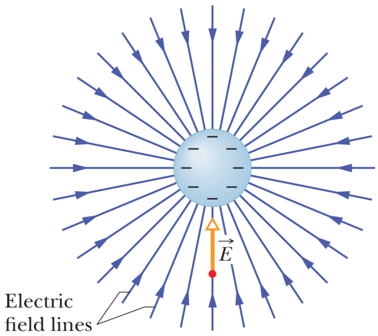


Figure 2.1: The electric field lines near a uniform sphere of negative charges.

If the sphere in Figure 2.1 were uniformly covered with positive charges, the electric field vectors at all points around it would be radially outward and thus so would the electric field lines. So, we observe the following rule:

Electric field lines extend away from positive charge (where they originate) and toward negative charge (where they terminate).

In Figure 2.1, the field lines originate on distant positive charges that are not shown.

Another feature of electric field lines is that *they never intersect each other* (think about why). For example, Figure 2.2 shows the field lines for two particles with equal positive charges.

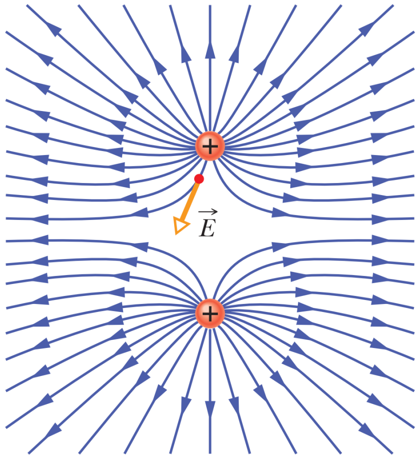


Figure 2.2: The electric field lines for two particles with equal positive charge.

We remark that a lot of different fields are used in science and engineering. The gravitational field mentioned in Section 1.6.5 is another important vector field. However, not all fields are vector fields. For example, a temperature field is a scalar field, which gives the distribution of temperatures at each point. In EM theory, the electric and magnetic fields are combined into an *electromagnetic tensor field*, and in Einstein's theory of General Relativity, the gravitational field is also a tensor field.

2.1.4 Electric fields due to charged objects

In this subsection, we examine some important examples of electric fields generated by several specific types of charged objects.

Example (Electric field due to an electric dipole)**.** *Consider a system of two particles that have the same charge magnitude q but opposite signs, a very common and important arrangement known as an* electric dipole. *Suppose the positive charge and negative charge are located at $\vec{d}/2$ and $-\vec{d}/2$, respectively. Calculate the electric field set up by this electric dipole.*

Solution: We evaluate the electric field at the point $\vec{r}$. Using (2.1.6), we have

$$\vec{E}(\vec{r}) = \frac{kq}{|\vec{r}-\vec{d}/2|^3}(\vec{r}-\vec{d}/2) - \frac{kq}{|\vec{r}+\vec{d}/2|^3}(\vec{r}+\vec{d}/2).$$

In applications, we are often interested in the electrical field of a dipole at distances that are large compared with the dimensions of the dipole, that is, at distances such that $r \gg d$. At such large distances, we can expand the above equation as

$$\begin{aligned}
\vec{E}(\vec{r}) &= kq\vec{r}\left(\frac{1}{|\vec{r}-\vec{d}/2|^3}-\frac{1}{|\vec{r}+\vec{d}/2|^3}\right)-kq\frac{\vec{d}}{2}\left(\frac{1}{|\vec{r}-\vec{d}/2|^3}+\frac{1}{|\vec{r}+\vec{d}/2|^3}\right)\\
&= kq\vec{r}\frac{|\vec{r}+\vec{d}/2|^3-|\vec{r}-\vec{d}/2|^3}{|\vec{r}-\vec{d}/2|^3|\vec{r}+\vec{d}/2|^3}-\frac{kq\vec{d}}{r^3}+O\left(\frac{kqd^2}{r^4}\right)\\
&= \frac{kqr^3\vec{r}}{|\vec{r}-\vec{d}/2|^3|\vec{r}+\vec{d}/2|^3}\left[\left(1+\frac{\vec{r}\cdot\vec{d}}{r^2}+\frac{d^2}{4r^2}\right)^{3/2}-\left(1-\frac{\vec{r}\cdot\vec{d}}{r^2}+\frac{d^2}{4r^2}\right)^{3/2}\right]-\frac{kq\vec{d}}{r^3}+O\left(\frac{kqd^2}{r^4}\right)\\
&= \frac{kq\vec{r}}{r^3}\frac{3\vec{r}\cdot\vec{d}}{r^2}-\frac{kq\vec{d}}{r^3}+O\left(\frac{kqd^2}{r^4}\right)=k\frac{3(\vec{p}\cdot\hat{r})\hat{r}-\vec{p}}{r^3}+O\left(\frac{kqd^2}{r^4}\right),
\end{aligned}$$

where $O(\cdot)$ denotes a *vector* with length of order at most $\frac{kqd^2}{r^4}$ and $\hat{r}=\vec{r}/r$ is the unit vector along $\vec{r}$ direction. The vector $\vec{p}=q\vec{d}$, which involves the two intrinsic properties q and $\vec{d}$ of the dipole, is an important vector quantity known as the *electric dipole moment* of the dipole. The unit of the dipole moment is $\mathrm{C}\cdot\mathrm{m}$. The magnitude and direction of $\vec{p}$ indicate the strength and orientation of a dipole, respectively. In an idealistic situation, we take the dimension of the dipole $d\to 0$, while keeping the dipole moment $\vec{p}$ unchanged (by increasing the charge $q=p/d\to\infty$). This limiting process results in a "point dipole", whose electric field is given by

$$\vec{E}(\vec{r})=k\frac{3(\vec{p}\cdot\hat{r})\hat{r}-\vec{p}}{r^3}.$$

An important feature of this field is that it decays with respect to r as $1/r^3$ in contrast to the $1/r^2$ decay in Coulomb's law, i.e., the electric field of a dipole decays faster than that of a point charge.

In Figure 2.3, we show the pattern of electric field lines for an electric dipole.

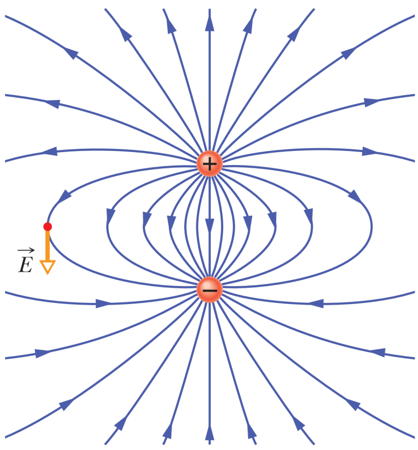
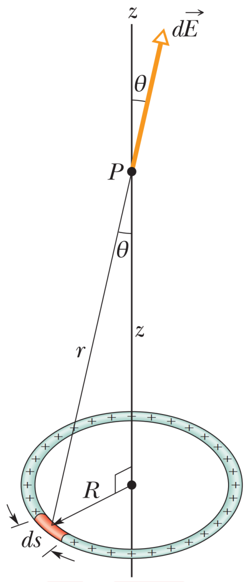


Figure 2.3: The pattern of electric field lines around an electric dipole.

Example (Electric field due to a circle)**.** *Consider a circle of radius R and with uniform linear* *charge density λ. Calculate the electric field on the central axis, i.e., the axis through the ring's center and perpendicular to the plane of the ring.*

Solution: Suppose the central axis is along z direction and the circle is on the xy-plane. At an arbitrary point $\vec{r} = (0, 0, z)$ on the axis, by the symmetry of the problem, we know that the electric field is along z-direction. We now calculate its magnitude.



Consider a different element ds on the circle. The magnitude of the electric field generated by this element is equal to

$$\frac{k\lambda}{z^2 + R^2}\mathrm{d}s.$$

Its component along the z direction is equal to

$$\frac{k\lambda}{z^2 + R^2}\cos\theta \mathrm{d}s = \frac{k\lambda z}{(z^2 + R^2)^{3/2}}\mathrm{d}s.$$

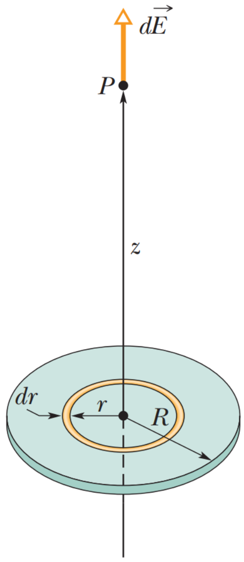
Summing over all these differential elements over ds, we get that the total electric field is along the z direction and has a magnitude

$$E(\vec{r}) = 2\pi R \frac{k\lambda z}{(z^2 + R^2)^{3/2}} = \frac{kqz}{(z^2 + R^2)^{3/2}},$$

where $q = 2\pi R\lambda$ is the total charge of the circle.

Example (Electric field due to a disk). *Consider a circular disk of radius R and with uniform surface charge density σ. Calculate the electric field on the central axis, i.e., the axis through the disk's center and perpendicular to the plane of the disk.*

Solution: Suppose the central axis is along z direction and the disk is on the xy-plane. At an arbitrary point $\vec{r} = (0, 0, z)$ on the axis, by the symmetry of the problem, we know that the electric field is along z-direction. We now calculate its magnitude.



By the previous example, we know that the thin ring with radial width dr sets up a differential electric field

$$\mathrm{d}E = \frac{kz(2\pi r\sigma \mathrm{d}r)}{(z^2 + r^2)^{3/2}} = \frac{zr\sigma}{2\varepsilon_0(z^2 + r^2)^{3/2}}\mathrm{d}r,$$

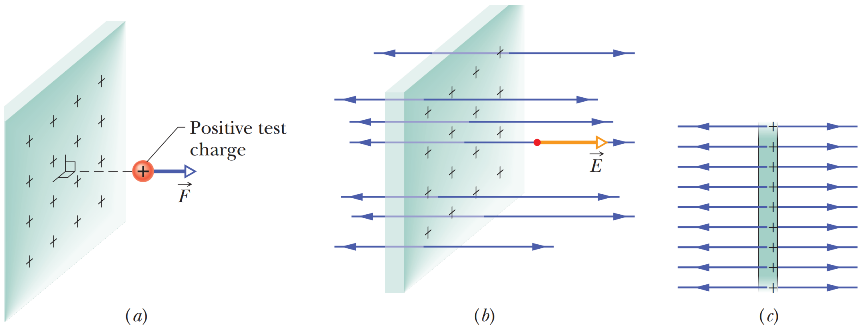
where $2\pi r\sigma dr$ is the total charge of the ring. Summing over all such rings, we get the total electric field as

$$E(\vec{r}) = \int_0^R \frac{zr\sigma}{2\varepsilon_0(z^2 + r^2)^{3/2}}\mathrm{d}r = -\frac{\sigma}{2\varepsilon_0}\frac{z}{\sqrt{z^2 + r^2}}\Bigg|_0^R = \frac{\sigma}{2\varepsilon_0}\left(1 - \frac{z}{\sqrt{z^2 + R^2}}\right).$$

If we let $R \to \infty$ while keeping z finite, the second term in the parentheses on the RHS approaches zero, and the above equation reduces to

$$E = \frac{\sigma}{2\varepsilon_0}.$$

For an infinite plane, any axis perpendicular to it can be regarded as a central axis. Hence, it sets up a uniform electric field around it, and the electric field lines are as follows:



If we insert another infinite charged plane that is parallel to the previous plane and has uniform surface charge density $-\sigma$, then by superposition of electric fields, the electric field between the two planes is uniform and equal to

$$E = \frac{\sigma}{\varepsilon_0},$$

while the electric field vanishes elsewhere. This is also how people generate uniform electric fields in the real world: a uniform electric field is produced by placing a potential (or voltage) difference across two parallel metal plates. As long as the separation between the two plates is small compared to the scale of the plates, then the electric field between them is approximately uniform at the points far away from the edges of the plates. (We will discuss the concept of potential and voltage in Section 2.2.)

2.1.5 Charged particles in electric fields

Given an electric field, it is simple to determine the electrostatic force acting on a particle of charge q:

$$\vec{F} = q\vec{E}, \tag{2.1.7}$$

where $\vec{E}$ is the electric field that *other charges* have produced at the location of the particle. The field is not the field set up by the particle itself; to distinguish the two fields, the field acting on the particle is often called the *external field*. A charged particle or object is not affected by its own electric field. Equation (2.1.7) played a key role in the measurement of the elementary charge e in the famous Millikan oil-drop experiment. You can read this wikipedia page about some interesting stories about this experiment.

Example. *Consider an electric dipole with electric dipole moment $\vec{p}$ in a uniform electric field $\vec{E}$. Calculate the net force and torque acting on the dipole, and find its potential energy.*

Solution: Suppose the electric dipole consists of two particles with charges q and $-q$ located at $\vec{r}_1$ and $\vec{r}_2$, respectively. Its electric dipole moment is equal to

$$\vec{p} = q(\vec{r}_1 - \vec{r}_2).$$

The total force on the dipole is

$$\vec{F} = q\vec{E} - q\vec{E} = 0.$$

The total force on the dipole is

$$\vec{\tau} = \vec{r}_1 \times q\vec{E} - \vec{r}_2 \times q\vec{E} = q(\vec{r}_1 - \vec{r}_2) \times \vec{E} = \vec{p} \times \vec{E}.$$

We claim that the potential energy of a charged particle in a uniform electric field $\vec{E}$ is equal to

$$U(\vec{r}) = -q\vec{E} \cdot \vec{r} + C,$$

where q is the charge of the particle and C is a constant depending on our choice of the zero reference potential energy. (We can take $C = 0$ if we let $U(0) = 0$.) To see this, we find that

$$\nabla_{\vec{r}} U(\vec{r}) = q\vec{E},$$

which is indeed the electric force on the charge. Thus, the potential energy of the dipole is

$$U = -q\vec{E} \cdot \vec{r}_1 + q\vec{E} \cdot \vec{r}_2 = -q(\vec{r}_1 - \vec{r}_2) \cdot \vec{E} = -\vec{p} \cdot \vec{E}.$$