PHYS 51: Electromagnetism and Optics

Connor Neely, Fall 2023

* Adapted from David J. Griffiths, *Introduction to Electrodynamics* (2013) and FA23 lectures.

1 Electrostatics

1.1 The Electrostatic Force

In mechanics, all of the quantities we worked with could be described using some combination of mass, length, and time. Starting now, though, we'll discuss a new kind of unit: charge, whose SI unit is the Coulomb (C). It is conserved in closed systems, relativistically invariant, and quantized. (The quantum is the charge q_e carried by a proton or electron, often called the elementary charge.)

Charge can be either positive or negative, and it's well know that charges with like charges repel while those with opposite charges attract. In particular, for two point charges q, q_0, the force that q exerts on q_0 is given by

$$\mathbf{F}_E(r) = \frac{1}{4\pi\epsilon_0}\frac{qq_0}{r^2}\hat{\mathbf{r}},$$

where $\mathbf{r}$ runs from q to q_0 and ϵ_0 is called the permittivity of free space. We call this relationship the electrostatic force, also known as Coulomb's law. Notice how it takes the exact same form as Newton's law of gravitation, simply with the added caveat that the particles may either attract or repel.

When we talk about gravity, the idea of a "gravitational field" often comes up to loosely describe how one massive body interacts with others around it. We can do the same with point charges—if a charge q exerts a force $\mathbf{F}_E$ on q_0, then the electric field due to q is given by

$$\mathbf{E}(r) = \frac{1}{q_0}\mathbf{F}_E(r) = \frac{1}{4\pi\epsilon_0}\frac{q}{r^2}\hat{\mathbf{r}}.$$

Notice how $\mathbf{E}$ is independent of our choice of q_0, depending only on the charge q and how far away it is. Also, electric fields obey the principle of superposition—in a space with several point charges, their net electric field is found by simply summing the individual charges' fields.

Example: Electric field due to a dipole

Electric dipoles frequently arise in scenarios involving charged objects. Consider two particles with equal and opposite charges q spaced a distance d apart; if we look at a point a distance x from the charges' midpoint, then by superposition we have

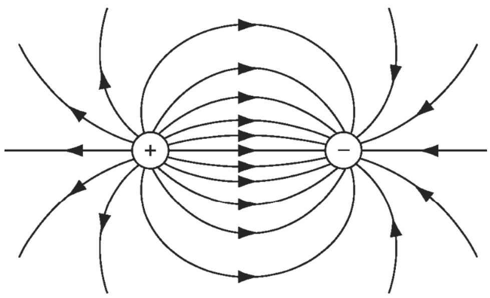
$$E_{\text{dip}} = 2 \cdot \frac{1}{4\pi\epsilon_0}\frac{q}{\mathcal{R}^2}\sin\theta,$$

where $\mathcal{R}$ is the distance from x to either of the charges and θ is the angle that the x-charge vector makes with the vertical. By geometry this gives

$$\mathbf{E}_{\text{dip}} = \frac{1}{4\pi\epsilon_0}\frac{qd}{(r^2 + (d/2)^2)^{3/2}}(-\hat{z}),$$

with $\hat{z}$ pointing from the charge midpoint toward $+q$. In general we define the dipole moment $p \equiv qd$ to describe the separation of two opposite charges. (Note that $\mathbf{E}_{\text{dip}} \propto r^{-3}$ for $r \gg d$.)

There's a nice way to visualize these fields using electric field lines. An example for the case of a dipole is provided at right. There are a few important features to note: the lines point in the direction of the electric field, never intersect, and are only drawn in the plane of the page. Also, the lines start and end at positive and negative charges, respectively, unless they go off to infinity. Finally, the strength of the field in a given region is indicated by the density of lines there.



All of the basic principles here apply not only to collections of point charges, but also to continuous charge distributions. In order to find the net electric field due to such a distribution, we can chop it up into a bunch of tiny pieces which can be treated as point charges. Then by the principle of superposition,

$$\mathbf{E} = \int_V d\mathbf{E} = \int_V \frac{dq}{4\pi\epsilon_0 r^2}\hat{\mathbf{r}},$$

where V is the volume inhabited by the charge distribution. Each dq is a function of the charge density at that point:

in 1D $dq = \lambda\, ds$, in 2D $dq = \sigma\, dA$, and in 3D $dq = \rho\, dV$,

where λ, σ, ρ are the charge densities in their respective dimensions. Still, in general the integral is very difficult no matter the dimension, and we usually rely on a high degree of symmetry to solve problems. There are three main considerations here.

- Different distributions lend themselves particularly well to different coordinate systems. For an infinite line of charge we may choose to work in cylindrical coordinates, while for a spherically symmetric distribution we may instead use spherical coordinates.
- For infinite charge distributions, it's useful to see if symmetry allows for the cancellation of any field components. In the case of an infinitely long line of charge, we'd notice that $\mathbf{E}$ must only have a radial component and that, consequentially, the half-infinite lines of charge on either side of the field point have the same contribution to the net electric field.
- It's also useful to break up distributions in clever ways, in hopes of reducing the problem to a one-dimensional integral. For example, we might treat a charged sphere as a series of thin, concentric shells, each carrying a charge $\rho\, dV$, and integrate with respect to the distribution's radius.

Given how similar gravity is to the electrostatic force, it should be no surprise that Newton's shell theorem also applies to spherically symmetric charge distributions. All of the charge "interior" to the field point can be treated as one big point charge, while all of that "exterior" to the field point cancels and can thus be ignored.

1.2 Gauss's Law

Calculating the electric field via direct integration can be quite clunky. Fortunately, one of the fundamental equations of electromagnetism provides a much cleaner alternative. Gauss's law states that the electric flux through a closed surface S is proportional to the charge enclosed by the surface. Symbolically,

$$\oiint_S \mathbf{E} \cdot d\mathbf{A} = \frac{q_{\text{enc}}}{\epsilon_0},$$

where $d\mathbf{A}$ is an area element pointing along the outward normal and q_{enc} is the net charge enclosed by S. Although this statement technically holds true for any charge distribution, it's really only practical in a select few kinds of scenarios. The idea is to construct a fictitious "Gaussian surface" that respects the symmetry of the electric field: for planar symmetry we may use a Gaussian box, for cylindrical symmetry a Gaussian cylinder, and for spherical symmetry a Gaussian sphere.

Example: Electric field due to a plane of charge

Consider an infinite plane of charge with charge density σ. The corresponding electric field exhibits planar symmetry—all field lines point in the $\hat{z}$ (upward) direction, and the magnitude depends only on z—so upon an arbitrary region of the plane we construct a Gaussian box with some side length s.

To evaluate the surface integral, we take advantage of the fact that $\mathbf{E} = E\,\hat{z}$ is orthogonal to the "caps" of the box and parallel to the lateral faces:

$$\oiint_{\text{box}} \mathbf{E} \cdot d\mathbf{A} = \iint_{\text{caps}} \mathbf{E} \cdot d\mathbf{A} + \cancel{\iint_{\text{lats}} \mathbf{E} \cdot d\mathbf{A}} = 2\iint_{\text{top}} E(|z|)\,\hat{z} \cdot dA\,\hat{z} = 2E(|z|)\iint_{\text{top}} dA,$$

so the integral evaluates to $E(|z|) \cdot 2s^2$. Since the box encloses a charge $q_{\text{enc}} = \sigma s^2$, by Gauss's law we have $E(|z|) \cdot 2s^2 = \sigma s^2$ and thus $\mathbf{E}(|z|) = (\sigma/2\epsilon_0)(\pm\hat{z})$, taking $+\hat{z}$ for positive z.

The integral form of Gauss's law states that electric flux is proportional to enclosed charge. This is useful, but there's another, more compact differential formulation. First note that, by the divergence theorem, we have

$$\iiint_V (\nabla \cdot \mathbf{E})\, dV = \oint\!\!\!\oint_{\partial V} \mathbf{E} \cdot d\mathbf{A}$$

for an arbitrary volume V and its boundary ∂V. By the integral form of Gauss's law the right-hand side turns into $q_{\text{enc}}/\epsilon_0$ or, alternatively,

$$\iiint_V (\nabla \cdot \mathbf{E})\, dV = \iiint_V \frac{\rho}{\epsilon_0} dV.$$

Since this equivalence is true for any volume we could think of, these integrands must be equal and we get

$$\nabla \cdot \mathbf{E} = \frac{\rho}{\epsilon_0},$$

the differential form of Gauss's law. The "flux density" at a point is proportional to the charge density there.

1.3 Electrostatic Potential

Having discussed forces and fields, the next step we'll take in building a theory of electricity is to see how energy fits in. To exploit our understanding of the electrostatic force, we'll begin by computing the amount of work $\mathbf{F}_E$ does in moving a charge through an electric field. In particular, if q_0 is moved between points a and b while in the presence of another charge q, then in a spherical coordinate system centered at q we have

$$W_E = \int_a^b \mathbf{F}_E \cdot d\mathbf{s} = \int_a^b \frac{q_0 q}{4\pi\varepsilon_0 r^2} \hat{r} \cdot (dr\,\hat{r} + r\,d\theta\,\hat{\theta} + r\sin\theta\,d\phi\,\hat{\phi}) = -\left(\frac{q_0 q}{4\pi\varepsilon_0 b} - \frac{q_0 q}{4\pi\varepsilon_0 a}\right).$$

The electrostatic force is, like gravity, a conservative force and thus has an associated potential energy related to its work by $W_E = -\Delta U_E$. If we define $U_E(\infty) = 0$ like we often do for gravity we get

$$U_E(r) = \frac{q_0 q}{4\pi\epsilon_0 r}.$$

We might interpret this as the amount of work that must be done against $\mathbf{F}_E$ in order to bring q and q_0 into close proximity from infinite separation. Generalizing a bit, the total potential energy in a larger system of charges is the work required to construct the system by bringing each charge in from infinite separation, one by one. In a sense, the work is "stored" in the interactions between pairs of charges, and these interaction energies may be summed to determine the system's total electrostatic potential energy.

What we've essentially done here is turn a vector problem into a scalar one. Since $\mathbf{F}_E = -\nabla U_E$, in order to study the behavior of a particle under the influence of a net electrostatic force, we can simply look at the corresponding potential energy function and draw conclusions from there. This is often much more convenient than having to deal with vectors, and in fact we can do something similar with electric fields.

If an electric field represents an electric force per unit charge, then we can define a new quantity V, simply called potential (measured in volts V), to represent an electric potential energy per unit charge. Again assuming an infinite reference point, this means

$$V(r) = \frac{U_E(r)}{q_0} = \frac{q}{4\pi\epsilon_0 r},$$

for a point charge. Like all of the other quantities we've discussed so far, electrostatic potential obeys the principle of superposition. To find the potential due to a continuous distribution of charge we might again break the distribution of little chunks, approximate them as point charges, and integrate their individual contributions to the potential.

This also provides a new way to visualize electric fields. In two dimensions we might draw equipotential lines, along which the electrostatic potential is constant. The electric field, the negative gradient of this potential field, points normal to these lines in the direction of steepest descent.

Now, the difference in potential between two points a and b is unsurprisingly given by

$$-\Delta V = \int_a^b \mathbf{E} \cdot d\mathbf{s},$$

meaning we have the relationship $\mathbf{E} = -\nabla V$. Thus $\mathbf{E}$ is a conservative vector field and

$$\oint_C \mathbf{E} \cdot d\mathbf{s} = 0$$

for any closed C. This is a simplified version of Faraday's law (which we'll discuss in more detail later), and in differential form we can write $\nabla \times \mathbf{E} = \mathbf{0}$.

As a mathematical side note, we can also write $-\nabla \cdot (\nabla V) = \rho/\epsilon_0$. This is often expressed as

$$\nabla^2 V = -\rho/\epsilon_0$$

where ∇^2 is called the Laplacian operator, simply the sum of the second partial derivatives of V. We might loosely interpret this as the "net concavity" of V at a point—$\nabla^2 V$ is positive at points that are generally concave up, and negative at ones that are generally concave down. Since in free space we have $\nabla^2 V = 0$, V cannot have any extrema there. It also happens that, under any physically allowed boundary conditions, solutions to this equation are unique.

1.4 Conductors and Dielectrics

Now we'll take a look at how electric fields interact with two different kinds of matter: conductors and dielectrics. We'll begin with the simpler case of a conductor, which is a material in which each atom has at least one electron that's free to roam about. Arguably the most important feature of such a material is that its interior has no net electric field—if a conductor is placed into an external field $\mathbf{E}_0$, then its positive and negative charges will move around and create their own field $\mathbf{E}_{\text{ind}}$ in the opposite direction of $\mathbf{E}_0$. Eventually there will be an equilibrium in which $\mathbf{E}_{\text{ind}} = \mathbf{E}_0$.

There's a few other characteristics of conductors which often come in handy.

- If there is a cavity inside of a conductor, then it has no electric field. (Other there would be a C such that $\oint_C \mathbf{E} \cdot d\mathbf{s} \neq 0$.) Charge distributes itself over the cavity's surface to make this happen.
- Since there is no electric field in the interior of a conductor, by Gauss's law there is no net charge anywhere inside. Everything is distributed over the surface.
- The lack of electric field also means that all points inside of a conductor are at the same potential. To "ground" a conductor (i.e., set $V = 0$) we could connect it to a huge, faraway object, like the interior of the Earth.
- The electric field just outside of the conductor is orthogonal to the surface. Any tangential component would redistribute charge so as to cancel it out.

In other materials like dielectrics, electrons are generally bound to their molecules. However, under the influence of an external electric field, positive and negative charges can still be slightly displaced from one another. The precise mechanism by which this occurs depends on the kind of dielectric we're talking about: polar dielectrics have a bunch of permanent dipole moments that become aligned when influenced by an electric field, while nonpolar dielectrics have no such moments until being influenced. (In the latter case, electrons are simply displaced slightly from their normal position.)

Just like with conductors, the above rearrangement of charge induces an electric field that opposes the external field. In the case of linear dielectrics, the magnitude of this induced field $\mathbf{E}_{\text{ind}}$ is proportional to that of the applied field $\mathbf{E}_0$; for such materials we can define a dielectric constant

$$\kappa_E = \frac{E_0}{E} = \frac{E_0}{E_0 - E_{\text{ind}}},$$

which is roughly a measure of how conductive a material is. A perfect conductor would have $\kappa_E = \infty$. We can also see that electric fields inside of dielectrics are diminished by a factor of $1/\kappa$, which lines up with our intuition about conducting materials.

1.5 Capacitors

To cap off our discussion of electrostatics, we'll introduce a quantity that is fundamental to a variety of practical applications: capacitance. Suppose we have an isolated conductor with charge Q and a potential

V_0; then the capacitance C of the conductor, measured in farads (F), is the proportionality constant relating these quantities via $Q = CV_0$. More relevant to our purposes, however, is a definition for systems of two conductors with opposite charges $\pm Q$. Here we have $Q = C\Delta V$, where ΔV is the potential difference between the conductors. In either case, C represents the amount of charge that a system can carry per volt.

We'll focus on a special kind of capacitor comprised of two parallel plates, separated by a distance D, with equal and opposite charges $\pm Q$ distributed over an area A. For sufficiently large plates, we can exploit our Gauss's law result from earlier—we find that the plates' individual electric fields completely cancel outside of the capacitor and add to

$$E_{||} = \frac{\sigma}{\epsilon_0} = \frac{Q}{\epsilon_0 A}$$

on the inside. This turns out to be a great approximation for points that are not near the capacitor's fringe. Now, to find the capacitance of this arrangement we must first compute the potential difference between the two plates:

$$\Delta V = \left| \int_{\text{pos.}}^{\text{neg.}} \mathbf{E}_{||} \cdot d\mathbf{s} \right| = \left| \int_{\text{pos.}}^{\text{neg.}} E_{||}\, ds \right| = \frac{Q}{\epsilon_0 A} D.$$

This conveniently gives us a linear relationship between Q and ΔV, so we immediately get

$$Q = \left(\frac{\epsilon_0 A}{D} \right) \Delta V \implies C_{||} = \frac{\epsilon_0 A}{D},$$

a result which should make geometric sense in the limits for A and D.

Finally, we'll take a look at how much energy is stored in this parallel-plate capacitor. We could compute this directly using $\Delta U = Q\Delta V$, but we'll take a more enlightening approach. Suppose the plates are initially neutral; we can charge the capacitor by moving charges between the plates, one at a time, in total doing some amount of work

$$W = \int_0^Q (\Delta v)\, dq = \int_0^Q \left(\frac{q}{C} \right) dq = \frac{Q^2}{2C}.$$

Since the electric field is conservative we have $W = \Delta U$, meaning this is the total amount of energy stored in the capacitor! Some useful equivalent expressions are

$$U = \frac{Q^2}{2C} = \frac{C(\Delta V)^2}{2} = \frac{Q\Delta V}{2}.$$

We can go a step further. If we substitute $C = \epsilon_0 A/D$ and $\Delta V = E_{||} D$ into the latter expression, we get $U = (\epsilon_0 E_{||}^2 Ad)/2$. But Ad is simply the volume in which the electric field exists between the plates! So we can define a new quantity

$$u_E = \frac{U}{\text{vol.}} = \frac{1}{2}\epsilon_0 E^2,$$

called the "energy density" of an electric field. This suggests that we might think of the capacitor's energy as actually being stored in the field itself!

So we now have two ways of qualifying the energy stored in an electric field:

- the configuration energy, which describes the work done in building the electric field, and
- the energy density, which describes the amount of energy stored in the field per unit volume.

This also provides us with two ways to compute the energy stored in the electric field due to some charge distribution. We may assemble the configuration charge-by-charge, or we may simply integrate the energy density over all space.

2 Magnetostatics

2.1 Current

In our study of electrostatics, we assumed that all charges were basically stationary. In this static equilibrium, we have $E = 0$ and a lack of net charge inside of conductors. But in dynamic situations, charge may move more persistently!

Consider a wire with some charge flowing through it. The rate at which charge passes through a given cross section of the wire is called the current

$$i = \frac{dq}{dt}$$

at that point, measured in amperes (A). To account for scenarios in which the flow of charge is not uniform over the full cross section, we can define a current density $\mathbf{j}$ via

$$i = \iint \mathbf{j} \cdot d\mathbf{A},$$

where $\mathbf{j}$ points in the direction of positive charge flow. Of course we now know that it's actually negative charges (electrons) that're moving, but we stick to this convention because it isn't really hurting anyone.

Speaking of which, when we say that negative charge moves throughout a wire, we don't actually mean that electrons are rocketing through the wire at relativistic speeds, all in the same direction. In reality their motion is much more gas-like, bouncing around with high velocities that mostly cancel each other out. The component that does not cancel is called the drift velocity $\mathbf{v}_d$, and it tends to be quite slow, on the order of one millimeter per second. This is what creates a net movement of charge.

Now, in order to maintain this steady current throughout the wire, the electrons must be generally drifting at a constant speed. There are two things making this happen: there is a nonzero electric field in the wire pushing the electrons along, and there is some "drag" force keeping the electrons from accelerating. The forces due to these two cancel at the terminal velocity $\mathbf{v}_d$.

The drift velocity is related to the current density via the equation $\mathbf{j} = \rho\mathbf{v}_d$, where ρ is the charge density in the wire. If we make the simplest possible assumption that the drift velocity is proportional to the electric field, we get

$$\mathbf{j} = \sigma\mathbf{E}.$$

This is called Ohm's law. Specifically, this is its intensive form since it does not depend on any of the wire's geometric properties like length or radius. The quantity σ is called conductivity, and it quantifies how well a material can conduct current per unit electric field. (Its inverse σ^{-1} is called resistivity.)

This relationship is useful, but we can use it to derive something that might be more familiar. Suppose, now, that our wire has length L and cross-sectional area A. Since there is an electric field $\mathbf{E}$ in the wire, there is a potential difference ΔV between the two ends; assuming constant $\mathbf{E}$ and σ,

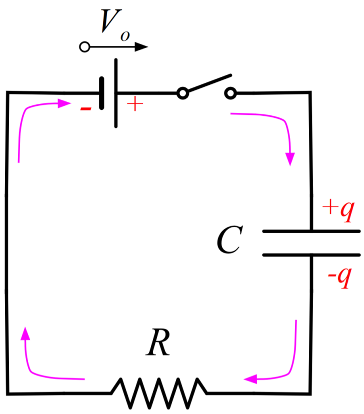
$$\Delta V = \left| \int_{x=0}^{x=L} \mathbf{E} \cdot d\mathbf{s} \right| = \left| \int_0^L \frac{j}{\sigma} dx \right| = \frac{jL}{\sigma}.$$

Since $j = i/A$, we have

$$\Delta V = i\left(\frac{L}{\sigma A}\right) = iR,$$

where we've defined the resistance R, measured in ohms (Ω). This is the extensive form of Ohm's law—note that R depends on the geometric properties of the wire. Devices that are created for the purpose of providing resistance are called resistors.

2.2 Circuits



Using wires, we can connect resistors, capacitors, and batteries (voltage sources) together to make current flow in ways that are practically useful to us. Schematics like the one at left are particularly useful abstractions when we want to work with circuits mathematically, like we will here.

Each circuit element is labeled with a corresponding quantity. The resistor has a resistance R, the capacitor a capacitance C, and the battery a voltage V_0. The battery dictates the direction of current flow—it pushes charge from low potential toward high potential, causing charge to move clockwise about the circuit. Right next to the battery is a switch; when closed current is allowed to flow, but when opened there is nowhere for the charge to go and the current ceases.

There are two key rules that we can exploit to "solve" a circuit.

- By conservation of energy, the net change in potential over any loop in the circuit must be zero. For these purposes we can ignore any drop in potential due to travel in a wire, since it will be negligibly small compared to the potential differences across the primary circuit elements.
- By conservation of charge, the total current flowing into any node (any place where two wires meet) must equal the current leaving it.

Using these two rules, we are usually able to construct a system of equations that, when solved, gives the current and potential difference across each circuit element present. We can also use them to find some convenient rules about combining like circuit elements together. For example, we could show that pairs of resistors in parallel or in series have respective equivalent resistances

$$R_{\text{par}} = \left(\frac{1}{R_1} + \frac{1}{R_2}\right)^{-1}, \qquad R_{\text{ser}} = R_1 + R_2.$$

We can be a bit more sophisticated, too, and examine the behavior of circuits over time. This is especially useful in seeing, for example, how the charge on a capacitor changes while charging or discharging.

Example: Charging an RC circuit

Consider the above circuit. By conservation of energy, when the switch is closed we have

$$V_0 = iR + \frac{q}{C}.$$

We can express this as a differential equation in $q(t)$, the charge on the capacitor:

$$\frac{dq}{dt} + \frac{1}{RC}q = \frac{V_0}{R}.$$

With the initial condition $q(0) = 0$, the solution is

$$q(t) = CV_0\left(1 - e^{-t/RC}\right).$$

As more charge gets deposited onto the positive plate of the capacitor, it becomes more difficult to push charge off of the other plate, around the circuit, and back onto the positive plate.

Finally, in order to keep the current flowing, the battery must do work on the charge to move them to higher potentials. This work is, of course, given by $W = qV_0$, corresponding to a power

$$P_{\text{battery}} = \frac{dW}{dt} = iV_0.$$

The resistor does an equal amount of work, so by Ohm's law we have

$$P_{\text{resistor}} = \frac{V_0^2}{R} = i^2R$$

dissipated as heat.

2.3 Magnetic Fields

Experimental evidence has shown that electric fields do not exist in isolation. There is another field, called the magnetic field $\mathbf{B}$ (measured in teslas T), which is generated by moving charge. If we were to "grab" the wire with our right hand, our thumb sticking in the direction of current, then our fingers would curl in the direction of the magnetic field. We might verify this by looking at how the wire deflects a compass needle.

Not only are magnetic fields generated by moving charge, they also only act upon moving charge. In particular, if a charge q moves with velocity $\mathbf{v}$ within a magnetic field $\mathbf{B}$, then the magnetic force on the charge is

$$\mathbf{F}_B = q\mathbf{v} \times \mathbf{B}.$$

There are a few things to note here. Most importantly, since $\mathbf{F}_B$ is always orthogonal to $\mathbf{v}$, the magnetic force can never do any work! This perpendicularity also allows $\mathbf{F}_B$ to act as a centripetal force when $\mathbf{v} \perp \mathbf{B}$, creating circular or helical trajectories. Finally, if there are both electric and magnetic fields present, we add their effects to get the Lorentz force $\mathbf{F} = q\mathbf{E} + q\mathbf{v} \times \mathbf{B}$.

Npw, since magnetic fields influence moving charge, they must also influence current-carrying wires. Suppose we have a segment of wire with current density $\mathbf{j}$ and electron drift velocity $\mathbf{v}_d$. If a constant magnetic field $\mathbf{B}$ acts on the wire, then the force on each electron is

$$\mathbf{F}_{B,e} = -q_e\mathbf{v} \times \mathbf{B} = -q_e\left(\frac{\mathbf{j}}{-q_e n}\right) \times \mathbf{B} = \frac{\mathbf{j} \times \mathbf{B}}{n},$$

where q_e is the elementary charge and n is the number density of electrons. If this segment of wire has N charges, then

$$\mathbf{F}_{B,L} = N\mathbf{F}_{B,e} = nAL\left(\frac{\mathbf{j} \times \mathbf{B}}{n}\right) = i\mathbf{L} \times \mathbf{B},$$

where L is the length of the segment, A is its cross-sectional area, and $\mathbf{L}$ points in the direction of current with magnitude L.

Now, everything we've done so far assumes that we already have $\mathbf{B}$ at all points in space. If we are instead only given a charge distribution, then we can find the magnetic field using the Biot-Savart law:

$$d\mathbf{B} = \frac{\mu_0}{4\pi}\frac{id\mathbf{l} \times \hat{r}}{r^2},$$

where i is the current in a small segment $d\mathbf{l}$ of wire and $\mathbf{r}$ points from the wire segment to the field point. μ_0 is called the permeability of free space. As we'll see now, applying the Biot-Savart law is very similar to calculating electric fields via direct integration.

Example: Magnetic field due to a current-carrying ring

Consider a ring of wire with radius R and current i, which flows counterclockwise when viewed from above. We'll determine the magnetic field at a distance z above the center of the ring.

By the right-hand rule, chunks of wire that are opposite each other on the ring create a net magnetic field that points only in the vertical direction. This means we can focus solely on the magnetic field's $\hat{z}$ component:

$$dB_z = |d\mathbf{B}|\cos\theta = |d\mathbf{B}|\frac{R}{\sqrt{R^2 + z^2}},$$

where θ is the angle of inclination when z is viewed from a point on the ring. (This is equivalent to the angle each $d\mathbf{B}$ makes with the vertical.) By the Biot-Savart law we have

$$|d\mathbf{B}| = \frac{\mu_0}{4\pi}\left|\frac{id\mathbf{l} \times \hat{r}}{r^2}\right| = \frac{\mu_0 i}{4\pi r^2}|d\mathbf{l} \times \hat{r}| = \frac{\mu_0 i}{4\pi r^2}dl.$$

Integrating dB_z to determine the total field:

$$\mathbf{B}_{\text{ring}} = \int_{\text{ring}} \frac{\mu_0 i\, dl}{4\pi(R^2 + z^2)}\frac{R}{(R^2 + z^2)^{1/2}}\hat{z} = \frac{\mu_0 iR}{4\pi(R^2 + z^2)^{3/2}}\hat{z}\int_{\text{ring}} dl = \frac{\mu_0 i\pi R^2}{2\pi(R^2 + z^2)^{3/2}}\hat{z}.$$

This current-carrying ring will actually become our model for a magnetic dipole. Just as we had the electric dipole moment qd, we have the magnetic dipole moment iA where A is the area enclosed by the ring. We can see that the magnetic field due to such a dipole obeys $B \propto iA/r^3$ for $r \gg R$, just like we saw with electric dipoles (in which case it was $E \propto qd/r^3$).

2.4 Ampere's Law

Just like with electric fields, this direct integration can get a little unwieldy. There's a cleaner way to determine magnetic fields in high-symmetry scenarios, but it'll be a bit different from Gauss's law. The main thing is that we have never detected any magnetic monopoles. This isn't a fundamental law of the universe or anything, it's just that every magnet we've encountered has had both a north pole and a south pole. Thus

$$\oint\!\!\!\oint_S \mathbf{B} \cdot d\mathbf{A} = 0, \qquad \nabla \cdot \mathbf{B} = 0.$$

Not very helpful for our purposes. But note that static magnetic fields are known for their twisty behavior about moving charges rather than the electric field's outward spray around point charges. So we can instead write down Ampere's law,

$$\oint_C \mathbf{B} \cdot d\mathbf{l} = \mu_0 i_{\text{enc}},$$

where i_{enc} is the net current puncturing an area enclosed by C. Once again, this statement holds for all scenarios involving magnetic fields, but it's only useful in a select few with the proper symmetries.

Example: Magnetic field due to a solenoid

Consider an infinite coil (or solenoid) of wire with radius R, current i, and n loops per unit length. We'll determine the magnetic field at some distance r from the central axis of the solenoid.

Let $\hat{z}$ represent the direction in which the magnetic field points, in accordance with the Biot-Savart law. The solenoid exhibits symmetry with respect to the cylindrical coordinates ϕ and z, so the magnetic field is only a function of r. Thus $\mathbf{B} = B(r)\,\hat{z}$.

Choose, as our Amperian loop, a length-r rectangle whose height h coincides with the solenoid's central axis. The loop is oriented such that it is parallel to $\hat{z}$ on the inside of the solenoid and antiparallel to $\hat{z}$ on the outside.

There are two regions to consider. For $r < R$, we have $i_{\text{enc}} = 0$. Also, since the lengths of the Amperian loop are orthogonal to the magnetic field, only the heights are relevant:

$$\oint_C \mathbf{B} \cdot d\mathbf{l} = \int_{\text{up}} B(0)\,\hat{z} \cdot dl\,\hat{z} + \int_{\text{down}} B(r)\,\hat{z} \cdot dl(-\hat{z})$$

Now, it can be shown by a direct integration of $\mathbf{B}_{\text{ring}}$ that the magnetic field at the center of an infinite solenoid is $\mathbf{B}(0) = \mu_0 ni$. Thus by Ampere's law

$$0 = \mu_0 nih - B(r)h \implies \mathbf{B}_{\text{in}}(r) = \mu_0 ni\,\hat{z}.$$

So the magnetic field inside of a solenoid is constant. For $r > R$ the setup is the same, only now we have $i_{\text{enc}} = nhi$. (i_{enc} is positive here due to how we've oriented our Amperian loop.) Thus by Ampere's law,

$$\mu_0 nih - B(r)h = \mu_0 nhi \implies \mathbf{B}_{\text{out}}(r) = \mathbf{0}.$$

2.5 Frames of Reference

At this point we have developed a complete theory of statics. Everything we've done so far can be derived from Maxwell's time-independent equations in their integral and differential forms:

$$\oiint \mathbf{E} \cdot d\mathbf{A} = \frac{q}{\epsilon_0} \qquad \nabla \cdot \mathbf{E} = \frac{\rho}{\epsilon_0}$$

$$\oint \mathbf{E} \cdot d\mathbf{l} = 0 \qquad \nabla \times \mathbf{E} = \mathbf{0}$$

$$\oiint \mathbf{B} \cdot d\mathbf{A} = 0 \qquad \nabla \cdot \mathbf{B} = 0$$

$$\oint \mathbf{B} \cdot d\mathbf{l} = \mu_0 \qquad \nabla \times \mathbf{B} = \mu_0 \mathbf{j}$$

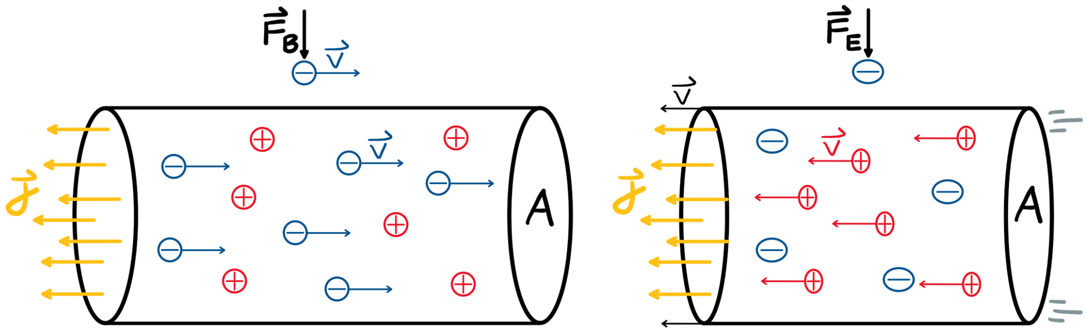
There is more to the theory of electromagnetism—we'll soon examine what happens when $\mathbf{E}$ and $\mathbf{B}$ are allowed to vary with time—but first we'll take a closer look at how these two fields are related, and why we so often see them lumped together as a unified electromagnetic field.

Consider the scenario illustrated below, on the left. In a neutral wire with cross-sectional area A electrons are free to move with drift velocity $\mathbf{v}$, generating a current density $\mathbf{j}$ in the opposite direction. We could use Ampere's law to show that the magnetic field due to this wire is $B = \mu_0 jA/2\pi r$, with direction determined by the right-hand rule.

Outside of the wire there is an electron that also has velocity $\mathbf{v}$. The force on this electron is purely magnetic in nature; specifically, taking the standard x-y coordinate system,

$$\mathbf{F}_B = -q_e \mathbf{v} \times \mathbf{B} = q_e v \left(\frac{\mu_0 j A}{2\pi r} \right) (-\hat{\mathbf{y}}) = \frac{q_e v^2 \mu_0 \rho_- A}{2\pi r} (-\hat{\mathbf{y}}),$$

where we've defined the density of negative charge ρ_-, a positive quantity, for later convenience. Let us call this scenario the S frame.



Now suppose we, the observer, started walking with velocity $\mathbf{v}$. We now observe the electrons to be (in net) at rest, and it is the positive charges that are actually moving. Reason dictates that the isolated electron should still experience the same downward force, but this force cannot be magnetic since it's no longer in motion. We must therefore uncover some sort of hidden electric force in this new S' frame.

Special relativity, specifically length contraction, provides an answer! When compared to the S frame, in the S' frame the positive charges are contracted while the electrons are uncontracted. So while in the unprimed frame we had $\rho_{\text{net}} = \rho_+ - \rho_- = 0$, in the primed frame we see

$$\rho'_{\text{net}} = \rho'_+ - \rho'_- = \gamma\rho_+ - \gamma^{-1}\rho_- = \gamma \frac{v^2}{c^2}\rho_-,$$

where γ is the Lorentz factor $\left(1 - v^2/c^2\right)^{-1/2}$. So the lone electron will at least experience an attractive force, which is promising. To find its magnitude we appeal to Gauss's law: a Gaussian cylinder with radius r' and length l' encloses a charge $\rho' A' l'/\epsilon_0$ which generates a flux $E'(2\pi r l')$, meaning

$$E' = \frac{\rho' A'}{2\pi\epsilon_0 r'} = \gamma \frac{v^2 \rho_- A}{2\pi\epsilon_0 c^2 r} \implies \mathbf{F}'_E = \gamma \frac{q_e v^2 \rho_- A}{2\pi\epsilon_0 c^2 r}(-\hat{\mathbf{y}}).$$

Now we can compare the two forces. We could use time dilation to argue that $\mathbf{F}_E = \gamma^{-1}\mathbf{F}'_E$, so we have

$$\mathbf{F}_B = \mu_0 \frac{q_e v^2 \rho_- A}{2\pi r}(-\hat{\mathbf{y}}), \qquad \mathbf{F}_E = \frac{1}{\epsilon_0 c^2}\frac{q_e v^2 \rho_- A}{2\pi r}(-\hat{\mathbf{y}}).$$

We can see that these expressions are equivalent if $\epsilon_0 \mu_0 = 1/c^2$ and, remarkably, experimental evidence suggests that this is actually the case! Thus a force that appeared as purely magnetic in the S frame is actually purely electric in the S' frame.

So it doesn't really make sense to speak about electric and magnetic fields as two separate entities. They are, in fact, two facets of a more fundamental electromagnetic field, related to each other via a Lorentz transformation:

$$\begin{aligned}
\mathbf{E}'_{||} &= \mathbf{E}_{||} \\
\mathbf{E}'_{\perp} &= \gamma\left(\mathbf{E}_{\perp} + \mathbf{v} \times \mathbf{B}\right) \\
\mathbf{B}'_{||} &= \mathbf{B}_{||} \\
\mathbf{B}'_{\perp} &= \gamma\left(\mathbf{B}_{\perp} - \frac{1}{c^2}\mathbf{v} \times \mathbf{E}\right)
\end{aligned}$$

Visually, at relativistic frame velocities the electric and magnetic fields get compressed along the direction of motion, similar to how lengths are contracted at these speeds.