

Part II Electrodynamics and Optics

William Royce

July 29, 2024

Part II Physics, The University of Cambridge

Preface

There are, to the best of our knowledge, four forces at play in the Universe. At the very largest scales – those of planets or stars or galaxies – the force of gravity dominates. At the very smallest distances, the two nuclear forces hold sway. For everything in between, it is force of electromagnetism that rules.

At the atomic scale, electromagnetism (admittedly in conjunction with some basic quantum effects) governs the interactions between atoms and molecules. It is the force that underlies the periodic table of elements, giving rise to all of chemistry and, through this, much of biology. It is the force which binds atoms together into solids and liquids. And it is the force which is responsible for the incredible range of properties that different materials exhibit.

At the macroscopic scale, electromagnetism manifests itself in the familiar phenomena that give the force its name. In the case of electricity, this means everything from rubbing a balloon on your head and sticking it on the wall, through to the fact that you can plug any appliance into the wall and be pretty confident that it will work. For magnetism, this means everything from the shopping list stuck to your fridge door, through to trains in Japan which levitate above the rail. Harnessing these powers through the invention of the electric dynamo and motor has transformed the planet and our lives on it.

As if this wasn't enough, there is much more to the force of electromagnetism for it is, quite literally, responsible for everything you've ever seen. It is the force that gives rise to light itself.

Rather remarkably, a full description of the force of electromagnetism is contained in four simple and elegant equations. These are known as the *Maxwell equations*. There are few places in physics, or indeed in any other subject, where such a richly diverse set of phenomena flows from so little. The purpose of this course is to introduce the Maxwell equations and to extract some of the many stories they contain.

However, there is also a second theme that runs through this course. The force of electromagnetism turns out to be a blueprint for all the other forces. There are various mathematical symmetries and structures lurking within the Maxwell equations, structures which Nature then repeats in other contexts. Understanding the mathematical beauty of the equations will allow us to see some of the principles that underly the laws of physics, laying the groundwork for future study of the other forces.

Contents

1	Introduction	1
1.1	Charge and Current	1
1.1.1	The Conservation Law	2
1.2	Forces and Fields	3
1.2.1	The Maxwell Equations	4
2	Electrostatics	7
2.1	Gauss' Law	7
2.1.1	The Coulomb Force	8
2.1.2	A Uniform Sphere	10
2.1.3	Line Charges	11
2.1.4	Surface Charges and Discontinuities	12
2.2	The Electrostatic Potential	15
2.2.1	The Point Charge	16
2.2.2	The Dipole	18
2.2.3	General Charge Distributions	19
2.2.4	Field Lines	21
2.2.5	Electrostatic Equilibrium	21
2.3	Electrostatic Energy	23
2.3.1	The Energy of a Point Particle	25
2.3.2	The Force Between Electric Dipoles	26
2.4	Conductors	27
2.4.1	Capacitors	29
2.4.2	Boundary Value Problems	30
2.4.3	Method of Images	32
2.4.4	Many Many More Problems	34
3	Magnetostatics	37
3.1	Ampère's Law	37
3.1.1	A Long Straight Wire	38
3.1.2	Surface Currents and Discontinuities	39
3.2	The Vector Potential	41
3.2.1	Magnetic Monopoles	42
3.2.2	Gauge Transformations	43
3.2.3	Biot-Savart Law	44
3.2.4	A Mathematical Diversion: The Linking Number	47
3.3	Units of Electromagnetism	48
4	Electrodynamics	49
5	Electromagnetism and Relativity	51
5.0.1	Gauge Invariance and Relativity	51
6	Light Scattering	53
7	Special Relativity	55
8	Radiation and Relativistic Electrodynamics	57

A	Appendix	A.1
A.1	Green’s Functions	A.1
A.2	Laplace and Poisson’s equations	A.1
A.3	uniqueness	A.1

List of Tables

List of Figures

1.1	Current flux.	2
1.2	The wire.	2
2.1	Left: The flux through \mathcal{S} and \mathcal{S}' is the same. Right: The flux through \mathcal{S} vanishes	8
2.2	Electric field produced by a spherically symmetric charge distribution, centered at the origin, contained within some radius R	9
2.3	Electric field inside a spherically symmetric charge distribution.	10
2.4	Radial dependence of electric field strength inside and outside a spherically symmetric charge distribution of radius R	11
2.5	The electric field \mathbf{E} of an infinite line charge with a uniform linear charge density η . Considering a Gaussian surface in the form of a cylinder at radius r , the electric field has the same magnitude at every point on the surface \mathcal{S}_3 and is directed outward, parallel to $\hat{\mathbf{n}}_3$	12
2.6	For an infinite sheet of charge with charge density σ , the electric field \mathbf{E} will be perpendicular to the surface. Therefore only the ends of a cylindrical Gaussian surface will contribute to the electric flux. In this case a cylindrical Gaussian surface perpendicular to the charge sheet is used. The resulting field is half that of a conductor at equilibrium with this surface charge density.	13
2.7	A pair of infinite planes at $z = 0$ and $z = a$, carrying uniform surface charge density $\pm\sigma$	14
2.8	Left: The Gaussian surface for a plane slab. Right: The resulting electric field.	14
2.9	A spherical shell of radius R , centered at the origin, with uniform surface charge density σ	15
2.10	Electric field at \mathbf{r} with $ \mathbf{r} \gg \mathbf{r}' $	19
2.11	Field lines for positive and negative point charges	21
2.12	Equipotentials for positive and negative point charges	22
2.13	Field lines and equipotentials for the dipole (left) and for a pair of charges of the same sign (on the right).	22
2.14	A spherical conductor placed inside the field generated by two charged plates.	28
2.15	Parallel plate capacitor.	30
2.16	Concentric sphere capacitor.	30
2.17	A particle near a conducting plane looks like a dipole.	32
2.18	A charged particle near a conducting sphere looks like an unbalanced dipole.	33
2.19	A charged particle near a conducting sphere looks like an unbalanced dipole.	34
3.1	An infinite, straight wire carrying current I and pointing in the $\hat{\mathbf{z}}$ direction.	39
3.2	A flat plane at $z = 0$ with a surface current density \mathbf{K} in the $\hat{\mathbf{x}}$ -direction.	39
3.3	Caption	39
3.4	A flat plane at $z = 0$ with a surface current density \mathbf{K} in the $\hat{\mathbf{x}}$ -direction.	40
3.5	Magnetic field \mathbf{B} inside a solenoid.	41

3.6 An infinite, straight wire carrying current I and pointing in the $\hat{\mathbf{z}}$ direction 46

3.7 Curves with linking number $n = 0$, $n = 1$ and $n = 2$ 47

CHAPTER 1

Introduction

1.1 Charge and Current

Each particle in the Universe carries with it a number of properties. These determine how the particle interacts with each of the four forces. For the force of gravity, this property is mass. For the force of electromagnetism, the property is called *electric charge*.

For the purposes of this course, we can think of electric charge as a real number, $q \in \mathbb{R}$. Importantly, charge can be positive or negative. It can also be zero, in which case the particle is unaffected by the force of electromagnetism.

The SI unit of charge is the *Coulomb*, denoted by C . It is, like all SI units, a parochial measure, convenient for human activity rather than informed by the underlying laws of the physics. (We'll learn more about how the Coulomb is defined in Section 3.3). At a fundamental level, Nature provides us with a better unit of charge. This follows from the fact that charge is quantised: the charge of any particle is an integer multiple of the charge carried by the electron which we denoted as $-e$, with

$$e = 1.602176634^{-19}\text{C}. \quad (1.1)$$

A much more natural unit would be to simply count charge as $q = ne$ with $n \in \mathbb{Z}$. Then electrons have charge -1 while protons have charge $+1$ and neutrons have charge 0 . Nonetheless, in this course, we will bow to convention and stick with SI units.¹

One of the key goals of this course is to move beyond the dynamics of point particles and onto the dynamics of continuous objects known as fields. To aid in this, it's useful to consider the *charge density*,

$$\rho(\mathbf{x}, t) \quad (1.2)$$

defined as charge per unit volume. The total charge Q in a given region \mathcal{V} is simply $Q = \int_{\mathcal{V}} d^3x \rho(\mathbf{x}, t)$. In most situations, we will consider smooth charge densities, which can be thought of as arising from averaging over many point-like particles. But, on occasion, we will return to the idea of a single particle of charge q , moving on some trajectory $\mathbf{r}(t)$, by writing $\rho = q\delta(\mathbf{x} - \mathbf{r}(t))$ where the delta-function ensures that all the charge sits at a point.

More generally, we will need to describe the movement of charge from one place to another. This is captured by a quantity known as the *current density* $\mathbf{J}(x, t)$, defined as follows: for every surface \mathcal{S} , the integral

$$I = \int_{\mathcal{S}} \mathbf{J} \cdot d\mathbf{S} \quad (1.3)$$

¹(An aside: the charge of quarks is actually $q = -e/3$ and $q = 2e/3$. This doesn't change the spirit of the above discussion since we could just change the basic unit. But, apart from in extreme circumstances, quarks are confined inside protons and neutrons so we rarely have to worry about this).

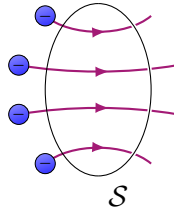


Fig. 1.1: Current flux.

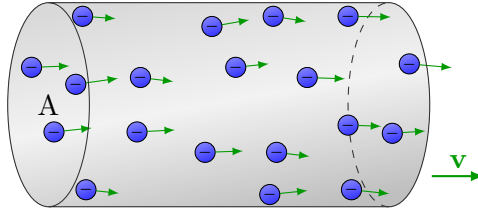


Fig. 1.2: The wire.

counts the charge per unit time passing through \mathcal{S} . (Here $d\mathbf{S}$ is the unit normal to \mathcal{S}). The quantity I is called the *current*. In this sense, the current density is the current-per-unit-area

The above is a rather indirect definition of the current density. To get a more intuitive picture, consider a continuous charge distribution in which the velocity of a small volume, at point \mathbf{x} , is given by $\mathbf{v}(\mathbf{x}, t)$. Then, neglecting relativistic effects, the current density is

$$\mathbf{J} = \rho \mathbf{v}. \quad (1.4)$$

In particular, if a single particle is moving with velocity $\mathbf{v} = \dot{\mathbf{r}}(t)$, the current density will be $\mathbf{J} = q\mathbf{v}\delta^3(\mathbf{x} - \mathbf{r}(t))$. This is illustrated in Fig. 1.1, where the underlying charged particles are shown as blue balls, moving through the surface \mathcal{S} .

As a simple example, consider electrons moving along a wire (see Fig. 1.2). We model the wire as a long cylinder of cross-sectional area A as shown below. The electrons move with velocity \mathbf{v} , parallel to the axis of the wire. (In reality, the electrons will have some distribution of speeds; we take \mathbf{v} to be their average velocity). If there are n electrons per unit volume, each with charge q , then the charge density is $\rho = nq$ and the current density is $\mathbf{J} = nq\mathbf{v}$. The current itself is $I = |\mathbf{J}|A$.

Throughout this course, the current density \mathbf{J} plays a much more prominent role than the current I . For this reason, we will often refer to \mathbf{J} simply as the “current” although we’ll be more careful with the terminology when there is any possibility for confusion.

1.1.1 The Conservation Law

The most important property of electric charge is that it’s conserved. This, of course, means that the total charge in a system can’t change. But it means much more than that because electric charge is conserved *locally*. An electric charge can’t just vanish from one part of the Universe and turn up somewhere else. It can only leave one point in space by moving to a neighbouring point.

The property of local conservation means that ρ can change in time only if there is a compensating current flowing into or out of that region. We express this in the *continuity equation*,

$$\frac{\partial \rho}{\partial t} + \nabla \cdot \mathbf{J} = 0. \quad (1.5)$$

This is an important equation. It arises in any situation where there is some quantity that is locally conserved.

To see why the continuity equation captures the right physics, it's best to consider the change in the total charge Q contained in some region \mathcal{V} .

$$\frac{dQ}{dt} = \int_{\mathcal{V}} d^3x \frac{\partial \rho}{\partial t} = - \int_{\mathcal{V}} d^3x \nabla \cdot \mathbf{J} = - \int_{\mathcal{S}} \mathbf{J} \cdot d\mathbf{S}. \quad (1.6)$$

From our previous discussion, $\int_{\mathcal{S}} \mathbf{J} \cdot d\mathbf{S}$ is the total current flowing out through the boundary \mathcal{S} of the region \mathcal{V} . (It is the total charge flowing *out*, rather than in, because $d\mathbf{S}$ is the outward normal to the region \mathcal{V}). The minus sign is there to ensure that if the net flow of current is outwards, then the total charge decreases.

If there is no current flowing out of the region, then $dQ/dt = 0$. This is the statement of (global) conservation of charge. In many applications we will take \mathcal{V} to be all of space, \mathbb{R}^3 , with both charges and currents localised in some compact region. This ensures that the total charge remains constant.

1.2 Forces and Fields

Any particle that carries electric charge experiences the force of electromagnetism. But the force does not act directly between particles. Instead, Nature chose to introduce intermediaries. These are *fields*.

In physics, a “field” is a dynamical quantity which takes a value at every point in space and time. To describe the force of electromagnetism, we need to introduce two fields, each of which is a three-dimensional vector. They are called the *electric field* \mathbf{E} and the *magnetic field* \mathbf{B} ,

$$\mathbf{E}(\mathbf{x}, t) \quad \text{and} \quad \mathbf{B}(\mathbf{x}, t). \quad (1.7)$$

When we talk about a “force” in modern physics, we really mean an intricate interplay between particles and fields. There are two aspects to this. First, the charged particles create both electric and magnetic fields. Second, the electric and magnetic fields guide the charged particles, telling them how to move. This motion, in turn, changes the fields that the particles create. We're left with a beautiful dance with the particles and fields as two partners, each dictating the moves of the other.

This dance between particles and fields provides a paradigm which all other forces in Nature follow. It feels like there should be a deep reason that Nature chose to introduce fields associated to all the forces. And, indeed, this approach does provide one overriding advantage: all interactions are local. Any object – whether particle or field – affects things only in its immediate neighbourhood. This influence can then propagate through the field to reach another point in space, but it does not do so instantaneously. It takes time for a

particle in one part of space to influence a particle elsewhere. This lack of instantaneous interaction allows us to introduce forces which are compatible with the theory of special relativity, something that we will explore in more detail in Section 5.

The purpose of this course is to provide a mathematical description of the interplay between particles and electromagnetic fields. In fact, you've already met one side of this dance: the position $\mathbf{r}(t)$ of a particle of charge q is dictated by the electric and magnetic fields through the Lorentz force law,

$$\mathbf{F} = q(\mathbf{E} + \dot{\mathbf{r}} \times \mathbf{B}). \quad (1.8)$$

The motion of the particle can then be determined through Newton's equation $\mathbf{F} = m\ddot{\mathbf{r}}$. Roughly speaking, an electric field accelerates a particle in the direction \mathbf{E} , while a magnetic field causes a particle to move in circles in the plane perpendicular to \mathbf{B} .

We can also write the Lorentz force law in terms of the charge distribution $\rho(\mathbf{x}, t)$ and the current density $\mathbf{J}(\mathbf{x}, t)$. Now we talk in terms of the *force density* $\mathbf{f}(\mathbf{x}, t)$, which is the force acting on a small volume at point \mathbf{x} . Now the Lorentz force law reads

$$\mathbf{f} = \rho\mathbf{E} + \mathbf{J} \times \mathbf{B}. \quad (1.9)$$

1.2.1 The Maxwell Equations

In this course, most of our attention will focus on the other side of the dance: the way in which electric and magnetic fields are created by charged particles. This is described by a set of four equations, known collectively as the *Maxwell equations*. They are:

$$\nabla \cdot \mathbf{E} = \frac{\rho}{\epsilon_0}, \quad (1.10)$$

$$\nabla \cdot \mathbf{B} = 0, \quad (1.11)$$

$$\nabla \times \mathbf{E} + \frac{\partial \mathbf{B}}{\partial t} = 0, \quad (1.12)$$

$$\nabla \times \mathbf{B} - \mu_0\epsilon_0 \frac{\partial \mathbf{E}}{\partial t} = \mu_0\mathbf{J}. \quad (1.13)$$

The equations involve two constants. The first is the *electric constant* (known also, in slightly old-fashioned terminology, as the *permittivity of free space*),

$$\epsilon_0 \approx 8.85 \times 10^{-12} \text{m}^{-3} \text{kg}^{-1} \text{s}^2 \text{C}^2. \quad (1.14)$$

It can be thought of as characterising the strength of the electric interactions. The other is the *magnetic constant* (or *permeability of free space*),

$$\begin{aligned} \mu_0 &= 4\pi \times 10^{-7} \text{m kg C}^{-2} \\ &\approx 1.25 \times 10^{-6} \text{m kg C}^{-2}. \end{aligned} \quad (1.15)$$

The presence of 4π in this formula isn't telling us anything deep about Nature, but simply reflects a rather outdated way in which this constant was first defined. (We will explain this in more detail in Section 3.3). Nonetheless, this can be thought of as characterising the strength of magnetic interactions (in units of Coulombs).

The Maxwell equations (1.10-1.13) will occupy us for the rest of the course. Rather than trying to understand all the equations at once, we'll proceed bit by bit, looking at situations where only some of the equations are important. By the end of the lectures, we will understand the physics captured by each of these equations and how they fit together.

However, equally importantly, we will also explore the mathematical structure of the Maxwell equations. At first glance, they look just like four random equations from vector calculus. Yet this couldn't be further from the truth. The Maxwell equations are special and, when viewed in the right way, are the essentially unique equations that can describe the force of electromagnetism. The full story of why these are the unique equations involves both quantum mechanics and relativity and will only be told in later courses. But we will start that journey here. The goal is that by the end of these lectures you will be convinced of the importance of the Maxwell equations on both experimental and aesthetic grounds.

CHAPTER 2

Electrostatics

In this section, we will be interested in electric charges at rest. This means that there exists a frame of reference in which there are no currents; only stationary charges. Of course, there will be forces between these charges but we will assume that the charges are pinned in place and cannot move. The question that we want to answer is: what is the electric field generated by these charges?

Since nothing moves, we are looking for time independent solutions to Maxwell's equations with $\mathbf{J} = \mathbf{0}$. This means that we can consistently set $\mathbf{B} = \mathbf{0}$ and we're left with two of Maxwell's equations to solve. They are

$$\nabla \cdot \mathbf{E} = \frac{\rho}{\epsilon_0}, \quad (2.1)$$

and

$$\nabla \times \mathbf{E} = 0. \quad (2.2)$$

If you fix the charge distribution ρ , Eqs. (2.1) and (2.2) have a unique solution. Our goal in this section is to find it.

2.1 Gauss' Law

Before we proceed, let's first present equation (2.1) in a slightly different form that will shed some light on its meaning. Consider some closed region $\mathcal{V} \subset \mathbb{R}^3$ of space. We'll denote the boundary of \mathcal{V} by $\mathcal{S} = \partial\mathcal{V}$. We now integrate both sides of (2.1) over \mathcal{V} . Since the left-hand side is a total derivative, we can use the divergence theorem to convert this to an integral over the surface \mathcal{S} . We have

$$\int_{\mathcal{V}} d^3x \nabla \cdot \mathbf{E} = \int_{\mathcal{S}} \mathbf{E} \cdot d\mathbf{S} = \frac{1}{\epsilon_0} \int_{\mathcal{V}} d^3x \rho. \quad (2.3)$$

The integral of the charge density over \mathcal{V} is simply the total charge contained in the region. We'll call it $Q = \int d^3x \rho$. Meanwhile, the integral of the electric field over \mathcal{S} is called the *flux* through \mathcal{S} . We learn that the two are related by

$$\int_{\mathcal{S}} \mathbf{E} \cdot d\mathbf{S} = \frac{Q}{\epsilon_0} \quad (2.4)$$

This is *Gauss's law*. However, because the two are entirely equivalent, we also refer to the original (2.1) as Gauss's law.

Notice that it doesn't matter what shape the surface \mathcal{S} takes. As long as it surrounds a total charge Q , the flux through the surface will always be Q/ϵ_0 . This is shown, for example, in Fig 2.1. A fancy way of saying this is that the integral of the flux doesn't depend on the geometry of the surface, but does depend on its topology since it must

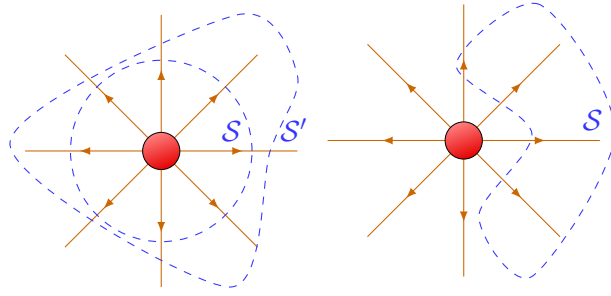


Fig. 2.1: Left: The flux through \mathcal{S} and \mathcal{S}' is the same. Right: The flux through \mathcal{S} vanishes

surround the charge Q . The choice of \mathcal{S} is called the *Gaussian surface*; often there's a smart choice that makes a particular problem simple.

Only charges that lie inside \mathcal{V} contribute to the flux. Any charges that lie outside will produce an electric field that penetrates through \mathcal{S} at some point, giving negative flux, but leaves through the other side of \mathcal{S} , depositing positive flux. The total contribution from these charges that lie outside of \mathcal{V} is zero, as illustrated in the right-hand figure above.

For a general charge distribution, we'll need to use both Gauss' law (2.1) and the extra equation (2.2). However, for rather special charge distributions – typically those with lots of symmetry – it turns out to be sufficient to solve the integral form of Gauss' law (2.4) alone, with the symmetry ensuring that (2.2) is automatically satisfied. We start by describing these rather simple solutions. We'll then return to the general case in Section (2.2).

2.1.1 The Coulomb Force

We'll start by showing that Gauss' law (2.4) reproduces the more familiar Coulomb force law that we all know and love. To do this, take a spherically symmetric charge distribution, centered at the origin, contained within some radius R . This will be our model for a particle. We won't need to make any assumption about the nature of the distribution other than its symmetry and the fact that the total charge is Q .

We want to know the electric field at some radius $r > R$. We take our Gaussian surface \mathcal{S} to be a sphere of radius r as shown in Fig. 2.2. Gauss' law states

$$\int_{\mathcal{S}} \mathbf{E} \cdot d\mathbf{S} = \frac{Q}{\epsilon_0}. \quad (2.5)$$

At this point we make use of the spherical symmetry of the problem. This tells us that the electric field must point radially outwards: $\mathbf{E}(\mathbf{x}) = E(r)\hat{\mathbf{r}}$. And, since the integral is only over the angular coordinates of the sphere, we can pull the function $E(r)$ outside. We have

$$\int_{\mathcal{S}} \mathbf{E} \cdot d\mathbf{S} = E(r) \int_{\mathcal{S}} \hat{\mathbf{r}} \cdot d\mathbf{S} = E(r)4\pi r^2 = \frac{Q}{\epsilon_0}, \quad (2.6)$$

where the factor of $4\pi r^2$ has arisen simply because it's the area of the Gaussian sphere.

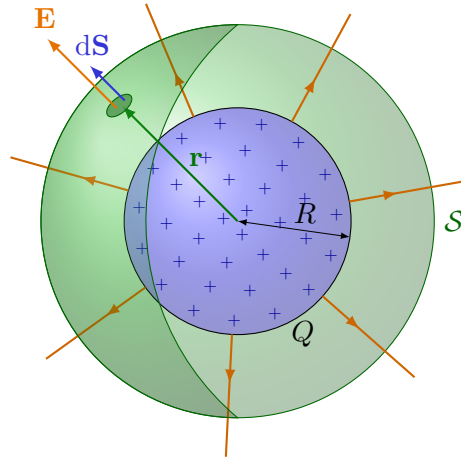


Fig. 2.2: Electric field produced by a spherically symmetric charge distribution, centered at the origin, contained within some radius R .

We learn that the electric field outside a spherically symmetric distribution of charge Q is

$$\mathbf{E}(\mathbf{x}) = \frac{Q}{4\pi\epsilon_0 r^2} \hat{\mathbf{r}}. \quad (2.7)$$

That's nice. This is the familiar result that we've seen before. The Lorentz force law (1.8) then tells us that a test charge q moving in the region $r > R$ experiences a force

$$\mathbf{F} = \frac{Qq}{4\pi\epsilon_0 r^2} \hat{\mathbf{r}}. \quad (2.8)$$

This, of course, is the *Coulomb force* between two static charged particles. Notice that, as promised, $1/\epsilon_0$ characterises the strength of the force. If the two charges have the same sign, so that $Qq > 0$, the force is repulsive, pushing the test charge away from the origin. If the charges have opposite signs, $Qq < 0$, the force is attractive, pointing towards the origin. We see that Gauss's law (2.1) reproduces this simple result that we know about charges.

Finally, note that the assumption of symmetry was crucial in our above analysis. Without it, the electric field $\mathbf{E}(\mathbf{x})$ would have depended on the angular coordinates of the sphere S and so been stuck inside the integral. In situations without symmetry, Gauss' law alone is not enough to determine the electric field and we need to also use $\nabla \times \mathbf{E} = 0$. We'll see how to do this in Section 2.2. If you're worried, however, it's simple to check that our final expression for the electric field (2.7) does indeed solve $\nabla \times \mathbf{E} = 0$.

2.1.1.1 Coulomb vs Newton

The inverse-square form of the force is common to both electrostatics and gravity. It's worth comparing the relative strengths of the two forces. For example, we can look at the relative strengths of Newtonian attraction and Coulomb repulsion between two electrons. These are point particles with mass m_e and charge $-e$ given by

$$e \approx 1.6 \times 10^{-19} \text{C} \quad \text{and} \quad m_e \approx 0.1 \times 10^{-31} \text{kg}. \quad (2.9)$$

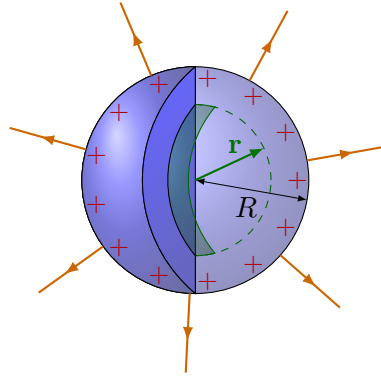


Fig. 2.3: Electric field inside a spherically symmetric charge distribution.

Regardless of the separation, we have

$$\frac{F_{\text{Coulomb}}}{F_{\text{Newton}}} = \frac{e^2}{4\pi\epsilon_0} \frac{1}{Gm_e^2}. \quad (2.10)$$

The strength of gravity is determined by Newton's constant $G \approx 6.7 \times 10^{-11} \text{ m}^3 \text{ kg}^{-1} \text{ s}^2$. Plugging in the numbers reveals something extraordinary:

$$\frac{F_{\text{Coulomb}}}{F_{\text{Newton}}} \approx 10^{42} \quad (2.11)$$

Gravity is puny. Electromagnetism rules. In fact you knew this already. The mere act of lifting up your arm is pitching a few electrical impulses up against the gravitational might of the entire Earth. Yet the electrical impulses win.

However, gravity has a trick up its sleeve. While electric charges come with both positive and negative signs, mass is only positive. It means that by the time we get to macroscopically large objects – stars, planets, cats – the mass accumulates while the charges cancel to good approximation. This compensates the factor of 10^{-42} suppression until, at large distance scales, gravity wins after all.

The fact that the force of gravity is so ridiculously tiny at the level of fundamental particles has consequence. It means that we can neglect gravity whenever we talk about the very small. (And indeed, we shall neglect gravity for the rest of this course). However, it also means that if we would like to understand gravity better on these very tiny distances – for example, to develop a quantum theory of gravity – then it's going to be tricky to get much guidance from experiment.

2.1.2 A Uniform Sphere

The electric field outside a spherically symmetric charge distribution is always given by (2.7). What about inside? This depends on the distribution in question. The simplest is a sphere of radius R with uniform charge distribution ρ (see Fig. 2.3). The total charge is

$$Q = \frac{4\pi}{3} R^3 \rho. \quad (2.12)$$

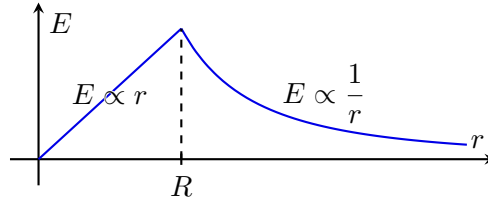


Fig. 2.4: Radial dependence of electric field strength inside and outside a spherically symmetric charge distribution of radius R .

Let's pick our Gaussian surface to be a sphere, centered at the origin, of radius $r < R$. The charge contained within this sphere is $4\pi\rho r^3/3 = Qr^3/R^3$, so Gauss' law gives

$$\int_S \mathbf{E} \cdot d\mathbf{S} = \frac{Qr^3}{\epsilon_0 R^3}. \quad (2.13)$$

Again, using the symmetry argument we can write $\mathbf{E}(\mathbf{r}) = E(r)\hat{\mathbf{r}}$ and compute

$$\int_S \mathbf{E} \cdot d\mathbf{S} = E(r) \int_S \hat{\mathbf{r}} \cdot d\mathbf{S} = E(r)4\pi r^2 = \frac{Qr^3}{\epsilon_0 R^3}. \quad (2.14)$$

This tells us that the electric field grows linearly inside the sphere

$$\mathbf{E}(\mathbf{x}) = \frac{Qr}{4\pi\epsilon_0 R^3} \hat{\mathbf{r}}, \quad r < R. \quad (2.15)$$

Outside the sphere we revert to the inverse-square $E \propto 1/r^2$ form (2.7). At the surface of the sphere, $r = R$, the electric field is continuous but the derivative, dE/dr , is not. This is shown in Fig. 2.4.

2.1.3 Line Charges

Consider, next, a charge smeared out along a line which we'll take to be the z -axis. We'll take uniform charge density η per unit length. (If you like you could consider a solid cylinder with uniform charge density and then send the radius to zero). We want to know the electric field due to this line of charge.

Our set-up now has cylindrical symmetry. We take the Gaussian surface to be a cylinder of length L and radius r (See Fig. 2.5). We have

$$\int_S \mathbf{E} \cdot d\mathbf{S} = \frac{\eta L}{\epsilon_0}. \quad (2.16)$$

Again, by symmetry, the electric field points in the radial direction, away from the line. We'll denote this vector in cylindrical polar coordinates as $\hat{\mathbf{r}}$ so that $\mathbf{E}(\mathbf{r}) = E(r)\hat{\mathbf{r}}$. The symmetry means that the two end caps of the Gaussian surface, \mathcal{S}_1 and \mathcal{S}_2 in Fig. 2.5, don't contribute to the integral because their normal points parallel to the $\hat{\mathbf{z}}$ direction and $\hat{\mathbf{z}} \cdot \hat{\mathbf{r}} = 0$. We're left only with a contribution from the curved side of the cylinder, \mathcal{S}_3 ,

$$\int_S \mathbf{E} \cdot d\mathbf{S} = E(r)2\pi rL = \frac{\eta L}{\epsilon_0}. \quad (2.17)$$

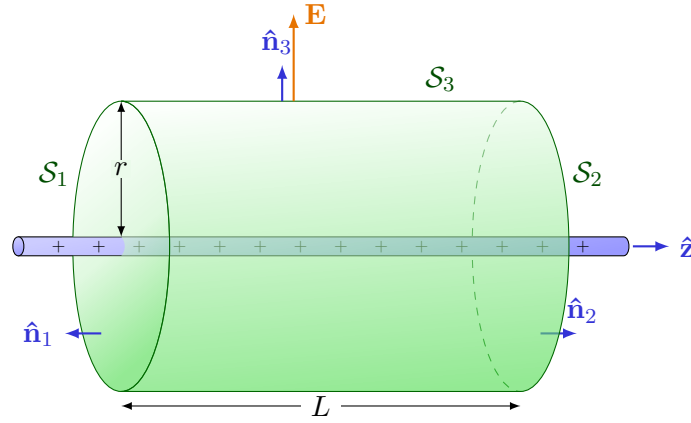


Fig. 2.5: The electric field \mathbf{E} of an infinite line charge with a uniform linear charge density η . Considering a Gaussian surface in the form of a cylinder at radius r , the electric field has the same magnitude at every point on the surface \mathcal{S}_3 and is directed outward, parallel to $\hat{\mathbf{n}}_3$.

So that the electric field is

$$\mathbf{E}(\mathbf{x}) = \frac{\eta}{2\pi\epsilon_0} \hat{\mathbf{r}}. \quad (2.18)$$

Note that, while the electric field for a point charge drops off as $1/r^2$ (with r the radial distance), the electric field for a line charge drops off more slowly as $1/r$. (Of course, the radial distance r means slightly different things in the two cases: it is $r = \sqrt{x^2 + y^2 + z^2}$ for the point particle, but is $r = \sqrt{x^2 + y^2}$ for the line).

2.1.4 Surface Charges and Discontinuities

Now consider an infinite plane, which we take to be $z = 0$, carrying uniform charge per unit area, σ . We again take our Gaussian surface to be a cylinder, this time with its axis perpendicular to the plane as shown in the figure. In this context, the cylinder is sometimes referred to as a Gaussian “pillbox” (on account of Gauss’ well known fondness for aspirin). On symmetry grounds, we have

$$\mathbf{E}(\mathbf{r}) = E(z)\hat{\mathbf{z}}. \quad (2.19)$$

Moreover, the electric field in the upper plane, $z > 0$, must point in the opposite direction from the lower plane, $z < 0$, so that $E(z) = -E(-z)$.

The surface integral now vanishes over the curved side of the cylinder and we only get contributions from the end caps, which we take to have area A . This gives

$$\int_S \mathbf{E} \cdot d\mathbf{S} = E(z)A - E(-z)A = 2E(z) = \frac{\sigma A}{\epsilon_0}. \quad (2.20)$$

The electric field above an infinite plane of charge is therefore

$$\mathbf{E}(\mathbf{x}) = \frac{\sigma}{2\epsilon_0} \hat{\mathbf{z}}. \quad (2.21)$$

Note that the electric field is independent of the distance from the plane! This is because the plane is infinite in extent: the further you move from it, the more comes into view.

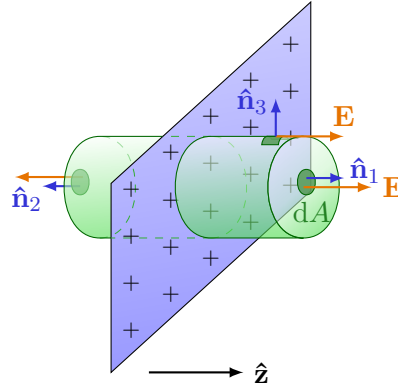


Fig. 2.6: For an infinite sheet of charge with charge density σ , the electric field \mathbf{E} will be perpendicular to the surface. Therefore only the ends of a cylindrical Gaussian surface will contribute to the electric flux. In this case a cylindrical Gaussian surface perpendicular to the charge sheet is used. The resulting field is half that of a conductor at equilibrium with this surface charge density.

There is another important point to take away from this analysis. The electric field is not continuous on either side of a surface of constant charge density. We have

$$E(z \rightarrow 0^+) - E(z \rightarrow 0^-) = \frac{\sigma}{\epsilon_0}. \quad (2.22)$$

For this to hold, it is not important that the plane stretches to infinity. It's simple to redo the above analysis for any arbitrary surface with charge density σ . There is no need for σ to be uniform and, correspondingly, there is no need for \mathbf{E} at a given point to be parallel to the normal to the surface $\hat{\mathbf{n}}$. At any point of the surface, we can take a Gaussian cylinder, as shown in the left-hand figure above, whose axis is normal to the surface at that point. Its cross-sectional area A can be arbitrarily small (since, as we saw, it drops out of the final answer). If \mathbf{E}_{\pm} denotes the electric field on either side of the surface, then

$$\hat{\mathbf{n}} \cdot \mathbf{E}_+ - \hat{\mathbf{n}} \cdot \mathbf{E}_- = \frac{\sigma}{\epsilon_0}. \quad (2.23)$$

In contrast, the electric field tangent to the surface is continuous. To see this, we need to do a slightly different calculation. Consider, again, an arbitrary surface with surface charge. Now we consider a loop \mathcal{C} with a length L which lies parallel to the surface and a length a which is perpendicular to the surface. We've drawn this loop in the right-hand figure above, where the surface is now shown side-on. We integrate \mathbf{E} around the loop. Using Stoke's theorem, we have

$$\oint_{\mathcal{C}} \mathbf{E} \cdot d\mathbf{r} = \int_{\mathcal{S}} \nabla \times \mathbf{E} \cdot d\mathbf{S}, \quad (2.24)$$

where \mathcal{S} is the surface bounded by \mathcal{C} . In the limit $a \rightarrow 0$, the surface \mathcal{S} shrinks to zero size so this integral gives zero. This means that the contribution to line integral must also vanish, leaving us with

$$\hat{\mathbf{n}} \times \mathbf{E}_+ - \hat{\mathbf{n}} \times \mathbf{E}_- = 0. \quad (2.25)$$

This is the statement that the electric field tangential to the surface is continuous.

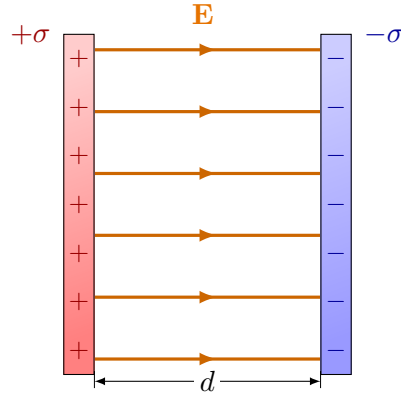


Fig. 2.7: A pair of infinite planes at $z = 0$ and $z = a$, carrying uniform surface charge density $\pm\sigma$

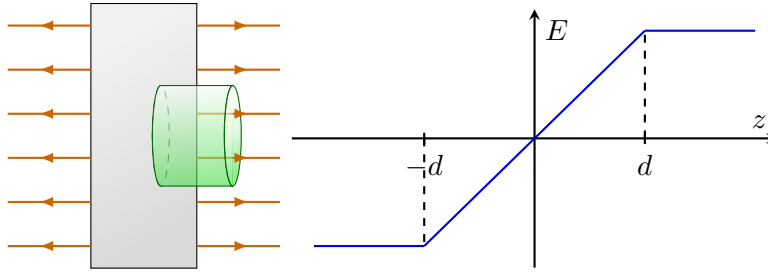


Fig. 2.8: Left: The Gaussian surface for a plane slab. Right: The resulting electric field.

2.1.4.1 A Pair of Planes

As a simple generalisation, consider a pair of infinite planes at $z = 0$ and $z = a$, carrying uniform surface charge density $\pm\sigma$ respectively as shown in Fig. 2.7. To compute the electric field we need only add the fields arising from two planes, each of which takes the form (2.21). We find that the electric field between the two planes is

$$\mathbf{E} = \frac{\sigma}{\epsilon_0} \hat{\mathbf{z}}, \quad 0 < z < a, \quad (2.26)$$

while $\mathbf{E} = 0$ outside the planes.

2.1.4.2 A Plane Slab

We can rederive the discontinuity (2.23) in the electric field by considering an infinite slab of thickness $2d$ and charge density per unit volume ρ . When our Gaussian pillbox lies inside the slab, with $z < d$, we have

$$2AE(z) = \frac{2zA\rho}{\epsilon_0} \implies E(z) = \frac{\rho z}{\epsilon_0}. \quad (2.27)$$

Meanwhile, for $z > d$ we get our earlier result (2.21). The electric field is now continuous as shown in the figure. Taking the limit $d \rightarrow 0$ and $\rho \rightarrow \infty$ such that the surface charge $\sigma = \rho d$ remains constant reproduces the discontinuity (2.22).

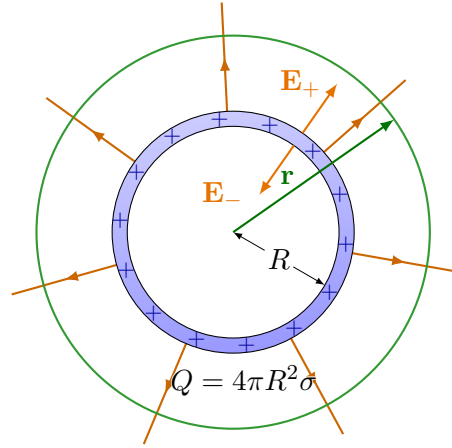


Fig. 2.9: A spherical shell of radius R , centered at the origin, with uniform surface charge density σ .

2.1.4.3 A Spherical Shell

Let's give one last example that involves surface charge and the associated discontinuity of the electric field. We'll consider a spherical shell of radius R , centered at the origin, with uniform surface charge density σ . The total charge is

$$Q = 4\pi R^2 \sigma \quad (2.28)$$

We already know that outside the shell, $r > R$, the electric field takes the standard inverse-square form (2.7). What about inside? Well, since any surface with $r < R$ doesn't surround a charge, Gauss' law tells us that we necessarily have $E = 0$ inside. That means that there is a discontinuity at the surface $r = R$,

$$\mathbf{E} \cdot \hat{\mathbf{r}}_+ - \mathbf{E} \cdot \hat{\mathbf{r}}_- = \frac{Q}{4\pi R^2 \epsilon_0} = \frac{\sigma}{\epsilon_0}, \quad (2.29)$$

in accord with the expectation (2.23).

2.2 The Electrostatic Potential

For all the examples in the last section, symmetry considerations meant that we only needed to consider Gauss' law. However, for general charge distributions Gauss' law is not sufficient. We also need to invoke the second equation, $\nabla \times \mathbf{E} = 0$.

In fact, this second equation is easily dispatched since $\nabla \times \mathbf{E} = 0$ implies that the electric field can be written as the gradient of some function,

$$\mathbf{E} = -\nabla \phi. \quad (2.30)$$

The scalar ϕ is called the *electrostatic potential* or *scalar potential* (or, sometimes, just the *potential*). To proceed, we revert to the original differential form of Gauss' law (2.1). This now takes the form of the *Poisson equation*

$$\nabla \cdot \mathbf{E} = \frac{\rho}{\epsilon_0} \implies \nabla^2 \phi = -\frac{\rho}{\epsilon_0}. \quad (2.31)$$

In regions of space where the charge density vanishes, we're left solving the Laplace equation

$$\nabla^2 \phi = 0. \quad (2.32)$$

Solutions to the Laplace equation are said to be *harmonic* functions.

A few comments:

- The potential ϕ is only defined up to the addition of some constant. This seemingly trivial point is actually the beginning of a long and deep story in theoretical physics known as *gauge invariance*. We'll come back to it in Subsection 5.0.1. For now, we'll eliminate this redundancy by requiring that $\phi(r) \rightarrow 0$ as $r \rightarrow \infty$.
- We know from our study of Newtonian mechanics that the electrostatic potential is proportional to the potential energy experienced by a test particle. Specifically, a test particle of mass m , position $\mathbf{r}(t)$ and charge q moving in a background electric field has conserved energy

$$E = \frac{1}{2} m \dot{\mathbf{r}} \cdot \dot{\mathbf{r}} + q \phi(\mathbf{r}). \quad (2.33)$$

- The Poisson equation is linear in both ϕ and ρ . This means that if we know the potential ϕ_1 for some charge distribution ρ_1 and the potential ϕ_2 for another charge distribution ρ_2 , then the potential for $\rho_1 + \rho_2$ is simply $\phi_1 + \phi_2$. What this really means is that the electric field for a bunch of charges is just the sum of the fields generated by each charge. This is called the *principle of superposition* for charges. This linearity of the equations is what makes electromagnetism easy compared to other forces of Nature.
- We stated above that $\nabla \times \mathbf{E} = 0$ is equivalent to writing $\mathbf{E} = -\nabla \phi$. This is true when space is \mathbb{R}^3 or, in fact, if we take space to be any open ball in \mathbb{R}^3 . But if our background space has a suitably complicated topology then there are solutions to $\nabla \times \mathbf{E} = 0$ which cannot be written in the form $\mathbf{E} = -\nabla \phi$. This is tied ultimately to the beautiful mathematical theory of de Rham cohomology. Needless to say, in this starter course we're not going to worry about these issues. We'll always take spacetime to have topology \mathbb{R}^4 and, correspondingly, any spatial hypersurface to be \mathbb{R}^3 .

2.2.1 The Point Charge

Let's start by deriving the Coulomb force law yet again. We'll take a particle of charge Q and place it at the origin. This time, however, we'll assume that the particle really is a point charge. This means that the charge density takes the form of a delta-function, $\rho(\mathbf{x}) = Q\delta^3(\mathbf{x})$. We need to solve the equation

$$\nabla^2 \phi = -\frac{Q}{\epsilon_0} \delta^3(\mathbf{x}). \quad (2.34)$$

The solution is essentially the Green's function (See Appendix A.1) for the Laplacian ∇^2 , an interpretation that we'll return to in Subsection 2.2.3. Let's recall how we find this solution. We first look away from the origin, $r \neq 0$, where there's no funny business going

on with delta-function. Here, we're looking for the spherically symmetric solution to the Laplace equation. This is

$$\phi = \frac{\alpha}{r}, \quad (2.35)$$

or some constant α . To see why this solves the Laplace equation, we need to use the result

$$\nabla r = \hat{\mathbf{r}} \quad (2.36)$$

where $\hat{\mathbf{r}}$ is the unit radial vector in spherical polar coordinates, so $\mathbf{x} = r\hat{\mathbf{r}}$. Using the chain rule, this means that $\nabla(1/r) = -\hat{\mathbf{r}}/r^2 = -\mathbf{x}/r^3$. This gives us

$$\nabla\phi = -\frac{\alpha}{r^3}\mathbf{x} \implies \nabla^2\phi = -\alpha\left(\frac{\nabla\cdot\mathbf{x}}{r^3} - \frac{3\mathbf{x}\cdot\mathbf{x}}{r^5}\right). \quad (2.37)$$

But $\nabla\cdot\mathbf{x} = 3$ and we find that $\nabla^2\phi = 0$ as required.

It remains to figure out what to do at the origin where the delta-function lives. This is what determines the overall normalisation α of the solution. At this point, it's simplest to use the integral form of Gauss' law to transfer the problem from the origin to the far flung reaches of space. To do this, we integrate (2.34) over some region \mathcal{V} which includes the origin. Integrating the charge density gives

$$\rho(\mathbf{x}) = Q\delta^3(\mathbf{x}) \implies \int_{\mathcal{V}} d^3x \rho = Q. \quad (2.38)$$

So, using Gauss' law (2.4), we require

$$\int_{\mathcal{S}} \nabla\phi \cdot d\mathbf{S} = -\frac{Q}{\epsilon_0}. \quad (2.39)$$

But this is exactly the kind of surface integral that we were doing in the last section. Substituting $\phi = \alpha/r$ into the above equation, and choosing \mathcal{S} to be a sphere of radius r , tells us that we must have $\alpha = Q/4\pi\epsilon_0$, or

$$\phi = \frac{Q}{4\pi\epsilon_0 r}. \quad (2.40)$$

Taking the gradient of this using (2.36) gives us Coulomb's law

$$\mathbf{E}(\mathbf{x}) = -\nabla\phi = \frac{Q}{4\pi\epsilon_0 r^2} \hat{\mathbf{r}}. \quad (2.41)$$

The derivation of Coulomb's law using the potential was somewhat more involved than the technique using Gauss' law alone that we saw in the last section. However, as we'll now see, introducing the potential allows us to write down the solution to essentially any problem.

A Note on Notation Throughout these lectures, we will use \mathbf{x} and \mathbf{r} interchangeably to denote position in space. For example, sometimes we'll write integration over a volume as $\int d^3x$ and sometimes as $\int d^3r$. The advantage of the \mathbf{r} notation is that it looks more natural when working in spherical polar coordinates. For example, we have $|\mathbf{r}| = r$ which is nice. The disadvantage is that it can lead to confusion when working in other coordinate systems, in particular cylindrical polar. For this reason, we'll alternate between the two notations, adopting the attitude that clarity is more important than consistency.

2.2.2 The Dipole

A *dipole* consists of two point charges, Q and $-Q$, a distance d apart. We place the first charge at the origin and the second at $\mathbf{r} = -\mathbf{d}$. The potential is simply the sum of the potential for each charge,

$$\phi = \frac{1}{4\pi\epsilon_0} \left(\frac{Q}{r} - \frac{Q}{|\mathbf{r} + \mathbf{d}|} \right). \quad (2.42)$$

Similarly, the electric field is just the sum of the electric fields made by the two point charges. This follows from the linearity of the equations and is a simple application of the principle of superposition that we mentioned earlier.

It will prove fruitful to ask what the dipole looks like far from the two point charges, at a distance $r \gg |\mathbf{d}|$. We need to Taylor expand the second term above. The vector version of the Taylor expansion for a general function $f(\mathbf{r})$ is given by

$$f(\mathbf{r} + \mathbf{d}) \approx f(\mathbf{r}) + \mathbf{d} \cdot \nabla f(\mathbf{r}) + \frac{1}{2} (\mathbf{d} \cdot \nabla)^2 f(\mathbf{r}) + \dots \quad (2.43)$$

Applying this to the function $1/|\mathbf{r} + \mathbf{d}|$ gives

$$\begin{aligned} \frac{1}{|\mathbf{r} + \mathbf{d}|} &\approx \frac{1}{r} + \mathbf{d} \cdot \nabla \frac{1}{r} + \frac{1}{2} (\mathbf{d} \cdot \nabla)^2 \frac{1}{r} + \dots \\ &= \frac{1}{r} - \frac{\mathbf{d} \cdot \mathbf{r}}{r^3} - \frac{1}{2} \left(\frac{\mathbf{d} \cdot \mathbf{d}}{r^3} - \frac{3(\mathbf{d} \cdot \mathbf{r})^2}{r^5} \right) + \dots \end{aligned} \quad (2.44)$$

(To derive the last term, it might be easiest to use index notation for $\mathbf{d} \cdot \nabla = d_i \partial_i$). For our dipole, we'll only need the first two terms in this expansion. They give the potential

$$\phi \approx \frac{Q}{4\pi\epsilon_0} \left(\frac{1}{r} - \frac{1}{r} + \mathbf{d} \cdot \nabla \frac{1}{r} + \dots \right) = \frac{Q}{4\pi\epsilon_0} \frac{\mathbf{d} \cdot \mathbf{r}}{r^3} + \dots \quad (2.45)$$

We see that the potential for a dipole falls off as $1/r^2$. Correspondingly, the electric field drops off as $1/r^3$; both are one power higher than the fields for a point charge.

The electric field is not spherically symmetric. The leading order contribution is governed by the combination

$$\mathbf{p} = Q\mathbf{d}. \quad (2.46)$$

This is called the electric *dipole moment*. By convention, it points from the negative charge to the positive. The dipole electric field is

$$\mathbf{E} = -\nabla\phi = \frac{1}{4\pi\epsilon_0} \left(\frac{3(\mathbf{p} \cdot \hat{\mathbf{r}})\hat{\mathbf{r}} - \mathbf{p}}{r^3} \right) + \dots \quad (2.47)$$

Notice that the sign of the electric field depends on where you sit in space. In some parts, the force will be attractive; in other parts repulsive.

It's sometimes useful to consider the limit $d \rightarrow 0$ and $Q \rightarrow \infty$ such that $\mathbf{p} = Q\mathbf{d}$ remains fixed. In this limit, all the \dots terms in (2.45) and (2.47) disappear since they contain higher powers of d . Often when people talk about the “dipole”, they implicitly mean taking this limit.

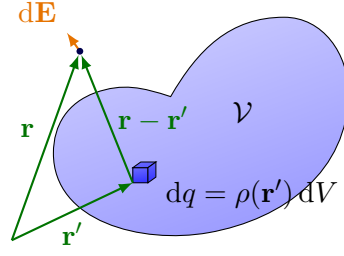


Fig. 2.10: Electric field at \mathbf{r} with $|\mathbf{r}| \gg |\mathbf{r}'|$.

2.2.3 General Charge Distributions

Our derivation of the potential due to a point charge (2.40), together with the principle of superposition, is actually enough to solve – at least formally – the potential due to any charge distribution. This is because the solution for a point charge is nothing other than the Green’s function for the Laplacian. The Green’s function is defined to be the solution to the equation

$$\phi(\mathbf{r}) = -\frac{1}{\epsilon_0} \int_{\mathcal{V}} d^3r' G(\mathbf{r}; \mathbf{r}') \rho(\mathbf{r}') = \frac{1}{4\pi\epsilon_0} \int_{\mathcal{V}} d^3r' \frac{\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} \quad (2.48)$$

(To check this, you just have to keep your head and remember whether the operators are hitting \mathbf{r} or \mathbf{r}' . The Laplacian acts on \mathbf{r} so, if we compute $\nabla^2 \phi$, it passes through the integral in the above expression and hits $G(\mathbf{r}; \mathbf{r}')$, leaving behind a delta-function which subsequently kills the integral).

Similarly, the electric field arising from a general charge distribution is

$$\begin{aligned} \mathbf{E}(\mathbf{r}) &= -\nabla \phi(\mathbf{r}) = -\frac{1}{4\pi\epsilon_0} \int_{\mathcal{V}} d^3r' \rho(\mathbf{r}') \nabla \frac{1}{|\mathbf{r} - \mathbf{r}'|} \\ &= -\frac{1}{4\pi\epsilon_0} \int_{\mathcal{V}} d^3r' \rho(\mathbf{r}') \frac{\mathbf{r} - \mathbf{r}'}{|\mathbf{r} - \mathbf{r}'|^3}. \end{aligned} \quad (2.49)$$

Given a very complicated charge distribution $\rho(\mathbf{r})$, this equation will give back an equally complicated electric field $\mathbf{E}(\mathbf{r})$. But if we sit a long way from the charge distribution, there’s a rather nice simplification that happens.

2.2.3.1 Long Distance Behaviour

Suppose now that you want to know what the electric field looks like far from the region \mathcal{V} . This means that we’re interested in the electric field at \mathbf{r} with $|\mathbf{r}| \gg |\mathbf{r}'|$ for all $\mathbf{r}' \in \mathcal{V}$. We can apply the same Taylor expansion (2.43), now replacing \mathbf{d} with $-\mathbf{r}'$ for each \mathbf{r}' in the charged region. This means we can write

$$\begin{aligned} \frac{1}{|\mathbf{r} - \mathbf{r}'|} &= \frac{1}{r} - \mathbf{r}' \cdot \nabla \frac{1}{r} + \frac{1}{2} (\mathbf{r}' \cdot \nabla)^2 \frac{1}{r} + \dots \\ &= \frac{1}{r} + \frac{\mathbf{r} \cdot \mathbf{r}'}{r^3} + \frac{1}{2} \left(\frac{3(\mathbf{r} \cdot \mathbf{r}')^2}{r^5} - \frac{\mathbf{r}' \cdot \mathbf{r}'}{r^3} \right) + \dots, \end{aligned} \quad (2.50)$$

and our potential becomes

$$\phi(\mathbf{r}) = \frac{1}{4\pi\epsilon_0} \int_{\mathcal{V}} d^3r' \rho(\mathbf{r}') \left(\frac{1}{r} + \frac{\mathbf{r}' \cdot \mathbf{r}'}{r^3} + \dots \right). \quad (2.51)$$

The leading term is just

$$\phi(\mathbf{r}) = \frac{Q}{4\pi\epsilon_0} + \dots, \quad (2.52)$$

where $Q = \int_{\mathcal{V}} d^3r' \rho(\mathbf{r}')$ is the total charge contained within \mathcal{V} . So, to leading order, if you're far enough away then you can't distinguish a general charge distribution from a point charge localised at the origin. But if you're careful with experiments, you can tell the difference. The first correction takes the form of a dipole,

$$\phi(\mathbf{r}) = \frac{1}{4\pi\epsilon_0} \left(\frac{Q}{r} + \frac{\mathbf{p} \cdot \hat{\mathbf{r}}}{r^2} + \dots \right), \quad (2.53)$$

where

$$\mathbf{p} = \int_{\mathcal{V}} d^3r' \mathbf{r}' \rho(\mathbf{r}') \quad (2.54)$$

is the dipole moment of the distribution. One particularly important situation is when we have a neutral object with $Q = 0$. In this case, the dipole is the dominant contribution to the potential.

We see that an arbitrarily complicated, localised charge distribution can be characterised by a few simple quantities, of decreasing importance. First comes the total charge Q . Next the dipole moment \mathbf{p} which contains some basic information about how the charges are distributed. But we can keep going. The next correction is called the quadrupole and is given by

$$\Delta\phi = \frac{1}{2} \frac{1}{4\pi\epsilon_0} \frac{r_i r_j Q_{ij}}{r^5}, \quad (2.55)$$

where Q_{ij} is a symmetric traceless tensor known as the quadrupole moment, given by

$$Q_{ij} = \int_{\mathcal{V}} d^3r' \rho(\mathbf{r}') (3r'_i r'_j - \delta_{ij} r'^2) \quad (2.56)$$

It contains some more refined information about how the charges are distributed. After this comes the octopole and so on. The general name given to this approach is the *multipole expansion*. It involves expanding the function ϕ in terms of spherical harmonics. A systematic treatment can be found, for example, in the book by Jackson.

2.2.3.2 A Comment on Infinite Charge Distribution

In the above, we assumed for simplicity that the charge distribution was restricted to some compact region of space, \mathcal{V} . The Green's function approach still works if the charge distribution stretches to infinity. However, for such distributions it's not always possible to pick $\phi(\mathbf{r}) \rightarrow 0$ as $r \rightarrow \infty$. In fact, we saw an example of this earlier. For an infinite line charge of density η , we computed the electric field in (2.18). It goes as

$$E(\mathbf{x}) = \frac{\eta}{2\pi\epsilon_0} \hat{\mathbf{r}}, \quad (2.57)$$

where now $r^2 = x^2 + y^2$ is the cylindrical radial coordinate perpendicular to the line. The potential ϕ which gives rise to this is

$$\phi(r) = -\frac{\eta}{2\pi\epsilon_0} \log\left(\frac{r}{r_0}\right). \quad (2.58)$$

Because of the log function, we necessarily have $\phi(r) \rightarrow 0$ as $r \rightarrow \infty$. Instead, we need to pick an arbitrary, but finite distance, r_0 at which the potential vanishes.

2.2.4 Field Lines

The usual way of depicting a vector is to draw an arrow whose length is proportional to the magnitude. For the electric field, there's a slightly different, more useful way to show what's going on. We draw continuous lines, tangent to the electric field \mathbf{E} , with the density of lines proportional to the magnitude of \mathbf{E} . This innovation, due to Faraday, is called the *field line*. (They are what we have been secretly drawing throughout these notes).

Field lines are continuous. They begin and end only at charges. They can never cross.

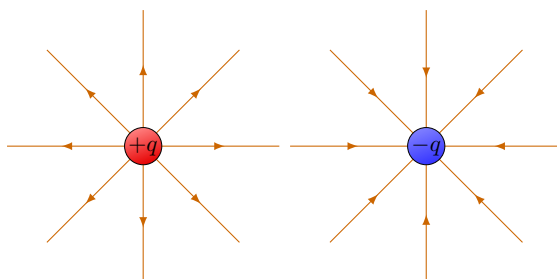


Fig. 2.11: Field lines for positive and negative point charges

By convention, the positive charges act as sources for the lines, with the arrows emerging. The negative charges act as sinks, with the arrows approaching.

It's also easy to draw the equipotentials – surfaces of constant ϕ – on this same figure. These are the surfaces along which you can move a charge without doing any work. The relationship $\mathbf{E} = -\nabla\phi$ ensures that the equipotentials cut the field lines at right angles. We usually draw them as dotted lines.

2.2.5 Electrostatic Equilibrium

Here's a simple question: can you trap an electric charge using only other charges? In other words, can you find some arrangements of charges such that a test charge sits in stable equilibrium, trapped by the fields of the others.

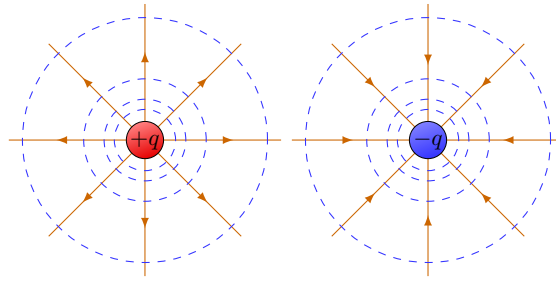


Fig. 2.12: Equipotentials for positive and negative point charges

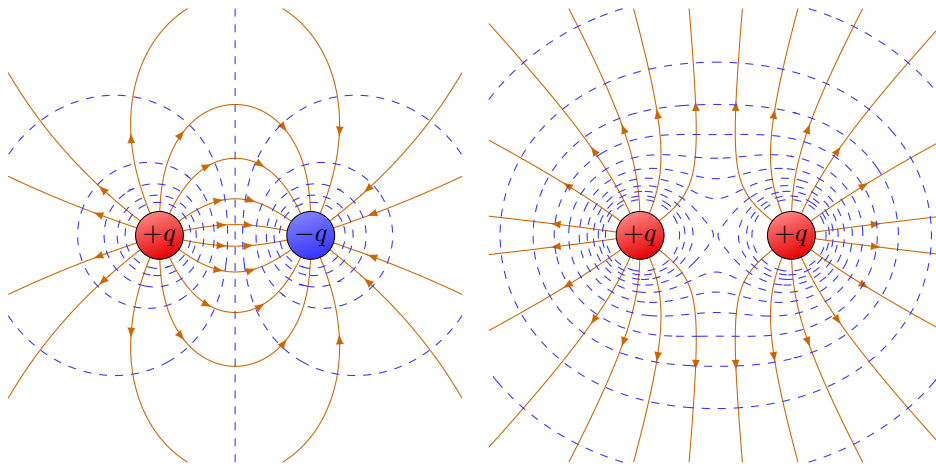


Fig. 2.13: Field lines and equipotentials for the dipole (left) and for a pair of charges of the same sign (on the right).

There's a trivial way to do this: just allow a negative charge to sit directly on top of a positive charge. But let's throw out this possibility. We'll ask that the equilibrium point lies away from all the other charges.

There are some simple set-ups that spring to mind that might achieve this. Maybe you could place four positive charges at the vertices of a pyramid; or perhaps 8 positive charges at the corners of a cube. Is it possible that a test positive charge trapped in the middle will be stable? It's certainly repelled from all the corners, so it might seem plausible.

The answer, however, is no. There is no electrostatic equilibrium. You cannot trap an electric charge using only other stationary electric charges, at least not in a stable manner. Since the potential energy of the particle is proportional to ϕ , mathematically, this is the statement that a harmonic function, obeying $\nabla^2\phi = 0$, can have no minimum or maximum.

To prove that there can be no electrostatic equilibrium, let's suppose the opposite: that there is some point in empty space \mathbf{r}_* that is stable for a particle of charge $q > 0$. By "empty space", we mean that $\rho(\mathbf{r}) = 0$ in a neighbourhood of \mathbf{r}_* . Because the point is stable, if the particle moves away from this point then it must always be pushed back. This, in turn, means that the electric field must always point inwards towards the point

\mathbf{r}_* ; never away. We could then surround \mathbf{r}_* by a small surface \mathcal{S} and compute

$$\int_{\mathcal{S}} \mathbf{E} \cdot d\mathbf{S} < 0. \quad (2.59)$$

But, by Gauss' law, the right-hand side must be the charge contained within \mathcal{S} which, by assumption, is zero. This is our contradiction: electrostatic equilibrium does not exist.

Of course, if you're willing to use something other than electrostatic forces then you can construct equilibrium situations. For example, if you restrict the test particle to lie on a plane then it's simple to check that equal charges placed at the corners of a polygon will result in a stable equilibrium point in the middle. But to do this you need to use other forces to keep the particle in the plane in the first place.

2.3 Electrostatic Energy

There is energy stored in the electric field. In this section, we calculate how much.

Suppose we have some test charge q moving in a background electrostatic potential ϕ . We'll denote the potential energy of the particle as $U(\mathbf{r})$. The potential $U(\mathbf{r})$ of the particle can be thought of as the work done bringing the particle in from infinity;

$$U(\mathbf{r}) = - \int_{\infty}^r \mathbf{F} \cdot d\mathbf{r} = +q \int_{\infty}^r \nabla \phi \cdot d\mathbf{r} = q\phi(\mathbf{r}), \quad (2.60)$$

where we've assumed our standard normalisation of $\phi(\mathbf{r}) \rightarrow 0$ as $r \rightarrow \infty$.

Consider a distribution of charges which, for now, we'll take to be made of point charges q_i at positions \mathbf{r}_i . The electrostatic potential energy stored in this configuration is the same as the work required to assemble the configuration in the first place. (This is because if you let the charges go, this is how much kinetic energy they will pick up). So how much work does it take to assemble a collection of charges?

Well, the first charge is free. In the absence of any electric field, you can just put it where you like - say, \mathbf{r}_1 . The work required is $W_1 = 0$.

To place the second charge at \mathbf{r}_2 takes work

$$W_2 = \frac{q_1 q_2}{4\pi\epsilon_0} \frac{1}{|\mathbf{r}_1 - \mathbf{r}_2|}. \quad (2.61)$$

Note that if the two charges have the same sign, so $q_1 q_2 > 0$, then $W_2 > 0$ which is telling us that we need to put work in to make them approach. If $q_1 q_2 < 0$ then $W_2 < 0$ where the negative work means that the particles wanted to be drawn closer by their mutual attraction.

The third charge has to battle against the electric field due to both q_1 and q_2 . The work required is

$$W_3 = \frac{q_3}{4\pi\epsilon_0} \left(\frac{q_2}{|\mathbf{r}_2 - \mathbf{r}_2|} + \frac{q_1}{|\mathbf{r}_1 - \mathbf{r}_3|} \right), \quad (2.62)$$

and so on. The total work needed to assemble all the charges is the potential energy stored in the configuration,

$$U = \sum_{i=1}^N W_i = \frac{1}{4\pi\epsilon_0} \sum_{i<j} \frac{q_i q_j}{|\mathbf{r}_i - \mathbf{r}_j|}, \quad (2.63)$$

where $\sum_{i<j}$ means that we sum over each pair of particles once. In fact, you probably could have just written down (2.63) as the potential energy stored in the configuration. The whole purpose of the above argument was really just to nail down a factor of 1/2: do we sum over all pairs of particles $\sum_{i<j}$ or all particles $\sum_{i \neq j}$? The answer, as we have seen, is all pairs.

We can make that factor of 1/2 even more explicit by writing

$$U = \frac{1}{2} \frac{1}{4\pi\epsilon_0} \sum_i \sum_{j \neq i} \frac{q_i q_j}{|\mathbf{r}_i - \mathbf{r}_j|}, \quad (2.64)$$

where now we sum over each pair twice.

There is a slicker way of writing (2.64). The potential at \mathbf{r}_i due to all the other charges q_j , $j \neq i$ is

$$\phi(\mathbf{r}_i) = \frac{1}{4\pi\epsilon_0} \sum_{j \neq i} \frac{q_j}{|\mathbf{r}_i - \mathbf{r}_j|} \quad (2.65)$$

which means that we can write the potential energy as

$$U = \frac{1}{2} \sum_{i=1}^N q_i \phi(\mathbf{r}_i). \quad (2.66)$$

This is the potential energy for a set of point charges. But there is an obvious generalisation to charge distributions $\rho(\mathbf{r})$. We'll again assume that $\rho(\mathbf{r})$ has compact support so that the charge is localised in some region of space. The potential energy associated to such a charge distribution should be

$$U = \frac{1}{2} \int d^3r \rho(\mathbf{r}) \phi(\mathbf{r}), \quad (2.67)$$

where we can quite happily take the integral over all of \mathbb{R}^3 , safe in the knowledge that anywhere that doesn't contain charge has $\rho(\mathbf{r}) = 0$ and so won't contribute.

Now this is in a form that we can start to play with. We use Gauss' law to rewrite it as

$$U = \frac{\epsilon_0}{2} \int d^3r (\nabla \cdot \mathbf{E}) \phi = \frac{\epsilon_0}{2} \int d^3r [\nabla \cdot (\mathbf{E} \phi) - \mathbf{E} \cdot \nabla \phi]. \quad (2.68)$$

But the first term is a total derivative. And since we're taking the integral over all of space and $\phi(\mathbf{r}) \rightarrow \infty$ as $r \rightarrow \infty$, this term just vanishes. In the second term we can replace $\nabla \phi = -\mathbf{E}$. We find that the potential energy stored in a charge distribution has an elegant expression solely in terms of the electric field that it creates,

$$U = \frac{\epsilon_0}{2} \int d^3r \mathbf{E} \cdot \mathbf{E}. \quad (2.69)$$

Isn't that nice!

2.3.1 The Energy of a Point Particle

There is a subtlety in the above derivation. In fact, I totally tried to pull the wool over your eyes. Here it's time to own up.

First, let me say that the final result (2.69) is right: this is the energy stored in the electric field. But the derivation above was dodgy. One reason to be dissatisfied is that we computed the energy in the electric field by equating it to the potential energy stored in a charge distribution that creates this electric field. But the end result doesn't depend on the charge distribution. This suggests that there should be a more direct way to arrive at (2.69) that only talks about fields and doesn't need charges. And there is. We will see it later.

But there is also another, more worrying problem with the derivation above. To illustrate this, let's just look at the simplest situation of a point particle. This has electric field

$$\mathbf{E} = \frac{q}{4\pi\epsilon_0 r^2} \hat{\mathbf{r}}. \quad (2.70)$$

So, by (2.69), the associated electric field should carry energy. But we started our derivation above by assuming that a single particle didn't carry any energy since it didn't take any work to put the particle there in the first place. What's going on?

Well, there was something of a sleight of hand in the derivation above. This occurs when we went from the expression $q\phi$ in (2.66) to $\rho\phi$ in (2.67). The former omits the “self-energy” terms; there is no contribution arising from $q_i\phi(\mathbf{r}_i)$. However, the latter includes them. The two expressions are not quite the same. This is also the reason that our final expression for the energy (2.69) is manifestly positive, while $q\phi$ can be positive or negative.

So which is right? Well, which form of the energy you use rather depends on the context. It is true that (2.69) is the correct expression for the energy stored in the electric field. But it is also true that you don't have to do any work to put the first charge in place since we're obviously not fighting against anything. Instead, the “self-energy” contribution coming from $\mathbf{E} \cdot \mathbf{E}$ in (2.70) should simply be thought of – using $E = mc^2$ – as a contribution to the mass of the particle.

We can easily compute this contribution for, say, an electron with charge $q = -e$. Let's call the radius of the electron a . Then the energy stored in its electric field is

$$\text{Energy} = \frac{\epsilon_0}{2} \int d^3r \mathbf{E} \cdot \mathbf{E} = \frac{e^2}{32\pi\epsilon_0} \int_a^\infty dr \frac{4\pi r^2}{r^4} = \frac{e^2}{8\pi\epsilon_0} \frac{1}{a}. \quad (2.71)$$

We see that, at least as far as the energy is concerned, we'd better not treat the electron as a point particle with $a \rightarrow 0$ or it will end up having infinite mass. And that will make it really hard to move.

So what is the radius of an electron? For the above calculation to be consistent, the energy in the electric field can't be greater than the observed mass of the electron m_e . In other words, we'd better have

$$m_e c^2 > \frac{e^2}{8\pi\epsilon_0} \frac{1}{a} \quad \implies \quad a > \frac{e^2}{8\pi\epsilon_0 m_e c^2}. \quad (2.72)$$

That, at least, puts a bound on the radius of the electron, which is the best we can do using classical physics alone. To give a more precise statement of the radius of the electron, we need to turn to quantum mechanics.

2.3.1.1 A Quick Foray into Quantum Electrodynamics

To assign a meaning of “radius” to seemingly point-like particles, we really need the machinery of quantum field theory. In that context, the size of the electron is called its *Compton wavelength*. This is the distance scale at which the electron gets surrounded by a swarm of electron-positron pairs which, roughly speaking, smears out the charge distribution. This distance scale is

$$a = \frac{\hbar}{m_e c}. \quad (2.73)$$

We see that the inequality (2.72) translates into an inequality on a bunch of fundamental constants. For the whole story to hang together, we require

$$\frac{e^2}{8\pi\epsilon_0\hbar c} < 1. \quad (2.74)$$

This is an almost famous combination of constants. It’s more usual to define the combination

$$\alpha = \frac{e^2}{4\pi\epsilon_0\hbar c}. \quad (2.75)$$

This is known as the *fine structure constant*. It is dimensionless and takes the value

$$\alpha \approx \frac{1}{137}. \quad (2.76)$$

Our discussion above requires $\alpha < 2$. We see that Nature happily meets this requirement.

2.3.2 The Force Between Electric Dipoles

As an application of our formula for electrostatic energy, we can compute the force between two, far separated dipoles. We place the first dipole, \mathbf{p}_1 , at the origin. It gives rise to a potential

$$\phi(\mathbf{r}) = \frac{1}{4\pi\epsilon_0} \frac{\mathbf{p}_1 \cdot \mathbf{r}}{r^3}. \quad (2.77)$$

Now, at some distance away, we place a second dipole. We’ll take this to consist of a charge Q at position \mathbf{r} and a charge $-Q$ at position $\mathbf{r} - \mathbf{d}$, with $d \ll r$. The resulting dipole moment is $\mathbf{p}_2 = Q\mathbf{d}$. We’re not interested in the energy stored in each individual dipole; only in the potential energy needed to bring the two dipoles together. This is given by (2.63),

$$\begin{aligned} U &= Q(\phi(\mathbf{r}) - \phi(\mathbf{r} - \mathbf{d})) = \frac{Q}{4\pi\epsilon_0} \left(\frac{\mathbf{p}_1 \cdot \mathbf{r}}{r^3} - \frac{\mathbf{p}_1 \cdot (\mathbf{r} - \mathbf{d})}{|\mathbf{r} - \mathbf{d}|^3} \right) \\ &= \frac{Q}{4\pi\epsilon_0} \left(\frac{\mathbf{p}_1 \cdot \mathbf{r}}{r^3} - \mathbf{p}_1 \cdot (\mathbf{r} - \mathbf{d}) \left(\frac{1}{r^3} + \frac{3\mathbf{d} \cdot \mathbf{r}}{r^5} + \dots \right) \right) \\ &= \frac{Q}{4\pi\epsilon_0} \left(\frac{\mathbf{p}_1 \cdot \mathbf{d}}{r^3} - \frac{3(\mathbf{p}_1 \cdot \mathbf{r})(\mathbf{d} \cdot \mathbf{r})}{r^5} \right), \end{aligned} \quad (2.78)$$

where, to get to the second line, we've Taylor expanded the denominator of the second term. This final expression can be written in terms of the second dipole moment. We find the nice, symmetric expression for the potential energy of two dipoles separated by distance \mathbf{r} ,

$$U = \frac{1}{4\pi\epsilon_0} \left(\frac{\mathbf{p}_1 \cdot \mathbf{p}_2}{r^3} - \frac{3(\mathbf{p}_1 \cdot \mathbf{r})(\mathbf{p}_2 \cdot \mathbf{r})}{r^5} \right). \quad (2.79)$$

But, we know from our first course on dynamics that the force between two objects is just given by $\mathbf{F} = -\nabla U$. We learn that the force between two dipoles is given by

$$\mathbf{F} = \frac{1}{4\pi\epsilon_0} \nabla \left(\frac{3(\mathbf{p}_1 \cdot \mathbf{r})(\mathbf{p}_2 \cdot \mathbf{r})}{r^5} - \frac{\mathbf{p}_1 \cdot \mathbf{p}_2}{r^3} \right). \quad (2.80)$$

The strength of the force, and even its sign, depends on the orientation of the two dipoles. If \mathbf{p}_1 and \mathbf{p}_2 lie parallel to each other and to \mathbf{r} then the resulting force is attractive. If \mathbf{p}_1 and \mathbf{p}_2 point in opposite directions, and lie parallel to \mathbf{r} , then the force is repulsive. The expression above allows us to compute the general force.

2.4 Conductors

Let's now throw something new into the mix. A *conductor* is a region of space which contains charges that are free to move. Physically, think “metal”. We want to ask what happens to the story of electrostatics in the presence of a conductor. There are a number of things that we can say straight away:

- Inside a conductor we must have $\mathbf{E} = 0$. If this isn't the case, the charges would move. But we're interested in electrostatic situations where nothing moves.
- Since $\mathbf{E} = 0$ inside a conductor, the electrostatic potential ϕ must be constant throughout the conductor.
- Since $\mathbf{E} = 0$ and $\nabla \cdot \mathbf{E} = \rho/\epsilon_0$, we must also have $\rho = 0$. This means that the interior of the conductor can't carry any charge.
- Conductors can be neutral, carrying both positive and negative charges which balance out. Alternatively, conductors can have net charge. In this case, any net charge must reside at the surface of the conductor.
- Since ϕ is constant, the surface of the conductor must be an equipotential. This means that any $\mathbf{E} = -\nabla\phi$ is perpendicular to the surface. This also fits nicely with the discussion above since any component of the electric field that lies tangential to the surface would make the surface charges move.
- If there is surface charge σ anywhere in the conductor then, by our previous discontinuity result (2.23), together with the fact that $\mathbf{E} = 0$ inside, the electric field just outside the conductor must be

$$\mathbf{E} = \frac{\sigma}{\epsilon_0} \hat{\mathbf{n}}. \quad (2.81)$$

Problems involving conductors are of a slightly different nature than those we've discussed up to now. The reason is that we don't know from the start where the charges are, so we don't know what charge distribution ρ that we should be solving for. Instead, the electric fields from other sources will cause the charges inside the conductor to shift around until they reach equilibrium in such a way that $\mathbf{E} = 0$ inside the conductor. In general, this will mean that even neutral conductors end up with some surface charge, negative in some areas, positive in others, just enough to generate an electric field inside the conductor that precisely cancels that due to external sources.

2.4.0.1 An Example: A Conducting Sphere

To illustrate the kind of problem that we have to deal with, it's probably best just to give an example. Consider a constant background electric field. (It could, for example, be generated by two charged plates of the kind we looked at in Subsection 2.1.4). Now place a neutral, spherical conductor inside this field. What happens?

We know that the conductor can't suffer an electric field inside it. Instead, the mobile charges in the conductor will move: the negative ones to one side; the positive ones to the other. The sphere now becomes *polarised*. These charges counteract the background electric field such that $\mathbf{E} = 0$ inside the conductor, while the electric field outside impinges on the sphere at right-angles. The end result must look qualitatively like Fig. 2.14.

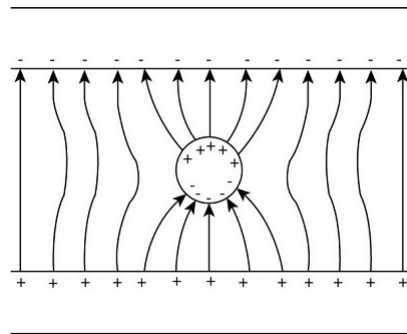


Fig. 2.14: A spherical conductor placed inside the field generated by two charged plates.

We'd like to understand how to compute the electric field in this, and related, situations. We'll give the answer in Subsection 2.4.4.

2.4.0.2 An Application: Faraday Cage

Consider some region of space that doesn't contain any charges, surrounded by a conductor. The conductor sits at constant $\phi = \phi_0$ while, since there are no charges inside, we must have $\nabla^2 \phi = 0$. But this means that $\phi = \phi_0$ everywhere. This is because, if it didn't then there would be a maximum or minimum of ϕ somewhere inside. And we know from the discussion in Subsection 2.2.5 that this can't happen. Therefore, inside a region surrounded by a conductor, we must have $\mathbf{E} = 0$.

This is a very useful result if you want to shield a region from electric fields. In this context, the surrounding conductor is called a *Faraday cage*. As an application, if you're worried that they're trying to read your mind with electromagnetic waves, then you need only wrap your head in tin foil and all concerns should be alleviated.

2.4.1 Capacitors

Let's now solve for the electric field in some conductor problems. The simplest examples are *capacitors*. These are a pair of conductors, one carrying charge Q , the other charge $-Q$.

2.4.1.1 Parallel Plate Capacitor

To start, we'll take the conductors to have flat, parallel surfaces as shown in the figure. We usually assume that the distance d between the surfaces is much smaller than \sqrt{A} , where A is the area of the surface. This means that we can neglect the effects that arise around the edge of plates and we're justified in assuming that the electric field between the two plates is the same as it would be if the plates were infinite in extent. The problem reduces to the same one that we considered in Subsection 2.1.4. The electric field necessarily vanishes inside the conductor while, between the plates we have the result (2.26),

$$\mathbf{E} = \frac{\sigma}{\epsilon_0} \hat{\mathbf{z}}, \quad (2.82)$$

where $\sigma = Q/A$ and we have assumed the plates are separated in the z -direction. We define the *capacitance* C to be

$$C = \frac{Q}{V}, \quad (2.83)$$

where V is the *voltage* or *potential difference* which is, as the name suggests, the difference in the potential ϕ on the two conductors. Since $E = -d\phi/dz$ is constant, we must have

$$\phi = -Ez + c \quad \implies \quad V = \phi(0) - \phi(d) = Ed = \frac{Qd}{A\epsilon_0}, \quad (2.84)$$

and the capacitance for parallel plates of area A , separated by distance d , is

$$C = \frac{A\epsilon_0}{d}. \quad (2.85)$$

Because V was proportional to Q , the charge has dropped out of our expression for the capacitance. Instead, C depends only on the geometry of the set-up. This is a general property; we will soon see another example.

Capacitors are usually employed as a method to store electrical energy. We can see how much. Using our result (2.69), we have

$$U = \frac{\epsilon_0}{2} \int d^3x \mathbf{E} \cdot \mathbf{E} = \frac{A\epsilon_0}{2} \int_0^d dz \left(\frac{\sigma}{\epsilon_0} \right)^2 = \frac{Q^2}{2C}. \quad (2.86)$$

This is the energy stored in a parallel plate capacitor.

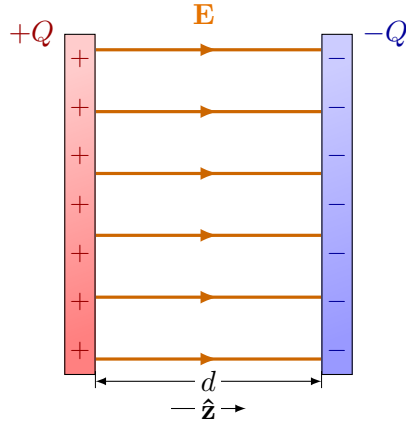


Fig. 2.15: Parallel plate capacitor.

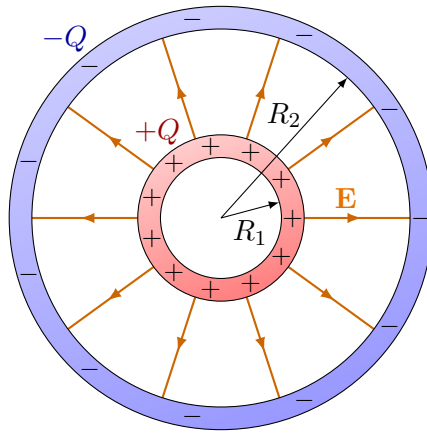


Fig. 2.16: Concentric sphere capacitor.

2.4.1.2 Concentric Sphere Capacitor

Consider a spherical conductor of radius R_1 . Around this we place another conductor in the shape of a spherical shell with inner surface lying at radius R_2 . We add charge $+Q$ to the sphere and $-Q$ to the shell. From our earlier discussion of charged spheres and shells, we know that the electric field between the two conductors must be

$$\mathbf{E} = \frac{Q}{4\pi\epsilon_0 r^2} \hat{\mathbf{r}}, \quad R_1 < r < R_2. \quad (2.87)$$

Correspondingly, the potential is

$$\phi = \frac{Q}{4\pi\epsilon_0 r}, \quad R_1 < r < R_2, \quad (2.88)$$

and the capacitance is given by $C = 4\pi\epsilon_0 R_1 R_2 / (R_2 - R_1)$.

2.4.2 Boundary Value Problems

Until now, we've thought of conductors as carrying some fixed charge Q . These conductors then sit at some constant potential ϕ . If there are other conductors in the vicinity that

carry a different charge then, as we've seen above, there will be some fixed potential difference, $V = \Delta\phi$ between them.

However, we can also think of a subtly different scenario. Suppose that we instead fix the potential ϕ in a conductor. This means that, whatever else happens, whatever other charges are doing all around, the conductor remains at a fixed ϕ . It never deviates from this value.

Now, this sounds a bit strange. We've seen above that the electric potential of a conductor depends on the distance to other conductors and also on the charge it carries. If ϕ remains constant, regardless of what objects are around it, then it must mean that the charge on the conductor is not fixed. And that's indeed what happens. Having conductors at fixed ϕ means that charge can flow in and out of the conductor. We implicitly assume that there is some background reservoir of charge which the conductor can dip into, taking and giving charge so that ϕ remains constant.

We can think of this reservoir of charge as follows: suppose that, somewhere in the background, there is a huge conductor with some charge Q which sits at some potential ϕ . To fix the potential of any other conductor, we simply attach it to one of this big reservoir-conductor. In general, some amount of charge will flow between them. The big conductor doesn't miss it, while the small conductor makes use of it to keep itself at constant ϕ .

The simplest example of the situation above arises if you connect your conductor to the planet Earth. By convention, this is taken to have $\phi = 0$ and it ensures that your conductor also sits at $\phi = 0$. Such conductors are said to be *grounded*. In practice, one may ground a conductor inside a chip in your cell phone by attaching it the metal casing.

Mathematically, we can consider the following problem. Take some number of objects, S_i . Some of the objects will be conductors at a fixed value of ϕ_i . Others will carry some fixed charge Q_i . This will rearrange itself into a surface charge σ_i such that $\mathbf{E} = 0$ inside while, outside the conductor, $\mathbf{E} = 4\pi\sigma\hat{\mathbf{n}}$. Our goal is to understand the electric field that threads the space between all of these objects. Since there is no charge sitting in this space, we need to solve the Laplace equation

$$\nabla^2\phi = 0, \tag{2.89}$$

subject to one of two boundary conditions

- Dirichlet Boundary Conditions: The value of ϕ is fixed on a given surface S_i .
- Neumann Boundary Conditions: The value of $\nabla\phi \cdot \hat{\mathbf{n}}$ is fixed perpendicular to a given surface S_i .

Notice that, for each S_i , we need to decide which of the two boundary conditions we want. We don't get to choose both of them. We then have that with either Dirichlet or Neumann boundary conditions chosen on each surface S_i , the Laplace equation has a unique solution. A proof is given in Appendix A.3.

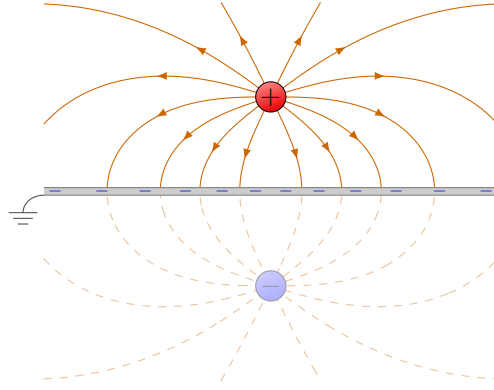


Fig. 2.17: A particle near a conducting plane looks like a dipole.

2.4.3 Method of Images

For particularly simple situations, there is a rather cute method that we can use to solve problems involving conductors. Although this technique is somewhat limited, it does give us some good intuition for what's going on. It's called the *method of images*.

2.4.3.1 A Charged Particle near a Conducting Plane

Consider a conductor which fills all of space $x < 0$. We'll ground this conductor so that $\phi = 0$ for $x < 0$. Then, at some point $x = d > 0$, we place a charge q . What happens?

We're looking for a solution to the Poisson equation with a delta-function source at $\mathbf{x} = \mathbf{d} = (d, 0, 0)$, together with the requirement that $\phi = 0$ on the plane $x = 0$. From our discussion in the previous section, there's a unique solution to this kind of problem. We just have to find it.

Here's the clever trick. Forget that there's a conductor at $x < 0$. Instead, suppose that there's a charge $-q$ placed opposite the real charge at $x = -d$. This is called the *image charge*. The potential for this pair of charges is just the potential

$$\phi = \frac{1}{4\pi\epsilon_0} \left(\frac{q}{\sqrt{(x-d)^2 + y^2 + z^2}} - \frac{q}{\sqrt{(x+d)^2 + y^2 + z^2}} \right). \quad (2.90)$$

By construction, this has the property that $\phi = 0$ for $x = 0$ and it has the correct source at $\mathbf{x} = (d, 0, 0)$. Therefore, this must be the right solution when $x \geq 0$. A cartoon of this is shown in Fig. 2.17. Of course, it's the wrong solution inside the conductor where the electric field vanishes. But that's trivial to fix: we just replace it with $\phi = 0$ for $x < 0$.

With the solution (2.90) in hand, we can now dispense with the image charge and explore what's really going on. We can easily compute the electric field from (2.90). If we focus on the electric field in the x -direction, it is

$$E_x = -\frac{\partial\phi}{\partial x} = \frac{q}{4\pi\epsilon_0} \left(\frac{x-d}{|\mathbf{r}-\mathbf{d}|^3} - \frac{x+d}{|\mathbf{r}+\mathbf{d}|^3} \right), \quad x \geq 0. \quad (2.91)$$

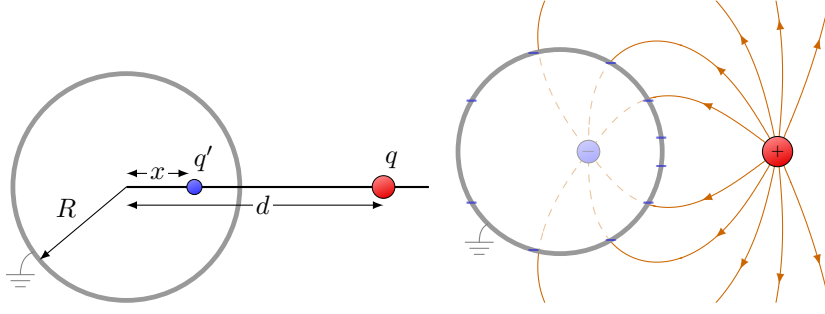


Fig. 2.18: A charged particle near a conducting sphere looks like an unbalanced dipole.

Meanwhile, $E_x = 0$ for $x < 0$. The discontinuity of E_x at the surface of the conductor determines the induced surface charge (2.81). It is

$$\sigma = E_x \epsilon_0|_{x=0} = -\frac{q}{2\pi} \frac{d}{(d^2 + y^2 + z^2)^{3/2}}. \quad (2.92)$$

We see that the surface charge is mostly concentrated on the plane at the point closest to the real charge. As you move away, it falls off as $1/(y^2 + z^2)^{3/2}$. We can compute the total induced surface charge by doing a simple integral,

$$q_{\text{induced}} = \int dy dz \sigma = -q. \quad (2.93)$$

The charge induced on the conductor is actually equal to the image charge. This is always true when we use the image charge technique.

Finally, as far as the real charge $+q$ is concerned, as long as it sits at $x > 0$, it feels an electric field which is identical in all respects to the field due to an image charge $-q$ embedded in the conductor. This means, in particular, that it will experience a force

$$\mathbf{F} = -\frac{q^2}{16\pi\epsilon_0 d^2} \hat{\mathbf{x}}. \quad (2.94)$$

This force is attractive, pulling the charge towards the conductor.

2.4.3.2 A Charged Particle Near a Conducting Sphere

We can play a similar game for a particle near a grounded, conducting sphere. The details are only slightly more complicated. We'll take the sphere to sit at the origin and have radius R . The particle has charge q and sits at $\mathbf{x} = \mathbf{d} = (d, 0, 0)$, with $d > R$. Our goal is to place an image charge q' somewhere inside the sphere so that $\phi = 0$ on the surface.

There is a way to derive the answer using conformal transformations. However, here we'll just state it. You should choose a particle of charge $q' = -$, placed at $x = R^2/d$ and, by symmetry, $y = z = 0$. A cartoon of this is shown in Fig. 2.18.

The resulting potential is

$$\phi = \frac{q}{4\pi\epsilon_0} \left(\frac{1}{\sqrt{(x-d)^2 + y^2 + z^2}} - \frac{R}{d} \frac{1}{\sqrt{(x-R^2/d)^2 + y^2 + z^2}} \right). \quad (2.95)$$

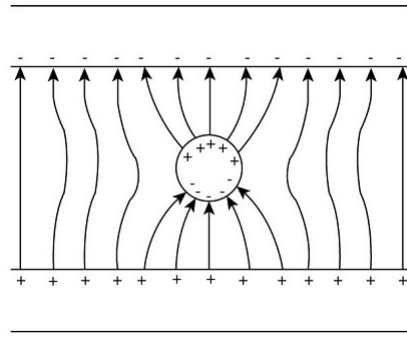


Fig. 2.19: A charged particle near a conducting sphere looks like an unbalanced dipole.

With a little algebra, you can check that $\phi = 0$ whenever $x^2 + y^2 + z^2 = R^2$. With a little more algebra, you can easily determine the induced surface charge and check that, when integrated over the sphere, we indeed have $q_{\text{induced}} = q'$. Once again, our charge experiences a force towards the conductor.

Above we've seen how to treat a grounded sphere. But what if we instead have an isolated conductor with some fixed charge, Q ? It's easy to adapt the problem above. We simply add the necessary excess charge $Q - q'$ as an image that sits at the origin of the sphere. This will induce an electric field which emerges radially from the sphere. Because of the principle of superposition, we just add this to the previous electric field and see that it doesn't mess up the fact that the electric field is perpendicular to the surface. This is now our solution.

2.4.4 Many Many More Problems

There are many more problems that you can cook up involving conductors, charges and electrostatics. Very few of them can be solved by the image charge method. Instead, you need to develop a number of basic tools of mathematical physics. A fairly comprehensive treatment of this can be found in the first 100 or so pages of Jackson.

For now, I would just like to leave you with the solution to the example that kicked off this section: what happens if you take a conducting sphere and place it in a constant electric field? This problem isn't quite solved by the image charge method. But it's solved by something similar: an image dipole.

We'll work in spherical polar coordinates and choose the original, constant electric field to point in the $\hat{\mathbf{z}}$ direction,

$$\mathbf{E}_0 = E_0 \hat{\mathbf{z}} \quad \implies \quad \phi_0 = -E_0 z = -E_0 r \cos \theta. \quad (2.96)$$

Take the conducting sphere to have radius R and be centered on the origin. Let's add to this an image dipole with potential (2.45). We'll place the dipole at the origin, and orient it along the \mathbf{z} axis like in Fig. 2.19.

The resulting potential is

$$\phi = -E_0 \left(r - \frac{R^3}{r^2} \right) \cos \theta. \quad (2.97)$$

Since we've added a dipole term, we can be sure that this still solves the Laplace equation outside the conductor (See Appendix A.2 for further details on the form of this solution). Moreover, by construction, $\phi = 0$ when $r = R$. This is all we wanted from our solution. The induced surface charge can again be computed by evaluating the electric field just outside

$$\sigma = -\epsilon_0 \frac{\partial \phi}{\partial r} = \epsilon_0 E_0 \left(1 + \frac{2R^3}{r^3} \right) \bigg|_{r=R} \cos \theta = 3\epsilon_0 E_0 \cos \theta. \quad (2.98)$$

We see that the surface charge is positive in one hemisphere and negative in the other. The total induced charge averages to zero.

CHAPTER 3

Magnetostatics

Charges give rise to electric fields. Current give rise to magnetic fields. In this chapter, we will study the magnetic fields induced by steady currents. This means that we are again looking for time independent solutions to the Maxwell equations. We will also restrict to situations in which the charge density vanishes, so $\rho = 0$. We can then set $\mathbf{E} = 0$ and focus our attention only on the magnetic field. We're left with two Maxwell equations to solve:

$$\nabla \times \mathbf{B} = \mu_0 \mathbf{J}, \quad (3.1)$$

and

$$\nabla \cdot \mathbf{B} = 0. \quad (3.2)$$

If you fix the current density \mathbf{J} , these equations have a unique solution. Our goal in this chapter is to find it.

3.0.0.1 Steady Currents

Before we solve (3.1) and (3.2), let's pause to think about the kind of currents that we're considering in this section. Because $\rho = 0$, there can't be any net charge. But, of course, we still want charge to be moving! This means that we necessarily have both positive and negative charges which balance out at all points in space. Nonetheless, these charges can move so there is a current even though there is no net charge transport.

This may sound artificial, but in fact it's exactly what happens in a typical wire. In that case, there is background of positive charge due to the lattice of ions in the metal. Meanwhile, the electrons are free to move. But they all move together so that at each point we still have $\rho = 0$. The continuity equation, which captures the conservation of electric charge, is

$$\frac{\partial \rho}{\partial t} + \nabla \cdot \mathbf{J} = 0. \quad (3.3)$$

Since the charge density is unchanging (and, indeed, vanishing), we have

$$\nabla \cdot \mathbf{J} = 0. \quad (3.4)$$

Mathematically, this is just saying that if a current flows into some region of space, an equal current must flow out to avoid the build up of charge. Note that this is consistent with (3.1) since, for any vector field, $\nabla \cdot (\nabla \times \mathbf{B}) = 0$.

3.1 Ampère's Law

The first equation of magnetostatics,

$$\nabla \times \mathbf{B} = \mu_0 \mathbf{J}, \quad (3.5)$$

is known as *Ampère's law*. As with many of these vector differential equations, there is an equivalent form in terms of integrals. In this case, we choose some open surface \mathcal{S} with boundary $\mathcal{C} = \partial\mathcal{S}$. Integrating (3.5) over the surface, we can use Stokes' theorem to turn the integral of $\nabla \times \mathbf{B}$ into a line integral over the boundary \mathcal{C} ,

$$\int_{\mathcal{S}} \nabla \times \mathbf{B} \cdot d\mathbf{S} = \oint_{\mathcal{C}} \mathbf{B} \cdot d\mathbf{r} = \mu_0 \int_{\mathcal{S}} \mathbf{J} \cdot d\mathbf{S}. \quad (3.6)$$

Recall that there's an implicit orientation in these equations. The surface \mathcal{S} comes with a normal vector $\hat{\mathbf{n}}$ which points away from \mathcal{S} in one direction. The line integral around the boundary is then done in the right-handed sense, meaning that if you stick the thumb of your right hand in the direction $\hat{\mathbf{n}}$ then your fingers curl in the direction of the line integral.

The integral of the current density over the surface \mathcal{S} is the same thing as the total current I that passes through \mathcal{S} . Ampère's law in integral form then reads

$$\oint_{\mathcal{C}} \mathbf{B} \cdot d\mathbf{r} = \mu_0 I. \quad (3.7)$$

For most examples, this isn't sufficient to determine the form of the magnetic field; we'll usually need to invoke (3.2) as well. However, there is one simple example where symmetry considerations mean that (3.7) is all we need.

3.1.1 A Long Straight Wire

Consider an infinite, straight wire carrying current I . We'll take it to point in the $\hat{\mathbf{z}}$ direction. The symmetry of the problem is jumping up and down telling us that we need to use cylindrical polar coordinates, (r, ϕ, z) , where $r = \sqrt{x^2 + y^2}$ is the radial distance away from the wire.

We take the open surface \mathcal{S} to lie in the $x - y$ plane, centered on the wire. For the line integral in (3.7) to give something that doesn't vanish, it's clear that the magnetic field has to have some component that lies along the circumference of the disc.

But, by the symmetry of the problem, that's actually the only component that \mathbf{B} can have: it must be of the form $\mathbf{B} = B(r)\hat{\phi}$. (If this was a bit too quick, we'll derive this more carefully soon). Any magnetic field of this form automatically satisfies the second Maxwell equation $\nabla \cdot \mathbf{B} = 0$. We need only worry about Ampère's law which tells us

$$\oint_{\mathcal{C}} \mathbf{B} \cdot d\mathbf{r} = B(r) \int_0^{2\pi} r d\phi = 2\pi r B(r) = \mu_0 I. \quad (3.8)$$

We see that the strength of the magnetic field is

$$\mathbf{B} = \frac{\mu_0 I}{2\pi r} \hat{\phi}. \quad (3.9)$$

The magnetic field circles the wire using the "right-hand rule": stick the thumb of your right hand in the direction of the current and your fingers curl in the direction of the magnetic field.

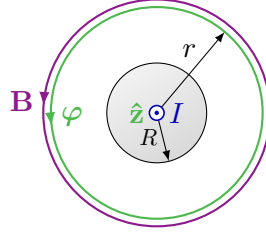


Fig. 3.1: An infinite, straight wire carrying current I and pointing in the \hat{z} direction.

Note that the simplest example of a magnetic field falls off as $1/r$. In contrast, the simplest example of an electric field – the point charge – falls off as $1/r^2$. You can trace this difference back to the geometry of the two situations. Because magnetic fields are sourced by currents, the simplest example is a straight line and the $1/r$ fall-off is because there are two transverse directions to the wire. Indeed, we saw in Subsection 2.1.3 that when we look at a line of charge, the electric field also drops off as $1/r$.

3.1.2 Surface Currents and Discontinuities

Consider the flat plane lying at $z = 0$ with a surface current density that we'll call \mathbf{K} . Note that \mathbf{K} is the current per unit length, as opposed to \mathbf{J} which is the current per unit area. You can think of the surface current as a bunch of wires, all lying parallel to each other.

We'll take the current to lie in the \hat{x} -direction: $\mathbf{K} = K\hat{x}$ as shown in Fig. 3.2.

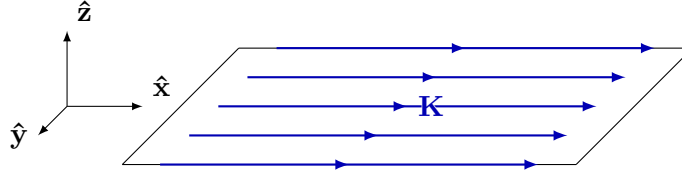


Fig. 3.2: A flat plane at $z = 0$ with a surface current density \mathbf{K} in the \hat{x} -direction.

From our previous result, we know that the \mathbf{B} field should curl around the current in the right-handed sense. But, with an infinite number of wires, this can only mean that \mathbf{B} is oriented along the \hat{y} direction. In fact, from the symmetry of the problem, it must look Fig. 3.3, with \mathbf{B} pointing in the $-\hat{y}$ direction when $z > 0$ and in the $+\hat{y}$ direction when

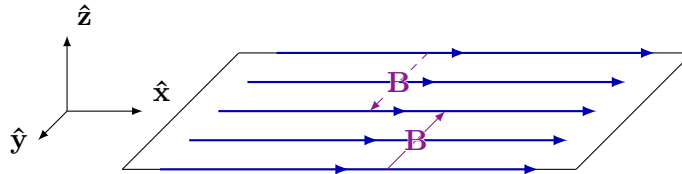


Fig. 3.3: Caption

$z < 0$. We write

$$\mathbf{B} = -B(z)\hat{y}, \quad (3.10)$$

with $B(z) = -B(-z)$. We invoke Ampère's law using the open surface in Fig. 3.4 with

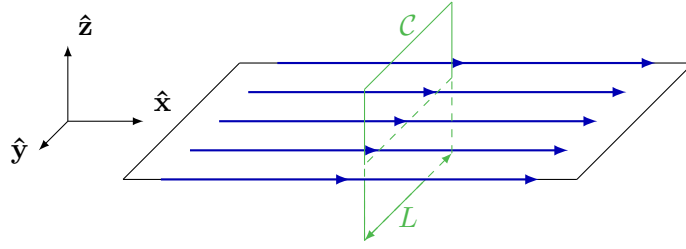


Fig. 3.4: A flat plane at $z = 0$ with a surface current density \mathbf{K} in the \hat{x} -direction.

length L in the y direction and extending to $\pm z$. We have

$$\oint_C \mathbf{B} \cdot d\mathbf{r} = LB(z) - LB(-z) = 2LB(z) = \mu_0 K L, \quad (3.11)$$

so we find that the magnetic field is constant above an infinite plane of surface current

$$B(z) = \frac{\mu_0 K}{2}, \quad z > 0. \quad (3.12)$$

This is rather similar to the case of the electric field in the presence of an infinite plane of surface charge.

The analogy with electrostatics continues. The magnetic field is not continuous across a plane of surface current. We have

$$B(z \rightarrow 0^+) - B(z \rightarrow 0^-) = \mu_0 K. \quad (3.13)$$

In fact, this is a general result that holds for any surface current \mathbf{K} . We can prove this statement by using the same curve that we used in Fig. 3.4 above and shrinking it until it barely touches the surface on both sides. If the normal to the surface is $\hat{\mathbf{n}}$ and \mathbf{B}_\pm denotes the magnetic field on either side of the surface, then

$$\hat{\mathbf{n}} \times \mathbf{B}_+ - \hat{\mathbf{n}} \times \mathbf{B}_- = \mu_0 \mathbf{K}. \quad (3.14)$$

Meanwhile, the magnetic field normal to the surface is continuous. (To see this, you can use a Gaussian pillbox, together with the other Maxwell equation $\nabla \cdot \mathbf{B} = 0$).

When we looked at electric fields, we saw that the normal component was discontinuous in the presence of surface charge (2.23) while the tangential component is continuous. For magnetic fields, it's the other way around: the tangential component is discontinuous in the presence of surface currents.

3.1.2.1 A Solenoid

A *solenoid* consists of a surface current that travels around a cylinder. It's simplest to think of a single current-carrying wire winding many times around the outside of the cylinder. (Strictly speaking, the cross-sectional shape of the solenoid doesn't have to be a circle - it can be anything. But we'll stick with a circle here for simplicity). To make life easy, we'll assume that the cylinder is infinitely long. This just means that we can neglect effects due to the ends.

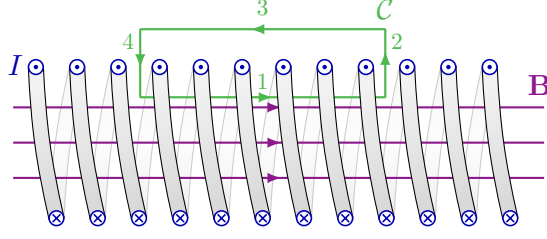


Fig. 3.5: Magnetic field \mathbf{B} inside a solenoid.

We'll again use cylindrical polar coordinates, (r, φ, z) , with the axis of the cylinder along $\hat{\mathbf{z}}$. By symmetry, we know that \mathbf{B} will point along the z -axis. Its magnitude can depend only on the radial distance: $\mathbf{B} = B(r)\hat{\mathbf{z}}$. Once again, any magnetic field of this form immediately satisfies $\nabla \cdot \mathbf{B} = 0$.

We solve Ampère's law in differential form. Anywhere other than the surface of the solenoid, we have $\mathbf{J} = 0$ and

$$\nabla \times \mathbf{B} = 0 \quad \implies \quad \frac{dB}{dr} = 0 \quad \implies \quad B(r) = \text{const.} \quad (3.15)$$

Outside the solenoid, we must have $B(r) = 0$ since $B(r)$ is constant and we know $B(r) \rightarrow 0$ as $r \rightarrow \infty$. To figure out the magnetic field inside the solenoid, we turn to the integral form of Ampère's law and consider the surface \mathcal{S} , bounded by the curve \mathcal{C} shown in Fig. 3.5. Only the line 1 that runs inside the solenoid contributes to the line integral. We have

$$\oint_{\mathcal{C}} \mathbf{B} \cdot d\mathbf{r} = BL = \mu_0 INL, \quad (3.16)$$

where N is the number of windings of wire per unit length. We learn that inside the solenoid, the constant magnetic field is given by

$$\mathbf{B} = \mu_0 IN \hat{\mathbf{z}}. \quad (3.17)$$

Note that, since $K = IN$, this is consistent with our general formula for the discontinuity of the magnetic field in the presence of surface currents (3.14).

3.2 The Vector Potential

For the simple current distributions of the last section, symmetry considerations were enough to lead us to a magnetic field which automatically satisfied

$$\nabla \cdot \mathbf{B} = 0. \quad (3.18)$$

But, for more general currents, this won't be the case. Instead we have to ensure that the second magnetostatic Maxwell equation is also satisfied.

In fact, this is simple to do. We are guaranteed a solution to $\nabla \cdot \mathbf{B} = 0$ if we write the magnetic field as the curl of some vector field,

$$\mathbf{B} = \nabla \times \mathbf{A} \quad (3.19)$$

Here \mathbf{A} is called the *vector potential*. While magnetic fields that can be written in the form (3.19) certainly satisfy $\nabla \cdot \mathbf{B} = 0$, the converse is also true; any divergence-free magnetic field can be written as (3.19) for some \mathbf{A} .

(Actually, this previous sentence is only true if our space has a suitably simple topology. Since we nearly always think of space as \mathbb{R}^3 or some open ball on \mathbb{R}^3 , we rarely run into subtleties. But if space becomes more interesting then the possible solutions to $\nabla \cdot \mathbf{B} = 0$ also become more interesting. This is analogous to the story of the electrostatic potential that we mentioned briefly in Section 2.2).

Using the expression (3.19), Ampère’s law becomes

$$\nabla \times \mathbf{B} = -\nabla^2 \mathbf{A} + \nabla(\nabla \cdot \mathbf{A}) = \mu_0 \mathbf{J}, \quad (3.20)$$

where, in the first equality, we’ve used a standard identity from vector calculus. This is the equation that we have to solve to determine \mathbf{A} and, through that, \mathbf{B} .

3.2.1 Magnetic Monopoles

Above, we dispatched with the Maxwell equation $\nabla \cdot \mathbf{B} = 0$ fairly quickly by writing $\mathbf{B} = \nabla \times \mathbf{A}$. But we never paused to think about what this equation is actually telling us. In fact, it has a very simple interpretation: it says that there are no magnetic charges. A point-like magnetic charge g would source the magnetic field, giving rise a $1/r^2$ fall-off

$$\mathbf{B} = \frac{g}{4\pi r^2} \hat{\mathbf{r}}. \quad (3.21)$$

An object with this behaviour is usually called a *magnetic monopole*. Maxwell’s equations says that they don’t exist. And we have never found one in Nature.

However, we could ask: how robust is this conclusion? Are we sure that magnetic monopoles don’t exist? After all, it’s easy to adapt Maxwell’s equations to allow for presence of magnetic charges: we simply need to change (3.18) to read $\nabla \cdot \mathbf{B} = \rho_m$ where ρ_m is the magnetic charge distribution. Of course, this means that we no longer get to use the vector potential \mathbf{A} . But is that such a big deal?

The twist comes when we turn to quantum mechanics. Because in quantum mechanics we’re *obliged* to use the vector potential \mathbf{A} . Not only is the whole framework of electromagnetism in quantum mechanics based on writing things using \mathbf{A} , but it turns out that there are experiments that actually detect certain properties of \mathbf{A} that are lost when we compute $\mathbf{B} = \nabla \times \mathbf{A}$. I won’t explain the details here, but if you’re interested then look up the “Aharonov-Bohm effect” in the lectures on Solid State Physics.

3.2.1.1 Monopoles After All?

To summarise, magnetic monopoles have never been observed. We have a law of physics (3.18) which says that they don’t exist. And when we turn to quantum mechanics we need

to use the vector potential \mathbf{A} which automatically means that (3.18) is true. It sounds like we should pretty much forget about magnetic monopoles, right?

Well, no. There are actually very good reasons to suspect that magnetic monopoles do exist. The most important part of the story is due to Dirac. He gave a beautiful argument which showed that it is in fact possible to introduce a vector potential \mathbf{A} which allows for the presence of magnetic charge, but only if the magnetic charge g is related to the charge of the electron e by

$$ge = 2\pi\hbar n, \quad \text{for } n \in \mathbb{Z}. \quad (3.22)$$

This is known as the *Dirac quantisation condition*.

Moreover, following work in the 1970s by 't Hooft and Polyakov, we now realise that magnetic monopoles are ubiquitous in theories of particle physics. Our best current theory – the Standard Model – does not predict magnetic monopoles. But every theory that tries to go beyond the Standard Model, whether Grand Unified Theories, or String Theory or whatever, always ends up predicting that magnetic monopoles should exist. They're one of the few predictions for new physics that nearly all theories agree upon.

These days most theoretical physicists think that magnetic monopoles probably exist and there have been a number of experiments around the world designed to detect them. However, while theoretically monopoles seem like a good bet, their future observational status is far from certain. We don't know how heavy magnetic monopoles will be, but all evidence suggests that producing monopoles is beyond the capabilities of our current (or, indeed, future) particle accelerators. Our only hope is to discover some that Nature made for us, presumably when the Universe was much younger. Unfortunately, here too things seem against us. Our best theories of cosmology, in particular inflation, suggest that any monopoles that were created back in the Big Bang have long ago been diluted. At a guess, there are probably only a few floating around our entire observable Universe. The chances of one falling into our laps seem slim. But I hope I'm wrong.

3.2.2 Gauge Transformations

The choice of \mathbf{A} in (3.19) is far from unique: there are lots of different vector potentials \mathbf{A} that all give rise to the same magnetic field \mathbf{B} . This is because the curl of a gradient is automatically zero. This means that we can always add any vector potential of the form $\nabla\chi$ for some function χ and the magnetic field remains the same,

$$\mathbf{A}' = \mathbf{A} + \nabla\chi \quad \implies \quad \nabla \times \mathbf{A}' = \nabla \times \mathbf{A}. \quad (3.23)$$

Such a change of \mathbf{A} is called a *gauge transformation*. As we will see in Subsection 5.0.1, it is closely tied to the possible shifts of the electrostatic potential ϕ . Ultimately, such gauge transformations play a key role in theoretical physics. But, for now, we're simply going to use this to our advantage. Because, by picking a cunning choice of χ , it's possible to simplify our quest for the magnetic field.

We can always find a gauge transformation χ such that \mathbf{A}' satisfies $\nabla \cdot \mathbf{A}' = 0$. Making this choice is usually referred to as *Coulomb gauge*. Suppose that we've found some \mathbf{A} which gives us the magnetic field that we want, so $\nabla \times \mathbf{A} = \mathbf{B}$, but when we take the

divergence we get some function $\nabla \cdot \mathbf{B} = \psi(\mathbf{x})$. We instead choose $\mathbf{A}' = \mathbf{A} + \nabla\chi$ which now has divergence

$$\nabla \cdot \mathbf{A}' = \nabla \cdot \mathbf{A} + \nabla^2\chi = \psi + \nabla^2\chi. \quad (3.24)$$

So if we want $\nabla \cdot \mathbf{A}' = 0$, we just have to pick our gauge transformation χ to obey

$$\nabla^2\chi = -\psi. \quad (3.25)$$

But this is just the Poisson equation again. And we know from our discussion in Section 2.2 that there is always a solution. (For example, we can write it down in integral form using the Green's function).

3.2.2.1 Something a Little Misleading: The Magnetic Scalar Potential

There is another quantity that is sometimes used called the *magnetic scalar potential*, Ω . The idea behind this potential is that you might be interested in computing the magnetic field in a region where there are no currents and the electric field is not changing with time. In this case, you need to solve $\nabla \times \mathbf{B} = 0$, which you can do by writing

$$\mathbf{B} = -\nabla\Omega. \quad (3.26)$$

Now calculations involving the magnetic field really do look identical to those involving the electric field.

However, you should be wary of writing the magnetic field in this way. As we'll see in more detail in Subsection 5.0.1, we can *always* solve two of Maxwell's equations by writing \mathbf{E} and \mathbf{B} in terms of the electric potential ϕ and vector potential \mathbf{A} and this formulation becomes important as we move onto more advanced areas of physics. In contrast, writing $\mathbf{B} = -\nabla\Omega$ is only useful in a limited number of situations. The reason for this really gets to the heart of the difference between electric and magnetic fields: electric charges exist; magnetic charges don't!

3.2.3 Biot-Savart Law

We're now going to use the vector potential to solve for the magnetic field \mathbf{B} in the presence of a general current distribution. From now, we'll always assume that we're working in Coulomb gauge and our vector potential obeys $\nabla \cdot \mathbf{A} = 0$. Then Ampère's law (3.20) becomes a whole lot easier: we just have to solve

$$\nabla^2\mathbf{A} = -\mu_0\mathbf{J}. \quad (3.27)$$

But this is just something that we've seen already. To see why, it's perhaps best to write it out in Cartesian coordinates. This then becomes three equations,

$$\nabla^2 A_i = -\mu_0 J_i, \quad (i = 1, 2, 3) \quad (3.28)$$

and each of these is the Poisson equation.

It's worth giving a word of warning at this point: the expression $\nabla^2\mathbf{A}$ is simple in Cartesian coordinates where, as we've seen above, it reduces to the Laplacian on each

component. But, in other coordinate systems, this is no longer true. The Laplacian now also acts on the basis vectors such as $\hat{\mathbf{r}}$ and $\hat{\boldsymbol{\varphi}}$. So in these other coordinate systems, $\nabla^2 \mathbf{A}$ is a little more of a mess. (You should probably use the identity $\nabla^2 \mathbf{A} = -\nabla \times (\nabla \times \mathbf{A}) + \nabla(\nabla \cdot \mathbf{A})$ if you really want to compute in these other coordinate systems).

Anyway, if we stick to Cartesian coordinates then everything is simple. In fact, the resulting equations (3.28) are of exactly the same form that we had to solve in electrostatics. And, in analogy to (2.48), we know how to write down the most general solution using Green's functions. It is

$$A_i(x) = \frac{\mu_0}{4\pi} \int_{\mathcal{V}} d^3x' \frac{J_i(\mathbf{x}')}{|\mathbf{x} - \mathbf{x}'|}. \quad (3.29)$$

Or, if you're feeling bold, you can revert back to vector notation and write

$$\mathbf{A}(\mathbf{x}) = \frac{\mu_0}{4\pi} \int_{\mathcal{V}} d^3x' \frac{\mathbf{J}(\mathbf{x}')}{|\mathbf{x} - \mathbf{x}'|}. \quad (3.30)$$

where you've just got to remember that the vector index on \mathbf{A} links up with that on \mathbf{J} (and not on \mathbf{x} or \mathbf{x}').

3.2.3.1 Checking Coulomb Gauge

We've derived a solution to (3.27), but this is only a solution to Ampère's equation (3.20) if the resulting \mathbf{A} obeys the Coulomb gauge condition, $\nabla \cdot \mathbf{A} = 0$. Let's now check that it does. We have

$$\nabla \cdot \mathbf{A}(\mathbf{x}) = \frac{\mu_0}{4\pi} \int_{\mathcal{V}} d^3x' \nabla \cdot \left(\frac{\mathbf{J}(\mathbf{x}')}{|\mathbf{x} - \mathbf{x}'|} \right), \quad (3.31)$$

where you need to remember that the index of ∇ is dotted with the index of \mathbf{J} , but the derivative in ∇ is acting on \mathbf{x} , not on \mathbf{x}' as that is the dummy variable of integration. We can write

$$\begin{aligned} \nabla \cdot \mathbf{A}(\mathbf{x}) &= \frac{\mu_0}{4\pi} \int_{\mathcal{V}} d^3x' \mathbf{J}(\mathbf{x}') \cdot \nabla \left(\frac{1}{|\mathbf{x} - \mathbf{x}'|} \right) \\ &= -\frac{\mu_0}{4\pi} \int_{\mathcal{V}} d^3x' \mathbf{J}(\mathbf{x}') \cdot \nabla' \left(\frac{1}{|\mathbf{x} - \mathbf{x}'|} \right). \end{aligned} \quad (3.32)$$

Here we've done something clever. Now our ∇' is differentiating with respect to \mathbf{x}' . To get this, we've used the fact that if you differentiate $1/|\mathbf{x} - \mathbf{x}'|$ with respect to \mathbf{x} then you get the negative of the result from differentiating with respect to \mathbf{x}' . But since ∇' sits inside an $\int d^3x'$ integral, it's ripe for integrating by parts. This gives

$$\text{div} \mathbf{A}(\mathbf{x}) = -\frac{\mu_0}{4\pi} \int_{\mathcal{V}} d^3x' \left[\nabla' \cdot \left(\frac{\mathbf{J}(\mathbf{x}')}{|\mathbf{x} - \mathbf{x}'|} \right) - \nabla' \cdot \mathbf{J}(\mathbf{x}') \left(\frac{1}{|\mathbf{x} - \mathbf{x}'|} \right) \right]. \quad (3.33)$$

The second term vanishes because we're dealing with steady currents obeying $\nabla \cdot \mathbf{J} = 0$. The first term also vanishes if we take the current to be localised in some region of space, $\hat{\mathcal{V}} \subset \mathcal{V}$ so that $\mathbf{J}(\mathbf{x}) = 0$ on the boundary $\partial\mathcal{V}$. We'll assume that this is the case. We conclude that

$$\nabla \cdot \mathbf{A} = 0, \quad (3.34)$$

and (3.30) is indeed the general solution to the Maxwell equations (3.1) and (3.2) as we'd hoped.

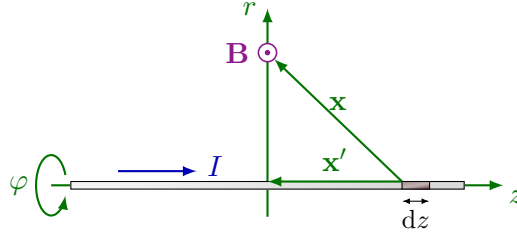


Fig. 3.6: An infinite, straight wire carrying current I and pointing in the $\hat{\mathbf{z}}$ direction

3.2.3.2 The Magnetic Field

From the solution (3.30), it is simple to compute the magnetic field $\mathbf{B} = \nabla \times \mathbf{A}$. Again, we need to remember that the ∇ acts on the \mathbf{x} in (3.30) rather than the \mathbf{x}' . We find

$$\mathbf{B}(\mathbf{x}) = \frac{\mu_0}{4\pi} \int_{\mathcal{V}} d^3x' \frac{\mathbf{J}(\mathbf{x}') \times (\mathbf{x} - \mathbf{x}')}{|\mathbf{x} - \mathbf{x}'|^3}. \quad (3.35)$$

This is known as the *Biot-Savart law*. It describes the magnetic field due to a general current density.

There is a slight variation on (3.35) which more often goes by the name of the Biot-Savart law. This arises if the current is restricted to a thin wire which traces out a curve \mathcal{C} . Then, for a current density \mathbf{J} passing through a small volume $\delta\mathcal{V}$, we write $\mathbf{J}\delta\mathcal{V} = (JA)\delta\mathbf{x}$ where A is the cross-sectional area of the wire and $\delta\mathbf{x}$ lies tangent to \mathcal{C} . Assuming that the cross-sectional area is constant throughout the wire, the current $I = JA$ is also constant. The Biot-Savart law becomes

$$\mathbf{B}(\mathbf{x}) = \frac{\mu_0}{4\pi} \int_{\mathcal{C}} \frac{d\mathbf{x}' \times (\mathbf{x} - \mathbf{x}')}{|\mathbf{x} - \mathbf{x}'|^3}. \quad (3.36)$$

This describes the magnetic field due to the current I in the wire.

3.2.3.3 An Example: The Straight Wire Revisited

Of course, we already derived the answer for a straight wire in (3.9) without using this fancy vector potential technology. Before proceeding, we should quickly check that the Biot-Savart law reproduces our earlier result. As before, we'll work in cylindrical polar coordinates. We take the wire to point along the $\hat{\mathbf{z}}$ axis and use $r^2 = x^2 + y^2$ as our radial coordinate. This means that the line element along the wire is parametrised by $d\mathbf{x}' = \hat{\mathbf{z}} dz$ and, for a point \mathbf{x} away from the wire, the vector $d\mathbf{x}' \times (\mathbf{x} - \mathbf{x}')$ points along the tangent to the circle of radius r ,

$$d\mathbf{x}' \times (\mathbf{x} - \mathbf{x}') = r\hat{\boldsymbol{\phi}} dz. \quad (3.37)$$

So we have

$$\mathbf{B} = \frac{\mu_0 I \hat{\boldsymbol{\phi}}}{4\pi} \int_{-\infty}^{+\infty} dz \frac{r}{(r^2 + z^2)^{3/2}} = \frac{\mu_0 I}{2\pi r} \hat{\boldsymbol{\phi}}, \quad (3.38)$$

which is the same result we found earlier (3.9).

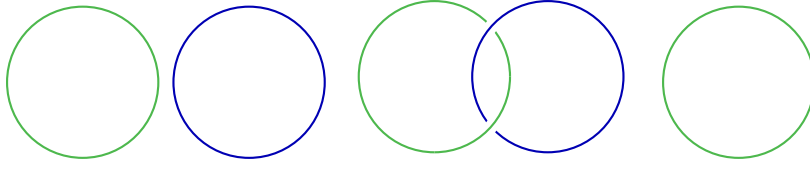


Fig. 3.7: Curves with linking number $n = 0$, $n = 1$ and $n = 2$.

3.2.4 A Mathematical Diversion: The Linking Number

There's a rather cute application of these ideas to pure mathematics. Consider two closed, non-intersecting curves, \mathcal{C} and \mathcal{C}' , in \mathbb{R}^3 . For each pair of curves, there is an integer $n \in \mathbb{Z}$ called the *linking number* which tells you how many times one of the curves winds around the other. For example, here are pairs of curves with linking number $|n| = 0, 1$ and 2 .

To determine the sign of the linking number, we need to specify the orientation of each curve. In the last two figures above, the linking numbers are negative, if we traverse both red and blue curves in the same direction. The linking numbers are positive if we traverse one curve in a clockwise direction, and the other in an anti-clockwise direction.

Importantly, the linking number doesn't change as you deform either curve, provided that the two curves never cross. In fancy language, the linking number is an example of a topological invariant.

There is an integral expression for the linking number, first written down by Gauss during his exploration of electromagnetism. The Biot-Savart formula (3.36) offers a simple physics derivation of Gauss' expression. Suppose that the curve \mathcal{C} carries a current I . This sets up a magnetic field everywhere in space. We will then compute $\oint_{\mathcal{C}'} \mathbf{B} \cdot d\mathbf{x}'$ around another curve \mathcal{C}' . (If you want a justification for computing $\oint_{\mathcal{C}'} \mathbf{B} \cdot d\mathbf{x}'$ then you can think of it as the work done when transporting a magnetic monopole of unit charge around \mathcal{C} , but this interpretation isn't necessary for what follows.) The Biot-Savart formula gives

$$\oint_{\mathcal{C}'} \mathbf{B}(\mathbf{x}') \cdot d\mathbf{x}' = \frac{\mu_0 I}{4\pi} \oint_{\mathcal{C}'} d\mathbf{x}' \cdot \oint_{\mathcal{C}} \frac{d\mathbf{x} \times (\mathbf{x}' - \mathbf{x})}{|\mathbf{x} - \mathbf{x}'|^3}, \quad (3.39)$$

where we've changed our conventions somewhat from (3.36): now \mathbf{x} labels coordinates on \mathcal{C} while \mathbf{x}' labels coordinates on \mathcal{C}' .

Meanwhile, we can also use Stokes' theorem, followed by Ampère's law, to write

$$\oint_{\mathcal{C}'} \mathbf{B}(\mathbf{x}') \cdot d\mathbf{x}' = \int_{S'} (\nabla \times \mathbf{B}) \cdot d\mathbf{S} = \mu_0 \int_{S'} \mathbf{J} \cdot d\mathbf{S}, \quad (3.40)$$

where S' is a surface bounded by \mathcal{C}' . The current is carried by the other curve, \mathcal{C} , which pierces S' precisely n times, so that

$$\oint_{\mathcal{C}'} \mathbf{B}(\mathbf{x}') \cdot d\mathbf{x}' = \mu_0 \int_{S'} \mathbf{J} \cdot d\mathbf{S} = n\mu_0 I. \quad (3.41)$$

Comparing the two equations above, we arrive at Gauss' double-line integral expression for the linking number n ,

$$n = \frac{1}{4\pi} \oint_{\mathcal{C}'} d\mathbf{x}' \cdot \oint_{\mathcal{C}} \frac{d\mathbf{x} \times (\mathbf{x}' - \mathbf{x})}{|\mathbf{x} - \mathbf{x}'|^3}. \quad (3.42)$$

Note that our final expression is symmetric in \mathcal{C} and \mathcal{C}' , even though these two curves played a rather different physical role in the original definition, with \mathcal{C} carrying a current, and \mathcal{C}' the path traced by some hypothetical monopole. To see that the expression is indeed symmetric, note that the triple product can be thought of as the determinant $\det(\mathbf{x}', \mathbf{x}, \mathbf{x}' - \mathbf{x})$. Swapping \mathbf{x} and \mathbf{x}' changes the order of the first two vectors and changes the sign of the third, leaving the determinant unaffected.

The formula (3.42) is rather pretty. It's not at all obvious that the right-hand-side doesn't change under (non-crossing) deformations of \mathcal{C} and \mathcal{C}' ; nor is it obvious that the right-hand-side must give an integer. Yet both are true, as the derivation above shows. This is the first time that ideas of topology sneak into physics. It's not the last.

3.3 Units of Electromagnetism

CHAPTER 4

Electrodynamics

CHAPTER 5

Electromagnetism and Relativity

5.0.1 Gauge Invariance and Relativity

CHAPTER 6

Light Scattering

CHAPTER 7

Special Relativity

CHAPTER 8

Radiation and Relativistic Electrodynamics

Appendix

A.1 Green's Functions

A.2 Laplace and Poisson's equations

A.3 uniqueness

p.43/236 However, if we've picked Dirichlet boundary conditions then $f = 0$ on the boundary, while Neumann boundary conditions ensure that $\nabla f = 0$ on the boundary. This means that the integral vanishes and, from (??), we must have $\nabla f = 0$ throughout space. But if we have imposed Dirichlet boundary conditions somewhere, then $f = 0$ on that boundary and so $f = 0$ everywhere. Alternatively, if we have Neumann boundary conditions on all surfaces than $\nabla f = 0$ everywhere and the two solutions ϕ_1 and ϕ_2 can differ only by a constant. But, as discussed in Section 2.2, this constant has no physical meaning.