Electromagnetism

James Sparks & Márk Mezei

Hilary Term 2026

$\nabla \cdot \mathbf{E} = \dfrac{\rho}{\epsilon_0}$ (Gauss' law)	$\nabla \cdot \mathbf{B} = 0$ (no magnetic monopoles)
$\nabla \times \mathbf{E} = -\dfrac{\partial \mathbf{B}}{\partial t}$ (Faraday's law)	$\nabla \times \mathbf{B} = \mu_0 \left(\mathbf{J} + \epsilon_0 \dfrac{\partial \mathbf{E}}{\partial t} \right)$ (Ampère-Maxwell law)

About these notes

These are lecture notes for the B7.2 Electromagnetism course, which is a third year course in the mathematics syllabus at the University of Oxford. Starred sections/paragraphs are not examinable, either because the material is slightly off-syllabus, or because it is more difficult. There are four problem sheets. Please send any questions/corrections/comments to mark.mezei@maths.ox.ac.uk.

These notes are largely unaltered from the version written by James Sparks for Hilary term 2022, built on earlier material from Xenia de la Ossa, Fernando Alday, and Paul Tod, and with additions by Erik Panzer.

Contents

Reading list

The lecture notes have grown too long over the years. I have decided to move some material that is either tangental or too advanced into a separate document, and I have expanded the discussion of the core material in the notes. I recommend that you read this additional document as a first step in going beyond the lecture notes.

There are many outstanding books on electromagnetism, often going into the subject in much more depth than we will have time for, especially physical applications. Those who wish to learn more are encouraged to dip into the following:

- R. P. Feynman, R. B. Leighton, M. Sands, *The Feynman Lectures on Physics, Volume 2: Electromagnetism*, Addison-Wesley. Online: `https://feynmanlectures.caltech.edu/`
- D. J. Griffiths, *Introduction to Electrodynamics*, Pearson.
- J. D. Jackson, *Classical Electrodynamics*, John Wiley.

The Feynman lectures are superb, especially for physical insight and applications. The book by Griffiths is perhaps closest to this course. The book by Jackson is comprehensive, although is closer to a graduate level text.

Preamble

In this course we take a first look at the *classical theory* of electromagnetism. Historically, this begins with Coulomb's inverse square law force between stationary point charges, dating from 1785, and culminates (for us, at least) with Maxwell's formulation of electromagnetism in his 1864 paper, *A Dynamical Theory of the Electromagnetic Field*. It was in this paper that the electromagnetic wave equation was first written down, and in which Maxwell first proposed that *"light is an electromagnetic disturbance propagated through the field according to electromagnetic laws"*. Maxwell's equations, which appear on the front page of these lecture notes, govern a diverse array of physical phenomena, and are valid over an enormous range of scales. It is the electromagnetic force that holds the negatively charged electrons in orbit around the positively charged nucleus of an atom.[1] Interactions between atoms and molecules are also electromagnetic, so that chemical forces are really electromagnetic forces. The electromagnetic force is then essentially responsible for almost all physical phenomena encountered in day-to-day experience, with the exception of

[1] Quantum mechanics also plays an important role.

gravity: friction, electricity (in homes, laptops, mobile phones, *etc*), electric motors, permanent magnets, electromagnets, lightning, electromagnetic radiation (radio waves, microwaves, X-rays, *etc*, as well as visible light), ... it's all electromagnetism.

Classical electromagnetism is an application of the three-dimensional vector calculus you learned in Prelims: div, grad, curl, and the Stokes and divergence theorems. Appendix A summarizes the main definitions and results, and I strongly encourage you to take a look at this at the start of the course. We'll then take a usual, fairly historical, route, starting with Coulomb's law in electrostatics, and eventually building up to Maxwell's equations on the front page. The disadvantage of this is that you'll begin by learning special cases of Maxwell's equations – having learned one equation, you will later find that more generally there are other terms in it. On the other hand, simply starting with Maxwell's equations and then deriving everything else from them is probably too abstract, and doesn't really give a feel for where the equations have come from. My advice is that after every few lectures you should take another look at the equations on the front page – each time you should find that you understand better what they mean.

> *From a long view of the history of mankind – seen from, say, ten thousand years from now – there can be little doubt that the most significant event of the 19th century will be judged as Maxwell's discovery of the laws of electrodynamics* – Richard Feynman

1 Electrostatics

1.1 Point charges and Coulomb's law

It is a fact of nature that elementary particles have a property called *electric charge*. In SI units[2] this is measured in *Coulombs* C, and the electron and proton carry equal and opposite charges $\mp e$, where $e \approx 1.602 \times 10^{-19}$ C.[3] Atoms consist of electrons orbiting a nucleus of protons and neutrons (with the latter carrying charge 0), and thus all charges in stable matter, made of atoms, arise from these electron and proton charges.

Electrostatics is the study of charges *at rest*. We model space by $\mathbb{R}^3$, or a subset thereof, and represent the position of a stationary point charge q by the position vector $\mathbf{r} \in \mathbb{R}^3$.

Coulomb's law Given two charges, q_1, q_2 at positions $\mathbf{r}_1$, $\mathbf{r}_2 \in \mathbb{R}^3$, respectively, the first charge experiences an *electrical force* $\mathbf{F}_1$ due to the second charge given by

$$\mathbf{F}_1 = \frac{1}{4\pi\epsilon_0} \frac{q_1 q_2}{|\mathbf{r}_1 - \mathbf{r}_2|^3} (\mathbf{r}_1 - \mathbf{r}_2) . \tag{1.1}$$

Note this only makes sense if $\mathbf{r}_1 \neq \mathbf{r}_2$, so that the charges are not on top of each other, which we thus assume. The constant ϵ_0 is called the *permittivity of free space*, which in SI units takes the value $\epsilon_0 \simeq 8.854 \times 10^{-12}\,\mathrm{C}^2\,\mathrm{N}^{-1}\,\mathrm{m}^{-2}$. Notice that by symmetry the second charge experiences an electrical force $\mathbf{F}_2$, due to the first charge, where $\mathbf{F}_2$ is given by the right hand side of (1.1) with the subscripts 1 and 2 swapped. In particular $\mathbf{F}_2 = -\mathbf{F}_1$, and Newton's third law is obeyed.

Without loss of generality, we might as well put the second charge at the origin $\mathbf{r}_2 = \mathbf{0}$, denote $\mathbf{r}_1 = \mathbf{r}$, $q_2 = q$, and equivalently rewrite (1.1) as

$$\mathbf{F}_1 = \frac{1}{4\pi\epsilon_0} \frac{q_1 q}{r^2} \hat{\mathbf{r}} , \tag{1.2}$$

where $\hat{\mathbf{r}} \equiv \mathbf{r}/r$ is a unit vector and $r \equiv |\mathbf{r}|$. This is *Coulomb's law* of electrostatics, and is an experimental fact. Note that:

> ***E1***: *The force is proportional to the product of the charges, so that opposite (different sign) charges attract, while like (same sign) charges repel.*
>
> ***E2***: *The force acts in the direction of the vector joining the two charges, and is inversely proportional to the square of the distance of separation.*

The above two laws of electrostatics are equivalent to Coulomb's law. Notice that if (1.2) is the *only* force acting on the first charge, by Newton's second law of motion it will necessarily accelerate and begin to move, and we are then no longer dealing with statics. We'll get a feel for electrostatics problems as the next two sections develop, and look at charges in motion starting in section 3.

The final law of electrostatics says what happens when there are more than two charges:

[2] where, for example, distance is measured in metres, time is measured in seconds, force is measured in Newtons.
[3] Since 2019, the Coulomb is *defined* such that the elementary charge is *exactly* $e = 1.602\,176\,634 \cdot 10^{-19}$ C.

E3: *Electrostatic forces obey the* Principle of Superposition.

This means that if we have N charges q_i at positions $\mathbf{r}_i$, $i = 1, \ldots, N$, then an additional charge q at position $\mathbf{r}$ experiences a force

$$\mathbf{F} = \sum_{i=1}^{N} \frac{1}{4\pi\epsilon_0} \frac{qq_i}{|\mathbf{r} - \mathbf{r}_i|^3} (\mathbf{r} - \mathbf{r}_i) \ . \tag{1.3}$$

That is, to get the total force on charge q due to all the other charges, we simply *add up* (superpose) the Coulomb force (1.1) from each charge q_i.

1.2 The electric field

The following looks trivial at first sight, but in fact it's an ingenious shift of viewpoint:

Definition Given a particular distribution of charges, as above, we define the *electric field* $\mathbf{E} = \mathbf{E}(\mathbf{r})$ to be the force on a *unit test charge* (*i.e.* $q = 1$) placed at position $\mathbf{r}$.

Here the nomenclature "test charge" indicates that the charge is not regarded as part of the distribution of charges that it is "probing". The force in (1.3) is thus

$$\mathbf{F} = q\,\mathbf{E} \ , \tag{1.4}$$

where by definition

$$\mathbf{E}(\mathbf{r}) = \frac{1}{4\pi\epsilon_0} \sum_{i=1}^{N} \frac{q_i}{|\mathbf{r} - \mathbf{r}_i|^3} (\mathbf{r} - \mathbf{r}_i) \tag{1.5}$$

is the electric field produced by the N charges. It is a vector field (here defined on $\mathbb{R}^3 \backslash \{\mathbf{r}_1, \ldots, \mathbf{r}_N\}$), depending on position $\mathbf{r}$.

As we have defined it, the electric field is just a mathematically convenient way of describing the force a unit test charge would feel if placed in some position in a fixed background of charges. In fact, the electric field will turn out to be a fundamental object in electromagnetic theory. Notice that $\mathbf{E}$ also satisfies the Principle of Superposition, and that it is measured in $\mathrm{N\,C^{-1}}$.

1.3 Gauss' law

Definition Given a surface $\Sigma \subset \mathbb{R}^3$ with *outward* unit normal vector $\mathbf{n}$, the integral $\int_\Sigma \mathbf{E} \cdot \mathbf{n}\,\mathrm{d}S$ is called the *flux* of the electric field $\mathbf{E}$ through Σ.[4]

Here one often uses the notation $\mathrm{d}\mathbf{S}$ for $\mathbf{n}\,\mathrm{d}S$. A central result in electrostatics is:

Theorem 1.1 (*Gauss' law*) *For any closed surface $\Sigma = \partial R$ bounding a region $R \subset \mathbb{R}^3$,*

$$\int_\Sigma \mathbf{E} \cdot \mathrm{d}\mathbf{S} = \frac{1}{\epsilon_0} \sum_{i=1}^{N} q_i = \frac{Q}{\epsilon_0} \ , \tag{1.6}$$

[4] See appendix A for vector calculus definitions and theorems.

where R contains the point charges $q_1, \ldots, q_N$, and $Q = q_1 + \cdots + q_N$ is the total charge *in R. In words: the flux of the electric field through a closed surface Σ is equal to $\frac{1}{\epsilon_0}$ times the total charge contained in the region bounded by Σ.*

Proof Consider first a point charge q at position $\mathbf{r}_0$. From (1.1), this produces an electric field

$$\mathbf{E}(\mathbf{r}) \;=\; \frac{q}{4\pi\epsilon_0}\,\frac{\mathbf{r}-\mathbf{r}_0}{|\mathbf{r}-\mathbf{r}_0|^3}\;. \tag{1.7}$$

Since $\nabla\cdot\mathbf{r} = 3$ and on $\mathbb{R}^3\setminus\{\mathbf{r}_0\}$ we have $\nabla\,|\mathbf{r}-\mathbf{r}_0| = (\mathbf{r}-\mathbf{r}_0)/|\mathbf{r}-\mathbf{r}_0|$ (see (A.15)), it follows that in this domain

$$\nabla\cdot\mathbf{E} \;=\; \frac{q}{4\pi\epsilon_0}\left(\frac{3}{|\mathbf{r}-\mathbf{r}_0|^3} - \frac{3\,(\mathbf{r}-\mathbf{r}_0)\cdot(\mathbf{r}-\mathbf{r}_0)}{|\mathbf{r}-\mathbf{r}_0|^5}\right) \;=\; 0\;. \tag{1.8}$$

The divergence of the electric field $\mathbf{E}$ produced by the point charge is thus zero everywhere, except at the location of the charge itself.

Consider next a sphere S of radius $a > 0$, centred on the point $\mathbf{r}_0$. Since the outward unit normal to S is $\mathbf{n} = (\mathbf{r}-\mathbf{r}_0)/|\mathbf{r}-\mathbf{r}_0|$, from (1.7) we have

$$\int_S \mathbf{E}\cdot\mathrm{d}\mathbf{S} \;=\; \frac{q}{4\pi\epsilon_0}\int_S \frac{1}{|\mathbf{r}-\mathbf{r}_0|^2}\,\mathrm{d}S \;=\; \frac{q}{4\pi a^2\epsilon_0}\int_S \mathrm{d}S \;=\; \frac{q}{\epsilon_0}\;. \tag{1.9}$$

Here we have used the fact that a sphere of radius a has surface area $4\pi a^2$, and $|\mathbf{r}-\mathbf{r}_0| = a$ on S.

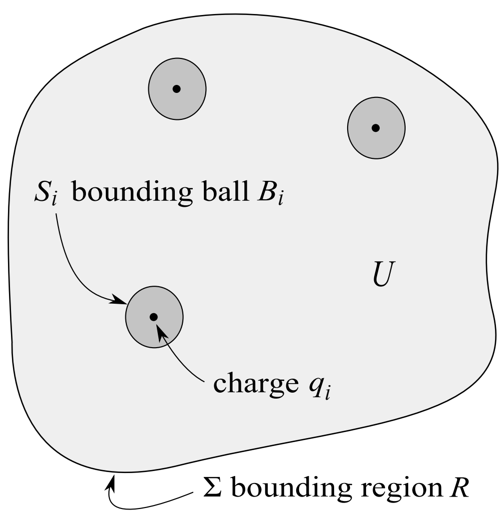


Figure 1: Region R, with boundary Σ, containing point charges $q_1, \ldots, q_N$. We divide $R = \bigcup_{i=1}^N B_i \cup U$ into a small balls B_i around each charge q_i, with boundary spheres $S_i = \partial B_i$, and a region U with boundary components Σ and $S_1, \ldots, S_N$.

With these results to hand, consider now a region R, with closed boundary $\Sigma = \partial R$, and assume that R contains the charges $q_1, \ldots, q_N$ at position vectors $\mathbf{r}_1, \ldots, \mathbf{r}_N \in R$. Introduce small balls B_i centred on each charge, such that each B_i contains only the charge q_i, and write $R = \bigcup_{i=1}^N B_i \cup U$, where the region U then contains no charges – see Figure 1. Each ball B_i has boundary sphere

S_i, and the *outward* unit normal to $S_i = \partial B_i$ is an *inward* unit normal to ∂U. The divergence theorem A.2 applied to U hence gives

$$\int_U \nabla\cdot\mathbf{E}\,\mathrm{d}V \;=\; \int_\Sigma \mathbf{E}\cdot\mathrm{d}\mathbf{S} - \sum_{i=1}^N \int_{S_i} \mathbf{E}\cdot\mathrm{d}\mathbf{S}\ . \tag{1.10}$$

The electric field generated by the configuration of charges is given by (1.5), and (1.8) shows that $\nabla\cdot\mathbf{E} = 0$ on $\mathbb{R}^3 \setminus \{\mathbf{r}_1, \ldots, \mathbf{r}_N\}$. Since the region U contains no charges, using (1.10) we have

$$\begin{aligned}
\int_\Sigma \mathbf{E}\cdot\mathrm{d}\mathbf{S} \;&=\; \left(\int_\Sigma \mathbf{E}\cdot\mathrm{d}\mathbf{S} - \sum_{i=1}^N \int_{S_i} \mathbf{E}\cdot\mathrm{d}\mathbf{S}\right) + \sum_{i=1}^N \int_{S_i} \mathbf{E}\cdot\mathrm{d}\mathbf{S} \\
&=\; \int_U \nabla\cdot\mathbf{E}\,\mathrm{d}V + \sum_{i=1}^N \int_{S_i} \mathbf{E}\cdot\mathrm{d}\mathbf{S} \qquad \text{(divergence theorem)} \\
&=\; \sum_{i=1}^N \int_{S_i} \mathbf{E}\cdot\mathrm{d}\mathbf{S} \;=\; \sum_{i=1}^N \frac{q_i}{\epsilon_0}\ .
\end{aligned} \tag{1.11}$$

Here the last step uses (1.9) for the i^{th} term in the sum (1.5): the remaining terms in the sum have zero divergence on the ball B_i, and so do not contribute to the integral. ■

Note that this proof starts by analysing a single charge, and the inclusion of multiple charges is an application of the Principle of Superposition, ***E3***.

1.4 Charge density and Gauss' law

For many problems it is not convenient to deal with point charges. If the point charges we have been discussing are, say, electrons, then a macroscopic object will consist of an absolutely enormous number of electrons, each with a very tiny charge.[5] We thus introduce the concept of *charge density* $\rho(\mathbf{r})$, which is a function giving the *charge per unit volume*. This means that, by definition, the total charge $Q = Q(R)$ in any region R is

$$Q \;=\; \int_R \rho\,\mathrm{d}V\ . \tag{1.12}$$

We shall generally assume that when describing a smooth three-dimensional distribution of charge, the function ρ is at least continuous. For the purposes of physical arguments, we shall often think of the Riemann integral as the limit of a sum (which is what it is). Thus, if $\delta R \subset \mathbb{R}^3$ is a small region centred around a point $\mathbf{r} \in \mathbb{R}^3$ in such a sum, that region contributes a charge $\rho(\mathbf{r})\,\delta V$, where δV is the volume of δR.

With this definition, the obvious limit of the sum in (1.5), replacing a point charge q' at position $\mathbf{r}'$ by $\rho(\mathbf{r}')\,\delta V'$, becomes a volume integral

$$\mathbf{E}(\mathbf{r}) \;=\; \frac{1}{4\pi\epsilon_0}\int_{\mathbf{r}'\in R} \frac{\rho(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|^3}\,(\mathbf{r}-\mathbf{r}')\,\mathrm{d}V'\ . \tag{1.13}$$

[5]A different issue is that, at the microscopic scale, the charge of an electron is not pointlike, but rather is effectively smeared out into a smooth distribution of charge. In fact in quantum mechanics the precise position of an electron cannot be measured *in principle*!

Here $\{\mathbf{r} \in \mathbb{R}^3 \mid \rho(\mathbf{r}) \neq 0\} \subset R$, so that all charge is contained in the (usually bounded) region R. (1.13) gives the electric field generated by a charge distribution described by the density ρ, where notice that we have again derived this from Coulomb's law and the Principle of Superposition.

Similarly, the limit of (1.6) becomes

$$\int_\Sigma \mathbf{E} \cdot \mathrm{d}\mathbf{S} = \frac{1}{\epsilon_0} \int_R \rho \,\mathrm{d}V , \tag{1.14}$$

for any region R with boundary $\partial R = \Sigma$. Using the divergence theorem we may rewrite this as

$$\int_R \left(\nabla \cdot \mathbf{E} - \frac{\rho}{\epsilon_0} \right) \mathrm{d}V = 0 . \tag{1.15}$$

Since this holds for all R, we conclude from Lemma A.3 another version of *Gauss' law*

$$\boxed{\nabla \cdot \mathbf{E} = \frac{\rho}{\epsilon_0} . \tag{1.16}}$$

We have derived the first of Maxwell's equations, on the front page, from the three simple laws of electrostatics. In fact this equation holds in general, *i.e.* even when there are magnetic fields and time dependence. We will directly show that (1.13) satisfies (1.16) in the next subsection.

As already mentioned, when writing a charge density ρ describing a smooth three-dimensional distribution of charge, we shall generally assume it is at least continuous. However, if (1.16) is to apply for the density ρ of a *point charge* q, say at the origin, notice that (1.8) implies that $\rho = 0$ on $\mathbb{R}^3 \setminus \{\mathbf{0}\}$, but still the integral of ρ over a neighbourhood of the origin is equal to the charge q. The "function" with this property is q times the Dirac delta function:

Definition The *Dirac delta function* $\delta(x)$ in one dimension is defined by:

$$\text{(i) } \delta(x) = 0 \text{ for } x \neq 0 , \qquad \text{(ii) } \int_I \delta(x)\,\mathrm{d}x = \begin{cases} 1 & \text{if } 0 \in I \subseteq \mathbb{R} , \\ 0 & \text{otherwise} . \end{cases} \tag{1.17}$$

One can define this rigorously as a *distribution*, which is a linear functional on an appropriate space of *test functions*. One can also view $\delta(x)$ as the limit of a sequence of "bump functions", supported around $x = 0$. For more details the reader might refer to the Part A Integral Transforms course, but we shall only need an informal understanding of the Dirac delta function for this course. The following Proposition summarizes some key properties:

Proposition 1.2 *The Dirac delta function satisfies:*

$$\begin{aligned} &\text{(i)} \quad \int_{-\infty}^{\infty} f(x)\delta(x - x')\,\mathrm{d}x = f(x') , \\ &\text{(ii)} \quad \delta[a(x - x')] = \frac{1}{|a|}\delta(x - x') , \\ &\text{(iii)} \quad \delta(g(x)) = \sum_{i=1}^{n} \frac{\delta(x - x_i)}{|g'(x_i)|} , \end{aligned} \tag{1.18}$$

where (i) $f(x)$ *is any continuous function,* (ii) $a \neq 0$ *is a real constant, and* (iii) *the differentiable function* $g(x)$ *has zeros* $g(x_i) = 0$, $i = 1, \ldots, n$, *with* $g'(x_i) \neq 0$.

The proofs of these essentially follow from the defining property (1.17), where in (iii) one changes variable $y = g(x)$, $\mathrm{d}y = g'(x)\,\mathrm{d}x$ in the integral. Notice from (ii) that $\delta(x - x') = \delta(x' - x)$.

The *Dirac delta function in three dimensions* is then defined in Cartesian coordinates $\mathbf{r} = (x_1, x_2, x_3)$ as the product

$$\delta(\mathbf{r}) \;\equiv\; \delta(x_1)\,\delta(x_2)\,\delta(x_3)\,. \tag{1.19}$$

From (1.17) this satisfies

$$\text{(i) } \delta(\mathbf{r} - \mathbf{r}') \;=\; 0 \text{ for } \mathbf{r} \neq \mathbf{r}'\ , \qquad \text{(ii) } \int_{\mathbf{r}\in R} \delta(\mathbf{r} - \mathbf{r}')\,\mathrm{d}V \;=\; \begin{cases} 1 & \text{if } \mathbf{r}' \in R\ , \\ 0 & \text{otherwise}\,. \end{cases} \tag{1.20}$$

The density for a point charge q at position $\mathbf{r}_0$ is then

$$\rho(\mathbf{r}) \;=\; q\,\delta(\mathbf{r} - \mathbf{r}_0)\,. \tag{1.21}$$

Example As an application of this, starting from (1.16) we may rederive Gauss' law (1.6) for N point charges $q_1, \ldots, q_N$ at positions $\mathbf{r}_1, \ldots, \mathbf{r}_N$ inside a region R, with boundary $\Sigma = \partial R$:

$$\int_\Sigma \mathbf{E}\cdot\mathrm{d}\mathbf{S} \;=\; \int_R \nabla\cdot\mathbf{E}\,\mathrm{d}V \;=\; \int_{\mathbf{r}\in R} \frac{1}{\epsilon_0}\sum_{i=1}^N q_i\,\delta(\mathbf{r} - \mathbf{r}_i)\,\mathrm{d}V \;=\; \frac{1}{\epsilon_0}\sum_{i=1}^N q_i\,. \tag{1.22}$$

Likewise, inserting $\rho(\mathbf{r}') = \sum_{i=1}^N q_i\,\delta(\mathbf{r}' - \mathbf{r}_i)$ into (1.13) gives

$$\mathbf{E}(\mathbf{r}) \;=\; \frac{1}{4\pi\epsilon_0}\int_{\mathbf{r}'\in R} \frac{1}{|\mathbf{r} - \mathbf{r}'|^3}\,(\mathbf{r} - \mathbf{r}')\sum_{i=1}^N q_i\,\delta(\mathbf{r}' - \mathbf{r}_i)\,\mathrm{d}V' \;=\; \frac{1}{4\pi\epsilon_0}\sum_{i=1}^N \frac{q_i}{|\mathbf{r} - \mathbf{r}_i|^3}(\mathbf{r} - \mathbf{r}_i)\,, \tag{1.23}$$

which is the point charge formula (1.5). ■

Proposition 1.3 *The three-dimensional Dirac delta function may be written as*

$$\delta(\mathbf{r} - \mathbf{r}_0) \;=\; -\frac{1}{4\pi}\nabla^2\left(\frac{1}{|\mathbf{r} - \mathbf{r}_0|}\right)\ . \tag{1.24}$$

Proof This essentially follows from the calculations we have done. On $\mathbb{R}^3 \setminus \{\mathbf{r}_0\}$ note that

$$\nabla\left(\frac{1}{|\mathbf{r} - \mathbf{r}_0|}\right) \;=\; -\frac{(\mathbf{r} - \mathbf{r}_0)}{|\mathbf{r} - \mathbf{r}_0|^3} \;=\; -\frac{4\pi\epsilon_0}{q}\,\mathbf{E}(\mathbf{r})\ , \tag{1.25}$$

where on the right hand side $\mathbf{E}(\mathbf{r})$ is the electric field (1.7) generated by a point charge at position $\mathbf{r}_0$. Taking the divergence of (1.25), from (1.8) we immediately deduce that $\nabla^2(1/|\mathbf{r} - \mathbf{r}_0|) = 0$ for $\mathbf{r} \neq \mathbf{r}_0$, and using the divergence theorem and then (1.9) we have

$$\int_{\mathbf{r}\in B} \nabla^2\left(\frac{1}{|\mathbf{r} - \mathbf{r}_0|}\right)\mathrm{d}V \;=\; \int_{S=\partial B} \nabla\left(\frac{1}{|\mathbf{r} - \mathbf{r}_0|}\right)\cdot\mathrm{d}\mathbf{S} \;=\; -\frac{4\pi\epsilon_0}{q}\int_S \mathbf{E}\cdot\mathrm{d}\mathbf{S} \;=\; -4\pi\ , \tag{1.26}$$

where B is a ball centred on $\mathbf{r}_0$. These properties establish the identification (1.24). ■

1.5 Electrostatic potential and Poisson's equation

Returning to our point charge q at position $\mathbf{r}_0$, note that on $\mathbb{R}^3 \setminus \{\mathbf{r}_0\}$ equation (1.25) implies that $\mathbf{E} = -\nabla\phi$ where

$$\phi(\mathbf{r}) \;=\; \frac{q}{4\pi\epsilon_0}\frac{1}{|\mathbf{r}-\mathbf{r}_0|}\;. \tag{1.27}$$

Definition The function ϕ with $\mathbf{E} = -\nabla\phi$ is called the *electrostatic potential.*

For a continuous charge density we then have the following:

Theorem 1.4 *Consider a continuous charge density $\rho(\mathbf{r})$ with support $\{\mathbf{r} \in \mathbb{R}^3 \mid \rho(\mathbf{r}) \neq 0\} \subset R$ with R a* bounded *region, and define*

$$\phi(\mathbf{r}) \;\equiv\; \frac{1}{4\pi\epsilon_0}\int_{\mathbf{r}'\in R}\frac{\rho(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}\,\mathrm{d}V'\;. \tag{1.28}$$

Then $-\nabla\phi = \mathbf{E}$, with the electric field $\mathbf{E}$ given by (1.13), which satisfies Gauss' law (1.16). Moreover, $\phi = O(1/r)$ as $r \to \infty$, so that this electrostatic potential is zero "at infinity".

Proof The formula (1.13) follows immediately from applying $-\nabla$ to (1.28):

$$\begin{aligned} -\nabla\phi(\mathbf{r}) \;&=\; \frac{1}{4\pi\epsilon_0}\int_{\mathbf{r}'\in R}\left[-\nabla\left(\frac{1}{|\mathbf{r}-\mathbf{r}'|}\right)\right]\rho(\mathbf{r}')\,\mathrm{d}V' \\ &=\; \frac{1}{4\pi\epsilon_0}\int_{\mathbf{r}'\in R}\frac{\rho(\mathbf{r}')(\mathbf{r}-\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|^3}\,\mathrm{d}V'\;, \end{aligned} \tag{1.29}$$

where in the second line we have used the first equality in (1.25). Instead applying ∇^2 to (1.28), and using Proposition 1.3, we deduce

$$\begin{aligned} \nabla^2\phi(\mathbf{r}) \;&=\; \frac{1}{4\pi\epsilon_0}\int_{\mathbf{r}'\in R}\nabla^2\left(\frac{1}{|\mathbf{r}-\mathbf{r}'|}\right)\rho(\mathbf{r}')\,\mathrm{d}V' \;=\; \frac{1}{4\pi\epsilon_0}\int_{\mathbf{r}'\in R}\left(-4\pi\,\delta(\mathbf{r}-\mathbf{r}')\right)\rho(\mathbf{r}')\,\mathrm{d}V' \\ &=\; -\frac{\rho(\mathbf{r})}{\epsilon_0}\;. \end{aligned} \tag{1.30}$$

It follows that $\nabla\cdot\mathbf{E} = \rho/\epsilon_0$, where $\mathbf{E} = -\nabla\phi$. Finally, R being bounded means it is contained inside some closed ball B of radius a, and ρ is continuous on this compact set so it is bounded. Thus $|\rho(\mathbf{r}')| \leq M$ for all $\mathbf{r}' \in B$, and

$$|\phi(\mathbf{r})| \;=\; \frac{1}{4\pi\epsilon_0}\left|\int_{\mathbf{r}'\in B}\frac{\rho(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}\,\mathrm{d}V'\right| \;\leq\; \frac{1}{4\pi\epsilon_0}\int_{\mathbf{r}'\in B}\frac{M}{|\mathbf{r}-\mathbf{r}'|}\,\mathrm{d}V' \;\leq\; \frac{M}{r-a}\,\mathrm{Vol}(B_a)\;, \tag{1.31}$$

where in the last step we have taken $r > a$, and used the reverse triangle inequality: $|\mathbf{r}-\mathbf{r}'| \geq ||\mathbf{r}|-|\mathbf{r}'|| = r - |\mathbf{r}'| \geq r - a$. ■

Corollary 1.5 *The electrostatic potential satisfies* Poisson's equation

$$\nabla^2\phi \;=\; -\frac{\rho}{\epsilon_0}\;. \tag{1.32}$$

Since the curl of a gradient is identically zero, we may also deduce from $\mathbf{E} = -\nabla\phi$ that

$$\boxed{\nabla \times \mathbf{E} = \mathbf{0}\ . \qquad (1.33)}$$

Equation (1.33) is another of Maxwell's equations from the front page, albeit only in the special case where $\partial\mathbf{B}/\partial t = \mathbf{0}$ (the magnetic field is time-independent).

* In fact $\nabla \times \mathbf{E} = \mathbf{0}$ *implies* that the vector field $\mathbf{E}$ is the gradient of a function, *provided* the domain of definition is *simply-connected*. Recall the latter means that every closed loop can be continuously contracted to a point. For example, this is true for $\mathbb{R}^3$ or in an open ball. For non-simply-connected domains, such as $\mathbb{R}^3$ minus a line (say, the z-axis), it is not always possible to write a vector field $\mathbf{E}$ with zero curl as a gradient. A systematic discussion of this is certainly beyond this course. The interested reader can find a proof for an open ball in appendix B of the book by Woodhouse in the reading list.

Recall from Prelims that forces $\mathbf{F}$ which are gradients are called *conservative forces*. Since $\mathbf{F} = q\,\mathbf{E}$, we see that the electrostatic force is conservative. The *work done* against the electrostatic force in moving a charge q along a curve C is then the line integral

$$W = -\int_C \mathbf{F}\cdot \mathrm{d}\mathbf{r} = -q\int_C \mathbf{E}\cdot \mathrm{d}\mathbf{r} = q\int_C \nabla\phi\cdot \mathrm{d}\mathbf{r} = q\,[\phi(\mathbf{r}_1) - \phi(\mathbf{r}_0)]\ . \qquad (1.34)$$

Here the curve C begins at $\mathbf{r}_0$ and ends at $\mathbf{r}_1$. The work done is of course independent of the choice of curve connecting the two points, because the force is conservative. Notice that one may add a constant to ϕ without changing $\mathbf{E}$. It is only the *difference* in values of ϕ that is physical, and this is called the *voltage*. If we fix some arbitrary point $\mathbf{r}_0$ and choose $\phi(\mathbf{r}_0) = 0$, then $\phi(\mathbf{r})$ has the interpretation of work done against the electric field in moving a unit charge from $\mathbf{r}_0$ to $\mathbf{r}$. Note that Theorem 1.4 says that the particular choice for ϕ in (1.28) is zero "at infinity". From the usual relation between work and energy, ϕ is also the *potential energy* per unit charge.

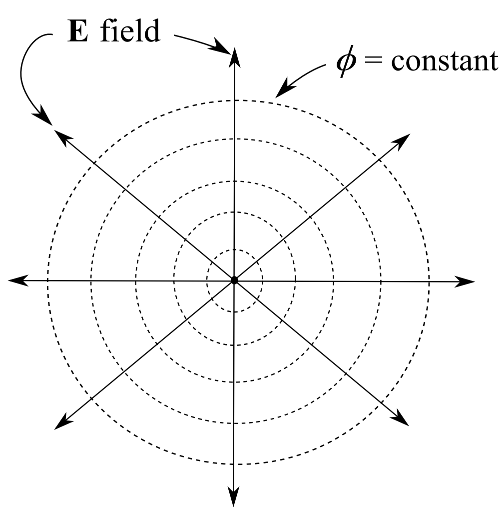


Figure 2: The field lines, which represent the direction of the electric field $\mathbf{E}$, and equipotentials around a positive point charge.

Definition Surfaces of constant ϕ are called *equipotentials.*

Proposition 1.6 *The electric field is always* normal *to an equipotential surface.*

Proof To see this, fix a point $\mathbf{r}$ and let $\mathbf{t}$ be a *tangent vector* to an equipotential for $\phi(\mathbf{r})$. By definition, the derivative of $\phi(\mathbf{r})$ in a direction tangent to such an equipotential surface is zero, and hence $\mathbf{t} \cdot \nabla\phi = 0$ at this point. Since this holds for all tangent vectors, this means that $\nabla\phi = -\mathbf{E}$ is normal to a surface of constant ϕ. ■

Example (Uniform charge density in a volume) Consider a slab of infinite extent in both the positive and negative x and y directions, but occupying the interval $-a \leq z \leq a$ in the z direction with uniform charge density ρ_0. We can determine the resulting electric field in two different ways.

Since the geometry and the charge density are translationally and rotationally invariant in the $x - y$ plane, $\mathbf{E} = E(z)\,\mathbf{e}_z$. Furthermore there is a reflection symmetry $z \to -z$ which sends $E(z) \to -E(-z)$, where the minus sign comes from the invariance of $\mathbf{E}$ combined with $\mathbf{e}_z \to -\mathbf{e}_z$.

Consider a cylinder of height h with disk area A centred at $z = 0$. The cylinder is invariant under reflections, and we can write the integrated form of Gauss' law

$$\begin{aligned} &\int_{\text{cylinder}} \mathbf{E} \cdot d\mathbf{S} = A\left[E\left(h\right) - E\left(-h\right)\right] = 2AE\left(h\right) \\ &= \int_{\text{cylinder}} \frac{\rho(z)}{\epsilon_0}\, dV = \frac{\rho_0 A}{\epsilon_0} = \frac{2\rho_0 A}{\epsilon_0} \begin{cases} h & (h \leq a)\ , \\ a & (h > a)\ . \end{cases} \end{aligned} \tag{1.35}$$

By equating the last expressions in the first and second lines, we get

$$E\left(z\right) = \frac{\rho_0}{\epsilon_0} \begin{cases} z & (0 \leq z \leq a)\ , \\ a & (z > a)\ . \end{cases} \tag{1.36}$$

Alternatively, we can obtain the electric field by first determining the electrostatic potential by solving Poisson's equation. By the same symmetry considerations as above the electrostatic potential only depends on z and Poisson's equation takes the form

$$\phi''(z) = -\frac{\rho}{\epsilon_0} = -\frac{\rho_0}{\epsilon_0} \begin{cases} 1 & (-a \leq z \leq a)\ , \\ 0 & (\text{otherwise})\ . \end{cases} \tag{1.37}$$

We focus on $0 \leq z$ and integrate

$$\phi(z) = -\frac{\rho_0}{\epsilon_0} \begin{cases} \frac{z^2}{2} + B\,z + C & (0 \leq z \leq a)\ , \\ D\,z + F & (a < z)\ . \end{cases} \tag{1.38}$$

Continuous differentiability of ϕ sets $D = a + B$, $F = C - a^2/2$. While a constant shift in ϕ is unphysical (hence we cannot hope to fix C), we are left with the unfixed parameter B. Taking the derivative gives

$$E(z) = -\phi'(z) = \frac{\rho_0}{\epsilon_0} \begin{cases} z + B & (0 \leq z \leq a)\ , \\ a + B & (a < z)\ . \end{cases} \tag{1.39}$$

This matches (1.36) only for $B = 0$. The $B \neq 0$ solutions represent a homogeneous electric field throughout $\mathbb{R}^3$ superposed with the field sourced by the charge density, which we discard. An alternative way to argue is that $B \neq 0$ breaks the reflection symmetry.

Note that unlike in most problems discussed in this course the electrostatic potential does not go to zero at infinity for any choice of the constant C. This does not contradict Theorem 1.4, as that theorem assumed that the charge density is localised to a bounded region, which is not satisfied here. ■

1.6 Conductors and surface charge

More interesting is when the distribution of charge is not described by a continuous charge density. We have already encountered point charges. For many problems it is useful to introduce the concepts of *surface charge density* σ on a surface S, say for a charge distribution on a thin metal sheet, and also *line charge density* λ on a curve C, say for a charge distribution in a thin wire. These will be taken to be appropriately well-behaved functions on S and C, representing charge per unit area and charge per unit length, respectively.

In fact the concept of surface charge density doesn't require a thin metal sheet to be useful, for the following reason:

Definition An *electrical conductor* is a material where some of the electrons ("conduction electrons") are free to move in the presence of an external electric field.

In a *static* situation, the electric field *inside* the conducting material must be zero. Why? Because if it weren't, then the conduction electrons in the interior would experience a force, and thus move by Newton's second law.

Imagine what happens if we now switch on an external electric field: a conduction electron will move in the opposite direction to the field (because it is negatively charged), until either (a) it gets to the boundary of the material, or (b) the electric field inside the material has relaxed to its equilibrium of zero. This way, one ends up with lots of electrons at, or very near, the surface of the material; their distribution (and the distribution of other immobile charges) throughout the material produces an electric field which precisely *cancels* the external field *inside* the material. Thus $\mathbf{E} = \mathbf{0}$ inside a conductor. More generally if the conductor is a surface or curve in $\mathbb{R}^3$, then similarly $\mathbf{t} \cdot \mathbf{E} = 0$ for any tangent vector $\mathbf{t}$ to the conductor. Since $\mathbf{E} = -\nabla\phi$ we deduce the important fact that:

Theorem 1.7 (*a "Physics Theorem"*) *A conductor in static equilibrium is always an* equipotential *for ϕ, i.e. $\phi =$ constant throughout the conducting material.*

Equation (1.16) furthermore implies that $\rho = 0$ inside a conducting material in equilibrium, and hence the charge must be described by a surface charge density.

Surface and line charge densities of course contribute to the total charge and electric field via surface and line integrals, respectively. For example, a surface S with surface charge density σ gives rise to an electric field

$$\mathbf{E}(\mathbf{r}) = \frac{1}{4\pi\epsilon_0} \int_{\mathbf{r}'\in S} \frac{\sigma(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|^3} (\mathbf{r}-\mathbf{r}')\, \mathrm{d}S' \ . \tag{1.40}$$

Notice that for $\mathbf{r} \in \mathbb{R}^3 \setminus S$ and σ smooth the integrand is smooth. However, it turns out that $\mathbf{E}$ is *not continuous* across S!

Proposition 1.8 *For the electric field given by (1.40), generated by a surface charge density σ on a surface S, the components of $\mathbf{E}$ tangent to S are continuous across S, but the* normal *component of $\mathbf{E}$ is not. Specifically, if $\mathbf{n}$ is a unit normal vector field to the surface pointing into what we'll call the "+ side", then*

$$\mathbf{E}^+ \cdot \mathbf{n} - \mathbf{E}^- \cdot \mathbf{n} = \frac{\sigma}{\epsilon_0} \ , \tag{1.41}$$

at every point on the surface.

Proof To see this, consider a surface S which has a surface charge density σ. Consider the cylindrical region R on left hand side of Figure 3, of height ε and cross-sectional area δA. Gauss' law gives

$$\int_{\partial R} \mathbf{E}\cdot \mathrm{d}\mathbf{S} = \frac{1}{\epsilon_0}\, (\text{total charge in } R) \ . \tag{1.42}$$

In the limit $\varepsilon \to 0$ the left hand side becomes $(\mathbf{E}^+ \cdot \mathbf{n} - \mathbf{E}^- \cdot \mathbf{n})\, \delta A$ for small δA, where $\mathbf{E}^\pm$ are the electric fields on the two sides of S and the unit normal $\mathbf{n}$ points into the + side. The right hand side, on the other hand, tends to $\sigma\, \delta A/\epsilon_0$. Thus there is necessarily a discontinuity in the component of $\mathbf{E}$ normal to S given by (1.41).

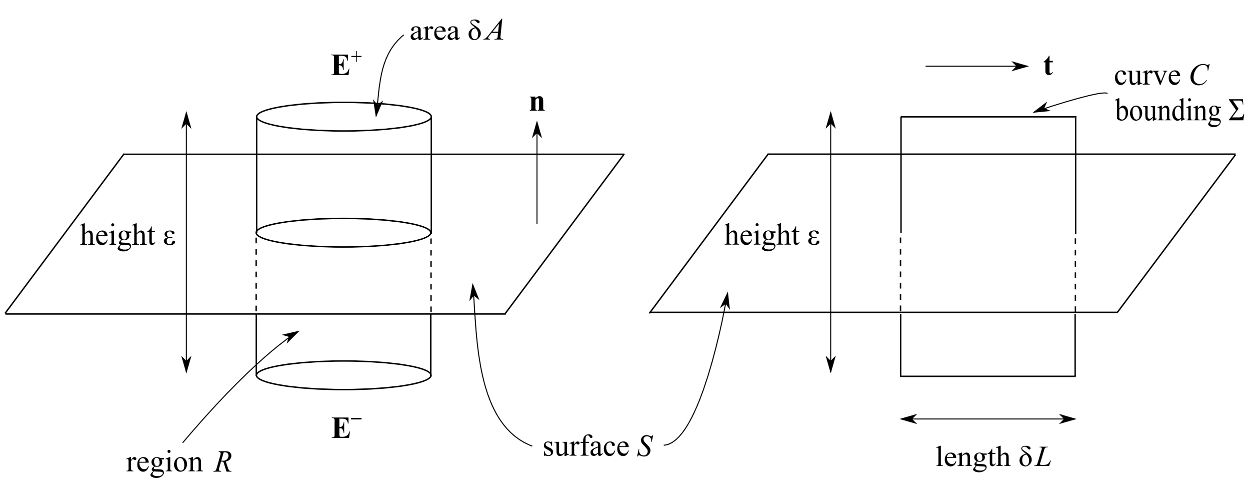


Figure 3: The surface S.

Consider, instead, the rectangular loop C on the right hand side of Figure 3, of height ε and length δL, bounding the rectangular surface Σ. By Stokes' theorem A.1 we have

$$\int_C \mathbf{E} \cdot \mathrm{d}\mathbf{r} \; = \; \int_\Sigma (\nabla \times \mathbf{E}) \cdot \mathrm{d}\mathbf{S} \; = \; 0 \; , \tag{1.43}$$

where we have used the electrostatic Maxwell equation (1.33) in the second equality. If $\mathbf{t}$ denotes a unit tangent vector along C on the + side, then in the limit $\varepsilon \to 0$ we obtain

$$\left(\mathbf{E}^+ \cdot \mathbf{t} - \mathbf{E}^- \cdot \mathbf{t}\right) \delta L \; = \; 0 \; , \tag{1.44}$$

for small δL. Thus the components of $\mathbf{E}$ tangent to S are continuous across S

$$\mathbf{E}^+ \cdot \mathbf{t} \; = \; \mathbf{E}^- \cdot \mathbf{t} \; . \qquad \blacksquare \tag{1.45}$$

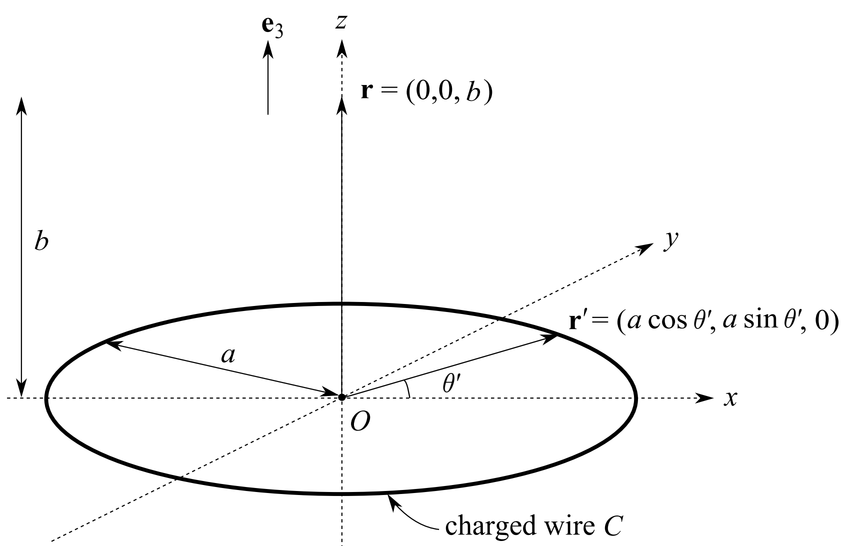


Figure 4: Charged plane circular wire C of radius a, centred on the origin O in the (x, y)-plane at $z = 0$.

Example (Line charge density) Consider a curve C with line charge density λ. This generates an electric field

$$\mathbf{E}(\mathbf{r}) \; = \; \frac{1}{4\pi\epsilon_0} \int_{\mathbf{r}' \in C} \frac{\lambda(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|^3} \, (\mathbf{r} - \mathbf{r}') \left| \frac{\mathrm{d}\mathbf{r}'}{\mathrm{d}s'} \right| \mathrm{d}s' \; , \tag{1.46}$$

where the curve $C \subset \mathbb{R}^3$ is parametrized by $\mathbf{r}' = \mathbf{r}'(s')$, with $s' \in [s_0, s_1] \subset \mathbb{R}$. For example, consider a plane circular wire of radius a, centred on the origin in the (x, y)-plane at $z = 0$, carrying a total charge Q that is uniformly distributed around the wire – see Figure 4. A point on the wire is $\mathbf{r}'(\theta') = (a\cos\theta', a\sin\theta', 0)$, parametrized by $\theta' \in [0, 2\pi]$, and so $|\mathrm{d}\mathbf{r}'/\mathrm{d}\theta'| = a$. The uniform line charge density is then $\lambda = Q/2\pi a$, and the electric field (1.46) is

$$\mathbf{E}(\mathbf{r}) \; = \; \frac{1}{4\pi\epsilon_0} \int_{\theta'=0}^{2\pi} \frac{Q}{2\pi a} \frac{\mathbf{r} - \mathbf{r}'(\theta')}{|\mathbf{r} - \mathbf{r}'(\theta')|^3} \, a \, \mathrm{d}\theta' \; . \tag{1.47}$$

It is difficult to evaluate this integral at a general point $\mathbf{r} = (x, y, z)$, so let us look at the point $\mathbf{r} = (0, 0, b)$ on the z-axis. In this case notice that $|\mathbf{r} - \mathbf{r}'(\theta')| = |(0 - a\cos\theta', 0 - a\sin\theta', b)| = \sqrt{a^2 + b^2}$ is independent of θ', and (1.47) easily integrates to give

$$\mathbf{E}((0,0,b)) = \frac{1}{4\pi\epsilon_0}\frac{Q}{2\pi}\int_{\theta'=0}^{2\pi}\frac{1}{(a^2+b^2)^{3/2}}\left(-a\cos\theta', -a\sin\theta', b\right)\mathrm{d}\theta' = \frac{Q}{4\pi\epsilon_0}\frac{b}{(a^2+b^2)^{3/2}}\,\mathbf{e}_3 \quad (1.48)$$

where $\mathbf{e}_3$ is a unit vector pointing along the z-axis. Here the integrals over $\cos\theta'$ and $\sin\theta'$ in the x and y components of (1.48) give zero, but we could also have anticipated this by the symmetry of the problem: the z-axis is an axis of symmetry, and on this locus the electric field then also necessarily points along the z-axis. ■

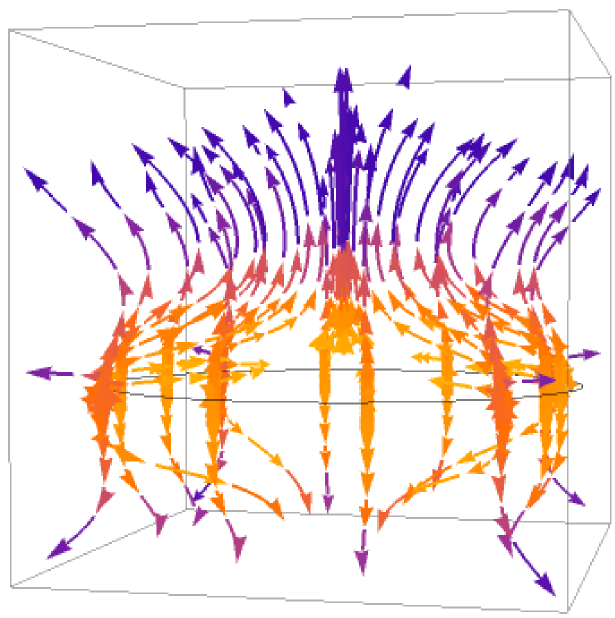


Figure 5: Electric field lines around the plane circular wire, given by (1.47).

1.7 Electrostatic energy

In this subsection we derive a formula for the *energy* of an electrostatic configuration as an integral of a *local energy density*. We shall return to this subject again in section 5.5.

We begin with a point charge q_1 at $\mathbf{r}_1$. This generates a potential

$$\phi^{(1)}(\mathbf{r}) = \frac{1}{4\pi\epsilon_0}\frac{q_1}{|\mathbf{r} - \mathbf{r}_1|} \ . \quad (1.49)$$

Consider now moving a charge q_2 from infinity to the point $\mathbf{r}_2$. From (1.34), the work done against the electric field in doing this is

$$W_2 = q_2\,\phi^{(1)}(\mathbf{r}_2) = \frac{1}{4\pi\epsilon_0}\frac{q_2 q_1}{|\mathbf{r}_2 - \mathbf{r}_1|} \ . \quad (1.50)$$

Next move in another charge q_3 from infinity to the point $\mathbf{r}_3$. We must now do work against the electric fields of *both* q_1 and q_2. By the Principle of Superposition, this work done is

$$W_3 \;=\; \frac{1}{4\pi\epsilon_0}\left(\frac{q_3 q_1}{|\mathbf{r}_3-\mathbf{r}_1|}\;+\;\frac{q_3 q_2}{|\mathbf{r}_3-\mathbf{r}_2|}\right)\;. \tag{1.51}$$

The total work done so far is thus W_2+W_3. We may continue this process and inductively deduce that the *total* work done in assembling charges $q_1,\ldots,q_N$ at $\mathbf{r}_1,\ldots,\mathbf{r}_N$ is

$$W \;=\; \frac{1}{4\pi\epsilon_0}\sum_{i=1}^{N}\sum_{j<i}\frac{q_i q_j}{|\mathbf{r}_i-\mathbf{r}_j|}\;=\;\frac{1}{2}\cdot\frac{1}{4\pi\epsilon_0}\sum_{i=1}^{N}\sum_{j\neq i}\frac{q_i q_j}{|\mathbf{r}_i-\mathbf{r}_j|}\;, \tag{1.52}$$

where notice that $q_i q_j/|\mathbf{r}_i-\mathbf{r}_j|$ is symmetric under swapping i and j.

We next rewrite (1.52) as

$$W \;=\; \frac{1}{2}\sum_{i=1}^{N} q_i\,\phi_i\;, \tag{1.53}$$

where we have defined

$$\phi_i \;\equiv\; \frac{1}{4\pi\epsilon_0}\sum_{j\neq i}\frac{q_j}{|\mathbf{r}_i-\mathbf{r}_j|}\;. \tag{1.54}$$

This is simply the electrostatic potential produced by all but the ith charge, evaluated at position $\mathbf{r}_i$. In the usual continuum limit, (1.53) becomes

$$W \;=\; \frac{1}{2}\int_R \rho\,\phi\,\mathrm{d}V\;, \tag{1.55}$$

where $\phi(\mathbf{r})$ is given by (1.28). Now, using Gauss' law (1.16) we may write

$$\phi\,\frac{\rho}{\epsilon_0}\;=\;\phi\,\nabla\cdot\mathbf{E}\;=\;\nabla\cdot(\phi\,\mathbf{E})\;-\;\nabla\phi\cdot\mathbf{E}\;=\;\nabla\cdot(\phi\,\mathbf{E})+\mathbf{E}\cdot\mathbf{E}\;, \tag{1.56}$$

where in the last step we used $\mathbf{E}=-\nabla\phi$. Inserting this into (1.55), we have

$$W \;=\; \frac{\epsilon_0}{2}\left[\int_{\Sigma=\partial R}\phi\,\mathbf{E}\cdot\mathrm{d}\mathbf{S}\;+\;\int_R\mathbf{E}\cdot\mathbf{E}\,\mathrm{d}V\right]\;, \tag{1.57}$$

where we have used the divergence theorem on the first term. Taking R to be a very large ball of radius r, enclosing all charge, the surface Σ is a sphere. From Theorem 1.4 we have $\phi = O(1/r)$ as $r\to\infty$, and one can check that this surface term is zero in the limit that the ball becomes infinitely large. We hence deduce the elegant formula

$$W \;=\; \frac{\epsilon_0}{2}\int_{\mathbb{R}^3}\mathbf{E}\cdot\mathbf{E}\,\mathrm{d}V\;. \tag{1.58}$$

Definition When the integral (1.58) exists the electrostatic configuration is said to have *finite energy* W. The formula (1.58) suggests that the energy is stored in a local *energy density*

$$\mathcal{E}\;\equiv\;\frac{\epsilon_0}{2}\,\mathbf{E}\cdot\mathbf{E}\;=\;\frac{\epsilon_0}{2}\,|\mathbf{E}|^2\;. \tag{1.59}$$

2 Boundary value problems in electrostatics

2.1 Boundary value problems

In the previous section we have seen that electrostatics reduces to solving the Poisson equation

$$\nabla^2\phi = -\frac{\rho}{\epsilon_0} . \tag{2.1}$$

For problems where we are simply *given* a charge distribution ρ everywhere in space, the solution to (2.1) for ϕ is given by the integral formula (1.28) in Theorem 1.4. We have then effectively *solved* this class of electrostatics problems.

In this section we consider a different class of problems. We still want to solve the Poisson equation (2.1), but now in a region of space R with boundary $\partial R = \Sigma$, where we *specify boundary conditions* for the fields on Σ. From a mathematical perspective, given a PDE such as (2.1) it is natural to ask what types of boundary condition we can impose. In this section we consider:

(i) *Dirichlet boundary conditions*, where ϕ is specified on Σ,

(ii) *Neumann boundary conditions*, where $\mathbf{E}\cdot\mathbf{n}$ is specified on Σ, where $\mathbf{n}$ is the outward unit normal to $\Sigma = \partial R$. Here the normal component of the electric field $\mathbf{E}\cdot\mathbf{n}$ on Σ is

$$\mathbf{E}\cdot\mathbf{n} = -\mathbf{n}\cdot\nabla\phi \equiv -\frac{\partial\phi}{\partial n} , \tag{2.2}$$

which is minus the *normal derivative* of ϕ.

From a mathematical perspective it makes sense to solve (2.1) *only* on the domain $R \subseteq \mathbb{R}^3$, *ignoring* the exterior (complement) of R. However, in physical problems there should be something that is effectively imposing these boundary conditions, and indeed then something outside the domain R. Here are two classes of examples:

(i) Consider a charge distribution described by a density ρ in the interior of a region R, where $\partial R = \Sigma$ and this boundary is surrounded by a thin *conducting material* – see Figure 6. As explained in section 1.6, the conduction electrons will distribute themselves in such a way to give $\phi =$ constant inside the thin layer of conducting material, but also induce a surface charge density σ on Σ, *that we do not know a priori*. However, given the solution ϕ to the Dirichlet problem, on Σ we may identify

$$\sigma = \epsilon_0 \left.\frac{\partial\phi}{\partial n}\right|_{\Sigma} . \tag{2.3}$$

This follows from (1.41) in Proposition 1.8, where inside the thin conducting material we have $\mathbf{E}^+ = \mathbf{0}$, and we have then used (2.2) to write $\mathbf{E}^-\cdot\mathbf{n} = -\partial\phi/\partial n$, where $\mathbf{E}^-$ is the electric field just inside the boundary Σ. Of course, the solution will determine $\mathbf{E} = -\nabla\phi$ in the whole interior of R.

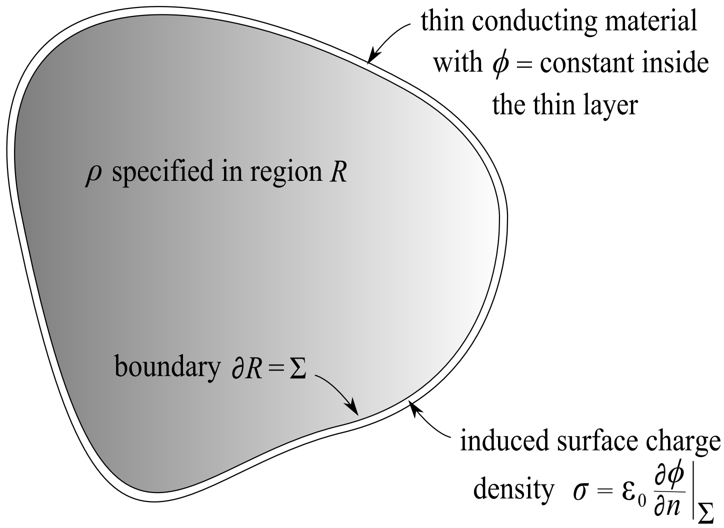


Figure 6: The Dirichlet problem with ϕ = constant on the boundary $\partial R = \Sigma$ of a region R containing a charge density ρ. Physically this boundary condition is enforced by wrapping a thin conducting material around Σ, with $\phi =$ constant inside this thin layer.

(ii) Consider the same set up, but where now Σ is replaced by a thin layer of *insulating* material, on which we place a *specific* surface charge density σ. By definition the insulator doesn't allow these charges to move. Assuming that $\mathbf{E}^+ = \mathbf{0}$ in the exterior of R, the same equation (2.3) still holds, but where we now interpret this as fixing the normal derivative of ϕ on Σ.

As we explain at the end of this subsection, the electric field will also be zero *outside* the thin conducting material, provided there are no charges outside of R and the conductor is *grounded*, meaning the potential ϕ on Σ is the same as that "at infinity". We shall see various physical examples later in this section, but for now we focus mainly on developing the general mathematical theory. We begin with the following:

Proposition 2.1 (*Green's identities*) *For any closed surface Σ bounding a region R, and for all suitably differentiable functions u, v in R, we have*

$$
\begin{aligned}
&\text{(a)} \qquad \int_R \left(u\,\nabla^2 v + \nabla u \cdot \nabla v\right) \mathrm{d}V \;=\; \int_\Sigma u\,\frac{\partial v}{\partial n}\,\mathrm{d}S\ , \\
&\text{(b)} \qquad \int_R \left(u\,\nabla^2 v - v\,\nabla^2 u\right) \mathrm{d}V \;=\; \int_\Sigma \left(u\,\frac{\partial v}{\partial n} - v\,\frac{\partial u}{\partial n}\right) \mathrm{d}S\ .
\end{aligned}
\tag{2.4}
$$

Proof For (a) apply the divergence theorem A.2 to $\mathbf{f} = u\,\nabla v$, where then $\nabla \cdot \mathbf{f} = u\,\nabla^2 v + \nabla u \cdot \nabla v$. Then for (b) simply take (a) and subtract the same equation with u and v interchanged. ■

Theorem 2.2 *The Poisson equation with either Dirichlet or Neumann boundary condition has a unique solution for ϕ, in the latter case up to an unphysical additive constant.*

Proof Let ϕ_1, ϕ_2 be two solutions to (2.1) in the interior of a region R with closed boundary surface $\Sigma = \partial R$, with the *same* boundary conditions on Σ. Define $\psi \equiv \phi_1 - \phi_2$, so that the

Laplace equation

$$\nabla^2\psi \;=\; \nabla^2\phi_1 - \nabla^2\phi_2 \;=\; -\frac{\rho}{\epsilon_0} - \left(-\frac{\rho}{\epsilon_0}\right) \;=\; 0\;, \tag{2.5}$$

holds inside R, with boundary conditions either (i) $\psi = 0$ on Σ, or (ii) $\partial\psi/\partial n = 0$ on Σ. Now apply Green's first identity (a) in Proposition 2.1, with $u = v = \psi$, to deduce

$$\int_R \left(\psi\,\nabla^2\psi + \nabla\psi\cdot\nabla\psi\right)\mathrm{d}V \;=\; \int_\Sigma \psi\,\frac{\partial\psi}{\partial n}\,\mathrm{d}S \;=\; 0\;. \tag{2.6}$$

Since $\nabla^2\psi = 0$ this gives $\int_R |\nabla\psi|^2\,\mathrm{d}V = 0$. Notice that $\nabla\psi$ needs to be differentiable in order for the Laplace equation to make sense, so the integrand $|\nabla\psi|^2$ is manifestly non-negative and continuous. We deduce that $\nabla\psi = \mathbf{0}$ and hence $\psi = \text{constant}$ in R, so that $\phi_1 = \phi_2 + \text{constant}$. Recall that the electrostatic potential is only determined up to such an additive constant, which drops out of the electric field $\mathbf{E} = -\nabla\phi$. However, in the Dirichlet case $\phi_1 = \phi_2$ on the boundary then forces this to hold everywhere in R, with $\psi \equiv 0$. ■

In principle we could consider *mixed* boundary conditions, writing $\Sigma = \Sigma_D \cup \Sigma_N$ as a disjoint union, with ϕ specified on Σ_D, and $\partial\phi/\partial n$ specified on Σ_N. The uniqueness proof in Theorem 2.2 still goes through, where on the right hand side of (2.6) either $\psi = 0$ or $\partial\psi/\partial n = 0$ on the respective components of Σ. The following is an important physical application of Theorem 2.2:

Theorem 2.3 *Consider a region R with zero charge density ρ inside R, which is bounded by a closed surface $\Sigma = \partial R$ consisting of a thin electrical conductor. Then the electric field is zero inside R.*

Proof From Theorem 1.7, the surface Σ is an equipotential for ϕ. On the other hand, since $\rho = 0$ inside R then $\phi = \text{constant}$ inside R solves the Poisson equation with Dirichlet boundary condition. Theorem 2.2 implies this is the unique solution, and we deduce that $\mathbf{E} = -\nabla\phi = \mathbf{0}$ inside R. ■

This is sometimes referred to as a *Faraday cage*: the thin conductor "shields" its interior from any external electric field. This is why it's hard to get a mobile phone signal inside a building where the walls have been reinforced with a steel mesh – that phone signal is an electromagnetic wave.

Finally, recall that in Figure 6 we had a thin conducting material wrapped around Σ. Assuming there is no charge in $R^c \equiv \mathbb{R}^3 \setminus R$, then the charge is confined to a bounded region in $\mathbb{R}^3$, and Theorem 1.4 says that $\phi = O(1/r)$ at infinity. By definition, the conductor is *grounded* if $\phi = 0$ on Σ, so that the potential is the same as that at infinity. We may then consider the Dirichlet problem for R^c, which has two boundary components: Σ and a sphere of very large radius $r \to \infty$. Since $\phi = 0$ on this boundary, then $\phi = 0$ inside R^c solves the Poisson equation with Dirichlet boundary condition, and it follows that $\mathbf{E} = \mathbf{0}$ on $R^c = \mathbb{R}^3 \setminus R$.

2.2 Green's functions

To investigate these boundary value problems further, we next introduce:

Definition A *Green's function* is a function $G(\mathbf{r},\mathbf{r}')$ satisfying

$$\nabla'^2 G(\mathbf{r},\mathbf{r}') \;=\; -4\pi\,\delta(\mathbf{r}-\mathbf{r}')\ , \tag{2.7}$$

where $\mathbf{r},\mathbf{r}' \in R$. Notice we have written the derivative with respect to the second, primed coordinates $\mathbf{r}'$ in $G(\mathbf{r},\mathbf{r}')$, and recall $\delta(\mathbf{r}-\mathbf{r}') = \delta(\mathbf{r}'-\mathbf{r})$.

From equation (1.24), we can write down the *particular* solution to (2.7)

$$G(\mathbf{r},\mathbf{r}') \;=\; \frac{1}{|\mathbf{r}-\mathbf{r}'|}\ . \tag{2.8}$$

Physically, this is $4\pi\epsilon_0$ times the electrostatic potential generated by a unit charge at position $\mathbf{r}'$. On the other hand, this solution is not unique: the *general* solution to (2.7) is

$$G(\mathbf{r},\mathbf{r}') \;=\; \frac{1}{|\mathbf{r}-\mathbf{r}'|} + F(\mathbf{r},\mathbf{r}')\ , \tag{2.9}$$

where $F(\mathbf{r},\mathbf{r}')$ satisfies the Laplace equation in R:

$$\nabla'^2 F(\mathbf{r},\mathbf{r}') \;=\; 0\ . \tag{2.10}$$

To see how this helps us solve the Poisson equation (2.1), take Green's second identity (b) in Proposition 2.1, with $u = \phi(\mathbf{r}')$ and $v = G(\mathbf{r},\mathbf{r}')$, where the integrals are over the primed coordinates $\mathbf{r}'$. Then

$$\int_R \left[\phi(\mathbf{r}')\,\nabla'^2 G(\mathbf{r},\mathbf{r}') - G(\mathbf{r},\mathbf{r}')\,\nabla'^2\phi(\mathbf{r}')\right]\mathrm{d}V' \;=\; \int_\Sigma \left[\phi(\mathbf{r}')\,\frac{\partial G(\mathbf{r},\mathbf{r}')}{\partial n'} - G(\mathbf{r},\mathbf{r}')\,\frac{\partial\phi(\mathbf{r}')}{\partial n'}\right]\mathrm{d}S'\ , \tag{2.11}$$

where this holds for all $\mathbf{r}\in R$. Using the definition (2.7) and the Poisson equation (2.1), the left hand side of (2.11) is

$$-4\pi\,\phi(\mathbf{r}) + \frac{1}{\epsilon_0}\int_R G(\mathbf{r},\mathbf{r}')\,\rho(\mathbf{r}')\,\mathrm{d}V'\ , \tag{2.12}$$

so that rearranging (2.11) we have proven

Proposition 2.4 *The electrostatic potential inside the region R may be expressed as*

$$\phi(\mathbf{r}) \;=\; \frac{1}{4\pi\epsilon_0}\int_R G(\mathbf{r},\mathbf{r}')\,\rho(\mathbf{r}')\,\mathrm{d}V' + \frac{1}{4\pi}\int_{\Sigma=\partial R}\left[G(\mathbf{r},\mathbf{r}')\,\frac{\partial\phi(\mathbf{r}')}{\partial n'} - \phi(\mathbf{r}')\,\frac{\partial G(\mathbf{r},\mathbf{r}')}{\partial n'}\right]\mathrm{d}S'\ . \tag{2.13}$$

This is a key formula, so let us make some remarks:

- The Green's function $G(\mathbf{r},\mathbf{r}')$ in (2.13) is *any* solution to (2.7). In particular, we may choose any solution F to the Laplace equation (2.10) in constructing $G(\mathbf{r},\mathbf{r}')$ via (2.9).

- The density ρ in R is specified as part of the data of the problem, while (i) in the Dirichlet problem it is ϕ that is specified on the boundary, while (ii) in the Neumann problem it is $\partial\phi/\partial n$ that is specified. However, *both* terms appear in the boundary intergal in (2.13).

- We immediately recognize the solution (1.28) as a special case of (2.13), where we take the Green's function (2.8). In (1.28) we took the charge density ρ to be supported in a bounded region of space, and $\phi(\mathbf{r})$ given by (1.28) then solves the Poisson equation in the whole of $\mathbb{R}^3$. The boundary terms in (2.13) are then effectively evaluated at infinity, where both the Green's function (2.8) and the potential ϕ are $O(1/r)$ and hence tend to zero.

Let us analyse (2.13) for the two boundary value problems in more detail:

(i) **Dirichlet** When ϕ is specified on Σ, it is convenient to choose the Green's function to also satisfy the Dirichlet boundary condition (denoted with a subscript D on G_D)

$$G_D(\mathbf{r},\mathbf{r}') \ = \ 0 \ , \qquad \text{for all } \mathbf{r}' \in \Sigma \ , \ \mathbf{r} \in R \ . \tag{2.14}$$

That is, the Green's function, viewed as a function of the second variable $\mathbf{r}'$, for any fixed $\mathbf{r}$, is zero on $\Sigma = \partial R$. Given such a function, (2.13) becomes

$$\phi(\mathbf{r}) \ = \ \frac{1}{4\pi\epsilon_0}\int_R G_D(\mathbf{r},\mathbf{r}')\,\rho(\mathbf{r}')\,\mathrm{d}V' - \frac{1}{4\pi}\int_\Sigma \phi(\mathbf{r}')\,\frac{\partial G_D(\mathbf{r},\mathbf{r}')}{\partial n'}\,\mathrm{d}S' \ . \tag{2.15}$$

Notice we have effectively reduced the problem to constructing the Green's function solution $G = G_D$ to (2.7), with Dirichlet boundary condition (2.14). This depends *only* on the region R, with boundary $\Sigma = \partial R$. Assuming we know G_D, then given *any* charge density ρ and *any* prescribed ϕ on Σ, the (unique by Theorem 2.2) solution to the Poisson equation with this Dirichlet boundary data is (2.15).

(ii) **Neumann** In this case we begin by using the divergence theorem to show that for $\mathbf{r} \in R$

$$\begin{aligned} -4\pi \ &= \ \int_R -4\pi\,\delta(\mathbf{r}-\mathbf{r}')\,\mathrm{d}V' \ = \ \int_R \nabla'^2\,G(\mathbf{r},\mathbf{r}')\,\mathrm{d}V' \ = \ \int_R \nabla'\cdot\nabla'\,G(\mathbf{r},\mathbf{r}')\,\mathrm{d}V' \\ &= \ \int_\Sigma \nabla'\,G(\mathbf{r},\mathbf{r}')\cdot\mathrm{d}\mathbf{S}' \ = \ \int_\Sigma \frac{\partial G(\mathbf{r},\mathbf{r}')}{\partial n'}\,\mathrm{d}S' \ . \end{aligned} \tag{2.16}$$

Thus we cannot simply set $\partial G(\mathbf{r},\mathbf{r}')/\partial n'$ to zero on Σ, analogously to (2.14). Instead a convenient condition to impose is

$$\frac{\partial G_N(\mathbf{r},\mathbf{r}')}{\partial n'} \ = \ -\frac{4\pi}{A} \ , \qquad \text{for all } \mathbf{r}' \in \Sigma \ , \ \mathbf{r} \in R \ , \tag{2.17}$$

where $A = \int_\Sigma \mathrm{d}S$ is the *area* of Σ. This is consistent with (2.16), and (2.13) becomes

$$\phi(\mathbf{r}) \ = \ \frac{1}{4\pi\epsilon_0}\int_R G_N(\mathbf{r},\mathbf{r}')\,\rho(\mathbf{r}')\,\mathrm{d}V' + \frac{1}{4\pi}\int_\Sigma G_N(\mathbf{r},\mathbf{r}')\,\frac{\partial\phi(\mathbf{r}')}{\partial n'}\,\mathrm{d}S' + \langle\,\phi\,\rangle \ , \tag{2.18}$$

where we have defined

$$\langle\,\phi\,\rangle \ \equiv \ \frac{1}{A}\int_\Sigma \phi(\mathbf{r}')\,\mathrm{d}S' \ , \tag{2.19}$$

which is an *average* of ϕ, per unit area, on the boundary. Notice that in the Neumann problem we are free to shift $\phi \to \phi +$ constant, without changing the boundary condition, and using this we may set $\langle \phi \rangle$ to any value (for example, $\langle \phi \rangle = 0$ is natural), thus fixing this non-uniqueness. Having found the Green's function G_N satisfying (2.17), again given any charge density ρ and any prescribed $\partial\phi/\partial n$ on Σ, the solution to the Poisson equation with this Neumann boundary data is (2.18).

In summary, we have reduced both problems to finding Green's functions (2.7) satisfying either the Dirichlet (2.14) or Neumann (2.17) boundary conditions. It is a general theorem in analysis that such Green's functions do in fact exist, assuming the boundary $\Sigma = \partial R$ is suitably well-behaved (for example, smooth is sufficient). Indeed, notice from (2.9) that finding such Green's functions is equivalent to solving the Laplace equation for $F(\mathbf{r}, \mathbf{r}')$, with appropriate boundary conditions. However, finding such functions *explicitly* is in general very hard. In the remainder of this section we will study methods that allow us to solve a variety of boundary value electrostatics problems, and find the Green's functions for the associated domains R. We focus mainly on the Dirichlet problem, where we note the following result:

Proposition 2.5 *The Dirichlet Green's function G_D, solving (2.7) subject to the boundary condition (2.14), is* symmetric*; that is, $G_D(\mathbf{r}_1, \mathbf{r}_2) = G_D(\mathbf{r}_2, \mathbf{r}_1)$ holds for all $\mathbf{r}_1, \mathbf{r}_2 \in R$.*

Proof We again make use of Green's second identity (b) in Proposition 2.1, this time with $u = G(\mathbf{r}_1, \mathbf{r}')$ and $v = G(\mathbf{r}_2, \mathbf{r}')$, where the integrals are over the primed coordinates $\mathbf{r}'$. We deduce

$$\begin{aligned} 4\pi \left[G_D(\mathbf{r}_2, \mathbf{r}_1) - G_D(\mathbf{r}_1, \mathbf{r}_2)\right] &= \int_R \left[G_D(\mathbf{r}_1, \mathbf{r}') \, \nabla'^2 \, G_D(\mathbf{r}_2, \mathbf{r}') - G_D(\mathbf{r}_2, \mathbf{r}') \, \nabla'^2 \, G_D(\mathbf{r}_1, \mathbf{r}')\right] \mathrm{d}V' \\ &= \int_\Sigma \left[G_D(\mathbf{r}_1, \mathbf{r}') \frac{\partial G_D(\mathbf{r}_2, \mathbf{r}')}{\partial n'} - G_D(\mathbf{r}_2, \mathbf{r}') \frac{\partial G_D(\mathbf{r}_1, \mathbf{r}')}{\partial n'}\right] \mathrm{d}S' \, , \\ &= 0 \, , \end{aligned} \tag{2.20}$$

where the boundary term is immediately zero due to the Dirichlet condition (2.14). ■

Remark Notice the Green's function $G(\mathbf{r}, \mathbf{r}') = 1/|\mathbf{r} - \mathbf{r}'|$ in (2.8) is the Dirichlet Green's function on $\mathbb{R}^3$ that is zero on the sphere at infinity, and is indeed symmetric.

2.3 Method of images

Consider the Dirichlet problem in a region $R \subset \mathbb{R}^3$ with boundary $\partial R = \Sigma$, where for simplicity we begin with a point charge distribution in R. Thus $\mathbf{r}_1, \ldots, \mathbf{r}_N \in R$ are the locations of point charges $q_1, \ldots, q_N$, and from (1.27) we know these generate an electrostatic potential

$$\phi(\mathbf{r})_{\text{point charges}} \equiv \frac{1}{4\pi\epsilon_0} \sum_{i=1}^{N} \frac{q_i}{|\mathbf{r} - \mathbf{r}_i|} \, . \tag{2.21}$$

This solves the Poisson equation with density $\rho(\mathbf{r}) = \sum_{i=1}^{N} q_i\, \delta(\mathbf{r}-\mathbf{r}_i)$ everywhere in $\mathbb{R}^3$, and hence in $R \subset \mathbb{R}^3$, but it will not in general satisfy the prescribed Dirichlet boundary condition for ϕ on Σ, which *a priori* is arbitrary.

We have already seen a hint of how to fix this in the last subsection: as in (2.9), write

$$\phi(\mathbf{r}) \;=\; \phi(\mathbf{r})_{\text{point charges}} + F(\mathbf{r}) \; . \tag{2.22}$$

If $F(\mathbf{r})$ satisfies the *Laplace equation* in R, then ϕ satisfies the same Poisson equation in R.

The method of images Take $F(\mathbf{r})$ to be the *electrostatic potential* generated by a set of charges $q_1^\star, \ldots, q_M^\star$ that lie *outside* the domain R, *i.e.* set

$$F(\mathbf{r}) \;=\; \frac{1}{4\pi\epsilon_0}\sum_{i=1}^{M}\frac{q_i^\star}{|\mathbf{r}-\mathbf{r}_i^\star|} \;\equiv\; \phi(\mathbf{r})_{\text{image charges}} \; , \tag{2.23}$$

with $\mathbf{r}_i^\star \in \mathbb{R}^3 \setminus R$, so that $\nabla^2 F = 0$ holds *inside* R. The idea is to choose these charges and their locations so that (2.22) satisfies the prescribed boundary condition for ϕ on $\Sigma = \partial R$.

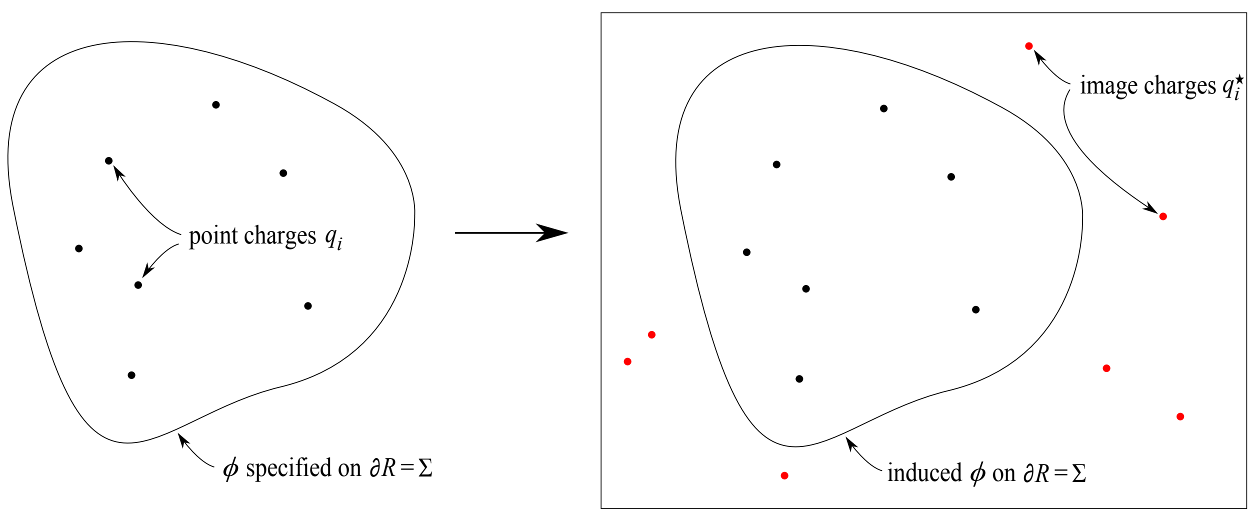


Figure 7: On the left hand side: the original Dirichlet problem, with point charges $q_1, \ldots, q_N$ inside R, and ϕ specified on $\partial R = \Sigma$. Replace with the right hand side: image charges $q_1^\star, \ldots, q_M^\star$ are added *outside* R, so that the solution to the Poisson equation on $\mathbb{R}^3$ induces the given ϕ on Σ.

The additional charges $q_1^\star, \ldots, q_M^\star$ are known as *image charges* (see the example below). They aren't part of the original problem, since they lie outside the domain R. Adding them to an enlarged domain is just a mathematical trick to solve the original boundary value problem on R. Of course, there is an art to this: how do we know where to put these image charges? There is no simple answer in general, but in certain problems it is clear from the geometry.

That is the mathematical problem, but what about in physical problems? Recall that in section 2.1 we described how the Dirichlet problem arises naturally when the boundary $\partial R = \Sigma$ is surrounded by a conducting material. This forces ϕ to be constant on Σ, and if the conductor is grounded then in fact $\phi\mid_{\mathbb{R}^3\setminus R}=0$, as explained at the end of section 2.1. As in Figure 6, this will result in a discontinuity of the normal derivative $\partial\phi/\partial n$ across Σ, which we then interpret physically in terms of a surface charge density on Σ, due to conduction electrons. The method of images gives us a mathematical way to solve this physics problem: forget about the conductor and the fact that we want $\phi=0$ in the region external to R, and instead introduce fictitious image charges in this region, as in (2.23), so that the potential in (2.22) has $\phi=0$ on Σ. Once we have this solution in $\mathbb{R}^3$, which by construction solves the correct Poisson equation inside R with the correct boundary condition $\phi\mid_\Sigma=0$, now simply set $\phi=0$ also outside R to obtain the solution to the original physics problem.

Example (Infinite conducting plane) Consider an infinite conducting plane at $\{z=0\}\subset\mathbb{R}^3$, and place a point charge q at position vector $\mathbf{r}_0=(x_0,y_0,z_0)$, with $z_0>0$. We want to solve the Poisson equation for ϕ in the region $R\equiv\{z\geq 0\}$, namely

$$\nabla^2\phi(\mathbf{r}) \;=\; -\frac{q}{\epsilon_0}\,\delta(\mathbf{r}-\mathbf{r}_0)\;, \qquad \text{for } z>0\;, \tag{2.24}$$

with the boundary condition that $\phi\mid_\Sigma=0$ on the conducting plane boundary $\Sigma\equiv\{z=0\}$.

As in (2.21) we can write down

$$\phi(\mathbf{r})_{\text{point charge}} \;\equiv\; \frac{1}{4\pi\epsilon_0}\frac{q}{|\mathbf{r}-\mathbf{r}_0|}\;. \tag{2.25}$$

This solves the Poisson equation in R, but it does not satisfy the boundary condition that $\phi\mid_\Sigma=0$. The problem is that for a point charge source the electric field points radially outwards, as in Figure 2, and the electric field lines then do not generically cross $\Sigma=\{z=0\}$ perpendicularly. However, we can solve this problem with the method of images: place an image charge $q^\star\equiv -q$ at the mirror image point to $\mathbf{r}_0$

$$\mathbf{r}_0^\star \;\equiv\; (x_0,y_0,-z_0)\;, \tag{2.26}$$

across the plane Σ. The electric field generated by this charge has the same magnitude as the original charge, but points in the opposite direction. If we superpose these electric fields generated by opposite charges at mirror image points across Σ, the net electric field tangent to Σ will be zero – see Figure 8. To see this explicitly, following (2.22), (2.23), put

$$\phi(\mathbf{r}) \;=\; \phi(\mathbf{r})_{\text{point charge}}+\phi(\mathbf{r})_{\text{image charge}} \;=\; \frac{1}{4\pi\epsilon_0}\frac{q}{|\mathbf{r}-\mathbf{r}_0|}-\frac{1}{4\pi\epsilon_0}\frac{q}{|\mathbf{r}-\mathbf{r}_0^\star|}\;. \tag{2.27}$$

This satisfies:

(i) $\phi\mid_\Sigma=0$, where restricting to Σ means setting $\mathbf{r}=(x,y,0)$. (Geometrically, this is because any point on Σ is equidistant from $\mathbf{r}_0$ and its mirror image point $\mathbf{r}_0^\star$ through Σ.)

(ii) The Poisson equation in $\mathbb{R}^3$:

$$\nabla^2\phi(\mathbf{r}) \;=\; -\frac{q}{\epsilon_0}\,\delta(\mathbf{r}-\mathbf{r}_0)+\frac{q}{\epsilon_0}\,\delta(\mathbf{r}-\mathbf{r}_0^\star)\;. \tag{2.28}$$

In particular this satisfies the correct Poisson equation inside $R = \{z \geq 0\}$, given in (2.24), precisely because the image charge lies *outside* of R. Having found this solution in R with the correct boundary condition on Σ, the solution to our original physics problem is given by taking $\phi(\mathbf{r})$ to be (2.27) for $z > 0$, and $\phi \equiv 0$ for $z \leq 0$.

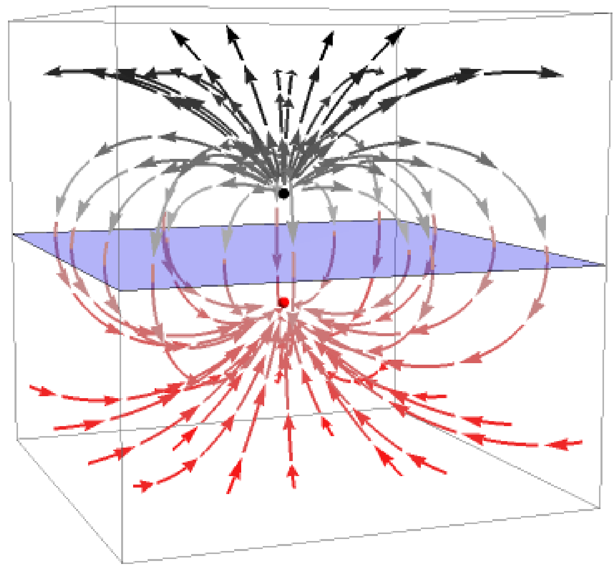


Figure 8: Electric field lines for $\mathbf{E} = -\nabla\phi$, with $\phi(\mathbf{r})$ given by (2.27). The original charge q is shown in black, and is taken to be positive, while the image charge $q^\star = -q$ is red. Note $\mathbf{E}$ is perpendicular to the (x, y)-plane $\Sigma = \{z = 0\}$, shown in blue. The solution to the Dirichlet problem for $z \geq 0$ simply discards the red part of the figure, beneath the conducting plane Σ.

Having found ϕ, we can now compute the induced surface charge density σ on the conducting plane Σ. This is given by (2.3), where the outward unit normal to Σ is $\mathbf{n} = -\mathbf{e}_3$. For simplicity we set $\mathbf{r}_0 = (0, 0, z_0)$, and compute (careful with the signs)

$$\begin{aligned} \sigma &= -\,\epsilon_0 \frac{\partial\phi}{\partial z}\bigg|_\Sigma = -\frac{q}{4\pi}\frac{\partial}{\partial z}\left[\frac{1}{\sqrt{x^2+y^2+(z-z_0)^2}} - \frac{1}{\sqrt{x^2+y^2+(z+z_0)^2}}\right]\Bigg|_{z=0} \\ &= -\frac{qz_0}{2\pi}\frac{1}{(x^2+y^2+z_0^2)^{3/2}} \ . \end{aligned} \tag{2.29}$$

Notice this is rotationally symmetric about the origin $\{x = y = 0\}$, where σ takes its maximum value, and tends to zero at infinity. The *total charge* Q induced on the conducting plane is

$$Q \equiv \int_\Sigma \sigma\,\mathrm{d}S = -\frac{qz_0}{2\pi}\int_{x=-\infty}^{\infty}\int_{y=-\infty}^{\infty}\frac{1}{(x^2+y^2+z_0^2)^{3/2}}\,\mathrm{d}x\,\mathrm{d}y = -q \ . \tag{2.30}$$

Notice this is the *same* as the image charge $q^\star = -q$.

Using the above results, we may now also construct the Dirichlet Green's function G_D in the region $R = \{z \geq 0\}$. Recall from section 2.2 this satisfies

$$\begin{cases} \nabla'^2 G_D(\mathbf{r},\mathbf{r}') = -4\pi\,\delta(\mathbf{r}-\mathbf{r}') \ , & \mathbf{r}',\mathbf{r} \in R \ , \\ \quad G_D(\mathbf{r},\mathbf{r}') = 0 \ , & \mathbf{r}' \in \Sigma \ , \mathbf{r} \in R \ . \end{cases} \tag{2.31}$$

But this is solved by taking G_D to be $4\pi\epsilon_0/q$ times the solution we have already obtained for ϕ, where $\mathbf{r}' = \mathbf{r}_0$ is the location of the point charge. We can hence simply write down the solution

$$G_D(\mathbf{r}, \mathbf{r}') \;=\; \frac{1}{|\mathbf{r}-\mathbf{r}'|} - \frac{1}{|\mathbf{r}-\mathbf{r}'_\star|} \;, \qquad (2.32)$$

where $\mathbf{r}' \equiv (x', y', z')$, and $\mathbf{r}'_\star \equiv (x', y', -z')$ is the reflection of $\mathbf{r}'$ across $\Sigma = \{z = 0\}$. Comparing to (2.9), the second term on the right hand side may be identified with $F(\mathbf{r}, \mathbf{r}')$, which satisfies the Laplace equation in R, as $\mathbf{r}'_\star$ lies outside of R.

Using (2.15) we may then write down the potential $\phi(\mathbf{r})$ for *any* distribution of charge $\rho(\mathbf{r})$ in the region R:

$$\phi(\mathbf{r}) \;=\; \frac{1}{4\pi\epsilon_0}\int_R G_D(\mathbf{r}, \mathbf{r}')\, \rho(\mathbf{r}')\, \mathrm{d}V' \;, \qquad (2.33)$$

where we have used the boundary condition $\phi\mid_\Sigma = 0$, so that the boundary term in (2.15) is zero. For example, for the original problem of a point charge q at position $\mathbf{r}_0$ we have $\rho(\mathbf{r}) = q\,\delta(\mathbf{r}-\mathbf{r}_0)$, and (2.33) gives $\phi(\mathbf{r}) = \frac{q}{4\pi\epsilon_0}G(\mathbf{r}, \mathbf{r}_0)$, which is the solution (2.27). For a *general* charge distribution ρ in R we may write (2.33) as

$$\phi(\mathbf{r}) \;=\; \frac{1}{4\pi\epsilon_0}\int_R \frac{\rho(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}\,\mathrm{d}V' - \frac{1}{4\pi\epsilon_0}\int_R \frac{\rho(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'_\star|}\,\mathrm{d}V' \;. \qquad (2.34)$$

The first term is the usual solution (1.28) for the electrostatic potential generated by the charge distribution ρ in R, while the second term is the electrostatic potential generated by the *image of minus this charge distribution*, after reflection across Σ. ■

Example (Conducting sphere) Consider a point charge q placed at a point $\mathbf{r}_0$ *outside* a grounded conducting sphere of radius a, centred on the origin. From Theorem 2.3, the electrostatic potential is zero on and inside the sphere, so that $\phi(\mathbf{r}) = 0$ for all $\mathbf{r}$ with $|\mathbf{r}| \le a$. Write $\mathbf{r}_0 = r_0\,\hat{\mathbf{r}}_0$, where $\hat{\mathbf{r}}_0$ is a unit vector pointing from the origin towards the charge q, with that latter a distance $r_0 > a$ from the origin.

It is less clear how to use the method of images in this case, but the *simplest* possibility would be to use a *single* image charge $q^\star$ at a point *inside* the sphere. By symmetry this image charge must lie on the line joining the origin to the original charge q, so we write the ansatz

$$\phi(\mathbf{r}) \;=\; \frac{1}{4\pi\epsilon_0}\left(\frac{q}{|\mathbf{r}-\mathbf{r}_0|} + \frac{q^\star}{|\mathbf{r}-\mathbf{r}_0^\star|}\right) \;, \qquad (2.35)$$

where $\mathbf{r}_0^\star = r_0^\star\,\hat{\mathbf{r}}_0$, with $0 \le r_0^\star < a$. We will have solved the problem if we can show there are $q^\star$ and $r_0^\star$ such that $\phi(\mathbf{r})$ given by (2.35) satisfies $\phi(\mathbf{r})\mid_\Sigma = 0$, where $\Sigma = \{|\mathbf{r}| = a\}$. We obtain two equations by imposing $\phi(\mathbf{r}) = 0$ on (2.35), where $\mathbf{r} = \pm a\,\hat{\mathbf{r}}_0$, namely

$$0 \;=\; \frac{q}{r_0 - a} + \frac{q^\star}{a - r_0^\star} \;, \qquad 0 \;=\; \frac{q}{a + r_0} + \frac{q^\star}{a + r_0^\star} \;, \qquad (2.36)$$

respectively, where note that $0 < r_0^\star < a < r_0$. These are easily solved to give

$$q^\star \;=\; -\frac{a}{r_0}\,q \;, \qquad r_0^\star \;=\; \frac{a}{r_0}\,a \;=\; \frac{a^2}{r_0} \;. \qquad (2.37)$$

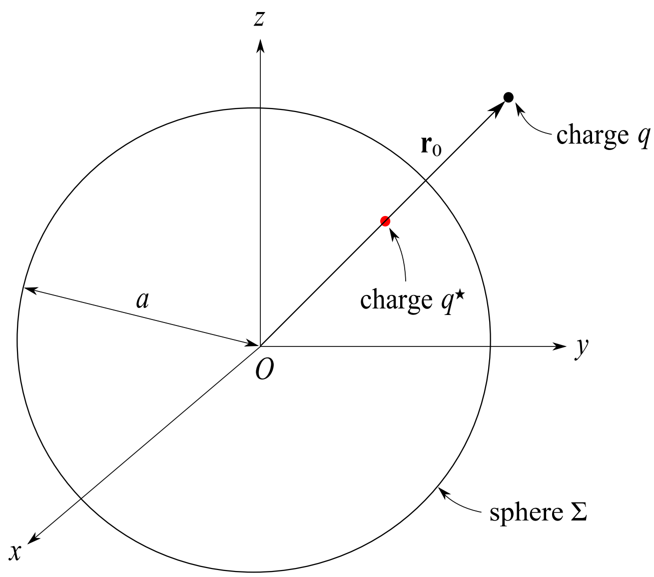


Figure 9: Conducting sphere Σ of radius a, centred on the origin O, with charge q at position $\mathbf{r}_0$ outside the sphere, and image charge $q^\star$ lying on the same line through O, but inside the sphere.

Notice that $\mathbf{r}_0$ and its image point $\mathbf{r}_0^\star$ are then geometrically *inverse points* for the sphere Σ. That is, they lie on the same straight line through the origin of the sphere, with $\mathbf{r}_0 \cdot \mathbf{r}_0^\star = a^2$. Substituting (2.37) back into (2.35), and writing $\mathbf{r} = r\,\hat{\mathbf{r}}$ with $\hat{\mathbf{r}}$ a unit vector, we have

$$\phi(\mathbf{r}) \;=\; \frac{q}{4\pi\epsilon_0}\left(\frac{1}{\sqrt{r^2+r_0^2-2rr_0\,\hat{\mathbf{r}}\cdot\hat{\mathbf{r}}_0}} - \frac{a}{r_0}\frac{1}{\sqrt{r^2+(\frac{a^2}{r_0})^2-2\frac{a^2}{r_0}r\,\hat{\mathbf{r}}\cdot\hat{\mathbf{r}}_0}}\right)\,, \tag{2.38}$$

It is then straightforward to see that putting $r = a$ on the right hand side of (2.38) indeed gives zero, for all directions $\hat{\mathbf{r}}$. Thus (2.38) is the required solution for ϕ, for $\mathbf{r} \in R \equiv \{|\mathbf{r}| \geq a\}$.

Since $\hat{\mathbf{r}}$ points away from the origin, it is the *inward* unit normal to the region $R = \{|\mathbf{r}| \geq a\}$, and thus (2.3) reads

$$\sigma \;=\; -\epsilon_0\;\hat{\mathbf{r}}\cdot\nabla\phi\,|_\Sigma \;\;=\; -\epsilon_0\,\frac{\partial\phi}{\partial r}\bigg|_{r=a} \;=\; -\frac{q}{4\pi}\,\frac{r_0^2-a^2}{a\left(a^2+r_0^2-2ar_0\,\hat{\mathbf{r}}\cdot\hat{\mathbf{r}}_0\right)^{3/2}}\,, \tag{2.39}$$

the last expression being obtained after a short computation using (2.38). For fixed $\mathbf{r}_0$, without loss of generality we may take this to point along the z-axis, so that $\hat{\mathbf{r}}\cdot\hat{\mathbf{r}}_0 = \cos\theta$, with θ the usual spherical polar coordinate giving the angle between $\hat{\mathbf{r}}$ and $\hat{\mathbf{r}}_0$. The total charge induced on the spherical conductor is then

$$\begin{aligned} Q \;\equiv\; \int_\Sigma \sigma\,\mathrm{d}S \;=\; \int_{\theta=0}^{\pi}\int_{\varphi=0}^{2\pi}\sigma(\theta)\,a^2\sin\theta\,\mathrm{d}\theta\,\mathrm{d}\varphi \;=\; 2\pi\cdot\frac{q}{4\pi}\left[\frac{r_0^2-a^2}{r_0\sqrt{a^2+r_0^2-2ar_0\cos\theta}}\right]_{\theta=0}^{\pi} \\ =\; -\frac{a}{r_0}\,q\,, \end{aligned} \tag{2.40}$$

which again is the same as the image charge $q^\star = -\frac{a}{r_0}\,q$.

Finally, we can write down the Dirichlet Green's function $G_D(\mathbf{r},\mathbf{r}')$ in the region $R = \{|\mathbf{r}| \geq a\}$. The discussion is analogous to that for the previous example, with this being $4\pi\epsilon_0/q$ times the

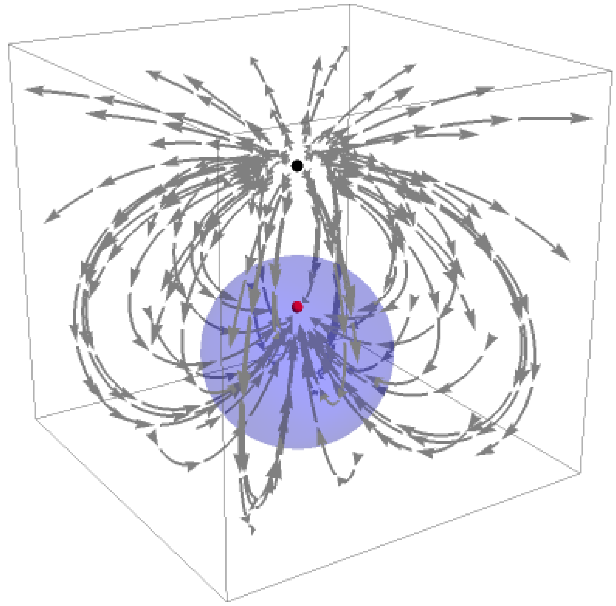


Figure 10: Electric field lines for $\mathbf{E} = -\nabla\phi$, with $\phi(\mathbf{r})$ given by (2.38). The original charge q is shown in black, and is taken to be positive, while the negative image charge $q^\star$ is red. These are located at *inverse points* for the sphere Σ, shown in blue. Note $\mathbf{E}$ is perpendicular to Σ.

solution (2.27) for ϕ, where $\mathbf{r}' = \mathbf{r}_0$ is the location of the point charge. Writing $\mathbf{r}' = r'\,\hat{\mathbf{r}}'$, we can identify $r_0 = r'$ and recall that $r_0^\star = a^2/r_0 = a^2/r'$, so that $\mathbf{r}'_\star = r_0^\star\,\hat{\mathbf{r}}' = (a^2/r'^2)\,\mathbf{r}'$. From (2.35) we then write down[6]

$$G_D(\mathbf{r},\mathbf{r}') \;=\; \frac{1}{|\mathbf{r}-\mathbf{r}'|} - \frac{a}{r'}\frac{1}{|\mathbf{r}-(a/r')^2\,\mathbf{r}'|}\;. \tag{2.41}$$

Consider the Dirichlet boundary problem where we fix $\phi(\mathbf{r})\mid_\Sigma = V(\theta,\varphi)$ to be a given function on the sphere Σ of radius a, and for simplicity suppose that there is no charge distribution, so that $\rho \equiv 0$. Only the second term in (2.15) now contributes, and recalling that our unit vectors $\hat{\mathbf{r}}$, $\hat{\mathbf{r}}'$ point *into* R, we compute

$$\left.\frac{\partial G_D(\mathbf{r},\mathbf{r}')}{\partial n'}\right|_{\mathbf{r}'\in\Sigma} \;=\; -\left.\frac{\partial G_D(\mathbf{r},\mathbf{r}')}{\partial r'}\right|_{r'=a} \;=\; -\frac{r^2-a^2}{a\,(r^2+a^2-2ar\,\hat{\mathbf{r}}\cdot\hat{\mathbf{r}}')^{3/2}}\;. \tag{2.42}$$

Here the last equality follows from the same calculation as (2.39). The solution for the potential (2.15) is then

$$\phi(\mathbf{r}) \;=\; \frac{1}{4\pi}\int_{\theta'=0}^{\pi}\int_{\varphi'=0}^{2\pi} V(\theta',\varphi')\,\frac{a(r^2-a^2)}{(r^2+a^2-2ar\,\hat{\mathbf{r}}\cdot\hat{\mathbf{r}}')^{3/2}}\sin\theta'\,\mathrm{d}\theta'\,\mathrm{d}\varphi'\;. \tag{2.43}$$

Here to perform the integral one should also write $\hat{\mathbf{r}}' = (\sin\theta'\cos\varphi', \sin\theta'\sin\varphi', \cos\theta')$ in spherical polars. Notice that the normal derivative of the Green's function (2.42) is acting as an integral kernel in (2.43), effectively "propagating" the prescribed electrostatic potential $\phi(\mathbf{r})\mid_\Sigma = V(\theta,\varphi)$ into the region outside this sphere. Physically we are holding the surface of the sphere Σ at fixed

[6]Notice this is symmetric $G_D(\mathbf{r},\mathbf{r}') = G_D(\mathbf{r}',\mathbf{r})$, as must be the case by Proposition 2.5: to see this for the second term write $|r'\mathbf{r} - (a^2/r')\,\mathbf{r}'|^2 = r'^2r^2 + a^4 - 2a^2\,\mathbf{r}\cdot\mathbf{r}'$.

voltage $V(\theta,\varphi)$, and (2.43) is the resulting electrostatic potential outside the sphere. In this case Σ couldn't be a conductor for non-constant V, but it could be an insulator. ■

2.4 Orthonormal functions

The electrostatic Maxwell equations (1.16), (1.33), or equivalently Poisson's equation (2.1), are *linear* – a manifestation of the Principle of Superposition. Solutions may then be expanded in terms of a convenient *basis of solutions*. This is a powerful technique, where the particular basis of functions is chosen according to the symmetries of the problem.

2.4.1 General theory

Definition A set of complex-valued functions $u_n : [a,b] \to \mathbb{C}$, defined on the interval $[a,b] \subset \mathbb{R}$ and labelled by a countable index $n \in I$ (usually $I = \mathbb{Z}^+$ or $I = \mathbb{Z}$), is said to be *orthonormal* if

$$\int_a^b \overline{u_m(x)}\, u_n(x)\, \mathrm{d}x \;=\; \delta_{mn}\ , \qquad \forall m,n \in I\ . \tag{2.44}$$

Suppose that a function $f : [a,b] \to \mathbb{C}$ may be expanded as a uniformly convergent series

$$f(x) \;=\; \sum_{n\in I} c_n\, u_n(x)\ , \tag{2.45}$$

for coefficients $c_n \in \mathbb{C}$. Then the coefficients may be determined via

$$\int_a^b \overline{u_m(x)}\, f(x)\, \mathrm{d}x \;=\; \sum_{n\in I} c_n \int_a^b \overline{u_m(x)}\, u_n(x)\, \mathrm{d}x \;=\; c_m\ , \tag{2.46}$$

where we have used (2.45) and uniform convergence in the first equality, and (2.44) in the second equality. Substituting (2.46) back into (2.45) then gives

$$f(x) \;=\; \sum_{n\in I} c_n\, u_n(x) \;=\; \sum_{n\in I} \int_a^b \overline{u_n(x')}\, f(x')\, u_n(x)\, \mathrm{d}x' \;=\; \int_a^b f(x') \left(\sum_{n\in I} \overline{u_n(x')}\, u_n(x) \right) \mathrm{d}x'\ . \tag{2.47}$$

If this is to hold for *any* function, then from the definition of the Dirac delta function we identify

$$\sum_{n\in I} \overline{u_n(x')}\, u_n(x) \;=\; \delta(x'-x)\ . \tag{2.48}$$

Definition The set of orthonormal functions $\{u_n(x)\}_{n\in I}$ is said to be *complete* if (2.48) holds.

Notice then that $f(x) = \int_a^b f(x')\,\delta(x'-x)\,\mathrm{d}x'$, and substituting (2.48) and reading equation (2.47) from right to left says that $f(x)$ may be expanded as in (2.45).

Example (Fourier sine series) Consider a function $f : [0,a] \to \mathbb{R}$ with $f(0) = f(a) = 0$. Then we may expand f in a *Fourier sine series*[7], with complete set of orthonormal functions

$$u_n(x) \;\equiv\; \sqrt{\frac{2}{a}}\, \sin\left(\frac{\pi n}{a} x\right)\ , \qquad n \in \mathbb{Z}^+\ . \tag{2.49}$$

[7]Particularly familiar to those who took Part A Quantum Theory.

That is, $f(x)$ may be expanded as in (2.45) with $I = \mathbb{Z}^+$, so

$$f(x) = \sqrt{\frac{2}{a}} \sum_{n=1}^{\infty} c_n \sin\left(\frac{\pi n}{a} x\right) , \tag{2.50}$$

where the coefficients are given by (2.46)

$$c_n = \int_0^a \overline{u_n(x)}\, f(x)\, \mathrm{d}x = \sqrt{\frac{2}{a}} \int_0^a \sin\left(\frac{\pi n}{a} x\right) f(x)\, \mathrm{d}x . \tag{2.51}$$

The series $\sum_{n=1}^{\infty} c_n\, u_n(x)$ converges absolutely and uniformly to $f(x)$, provided the latter is continuously differentiable. More generally, for *any* function that is square-integrable, meaning $\int_0^a |f(x)|^2\, \mathrm{d}x < \infty$, the Fourier series converges almost everywhere to f. This general subject is put on a rigorous footing in functional analysis. We also have

$$\delta_N(x, x') \equiv \sum_{n=1}^{N} u_n(x')\, u_n(x) = \frac{2}{a} \sum_{n=1}^{N} \sin\left(\frac{\pi n}{a} x'\right) \sin\left(\frac{\pi n}{a} x\right) \overset{N\to\infty}{\longrightarrow} \delta(x' - x) , \tag{2.52}$$

(understood in terms of distributions). The function $\delta_N(x, x')$ is plotted in Figure 11. ■

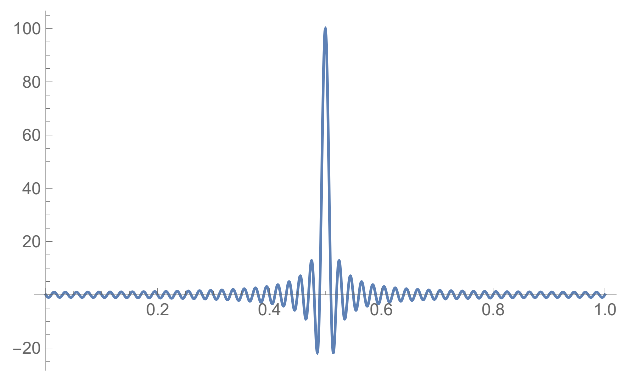


Figure 11: The function $\delta_N(x, x')$ in (2.52), with $x' = \frac{1}{2}$ the midpoint of $[0, 1]$, and $N = 100$.

Example (Complex exponential Fourier series) More generally, from Prelims you know that a function on $[-\frac{L}{2}, \frac{L}{2}]$ may be expanded in terms of Fourier modes involving both $\cos\left(\frac{2\pi n}{L} x\right)$ and $\sin\left(\frac{2\pi n}{L} x\right)$. The Fourier sine series arises in the special case that $f : [-\frac{L}{2}, \frac{L}{2}] \to \mathbb{R}$ is an *odd* function, and in the previous example $L = 2a$ and we then restrict this function to $x \geq 0$. It is sometimes convenient to rewrite the sine and cosine functions in terms of complex exponentials, leading to a basis of functions

$$u_n(x) \equiv \frac{1}{\sqrt{L}}\, \mathrm{e}^{\mathrm{i}\left(\frac{2\pi n}{L}\right)x} , \qquad n \in \mathbb{Z} . \tag{2.53}$$

Notice that now $n \in I = \mathbb{Z}$, and it is straightforward to check the orthonormal property (2.44). The usual Fourier expansion may be written as

$$f(x) = \frac{1}{\sqrt{L}} \sum_{n=-\infty}^{\infty} c_n\, \mathrm{e}^{\mathrm{i}\left(\frac{2\pi n}{L}\right)x} = \sum_{n=-\infty}^{\infty} c_n\, u_n(x) , \tag{2.54}$$

as in (2.45), where the coefficients are

$$c_n \;=\; \frac{1}{\sqrt{L}}\int_{-L/2}^{L/2} \mathrm{e}^{-\mathrm{i}\left(\frac{2\pi n}{L}\right)x}\, f(x)\,\mathrm{d}x \;=\; \int_{-L/2}^{L/2} \overline{u_n(x)}\, f(x)\,\mathrm{d}x\ . \qquad \blacksquare \qquad (2.55)$$

The last example has a natural "continuum limit", where formally we take the size of the interval $L\to\infty$. In this case the discrete set of orthonormal functions $\{u_n(x)\}_{n\in\mathbb{Z}}$ becomes a *continuous* set of functions $\{u_k(x)\}_{k\in\mathbb{R}}$, and correspondingly sums such as (2.45) are replaced by integrals, and the Kronecker delta symbol in the orthonormality relation (2.44) is replaced by a Dirac delta function. The main example is:

Example (Exponential Fourier expansion/Fourier transform) Analogously to (2.53), consider

$$u_k(x) \;\equiv\; \frac{1}{\sqrt{2\pi}}\,\mathrm{e}^{\mathrm{i}kx}\ , \qquad (2.56)$$

where now the index $k\in\mathbb{R}$ is a *continuous* variable, rather than discrete. Note that the prefactor has been changed compared to $u_n(x)$ in (2.53). A general integrable function on $\mathbb{R}$ may be expanded as

$$f(x) \;=\; \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} C(k)\,\mathrm{e}^{\mathrm{i}kx}\,\mathrm{d}k\ , \qquad (2.57)$$

which is usually called the *exponential Fourier expansion*. The "coefficients" $C(k)$, which are now complex-valued functions on $\mathbb{R}$, are given by (*cf.* (2.46))

$$C(k) \;=\; \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} \mathrm{e}^{-\mathrm{i}kx}\, f(x)\,\mathrm{d}x \;=\; \int_{-\infty}^{\infty} \overline{u_k(x)}\, f(x)\,\mathrm{d}x\ , \qquad (2.58)$$

which is also called the *Fourier transform* of $f(x)$. The orthonormality condition (2.44) now reads

$$\int_{-\infty}^{\infty} \overline{u_{k'}(x)}\, u_k(x)\,\mathrm{d}x \;=\; \frac{1}{2\pi}\int_{-\infty}^{\infty} \mathrm{e}^{\mathrm{i}(k-k')x}\,\mathrm{d}x \;=\; \delta(k-k')\ , \qquad (2.59)$$

while the completeness relation (2.48) takes the similar-looking form

$$\int_{-\infty}^{\infty} \overline{u_k(x')}\, u_k(x)\,\mathrm{d}k \;=\; \frac{1}{2\pi}\int_{-\infty}^{\infty} \mathrm{e}^{\mathrm{i}k(x-x')}\,\mathrm{d}k \;=\; \delta(x-x')\ . \qquad (2.60)$$

These give useful representations of the Dirac delta function. $\blacksquare$

Remark Note that if two expansions are equal to each other, we may identify their coefficients. For example, if $\int_{-\infty}^{\infty} C(k)\,u_k(x)\,\mathrm{d}k = \int_{-\infty}^{\infty} D(k)\,u_k(x)\,\mathrm{d}k$ then using (2.59)

$$\begin{aligned}\int_{x=-\infty}^{\infty}\left[\int_{k=-\infty}^{\infty} C(k)\,u_k(x)\,\mathrm{d}k\right]\overline{u_{k'}(x)}\,\mathrm{d}x \;&=\; \int_{x=-\infty}^{\infty}\left[\int_{k=-\infty}^{\infty} D(k)\,u_k(x)\,\mathrm{d}k\right]\overline{u_{k'}(x)}\,\mathrm{d}x\\ \Longrightarrow\quad \int_{-\infty}^{\infty} C(k)\,\delta(k-k')\,\mathrm{d}k \;&=\; \int_{-\infty}^{\infty} D(k)\,\delta(k-k')\,\mathrm{d}k \quad\Longrightarrow\quad C(k') \;=\; D(k')\ . \qquad (2.61)\end{aligned}$$

2.4.2 Cartesian coordinates

Consider the Laplace equation in Cartesian coordinates

$$\nabla^2\phi \;=\; \frac{\partial^2\phi}{\partial x^2}+\frac{\partial^2\phi}{\partial y^2}+\frac{\partial^2\phi}{\partial z^2} \;=\; 0\;. \tag{2.62}$$

This is of course Poisson's equation (2.1) with zero charge density $\rho\equiv 0$. We may seek separable solutions by substituting $\phi(x,y,z)=X(x)\,Y(y)\,Z(z)$ into (2.62), and dividing through by ϕ, giving

$$\frac{1}{X}\frac{\mathrm{d}^2X}{\mathrm{d}x^2}+\frac{1}{Y}\frac{\mathrm{d}^2Y}{\mathrm{d}y^2}+\frac{1}{Z}\frac{\mathrm{d}^2Z}{\mathrm{d}z^2} \;=\; 0\;. \tag{2.63}$$

The usual separation of variable argument implies that each term is separately constant, so that

$$\frac{1}{X}\frac{\mathrm{d}^2X}{\mathrm{d}x^2} \;=\; -\alpha^2\;, \qquad \frac{1}{Y}\frac{\mathrm{d}^2Y}{\mathrm{d}y^2} \;=\; -\beta^2\;, \qquad \frac{1}{Z}\frac{\mathrm{d}^2Z}{\mathrm{d}x^2} \;=\; \alpha^2+\beta^2\;. \tag{2.64}$$

Here *a priori* the constants α^2,β^2 are any real numbers, so that α,β can be real or purely imaginary, although in the example below we will take $\alpha,\beta\in\mathbb{R}$. The two solutions to the first ODE in (2.64) are $X(x)=A_\pm\,\mathrm{e}^{\pm\mathrm{i}\alpha x}$, with $A_\pm$ integration constants, and linearity of the original Laplace equation (2.62) implies that any linear combination of

$$\phi(x,y,z) \;=\; \mathrm{e}^{\pm\mathrm{i}\alpha x}\,\mathrm{e}^{\pm\mathrm{i}\beta y}\,\mathrm{e}^{\pm\sqrt{\alpha^2+\beta^2}z} \tag{2.65}$$

is a solution. Imposing *boundary conditions* on ϕ will then restrict the possible values of α,β, and also the coefficients in the series.

Example (Electrostatic potential for a rectangular box) Consider a rectangular box $R=\{(x,y,z)\mid x\in[0,a]\;,\;y\in[0,b]\;,\;z\in[0,c]\}$, and consider the Dirichlet problem in which $\phi=0$ on all boundary surfaces except the face $\{z=c\}$, where we impose $\phi(x,y,c)=V(x,y)$, with charge density $\rho\equiv 0$.[8]

The resulting Laplace equation in R has separable solutions which are linear combinations of (2.65). Let us look at a solution for fixed α,β, and impose the boundary conditions

$$\phi\,|_{x=0} \;=\; 0\;, \qquad \phi\,|_{y=0} \;=\; 0\;, \qquad \phi\,|_{z=0} \;=\; 0\;. \tag{2.66}$$

Writing $\mathrm{e}^{\pm\mathrm{i}\alpha x}=\cos\alpha x\pm\mathrm{i}\sin\alpha x$, (2.66) sets the coefficient of the cosine term to zero, so that the solutions satisfying (2.66) take the form

$$\sin\alpha x\,\sin\beta y\,\sinh(\sqrt{\alpha^2+\beta^2}\,z)\;. \tag{2.67}$$

On the other hand, the boundary conditions $\phi\,|_{x=a}=0$, $\phi\,|_{y=b}=0$ then imply that $\sin\alpha a=0=\sin\beta b$, which implies

$$\alpha \;=\; \frac{m\pi}{a}\;, \qquad \beta \;=\; \frac{n\pi}{b}\;, \qquad m,n\in\mathbb{Z}^+\;. \tag{2.68}$$

[8] *cf.* the discussion of the Dirichlet Green's function outside a sphere, after equation (2.41).

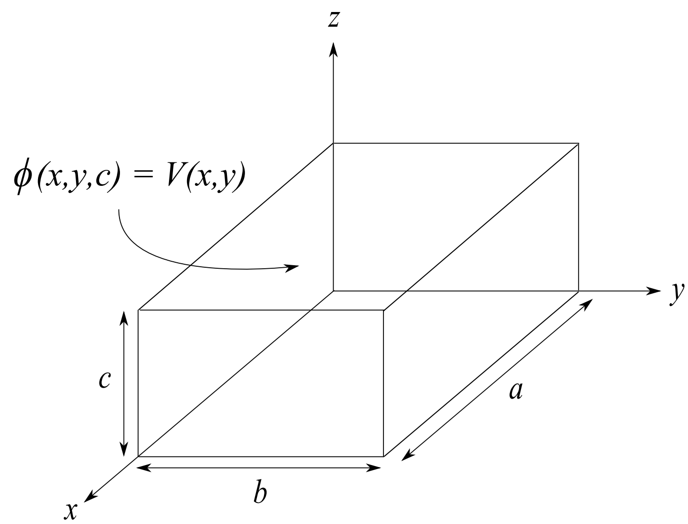


Figure 12: Rectangular box $R = \{(x,y,z) \mid x \in [0,a]\,,\, y \in [0,b]\,,\, z \in [0,c]\}$, with boundary condition that ϕ is zero on all boundary faces except $\{z = c\}$, where $\phi(x,y,c) = V(x,y)$.

Notice here that we have taken $m, n > 0$ without loss of generality, since sine is an odd function. We may then write

$$\phi(x,y,z) \;=\; \sum_{m,n=1}^{\infty} c_{m,n} \sin\left(\frac{m\pi}{a}x\right) \sin\left(\frac{n\pi}{b}y\right) \sinh\left(\pi\sqrt{\frac{m^2}{a^2}+\frac{n^2}{b^2}}\, z\right) , \qquad (2.69)$$

with coefficients $c_{m,n}$. From section 2.4.1, we recognize this as a Fourier sine series in the x and y variables.

Finally, we just need to impose the boundary condition $\phi\,|_{z=c}(x,y) = V(x,y)$. Setting $z = c$ in (2.69) gives the Fourier sine series expansion

$$\phi(x,y,c) \;=\; V(x,y) \;=\; \sum_{m,n=1}^{\infty} \left[c_{m,n} \sinh\left(\pi\sqrt{\frac{m^2}{a^2}+\frac{n^2}{b^2}}\, c\right)\right] \sin\left(\frac{m\pi}{a}x\right) \sin\left(\frac{n\pi}{b}y\right) . \qquad (2.70)$$

We can then read off the coefficients, in square brackets, from (2.50), (2.51), which gives

$$c_{m,n} \;=\; \frac{4}{ab \sinh\left(\pi\sqrt{\frac{m^2}{a^2}+\frac{n^2}{b^2}}\, c\right)} \int_{x=0}^{a}\int_{y=0}^{b} \sin\left(\frac{m\pi}{a}x\right) \sin\left(\frac{n\pi}{b}y\right) V(x,y)\, \mathrm{d}x\, \mathrm{d}y \;. \qquad \blacksquare \quad (2.71)$$

We know from the general theory in section 2.2 that the above problem must also have a solution in terms of the Dirichlet Green's function $G_D(\mathbf{r},\mathbf{r}')$. Recall this satisfies (2.7) for $\mathbf{r},\mathbf{r}' \in R$ inside the rectangular box, with Dirichlet boundary (2.14) that $G_D(\mathbf{r},\mathbf{r}')\,|_{\mathbf{r}'\in\Sigma} = 0$, where $\Sigma = \partial R$ is the boundary of the box. The solution for the electrostatic potential is then (2.15). Since here $\rho \equiv 0$ and $\phi\,|_{\Sigma}$ is zero except on the face $\phi\,|_{z=c} = V$, this solution reads

$$\phi(\mathbf{r}) \;=\; -\frac{1}{4\pi}\int_{z'=c} V(x',y')\, \frac{\partial G_D(\mathbf{r},\mathbf{r}')}{\partial n'}\, \mathrm{d}S' \;. \qquad (2.72)$$

The last example suggests making an orthonormal series expansion for G_D. This Green's function is then effectively constructed from the solution to the following physical problem:

Example (Point charge inside a rectangular box) Consider a point charge q inside a rectangular box that consists of a grounded conductor. Thus $\phi\,|_\Sigma = 0$, so that ϕ is zero on all six faces, and

$$\nabla^2\phi(\mathbf{r}) \;=\; -\frac{q}{\epsilon_0}\,\delta(\mathbf{r}-\mathbf{r}_0)\;, \tag{2.73}$$

for $\mathbf{r} = (x, y, z)$ inside the box R, where $\mathbf{r}_0 = (x_0, y_0, z_0) \in R$ is the location of the point charge.

Following the last example, we can solve this via separation of variables. Due to the boundary conditions we can expand each variable as a Fourier sine series, leading to

$$\phi(\mathbf{r}) \;=\; \sqrt{\frac{8}{abc}}\sum_{m,n,\ell=1}^{\infty} c_{m,n,\ell}\,\sin\left(\frac{m\pi}{a}x\right)\sin\left(\frac{n\pi}{b}y\right)\sin\left(\frac{\ell\pi}{c}z\right)\;, \tag{2.74}$$

with coefficients $c_{m,n,\ell}$. Substituting (2.74) into (2.73) then gives for the left hand side

$$\nabla^2\phi \;=\; \sqrt{\frac{8}{abc}}\sum_{m,n,\ell=1}^{\infty}\left[c_{m,n,\ell}\left(-\frac{m^2\pi^2}{a^2}-\frac{n^2\pi^2}{b^2}-\frac{\ell^2\pi^2}{c^2}\right)\right]\sin\left(\frac{m\pi}{a}x\right)\sin\left(\frac{n\pi}{b}y\right)\sin\left(\frac{\ell\pi}{c}z\right). \tag{2.75}$$

On the other hand, from (2.52) we can write the Dirac delta function in one dimension as $\delta(x-x_0) = \frac{2}{a}\sum_{m=1}^{\infty}\sin\left(\frac{m\pi}{a}x_0\right)\sin\left(\frac{m\pi}{a}x\right)$, and then (1.19) gives the right hand side of (2.73) as

$$\begin{aligned}-\frac{q}{\epsilon_0}\,\delta(\mathbf{r}-\mathbf{r}_0) \;=\; -\frac{q}{\epsilon_0}\,\frac{8}{abc}\sum_{m,n,\ell=1}^{\infty}\Bigg[&\sin\left(\frac{m\pi}{a}x_0\right)\sin\left(\frac{m\pi}{a}x\right)\sin\left(\frac{n\pi}{b}y_0\right)\sin\left(\frac{n\pi}{b}y\right)\\ &\times\sin\left(\frac{\ell\pi}{c}z_0\right)\sin\left(\frac{\ell\pi}{c}z\right)\Bigg]\;.\end{aligned} \tag{2.76}$$

Equating (2.76) with (2.75), the coefficients of $\sin\left(\frac{m\pi}{a}x\right)\sin\left(\frac{n\pi}{b}y\right)\sin\left(\frac{\ell\pi}{c}z\right)$ must be equal, giving

$$c_{m,n,\ell} \;=\; \frac{q}{\epsilon_0}\sqrt{\frac{8}{abc}}\,\frac{\sin\left(\frac{m\pi}{a}x_0\right)\sin\left(\frac{n\pi}{b}y_0\right)\sin\left(\frac{\ell\pi}{c}z_0\right)}{\left(\frac{m^2\pi^2}{a^2}+\frac{n^2\pi^2}{b^2}+\frac{\ell^2\pi^2}{c^2}\right)}\;. \tag{2.77}$$

Substituting this into (2.74) then gives the solution, expressed as a Fourier sine series. As for the examples in section 2.3, the Dirichlet Green's function $G_D(\mathbf{r},\mathbf{r}')$ for this problem is then simply $(4\pi\epsilon_0/q)\,\phi(\mathbf{r})$, with $\phi(\mathbf{r})$ given by (2.74), (2.77), and where we identify $\mathbf{r}_0 = \mathbf{r}'$. ■

In the above examples we have expanded the functions using the orthonormal basis of Fourier sine modes (2.49). This was a convenient basis to use due to the rectangular boundary conditions, which took a simple form after separating variables in Cartesian coordinates. However, in other problems it is convenient to use a different orthonormal basis.

Example (Fourier transform of the Dirichlet Greens function on $\mathbb{R}^3$) Consider the Dirichlet Green's function on $\mathbb{R}^3$ given by $G_D(\mathbf{r},\mathbf{r}') = 1/|\mathbf{r}-\mathbf{r}'|$. This satisfies (2.7), where from (2.60) and setting $\mathbf{r} = (x_1, x_2, x_3)$, $\mathbf{r}' = (x_1', x_2', x_3')$ we may write

$$\begin{aligned}\delta(\mathbf{r}-\mathbf{r}') \;&\equiv\; \delta(x_1-x_1')\,\delta(x_2-x_2')\,\delta(x_3-x_3')\\ &=\; \frac{1}{(2\pi)^3}\int_{-\infty}^{\infty}\mathrm{d}k_1\int_{-\infty}^{\infty}\mathrm{d}k_2\int_{-\infty}^{\infty}\mathrm{d}k_3\,\mathrm{e}^{\mathrm{i}k_1(x_1-x_1')}\,\mathrm{e}^{\mathrm{i}k_2(x_2-x_2')}\,\mathrm{e}^{\mathrm{i}k_3(x_3-x_3')}\\ &=\; \frac{1}{(2\pi)^3}\int_{\mathbf{k}\in\mathbb{R}^3}\mathrm{e}^{\mathrm{i}\mathbf{k}\cdot(\mathbf{r}-\mathbf{r}')}\,\mathrm{d}^3k\;,\end{aligned} \tag{2.78}$$

where $\mathbf{k} = (k_1, k_2, k_3)$, and $\mathrm{d}^3k \equiv \mathrm{d}k_1\,\mathrm{d}k_2\,\mathrm{d}k_3$. We may then similarly expand the Green's function as in (2.57), which using the compact notation on the last line of (2.78) reads

$$\frac{1}{|\mathbf{r}-\mathbf{r}'|} \;=\; \frac{1}{(2\pi)^{3/2}} \int_{\mathbf{k}\in\mathbb{R}^3} C(\mathbf{k})\,\mathrm{e}^{\mathrm{i}\mathbf{k}\cdot(\mathbf{r}-\mathbf{r}')}\,\mathrm{d}^3k\;, \tag{2.79}$$

with coefficient function $C(\mathbf{k})$, which is a function on $\mathbb{R}^3$. Noting that, analogously to the computation in (2.75), we have

$$\begin{aligned}\nabla'^2\left(\mathrm{e}^{\mathrm{i}\mathbf{k}\cdot(\mathbf{r}-\mathbf{r}')}\right) \;&=\; \left(\frac{\partial^2}{\partial x_1'^2}+\frac{\partial^2}{\partial x_2'^2}+\frac{\partial^2}{\partial x_3'^2}\right)\mathrm{e}^{\mathrm{i}\left[k_1(x_1-x_1')+k_2(x_2-x_2')+k_3(x_3-x_3')\right]}\\ &=\; \left(-k_1^2-k_2^2-k_3^2\right)\mathrm{e}^{\mathrm{i}\mathbf{k}\cdot(\mathbf{r}-\mathbf{r}')} \;=\; -|\mathbf{k}|^2\,\mathrm{e}^{\mathrm{i}\mathbf{k}\cdot(\mathbf{r}-\mathbf{r}')}\;,\end{aligned} \tag{2.80}$$

applying ∇'^2 to (2.79) gives

$$\nabla'^2\left(\frac{1}{|\mathbf{r}-\mathbf{r}'|}\right) \;=\; -\frac{1}{(2\pi)^{3/2}}\int_{\mathbf{k}\in\mathbb{R}^3} C(\mathbf{k})\,|\mathbf{k}|^2\,\mathrm{e}^{\mathrm{i}\mathbf{k}\cdot(\mathbf{r}-\mathbf{r}')}\,\mathrm{d}^3k\;. \tag{2.81}$$

Since $\nabla'^2(1/|\mathbf{r}-\mathbf{r}'|) = -4\pi\,\delta(\mathbf{r}-\mathbf{r}')$, equating (2.81) with -4π times (2.78) leads to the identification of coefficient functions (*cf.* the remark at the end of section 2.4.1)

$$C(\mathbf{k}) \;=\; \frac{4\pi}{(2\pi)^{3/2}}\,\frac{1}{|\mathbf{k}|^2}\;. \tag{2.82}$$

Thus the Dirichlet Green's function on $\mathbb{R}^3$ is

$$G_D(\mathbf{r},\mathbf{r}') \;=\; \frac{1}{|\mathbf{r}-\mathbf{r}'|} \;=\; \frac{1}{2\pi^2}\int_{\mathbf{k}\in\mathbb{R}^3}\frac{1}{|\mathbf{k}|^2}\,\mathrm{e}^{\mathrm{i}\mathbf{k}\cdot(\mathbf{r}-\mathbf{r}')}\,\mathrm{d}^3k\;. \tag{2.83}$$

We shall return to this formula in section 5.6. ■

2.4.3 Spherical polar coordinates

For problems with spherical symmetry, it is often more convenient to use a complete set of orthonormal functions adapted to that symmetry.

The Laplacian in spherical polar coordinates (r,θ,φ) is

$$\nabla^2\phi \;=\; \frac{1}{r}\frac{\partial^2(r\phi)}{\partial r^2} + \frac{1}{r^2}\left[\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\,\frac{\partial\phi}{\partial\theta}\right)+\frac{1}{\sin^2\theta}\frac{\partial^2\phi}{\partial\varphi^2}\right] \;\equiv\; \frac{1}{r}\frac{\partial^2(r\phi)}{\partial r^2}+\frac{1}{r^2}\nabla^2_{\theta,\varphi}\,\phi\;, \tag{2.84}$$

where the second equality defines the *angular Laplacian* $\nabla^2_{\theta,\varphi}\phi$ as the term in square brackets. Recall here that $r\geq 0$, $\theta\in[0,\pi]$, $\varphi\in[0,2\pi)$, with the latter being periodically identified. Following the start of section 2.4.2, we may again use separation of variables, but this time we write $\phi(r,\theta,\varphi) = R(r)\,Y(\theta,\varphi)$. Substituting this into (2.84) and dividing through by $\phi/r^2 = R(r)\,Y(\theta,\varphi)/r^2$, the Laplace equation reads

$$\frac{r}{R}\frac{\mathrm{d}^2(rR)}{\mathrm{d}r^2} + \frac{1}{Y}\nabla^2_{\theta,\varphi}Y \;=\; 0\;. \tag{2.85}$$

Both terms must be constant, so that

$$\frac{\mathrm{d}^2}{\mathrm{d}r^2}(rR(r)) \;=\; \lambda\,\frac{R(r)}{r}\;,\qquad \nabla^2_{\theta,\varphi}Y(\theta,\varphi) \;=\; -\lambda\,Y(\theta,\varphi)\;,\qquad \text{where } \lambda \;=\; \text{constant}\;. \tag{2.86}$$

Let us focus first on the second equation in (2.86), for the angular coordinates. We may also separate variables here, writing $Y(\theta,\varphi) = P(\theta)\,\Phi(\varphi)$. Dividing through by $Y/\sin^2\theta = P(\theta)\,\Phi(\varphi)/\sin^2\theta$, this equation then reads

$$\left[\frac{\sin\theta}{P}\frac{\mathrm{d}}{\mathrm{d}\theta}\left(\sin\theta\,\frac{\mathrm{d}P}{\mathrm{d}\theta}\right) + \lambda\,\sin^2\theta\right] + \frac{1}{\Phi}\,\frac{\mathrm{d}^2\Phi}{\mathrm{d}\varphi^2} \;=\; 0\ . \tag{2.87}$$

Again, both terms must be constant, with the second leading to the equation $\Phi''(\varphi) = -c\,\Phi(\varphi)$, with c constant. Since φ is a periodic coordinate, in order to be single-valued we must have $\Phi(\varphi) = \Phi(\varphi + 2\pi)$ for all φ, and this fixes $c = m^2$ with $m \in \mathbb{Z}$ an *integer*, with solutions $\Phi(\varphi) = \mathrm{e}^{\pm\mathrm{i}m\varphi}$. Substituting this back into (2.87) leads to a second order ODE for $P(\theta)$, depending on the constants m^2 and $\lambda \in \mathbb{R}$, called the *general or associated Legendre equation*. The solutions to this equation are typically singular in the limits $\theta \to 0, \pi$. To find *smooth* solutions $Y(\theta,\varphi) = Y(\hat{\mathbf{r}})$ on the unit sphere $\mathbb{S}^2 \equiv \{r = 1\} =\subset \mathbb{R}^3$, we need a condition on λ:[9]

Theorem 2.6 *In the space of smooth functions $Y\colon \mathbb{S}^2 \longrightarrow \mathbb{C}$, the eigenvalue equation*

$$\nabla^2_{\theta,\varphi}\,Y(\theta,\varphi) \;=\; -\lambda\,Y(\theta,\varphi) \tag{2.88}$$

has non-trivial solutions only for λ of the form $\lambda = \ell(\ell+1)$ with some non-negative integer $\ell \in \mathbb{Z}_{\geq 0}$. For such ℓ, the eigenspace has dimension $2\ell+1$ and is spanned by separable solutions of the form $Y(\theta,\varphi) = P(\theta)\mathrm{e}^{\mathrm{i}m\varphi}$ with $m \in \{-\ell, -\ell+1, \dots, \ell-1, \ell\}$.

We can write down such a basis of Eigenfunctions explicitly:

Definition The *spherical harmonics* $Y_{\ell,m}\colon \mathbb{S}^2 \longrightarrow \mathbb{C}$ are defined for integers $\ell \geq m \geq 0$ by

$$Y_{\ell,m}(\theta,\varphi) = \sqrt{\frac{2\ell+1}{4\pi}\frac{(\ell-m)!}{(\ell+m)!}}\cdot(-1)^m\cdot P_{\ell,m}(\cos\theta)\cdot\mathrm{e}^{\mathrm{i}m\varphi} \tag{2.89}$$

and $Y_{\ell,-m} = (-1)^m\overline{Y_{\ell,m}}$. Here, $P_{\ell,m}$ are the *associated Legendre functions*, defined by

$$\begin{aligned} P_{\ell,0}(x) &= \frac{1}{2^\ell\cdot\ell!}\frac{\mathrm{d}^\ell}{\mathrm{d}x^\ell}(x^2-1)^\ell \quad \text{and} \\ P_{\ell,m}(x) &= (1-x^2)^{m/2}\frac{\mathrm{d}^m}{\mathrm{d}x^m}P_{\ell,0}(x). \end{aligned} \tag{2.90}$$

The polynomials $P_\ell(x) = P_{\ell,0}(x)$ are known as *Legendre polynomials.*

Theorem 2.7 *The spherical harmonics $\{Y_{\ell,m}\}_{\ell\geq|m|\geq 0}$ are a complete set of orthonormal Eigenfunctions on the unit sphere $\mathbb{S}^2 = \{r = 1\} \subset \mathbb{R}^3$, for the operator $\nabla^2_{\theta,\varphi}$. They satisfy:*

$$\begin{aligned}
&\text{(i)}\quad Y_{\ell,-m}(\theta,\varphi) \;=\; (-1)^m\,\overline{Y_{\ell,m}(\theta,\varphi)}\ ,\\
&\text{(ii)}\quad \int_{\theta=0}^{\pi}\int_{\varphi=0}^{2\pi}\overline{Y_{\ell,m}(\theta,\varphi)}\,Y_{\ell',m'}(\theta,\varphi)\,\sin\theta\,\mathrm{d}\theta\,\mathrm{d}\varphi \;=\; \delta_{\ell,\ell'}\,\delta_{m,m'} \qquad \text{(orthonormal)}\ ,\\
&\text{(iii)}\quad \sum_{\ell=0}^{\infty}\sum_{m=-\ell}^{\ell}\overline{Y_{\ell,m}(\theta',\varphi')}\,Y_{\ell,m}(\theta,\varphi) \;=\; \frac{1}{\sin\theta}\,\delta(\varphi'-\varphi)\,\delta(\theta'-\theta) \qquad \text{(completeness)}\ ,\\
&\text{(iv)}\quad \nabla^2_{\theta,\varphi}Y_{\ell,m} = -\ell(\ell+1)Y_{\ell,m}\,.
\end{aligned} \tag{2.91}$$

[9]We state this here without proof. This material will be familar to those who took Part A Quantum Theory.

In particular, any square-integrable function $f(\theta, \varphi)$ on the sphere $\mathbb{S}^2$ can be uniquely written as

$$f(\theta, \varphi) = \sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{\ell} c_{\ell,m} Y_{\ell,m}(\theta, \varphi) , \tag{2.92}$$

where the coefficients $c_{\ell,m} \in \mathbb{C}$ are constants.

Compare the orthonormal and completeness conditions in (2.91) to the one-dimensional conditions in equations (2.44) and (2.48), respectively. The first few spherical harmonics are

$$\begin{aligned}
Y_{0,0} &= \sqrt{\frac{1}{4\pi}} , \\
Y_{1,0} &= \sqrt{\frac{3}{4\pi}} \cos\theta , \qquad & Y_{1,1} &= -\sqrt{\frac{3}{8\pi}} \sin\theta \, \mathrm{e}^{\mathrm{i}\varphi} , \\
Y_{2,0} &= \frac{1}{2}\sqrt{\frac{5}{4\pi}} \left(3\cos^2\theta - 1\right) , \quad & Y_{2,1} &= -\sqrt{\frac{15}{8\pi}} \sin\theta \, \cos\theta \, \mathrm{e}^{\mathrm{i}\varphi} , \quad Y_{2,2} = \frac{1}{4}\sqrt{\frac{15}{2\pi}} \sin^2\theta \, \mathrm{e}^{2\mathrm{i}\varphi} .
\end{aligned} \tag{2.93}$$

Notice we have only listed the functions with $m \geq 0$, with those for $m < 0$ determined by property (i) in (2.91).

We now return to the first, radial equation in (2.86), which reads

$$\frac{\mathrm{d}^2}{\mathrm{d}r^2}(rR(r)) = \ell(\ell+1)\frac{R(r)}{r} . \tag{2.94}$$

We may try to solve this by setting $R = r^\alpha$. Substituting this into (2.94) gives

$$\alpha(\alpha+1)\, r^{\alpha-1} = \ell(\ell+1)\, r^{\alpha-1} \quad \Longrightarrow \quad \alpha(\alpha+1) = \ell(\ell+1) , \tag{2.95}$$

which has solutions $\alpha = \ell$, $\alpha = -(\ell+1)$. The general solution is hence $R(r) = A\, r^\ell + B\, r^{-(\ell+1)}$, with A and B integration constants.

Putting everything together, a general solution to the Laplace equation $\nabla^2\phi = 0$ may be expanded in spherical harmonics as

$$\phi(\mathbf{r}) = \sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{\ell} \left(A_{\ell,m}\, r^\ell + B_{\ell,m}\, r^{-(\ell+1)}\right) Y_{\ell,m}(\theta, \varphi) , \tag{2.96}$$

with constants $A_{\ell,m}$, $B_{\ell,m}$.

Example (Electrostatic potential *inside* a sphere, with Dirichlet boundary condition) Consider the Dirichlet boundary problem inside a sphere of radius a, with $\phi(a, \theta, \varphi) = V(\theta, \varphi)$ prescribed on the boundary, and zero charge density $\rho \equiv 0$.

We thus want to solve the Laplace equation for $r < a$, with the above boundary condition. General solutions to the Laplace equation may be expanded as in (2.96), but notice that we should set the constants $B_{\ell,m} = 0$, as the negative powers $r^{-(\ell+1)}$ all diverge at the centre of the sphere $r = 0$. Imposing the boundary condition at $r = a$ then sets

$$\phi(a, \theta, \varphi) = V(\theta, \varphi) = \sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{\ell} a^\ell A_{\ell,m} Y_{\ell,m}(\theta, \varphi) . \tag{2.97}$$

Comparing to (2.92), this is simply the spherical harmonic expansion of the function $V(\theta,\varphi)$. We may compute the coefficients via

$$\begin{aligned}
& \int_{\theta=0}^{\pi}\int_{\varphi=0}^{2\pi} \overline{Y_{\ell,m}(\theta,\varphi)}\, V(\theta,\varphi)\, \sin\theta\, \mathrm{d}\theta\, \mathrm{d}\varphi \\
&= \sum_{\ell'=0}^{\infty}\sum_{m'=-\ell'}^{\ell} a^{\ell'} A_{\ell',m'} \int_{\theta=0}^{\pi}\int_{\varphi=0}^{2\pi} \overline{Y_{\ell,m}(\theta,\varphi)}\, Y_{\ell',m'}(\theta,\varphi)\, \sin\theta\, \mathrm{d}\theta\, \mathrm{d}\varphi \\
&= a^{\ell} A_{\ell,m} \ , \qquad (2.98)
\end{aligned}$$

where in the second line we have substituted the expansion (2.97), and the last line uses orthonormality in (2.91). Notice here that the factor of $1/\sin\theta$ on the right hand side of (iii) in (2.91) cancels the $\sin\theta$ factor in the area element $\sin\theta\, \mathrm{d}\theta\, \mathrm{d}\varphi$ for the sphere (see also the following example). Substituting for $A_{\ell,m}$ given by (2.98) into (2.96) then gives the solution. ■

Finally, comparing to the last example in section 2.3, we examine the following:

Example (Dirichlet Green's function *outside* a sphere, in spherical harmonics) Let $R = \{r \geq a\}$ be the region outside a sphere $\Sigma = \{r = a\}$ of radius a, centred on the origin, and consider the Dirichlet Green's function $G_D(\mathbf{r},\mathbf{r}')$ in R, which recall satisfies

$$\begin{aligned}
\nabla'^2 G_D(\mathbf{r},\mathbf{r}') &= -4\pi\,\delta(\mathbf{r}-\mathbf{r}') \ , \quad \text{and} \quad G_D(\mathbf{r},\mathbf{r}') = G_D(\mathbf{r}',\mathbf{r}) \ , \qquad \forall\ \mathbf{r},\mathbf{r}' \in R \ , \\
G_D(\mathbf{r},\mathbf{r}') &= 0 \ , \qquad \forall\ \mathbf{r} \in R,\ \mathbf{r}' \in \Sigma \ . \qquad (2.99)
\end{aligned}$$

Focusing first on the dependence on $\mathbf{r}$, which we write in spherical polars (r,θ,φ), we may expand in spherical harmonics as in (2.92) by writing

$$G_D(\mathbf{r},\mathbf{r}') = \sum_{\ell=0}^{\infty}\sum_{m=-\ell}^{\ell} c_{\ell,m}(r,\mathbf{r}')\, Y_{\ell,m}(\theta,\varphi) \ . \qquad (2.100)$$

Note the coefficients $c_{\ell,m}(r,\mathbf{r}')$ depend on the *scalar distance* r to the origin, and the *vector* position $\mathbf{r}'$. Acting on (2.100) with the Laplacian in spherical coordinates (2.84) gives

$$\nabla^2 G_D(\mathbf{r},\mathbf{r}') = \sum_{\ell=0}^{\infty}\sum_{m=-\ell}^{\ell} \left[\frac{1}{r}\frac{\mathrm{d}^2(r\, c_{\ell,m}(r,\mathbf{r}'))}{\mathrm{d}r^2} - \frac{1}{r^2}\,\ell(\ell+1)\, c_{\ell,m}(r,\mathbf{r}')\right] Y_{\ell,m}(\theta,\varphi) \ , \qquad (2.101)$$

where we have used $\nabla^2_{\theta,\varphi} Y_{\ell,m}(\theta,\varphi) = -\ell(\ell+1)\, Y_{\ell,m}(\theta,\varphi)$. The expression (2.101) should equal $-4\pi\,\delta(\mathbf{r}-\mathbf{r}')$, where notice that it doesn't matter whether we act with ∇^2 or ∇'^2, due to the symmetry $G_D(\mathbf{r},\mathbf{r}') = G_D(\mathbf{r}',\mathbf{r})$.

Lemma 2.8 *The Dirac delta function in spherical polar coordinates is*

$$\delta(\mathbf{r}-\mathbf{r}') = \frac{1}{r^2\sin\theta}\,\delta(r-r')\,\delta(\theta-\theta')\,\delta(\varphi-\varphi') \ . \qquad (2.102)$$

Proof Notice the denominator on the right hand side is precisely the *Jacobian* in going from Cartesian coordinates to spherical polar coordinates, and indeed this is also the origin of the $1/\sin\theta$ term on the right hand side of (iii) in equation (2.91). Starting with the defining property of the Dirac delta function we have

$$f(\mathbf{r}') = \int_{\mathbb{R}^3} f(\mathbf{r})\, \delta(\mathbf{r}-\mathbf{r}')\, \mathrm{d}x_1\, \mathrm{d}x_2\, \mathrm{d}x_3 \;=\; \int_{r=0}^{\infty}\int_{\theta=0}^{\pi}\int_{\varphi=0}^{2\pi} \left[f(\mathbf{r})\, \delta(\mathbf{r}-\mathbf{r}')\, r^2 \sin\theta\right] \mathrm{d}r\, \mathrm{d}\theta\, \mathrm{d}\varphi \tag{2.103}$$

In order for the right hand side to give $f(\mathbf{r}')$, thus picking out $r = r'$, $\theta = \theta'$, $\varphi = \varphi'$ in the integral, we precisely have to make the identification (2.102). ■

The completeness relation (iii) in (2.91) for the spherical harmonics in turn allows us to write

$$\delta(\mathbf{r}-\mathbf{r}') \;=\; \frac{1}{r^2}\, \delta(r-r') \sum_{\ell=0}^{\infty}\sum_{m=-\ell}^{\ell} \overline{Y_{\ell,m}(\theta',\varphi')}\, Y_{\ell,m}(\theta,\varphi)\ . \tag{2.104}$$

Equating (2.101) with -4π times (2.104) then allows us to equate the coefficients

$$\frac{1}{r}\frac{\mathrm{d}^2(r\, c_{\ell,m}(r,\mathbf{r}'))}{\mathrm{d}r^2} - \frac{1}{r^2}\,\ell(\ell+1)\, c_{\ell,m}(r,\mathbf{r}') \;=\; -\frac{4\pi}{r^2}\, \delta(r-r')\, \overline{Y_{\ell,m}(\theta',\varphi')}\ . \tag{2.105}$$

In particular, focusing on the $\mathbf{r}'$ dependence we may read off

$$c_{\ell,m}(r,\mathbf{r}') \;=\; A_{\ell,m}(r,r')\, \overline{Y_{\ell,m}(\theta',\varphi')}\ . \tag{2.106}$$

The Green's function (2.100) now reads

$$G_D(\mathbf{r},\mathbf{r}') \;=\; \sum_{\ell=0}^{\infty}\sum_{m=-\ell}^{\ell} A_{\ell,m}(r,r')\, \overline{Y_{\ell,m}(\theta',\varphi')}\, Y_{\ell,m}(\theta,\varphi)\ , \tag{2.107}$$

where from (2.105) the coefficient functions $A_{\ell,m}(r,r')$ satisfy

$$\frac{1}{r}\frac{\mathrm{d}^2(r\, A_{\ell,m}(r,r'))}{\mathrm{d}r^2} - \frac{1}{r^2}\,\ell(\ell+1)\, A_{\ell,m}(r,r') \;=\; -\frac{4\pi}{r^2}\, \delta(r-r')\ . \tag{2.108}$$

It remains to solve this second order ODE. When $r \neq r'$, so either $r < r'$ or $r > r'$, then (2.108) reduces to

$$\frac{\mathrm{d}^2(r\, A_{\ell,m}(r,r'))}{\mathrm{d}r^2} \;=\; \frac{1}{r}\,\ell(\ell+1)\, A_{\ell,m}(r,r')\ . \tag{2.109}$$

We already solved this equation in (2.94), where the two independent solutions are proportional to r^ℓ and $r^{-(\ell+1)}$. We may thus write down

$$A_{\ell,m}(r,r') \;=\; \begin{cases} A_\ell(r')\, r^\ell + B_\ell(r')\, r^{-(\ell+1)}\ , & a \le r < r'\ , \\ C_\ell(r')\, r^\ell + D_\ell(r')\, r^{-(\ell+1)}\ , & r > r' \ge a\ , \end{cases} \tag{2.110}$$

anticipating that the remaining coefficients will depend on ℓ, but not on m. In order for the Green's function to be bounded as $r \to \infty$, notice we must set $C_\ell(r') \equiv 0$. Next we impose the Dirichlet boundary condition at $r = a$ (or equivalently $r' = a$ due to the symmetry property on the first line of (2.99)). This sets

$$0 \;=\; A_{\ell,m}(a, r') \;=\; A_\ell(r')\, a^\ell + B_\ell(r')\, a^{-(\ell+1)} \quad\Longrightarrow\quad B_\ell(r') \;=\; -a^{2\ell+1}\, A_\ell(r') \ . \tag{2.111}$$

Then (2.110) reads

$$A_{\ell,m}(r, r') \;=\; \begin{cases} A_\ell(r')\left(r^\ell - \frac{a^{2\ell+1}}{r^{\ell+1}}\right) , & a \leq r < r' \ , \\ D_\ell(r')\, \frac{1}{r^{\ell+1}} \ , & r > r' \geq a \ , \end{cases} \tag{2.112}$$

Next we impose the symmetry property $A_{\ell,m}(r, r') = A_{\ell,m}(r', r)$.[10] From (2.112), this relates

$$D_\ell(r')\, \frac{1}{r^{\ell+1}} \;=\; A_\ell(r)\left(r'^\ell - \frac{a^{2\ell+1}}{r'^{\ell+1}}\right) \quad\Longrightarrow\quad \frac{D_\ell(r')}{\left(r'^\ell - \frac{a^{2\ell+1}}{r'^{\ell+1}}\right)} \;=\; A_\ell(r)\, r^{\ell+1} \;=\; K_\ell \ , \tag{2.113}$$

where in the last step notice both sides must be a *constant* K_ℓ. Substituting into (2.112) gives

$$A_{\ell,m}(r, r') \;=\; \begin{cases} \dfrac{K_\ell\, r^\ell}{r'^{\ell+1}}\left[1 - \left(\dfrac{a}{r}\right)^{2\ell+1}\right] , & a \leq r < r' \ , \\ \dfrac{K_\ell\, r'^\ell}{r^{\ell+1}}\left[1 - \left(\dfrac{a}{r'}\right)^{2\ell+1}\right] , & r > r' \geq a \ , \end{cases} \tag{2.114}$$

which is indeed symmetric under $r \leftrightarrow r'$, and continuous at $r = r'$.

It remains only to determine the constants K_ℓ, where notice that we haven't yet used the normalization of the Dirac delta function on the right hand side of (2.108). Multiplying this equation by r and integrating from $r = r' - \varepsilon$ to $r = r' + \varepsilon$ implies

$$\lim_{\varepsilon\to 0}\left[\frac{\mathrm{d}}{\mathrm{d}r}\left(r\, A_{\ell,m}(r, r')\right)\right]_{r'-\varepsilon}^{r'+\varepsilon} \;=\; -\frac{4\pi}{r'} \ . \tag{2.115}$$

Notice that the second term on the left hand side of (2.108) doesn't contribute in this limit, as $A_{\ell,m}(r, r')$ is continuous at $r = r'$. Note also that in the upper limit $r = r' + \varepsilon > r'$, while in the lower limit $r = r' - \varepsilon < r'$. Using (2.114) thus gives

$$\begin{aligned} -\frac{4\pi}{r'} \;&=\; K_\ell \lim_{\varepsilon\to 0}\left\{ \frac{\mathrm{d}}{\mathrm{d}r}\left(\frac{r'^\ell}{r^\ell}\left[1 - \left(\frac{a}{r'}\right)^{2\ell+1}\right]\right)\Bigg|_{r=r'+\varepsilon} - \frac{\mathrm{d}}{\mathrm{d}r}\left(\frac{r^{\ell+1}}{r'^{\ell+1}}\left[1 - \left(\frac{a}{r}\right)^{2\ell+1}\right]\right)\Bigg|_{r=r'-\varepsilon} \right\} , \\ &=\; K_\ell \lim_{\varepsilon\to 0}\left\{ -\,\ell\, \frac{r'^\ell}{(r'+\varepsilon)^{\ell+1}}\left[1 - \left(\frac{a}{r'}\right)^{2\ell+1}\right] - (\ell+1)\,\frac{(r'-\varepsilon)^\ell}{r'^{\ell+1}} - \ell\,\frac{(r'-\varepsilon)^\ell}{r'^{\ell+1}}\left(\frac{a}{r'-\varepsilon}\right)^{2\ell+1} \right\} \\ &=\; -K_\ell\, \frac{2\ell+1}{r'} \ , \end{aligned} \tag{2.116}$$

so that

$$K_\ell \;=\; \frac{4\pi}{2\ell+1} \ . \tag{2.117}$$

[10]This symmetry property was *derived* in Proposition 2.5, while here it is convenient to *impose* this to simplify finding the Green's function.

We can now write down the final form of the Dirichlet Green's function outside the sphere. Combining (2.107), (2.114) and (2.117) gives

$$G_D(\mathbf{r},\mathbf{r}') = \sum_{\ell=0}^{\infty}\sum_{m=-\ell}^{\ell} A_{\ell,m}(r,r')\,\overline{Y_{\ell,m}(\theta',\varphi')}\,Y_{\ell,m}(\theta,\varphi) \ , \tag{2.118}$$

Here we may write the compact expression

$$A_{\ell,m}(r,r') = \frac{4\pi}{2\ell+1}\frac{r_<^{\ell}}{r_>^{\ell+1}}\left[1-\left(\frac{a}{r_<}\right)^{2\ell+1}\right] \ , \tag{2.119}$$

where we have defined

$$\begin{cases} r_> \equiv r' \ , \quad r_< \equiv r \ , & r < r' \ , \\ r_> \equiv r \ , \quad r_< \equiv r' \ , & r > r' \ . \end{cases} \tag{2.120}$$

That is, $r_<$ is the smaller of r or r', while $r_>$ is the larger of r or r'. ■

3 Magnetostatics

3.1 Electric currents

So far we have been dealing with stationary charges. In this subsection we consider how to describe charges in motion.

Recall that in an electrical conductor there are electrons (the "conduction electrons") which are free to move when an external electric field is applied. Although these electrons move around fairly randomly, with typically large velocities, in the presence of a macroscopic electric field there is an induced *average drift velocity* $\mathbf{v} = \mathbf{v}(\mathbf{r})$. This is the average velocity of a particle at position $\mathbf{r}$. In fact, we might as well simply ignore the random motion, and regard the electrons as moving through the material with velocity vector field $\mathbf{v}(\mathbf{r})$.

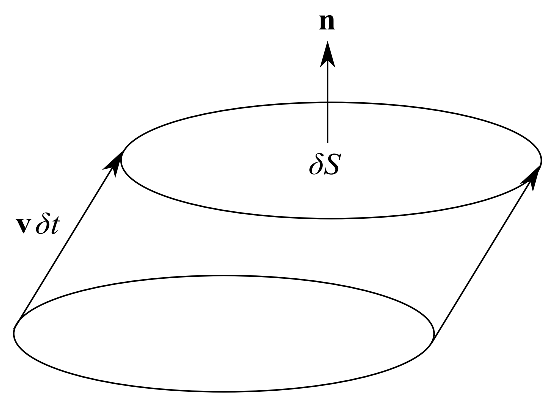


Figure 13: The current flow through a surface element $\delta\Sigma$ of area δS.

Definition Given a distribution of charge with density ρ and velocity vector field $\mathbf{v}$, the *electric current density* $\mathbf{J}$ is defined as

$$\mathbf{J} \equiv \rho\,\mathbf{v} \ . \tag{3.1}$$

To interpret this, imagine a small surface $\delta\Sigma$ of area δS at position $\mathbf{r}$, as shown in Figure 13. Recall that we define $\delta\mathbf{S} = \mathbf{n}\,\delta S$, where $\mathbf{n}$ is the unit normal vector to the surface $\delta\Sigma$. The volume of the oblique cylinder in Figure 13 is $\mathbf{v}\,\delta t \cdot \delta\mathbf{S}$, which thus contains the charge $\rho\,\mathbf{v}\,\delta t \cdot \delta\mathbf{S} = \mathbf{J} \cdot \delta\mathbf{S}\,\delta t$. In the time δt this is the total charge passing through $\delta\Sigma$. Thus $\mathbf{J}$ is a vector field in the direction of the flow, and its magnitude is the amount of charge flowing per unit time per unit perpendicular cross-section to the flow.

Definition The *electric current* $I = I(\Sigma)$ through a surface Σ is defined to be

$$I \equiv \int_\Sigma \mathbf{J} \cdot \mathrm{d}\mathbf{S} \ . \tag{3.2}$$

This is the rate of flow of charge through Σ. The units of electric current are $\mathrm{C\,s^{-1}}$, which is also called the *Ampère* A.

3.2 Continuity equation

An important property of electric charge is that it is *conserved*, *i.e.* it is neither created nor destroyed. There is a differential equation that expresses this experimental fact called the *continuity equation*.

Suppose that Σ is a closed surface bounding a region R, so $\partial R = \Sigma$. From the discussion of current density $\mathbf{J}$ in the previous subsection, we see that the rate of flow of electric charge passing *out* of Σ is given by the current (3.2) through Σ. On the other hand, the total charge in R is

$$Q = \int_R \rho \, \mathrm{d}V \ . \tag{3.3}$$

If electric charge is conserved, then the rate of charge passing out of Σ must equal *minus* the rate of change of Q:

$$\int_\Sigma \mathbf{J} \cdot \mathrm{d}\mathbf{S} = -\frac{\mathrm{d}Q}{\mathrm{d}t} = -\int_R \frac{\partial \rho}{\partial t} \mathrm{d}V \ . \tag{3.4}$$

Here we have allowed time dependence in $\rho = \rho(\mathbf{r}, t)$. Using the divergence theorem A.2 this becomes

$$\int_R \left(\frac{\partial \rho}{\partial t} + \nabla \cdot \mathbf{J} \right) \mathrm{d}V = 0 \ , \tag{3.5}$$

which holds for all R, and thus

$$\boxed{\frac{\partial \rho}{\partial t} + \nabla \cdot \mathbf{J} = 0 \ .} \tag{3.6}$$

This is the *continuity equation*.

In magneto*statics* we shall impose $\partial \rho / \partial t = 0$, and thus

$$\nabla \cdot \mathbf{J} = 0 \ . \tag{3.7}$$

Definition Currents $\mathbf{J}$ satisfying (3.7) are called *steady currents*.

3.3 Lorentz force and the magnetic field

The force on a point charge q at *rest* in an electric field $\mathbf{E}$ is simply $\mathbf{F} = q\,\mathbf{E}$. We used this to *define* $\mathbf{E}$ in fact. When the charge is moving the force law is more complicated. From experiments one finds that if q at position $\mathbf{r}$ is moving with velocity $\mathbf{u} = \mathrm{d}\mathbf{r}/\mathrm{d}t$ it experiences a force

$$\boxed{\mathbf{F} = q\,\mathbf{E}(\mathbf{r}) + q\,\mathbf{u} \times \mathbf{B}(\mathbf{r}) \ .} \tag{3.8}$$

Here $\mathbf{B} = \mathbf{B}(\mathbf{r})$ is a vector field, called the *magnetic field*, and we may similarly regard the *Lorentz force* $\mathbf{F}$ in (3.8) as defining $\mathbf{B}$. The magnetic field is measured in SI units in *Teslas*, which is the same as $\mathrm{N\,s\,m^{-1}\,C^{-1}}$.

Since (3.8) may look peculiar at first sight, it is worthwhile discussing it a little further. The magnetic component may be written as

$$\mathbf{F}_{\text{mag}} = q\,\mathbf{u} \times \mathbf{B} \ . \tag{3.9}$$

In experiments, the magnetic force on q is found to be proportional to q, proportional to the magnitude $|\mathbf{u}|$ of $\mathbf{u}$, and is perpendicular to $\mathbf{u}$. Note this latter point means that the magnetic force *does no work* on the charge. One also finds that the magnetic force at each point is perpendicular to a particular fixed direction at that point, and is also proportional to the sine of the angle between $\mathbf{u}$ and this fixed direction. The vector field that describes this direction is called the magnetic field $\mathbf{B}$, and the above, rather complicated, experimental observations are summarized by the simple formula (3.9).

In practice (3.8) was deduced not from moving test charges, but rather from *currents* in *test wires*. A current of course consists of moving charges, and (3.8) was deduced from the forces on these test wires.

3.4 Biot-Savart law

If electric charges produce the electric field, what produces the magnetic field? The answer is that *electric currents* produce magnetic fields! Note carefully the distinction here: currents produce magnetic fields, but, by the Lorentz force law just discussed, magnetic fields exert a force on moving charges, and hence currents.

The usual discussion of this involves currents in wires, since this is what Ampère actually did in 1820. One has a wire with steady current I flowing through it. Here the latter is defined in terms of the current density via (3.2), where Σ is any cross-section of the wire. This is independent of the choice of cross-section, and thus makes sense, because of the steady current condition (3.7).[11] One finds that another wire, the test wire, with current I' experiences a force. This force is conveniently summarized by introducing the concept of a magnetic field: the first wire produces a magnetic field, which generates a force on the second wire via the Lorentz force (3.8) acting on the charges that make up the current I'.

Rather than describe this in detail, we shall instead simply note that if currents produce magnetic fields, then fundamentally it is *charges in motion* that produce magnetic fields. One may summarize this by an analogous formula to the Coulomb formula (1.7) for the electric field due to a point charge:

Biot-Savart law A charge q at position $\mathbf{r}_0$ moving with velocity $\mathbf{v}$ produces a magnetic field

$$\mathbf{B}(\mathbf{r}) = \frac{\mu_0\, q}{4\pi}\,\frac{\mathbf{v} \times (\mathbf{r} - \mathbf{r}_0)}{|\mathbf{r} - \mathbf{r}_0|^3} \ . \tag{3.10}$$

[11] To see this, use the divergence theorem for a cylindrical region bounded by any two cross-sections and the surface of the wire.

The constant μ_0 is called *permeability of free space*. It is approximately $\mu_0 \approx 1.257 \cdot 10^{-6}\,\mathrm{N\,A^{-2}}$.[12] Compare (3.10) to (1.7).

As in electrostatics we also have: *the magnetic field obeys the Principle of Superposition.*

Using the above laws of magnetostatics we may compute the magnetic field due to the steady current I in a wire C by summing contributions of the form (3.10):

$$\mathbf{B}(\mathbf{r}) \;=\; \frac{\mu_0 I}{4\pi}\int_C \frac{\mathrm{d}\mathbf{r}' \times (\mathbf{r}-\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|^3}\;. \tag{3.11}$$

To derive this, imagine dividing the wire into segments, with the segment $\delta\mathbf{r}'$ at position $\mathbf{r}'$. Suppose this segment contains a charge $q(\mathbf{r}')$ with velocity vector $\mathbf{v}(\mathbf{r}')$ – notice here that $\mathbf{v}(\mathbf{r}')$ points in the same direction as $\delta\mathbf{r}'$. From (3.10) this segment contributes

$$\delta\mathbf{B}(r) \;=\; \frac{\mu_0\, q}{4\pi}\,\frac{\mathbf{v}\times(\mathbf{r}-\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|^3} \tag{3.12}$$

to the magnetic field. Now by definition of the current I we have $I\,\delta\mathbf{r}' = \mathbf{J}\cdot\delta\mathbf{S}\,\delta\mathbf{r}' = \rho\,\mathbf{v}\cdot\delta\mathbf{S}\,\delta\mathbf{r}'$, where we may take $\delta\Sigma$ of area δS to be a perpendicular cross-section of the wire. But the total charge q in the cylinder of cross-sectional area δS and length $|\delta\mathbf{r}'|$ is $q = \rho\,\delta S\,|\delta\mathbf{r}'|$. We thus deduce that $I\,\delta\mathbf{r}' = q\,\mathbf{v}$ and hence that (3.11) holds.

Remark The formula (3.11) for the magnetic field produced by the steady current in a wire is also often called the *Biot-Savart law.*

Example (Magnetic field produced by a current in a long straight wire) As an example of (3.11), let us compute the magnetic field due to a steady current I in an infinitely long straight wire. Place the wire along the z-axis, and let P be a point in the (x,y)-plane at distance s from the origin O, as in Figure 14. This is the point at which we compute the magnetic field $\mathbf{B}(\mathbf{r})$, so that $\overrightarrow{OP} = \mathbf{r}$. Let Q be a point on the wire at distance z' from the origin, and write $\overrightarrow{OQ} = \mathbf{r}'$. Then in the integral Biot-Savart formula (3.11) we have $\overrightarrow{OQ} = \mathbf{r}' = \mathbf{e}_3\, z'$, where $\mathbf{e}_3$ is a unit vector in the z-direction, *i.e.* along the wire, and so $\mathrm{d}\mathbf{r}' = \mathbf{e}_3\,\mathrm{d}z'$. Notice that $\mathbf{r}-\mathbf{r}' = \overrightarrow{QP}$.

Next notice that the vector $\mathrm{d}\mathbf{r}'\times(\mathbf{r}-\mathbf{r}')$ points in a direction that is independent of $\overrightarrow{OQ} = \mathbf{r}'$: it is always tangent to the circle of radius s in the (x,y)-plane. To evaluate the integral in (3.11) we thus simply have to compute the magnitude $|\mathrm{d}\mathbf{r}'\times(\mathbf{r}-\mathbf{r}')| = \mathrm{d}z'\,|\overrightarrow{QP}|\sin\theta = s\,\mathrm{d}z'$, where θ is the angle between $\overrightarrow{OQ}$ and $\overrightarrow{QP}$, as shown. Since also $|\mathbf{r}-\mathbf{r}'| = |\overrightarrow{QP}| = \sqrt{s^2+z'^2}$, from (3.11) we compute the magnitude $B(s) \equiv |\mathbf{B}|$ of the magnetic field to be

$$\begin{aligned} B(s) \;&=\; \frac{\mu_0 I}{4\pi}\int_{-\infty}^{\infty}\frac{s}{(s^2+z'^2)^{3/2}}\,\mathrm{d}z' \\ &=\; \frac{\mu_0 I}{2\pi s}\;, \end{aligned} \tag{3.13}$$

[12] The recommended 2018 CODATA value is $\mu_0 = (1.256\,637\,062\,12 \pm 0.000\,000\,000\,19)\cdot 10^{-6}\,\mathrm{N\,A^{-2}}$, so it is measured to about 10 digits. Until 2018, the value was *exactly* $\mu_0 = 4\pi\cdot 10^{-7}\,\mathrm{N\,A^{-2}} = 1.256\,637\,061\,43\ldots\cdot 10^{-6}\,\mathrm{N\,A^{-2}}$, by the old definition of the unit Ampère in terms of the force between two parallel wires.

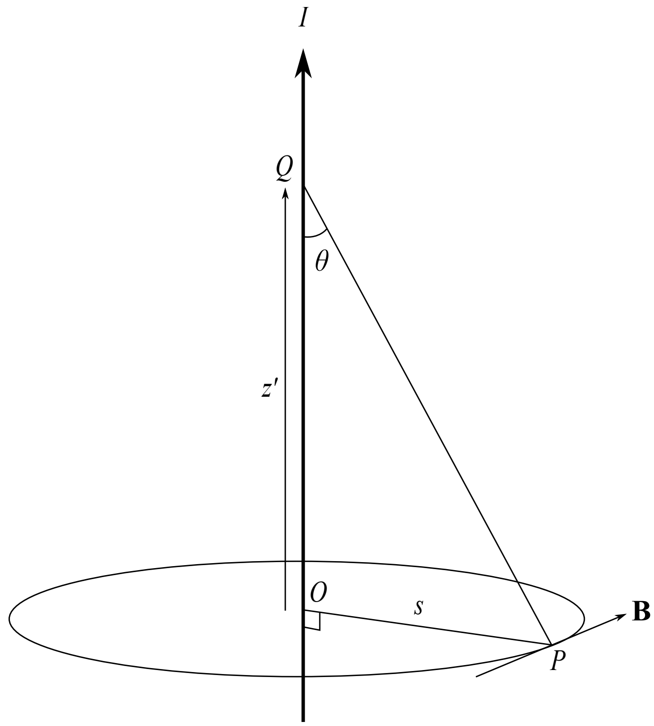


Figure 14: Computing the magnetic field around a straight wire carrying a steady current I.

where the integral may be evaluated by setting $z' = s \tan \xi$. The magnetic field is thus tangent to circles in the plane perpendicular to the wire, with a magnitude which decreases inversely with the perpendicular distance to the wire in this plane. ■

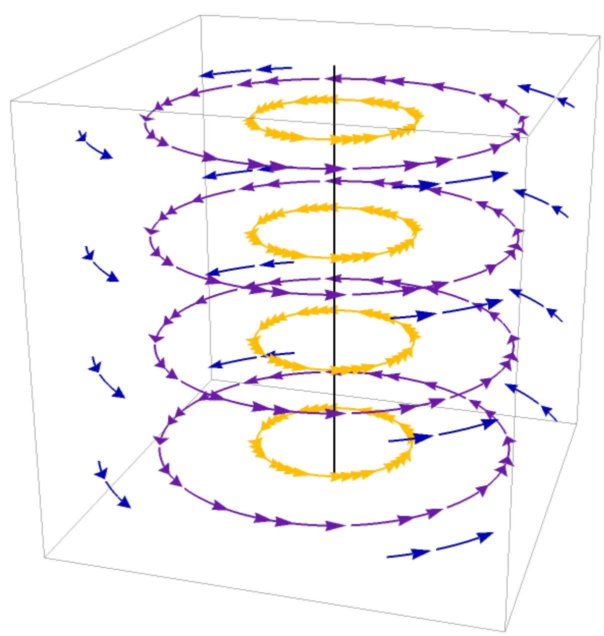


Figure 15: Magnetic field lines around the infinite straight wire. You might have seen a demonstration of this with iron filings, where the latter line up with the direction of the magnetic field.

3.5 Magnetic monopoles?

Let us return to the point charge q at position vector $\mathbf{r}_0$ with velocity $\mathbf{v}$, generating the magnetic field (3.10). We might as well put the charge at the origin, so $\mathbf{r}_0 = \mathbf{0}$. Then note that since $\mathbf{r}/r^3 = -\nabla(1/r)$ we have

$$\nabla \cdot \left(\frac{\mathbf{v} \times \mathbf{r}}{r^3}\right) = \mathbf{v} \cdot [\nabla \times \nabla (1/r)] = 0 \, . \tag{3.14}$$

Here we have used the identity (A.12) in the first equality, and the fact that the curl of a gradient is zero in the second equality. Thus we have shown that $\nabla \cdot \mathbf{B} = 0$, except at the origin $\mathbf{r} = \mathbf{0}$. However, unlike the case of the electric field and Gauss' law, the integral of the magnetic field around the point charge is zero. To see this, let Σ be a sphere of radius a centred on the charge, so that the outward unit normal is $\mathbf{n} = \mathbf{r}/r$, and compute

$$\int_\Sigma \mathbf{B} \cdot \mathrm{d}\mathbf{S} = \frac{\mu_0 \, q}{4\pi} \int_\Sigma \left(\frac{\mathbf{v} \times \mathbf{r}}{r^3}\right) \cdot \left(\frac{\mathbf{r}}{r}\right) \mathrm{d}S = 0 \, , \tag{3.15}$$

since $\mathbf{v} \times \mathbf{r}$ is perpendicular to $\mathbf{r}$. By the divergence theorem, it follows that

$$\int_R \nabla \cdot \mathbf{B} \, \mathrm{d}V = 0 \, , \tag{3.16}$$

for *any* region R, and thus

$$\boxed{\nabla \cdot \mathbf{B} = 0 \, .} \tag{3.17}$$

This is another of Maxwell's equations on the front cover. It says that there are *no magnetic monopoles* (*i.e.* magnetic point charges) that generate magnetic fields, analogous to the way that electric charges generate electric fields. Instead magnetic fields are produced by electric currents. Although we have only deduced (3.17) above for the magnetic field of a moving point charge, the general case follows from the Principle of Superposition.

You might wonder what produces the magnetic field in *permanent* magnets, such as bar magnets. Where is the electric current? We discuss this in section 4.3.

* Mathematically, it is certainly possible to allow for *magnetic monopoles* and *magnetic currents* in Maxwell's equations. In fact the equations then become completely symmetric under the interchange of $\mathbf{E}$ with $-c\,\mathbf{B}$, $c\,\mathbf{B}$ with $\mathbf{E}$, and corresponding interchanges of electric with magnetic charge densities and currents. Here $c^2 \equiv 1/\epsilon_0\mu_0$. There are also theoretical reasons for introducing magnetic monopoles. For example, the quantization of electric charge – that all electric charges are an *integer* multiple of some fixed fundamental quantity of charge – may be understood in quantum mechanics using magnetic monopoles. This is a beautiful argument due to Dirac. However, no magnetic monopoles have ever been observed in nature. If they existed, (3.17) would need correcting.

3.6 Ampère's law

There is just one more static Maxwell equation to discuss, namely the equation involving the curl of $\mathbf{B}$. This is *Ampère's law*. In many treatments of magnetostatics, this is often described as an additional experimental result, and/or is derived for special symmetric configurations, having first solved for $\mathbf{B}$ using the Biot-Savart law (3.10) or (3.11). However, it is possible to *derive* Ampère's law directly from (3.10), as we now show.

We first use (3.10) and the Principle of Superposition to write $\mathbf{B}$ generated by a current density $\mathbf{J}$ as a volume integral

$$\mathbf{B}(\mathbf{r}) \;=\; \frac{\mu_0}{4\pi}\int_{\mathbf{r}'\in R}\frac{\mathbf{J}(\mathbf{r}')\times(\mathbf{r}-\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|^3}\,\mathrm{d}V'\ . \tag{3.18}$$

This follows directly from the definition $\mathbf{J}=\rho\,\mathbf{v}$ in (3.1) and taking the limit of a sum of terms of the form (3.10). $R\subset\mathbb{R}^3$ is by definition a region containing the set of points with $\mathbf{J}\neq\mathbf{0}$. We next define the vector field

$$\mathbf{A}(\mathbf{r}) \;\equiv\; \frac{\mu_0}{4\pi}\int_R\frac{\mathbf{J}(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}\,\mathrm{d}V'\ . \tag{3.19}$$

One then computes (*cf.* the proof of Theorem 1.4)

$$\frac{\partial A_i}{\partial x_j}(\mathbf{r}) \;=\; -\frac{\mu_0}{4\pi}\int_R\frac{J_i(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|^3}\,(x_j\,-x_j')\,\mathrm{d}V'\ , \tag{3.20}$$

where $\mathbf{r}=(x_1,x_2,x_3)$, $\mathbf{r}'=(x_1',x_2',x_3')$. From the first line of (A.9) we have

$$\nabla\times\mathbf{A}\;\equiv\;\sum_{j=1}^{3}\mathbf{e}_j\times\frac{\partial\mathbf{A}}{\partial x_j}\;=-\sum_{i,j=1}^{3}\left(\mathbf{e}_i\times\mathbf{e}_j\right)\frac{\partial A_i}{\partial x_j}\ , \tag{3.21}$$

so that comparing (3.20) with (3.18) we see that

$$\mathbf{B}(\mathbf{r})\;=\;\nabla\times\mathbf{A}(\mathbf{r})\ . \tag{3.22}$$

Note that (3.22) immediately implies $\nabla\cdot\mathbf{B}=0$, as the divergence of a curl is zero.

Lemma 3.1 *For steady currents* $\mathbf{J}$*, supported inside a bounded region* R*, the vector field* $\mathbf{A}$ *defined by* (3.19) *satisfies* $\nabla\cdot\mathbf{A}=0$.

Proof Using

$$\nabla\left(\frac{1}{|\mathbf{r}-\mathbf{r}'|}\right)\;=\;-\nabla'\left(\frac{1}{|\mathbf{r}-\mathbf{r}'|}\right)\ ,\qquad \mathbf{r}\neq\mathbf{r}'\ , \tag{3.23}$$

where ∇' denotes derivative with respect to $\mathbf{r}'$, from (3.19) we compute

$$\begin{aligned}\nabla\cdot\mathbf{A}\;&=\;\frac{\mu_0}{4\pi}\int_R\mathbf{J}(\mathbf{r}')\cdot\nabla\left(\frac{1}{|\mathbf{r}-\mathbf{r}'|}\right)\mathrm{d}V'\;=\;-\frac{\mu_0}{4\pi}\int_R\mathbf{J}(\mathbf{r}')\cdot\nabla'\left(\frac{1}{|\mathbf{r}-\mathbf{r}'|}\right)\mathrm{d}V'\\ &=\;-\frac{\mu_0}{4\pi}\int_R\nabla'\cdot\left(\frac{\mathbf{J}(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}\right)\mathrm{d}V'+\frac{\mu_0}{4\pi}\int_R\frac{1}{|\mathbf{r}-\mathbf{r}'|}\left(\nabla'\cdot\mathbf{J}(\mathbf{r}')\right)\mathrm{d}V'\ .\end{aligned} \tag{3.24}$$

The second term on the right hand side of this equation is zero for *steady currents*, satisfying (3.7). Moreover, we may use the divergence theorem on the first term to obtain a surface integral on ∂R. But by assumption $\mathbf{J}$ vanishes on this boundary, so this term is also zero. ■

We next take the curl of (3.22). Using the identity

$$\nabla \times (\nabla \times \mathbf{A}) = \nabla (\nabla \cdot \mathbf{A}) - \nabla^2 \mathbf{A} \ , \tag{3.25}$$

which holds for any vector field $\mathbf{A}$, together with Lemma 3.1, we deduce that

$$\nabla \times \mathbf{B} = -\nabla^2 \mathbf{A} \ . \tag{3.26}$$

On the other hand, from (3.19)

$$\nabla^2 \mathbf{A} = \frac{\mu_0}{4\pi} \int_R \mathbf{J}(\mathbf{r}') \, \nabla^2 \left(\frac{1}{|\mathbf{r} - \mathbf{r}'|} \right) \mathrm{d}V' = \frac{\mu_0}{4\pi} \int_R \mathbf{J}(\mathbf{r}') \, \big(-4\pi \, \delta(\mathbf{r} - \mathbf{r}') \big) \, \mathrm{d}V' = -\mu_0 \, \mathbf{J}(\mathbf{r}) \ , \tag{3.27}$$

where we have used Proposition 1.3 in the middle step. We hence deduce

$$\boxed{\nabla \times \mathbf{B} = \mu_0 \, \mathbf{J} \ .} \tag{3.28}$$

This is *Ampère's law* for magnetostatics. It is the final Maxwell equation on the front page, albeit in the special case where the electric field is independent of time, so $\partial \mathbf{E} / \partial t = \mathbf{0}$. Notice this equation is consistent with the steady current assumption (3.7).

We may equivalently rewrite (3.28) using Stokes' theorem as

Ampère's law *For any simple closed curve $C = \partial \Sigma$ bounding a surface Σ*

$$\int_{C = \partial \Sigma} \mathbf{B} \cdot \mathrm{d}\mathbf{r} = \mu_0 \int_\Sigma \mathbf{J} \cdot \mathrm{d}\mathbf{S} = \mu_0 I \ , \tag{3.29}$$

where I is the current through Σ.

Example Notice that integrating the magnetic field $\mathbf{B}$ given by (3.13) around a circle C in the (x, y)-plane of radius s, centred on the z-axis, indeed gives $\mu_0 I$. ■

3.7 Magnetostatic vector potential

Definition In magnetostatics the *magnetic vector potential* is a vector field $\mathbf{A}$ such that the magnetic field $\mathbf{B}$ is given by

$$\mathbf{B} = \nabla \times \mathbf{A} \ . \tag{3.30}$$

In fact we have already introduced such a vector potential in equations (3.19) and (3.22). It is analogous to the electrostatic potential ϕ in electrostatics.

* Notice that (3.30) is a *sufficient* condition for the Maxwell equation (3.17) to hold, since the divergence of a curl is zero. It is also *necessary* if we work in a domain with simple enough topology, such as $\mathbb{R}^3$ or an open ball. An example of a domain where not every vector field $\mathbf{B}$ with zero divergence may be written as a curl is $\mathbb{R}^3 \setminus \{\text{point}\}$. Compare this to the corresponding starred paragraph in section 1.5. Again, a proof for an open ball is contained in appendix B of the book by Woodhouse.

The vector field $\mathbf{A}$ in (3.30) is far from unique: since the curl of a gradient is zero, we may add $\nabla\psi$ to $\mathbf{A}$, for any function ψ, without changing $\mathbf{B}$:

$$\mathbf{A} \to \widehat{\mathbf{A}} \equiv \mathbf{A} + \nabla\psi \ . \tag{3.31}$$

That is, $\mathbf{B} = \nabla \times \mathbf{A} = \nabla \times \widehat{\mathbf{A}}$.

Definition The transformation (3.31) is called a *gauge transformation* of $\mathbf{A}$.

We may fix this *gauge freedom* by imposing additional conditions on $\mathbf{A}$. For example, suppose we have chosen a particular $\mathbf{A}$ satisfying (3.30). Then if ψ is a solution of the Poisson equation

$$\nabla^2\psi = -\nabla \cdot \mathbf{A} \ , \tag{3.32}$$

it follows that $\widehat{\mathbf{A}}$ in (3.31) satisfies

$$\nabla \cdot \widehat{\mathbf{A}} = 0 \ . \tag{3.33}$$

Definition The condition (3.33) on the magnetostatic vector potential is called the *Lorenz gauge.*[13]

From Theorem 2.2 we know that solutions ψ to the Poisson equation (3.32) are unique, for fixed boundary conditions.

Many equations simplify with this gauge choice for $\mathbf{A}$. For example, Ampère's law (3.28) is

$$\mu_0 \mathbf{J} = \nabla \times (\nabla \times \mathbf{A}) = \nabla(\nabla \cdot \mathbf{A}) - \nabla^2\mathbf{A} \ , \tag{3.34}$$

so that in Lorenz gauge $\nabla \cdot \mathbf{A} = 0$ this becomes

$$\nabla^2\mathbf{A} = -\mu_0 \mathbf{J} \ . \tag{3.35}$$

Compare with Poisson's equation (1.32) in electrostatics. Notice that Lemma 3.1 shows that the vector potential $\mathbf{A}$ given by (3.19) is in Lorenz gauge.

* Gauge invariance and the vector potential $\mathbf{A}$ play a fundamental role in more advanced formulations of electromagnetism. The magnetic potential also plays an essential *physical* role in the *quantum* theory of electromagnetism.

3.8 Multipole expansion

In electrostatics and magnetostatics we have now derived the similar formulae

$$\phi(\mathbf{r}) = \frac{1}{4\pi\epsilon_0}\int_{\mathbf{r}'\in R} \frac{\rho(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}\,\mathrm{d}V' \ , \qquad \mathbf{A}(\mathbf{r}) = \frac{\mu_0}{4\pi}\int_{\mathbf{r}'\in R} \frac{\mathbf{J}(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}\,\mathrm{d}V' \ . \tag{3.36}$$

In both cases $R \subset \mathbb{R}^3$ is a bounded region, and the first formula gives the electrostatic potential ϕ generated by a charge density ρ supported inside R, while the second formula gives the magnetostatic vector potential $\mathbf{A}$ generated by a steady current density $\mathbf{J}$ supported inside R. The

[13] Mr Lorenz and Mr Lorentz were two different people.

electric and magnetic fields are obtained from these via $\mathbf{E} = -\nabla\phi$, $\mathbf{B} = \nabla \times \mathbf{A}$, respectively. In this subsection we want to examine what these fields look like far away from the localized source region R. Here "far away" means that the observation point P, with position vector $\mathbf{r}$ measured from the origin O, is at a distance that is large compared to the size of R – see Figure 16.[14]

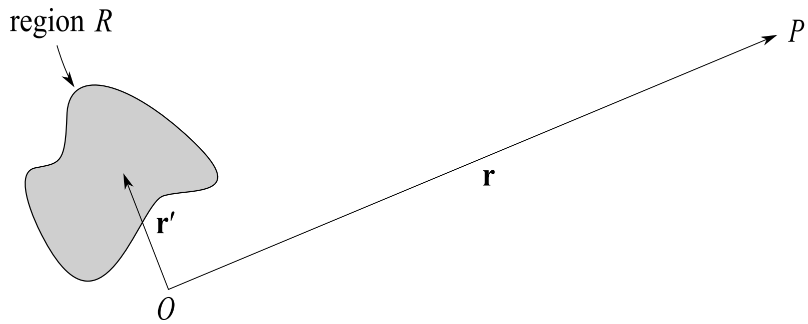


Figure 16: Observing the fields, generated by sources supported in a region R, at a point P far away from R.

Our starting point is the Taylor expansion

$$\frac{1}{|\mathbf{r}-\mathbf{r}'|} = \frac{1}{r\sqrt{1-\frac{2}{r^2}\,\mathbf{r}\cdot\mathbf{r}' + \frac{1}{r^2}\,\mathbf{r}'\cdot\mathbf{r}'}} = \frac{1}{r} + \frac{1}{r^3}\,\mathbf{r}\cdot\mathbf{r}' + O(1/r^3)\ . \qquad (3.37)$$

Here as usual $r \equiv |\mathbf{r}| = \sqrt{\mathbf{r}\cdot\mathbf{r}}$, and we have expanded for $r \gg r'$. In particular the Cauchy-Schwarz inequality gives $|\mathbf{r}\cdot\mathbf{r}'| \le |\mathbf{r}|\,|\mathbf{r}'| = r\,r'$, so that the second term on the right hand side of (3.37) is $O(1/r^2)$. Applying (3.37) to the electrostatic potential in (3.36) immediately gives

$$\phi(\mathbf{r}) = \frac{Q}{4\pi\epsilon_0}\frac{1}{r} + \frac{1}{4\pi\epsilon_0}\frac{\mathbf{p}\cdot\mathbf{r}}{r^3} + O(1/r^3)\ , \qquad (3.38)$$

where we have defined

$$Q \equiv \int_{\mathbf{r}'\in R} \rho(\mathbf{r}')\,\mathrm{d}V'\ , \qquad \mathbf{p} \equiv \int_{\mathbf{r}'\in R} \mathbf{r}'\rho(\mathbf{r}')\,\mathrm{d}V'\ . \qquad (3.39)$$

Here Q is simply the total charge, and (3.38) says that far away from the localized charge distribution, the electrostatic potential to leading order looks like that generated by a *point charge* Q (at the origin $r = 0$). This is, of course, entirely sensible.

Definition The vector $\mathbf{p}$ defined in (3.39) is called the *electric dipole moment.*

This governs the next-to-leading order term in the general expansion (3.38), and is the *leading order term when the total charge* $Q = 0$. On the first problem sheet you will have computed the electrostatic potential of an *electric dipole*. This is a configuration of two point charges $\pm q$, with

[14] The methods developed in this subsection apply also to other areas of theoretical physics, perhaps most notably to gravitational waves in General Relativity, that might have been generated by a distant collision and subsequent merger of two black holes.

separation vector $\mathbf{d}$. Defining $\mathbf{p} \equiv q\,\mathbf{d}$, then in the limit that $q \to \infty$, $d \equiv |\mathbf{d}| \to 0$, with $q\,d = |\mathbf{p}|$ held fixed and the charges at the origin, you showed that the resulting electrostatic potential is

$$\phi(\mathbf{r})_{\text{dipole}} = \frac{1}{4\pi\epsilon_0}\frac{\mathbf{p}\cdot\mathbf{r}}{r^3} . \tag{3.40}$$

This is precisely the second term on the right hand side of (3.38). It arises when there is asymmetry in the charge distribution around the origin.

Remark One might be tempted to define a *centre of charge* for a general charge distribution as $\mathbf{d} \equiv \mathbf{p}/Q$, with Q and $\mathbf{p}$ defined in (3.39), in precise analogy with centre of mass for a mass distribution. However, charge distributions are often *neutral*, with $Q = 0$, and the electric dipole moment then governs the leading order behaviour of the resulting electric field. For example, a neutral molecule, such as a water molecule, will involve shared electrons between atoms. These are not distributed uniformly around the molecule, which thus behaves like a tiny electric dipole. The behaviour of a large number of small electric dipoles is discussed in section 4.

Using (3.37) we may perform a similar expansion of the magnetostatic vector potential in (3.36):

$$\mathbf{A}(\mathbf{r}) = \frac{\mu_0}{4\pi}\left[\frac{1}{r}\int_R \mathbf{J}(\mathbf{r}')\,\mathrm{d}V' + \frac{1}{r^3}\int_R (\mathbf{r}\cdot\mathbf{r}')\,\mathbf{J}(\mathbf{r}')\,\mathrm{d}V' + \cdots\right] . \tag{3.41}$$

The following result is useful for manipulating this expression:

Lemma 3.2 *For any vector field $\mathbf{J}$ satisfying $\nabla\cdot\mathbf{J} = 0$ in a region R, with $\mathbf{J}$ being zero on the closed boundary $\Sigma = \partial R$ of R, and any functions f, g, we have*

$$\int_R (f\,\mathbf{J}\cdot\nabla g + g\,\mathbf{J}\cdot\nabla f)\,\mathrm{d}V = 0 . \tag{3.42}$$

Proof Using the product rule $\nabla\cdot(fg\,\mathbf{J}) = fg\,\nabla\cdot\mathbf{J} + \mathbf{J}\cdot\nabla(fg)$ we may write the first term in (3.42) as

$$\begin{aligned}\int_R f\,\mathbf{J}\cdot\nabla g\,\mathrm{d}V &= \int_R \nabla\cdot(fg\,\mathbf{J})\,\mathrm{d}V - \int_R f\,g\,\nabla\cdot\mathbf{J}\,\mathrm{d}V - \int_R g\,\mathbf{J}\cdot\nabla f\,\mathrm{d}V \\ &= \int_\Sigma fg\,\mathbf{J}\cdot\mathrm{d}\mathbf{S} - 0 - \int_R g\,\mathbf{J}\cdot\nabla f\,\mathrm{d}V = -\int_R g\,\mathbf{J}\cdot\nabla f\,\mathrm{d}V .\end{aligned} \tag{3.43}$$

Here in the second equality we have used the divergence theorem for the first term, and $\nabla\cdot\mathbf{J} = 0$ for the second term, and the last equality follows since $\mathbf{J}$ is assumed to be zero on $\Sigma = \partial R$. ■

Corollary 3.3 *For a steady current density $\mathbf{J}$ supported inside a region R*

$$\int_R \mathbf{J}(\mathbf{r})\,\mathrm{d}V = \mathbf{0} . \tag{3.44}$$

Proof Applying Lemma 3.2 with $f \equiv 1$ and $g = x_i$, with $\mathbf{r} = (x_1, x_2, x_3)$, gives $\int_R \mathbf{J}\cdot\nabla x_i\,\mathrm{d}V = \int_R J_i\,\mathrm{d}V = 0$, where in the first equality note $(\nabla x_i)_j = \delta_{ij}$. ■

It follows that the first term on the right hand side of the expansion (3.41) is *always zero*! This is ultimately a reflection of the fact there are no magnetic point charges, analogous to electric point charges in electrostatics. For the second term in (3.41) we may again use Lemma 3.2, this time with $f = x'_i$, $g = x'_j$, and we take the integration to be over the primed variable $\mathbf{r}'$. From (3.42) we then deduce

$$0 \;=\; \int_R \left[x'_i\, \mathbf{J}(\mathbf{r}')\cdot\nabla' x'_j + x'_j\, \mathbf{J}(\mathbf{r}')\cdot\nabla' x'_i\right] \mathrm{d}V' \;=\; \int_R \left[x'_i\, J_j(\mathbf{r}') + x'_j\, J_i(\mathbf{r}')\right] \mathrm{d}V' \;, \qquad (3.45)$$

and hence, for any vector $\mathbf{f}$, we have

$$\begin{aligned}\int_R (\mathbf{f}\cdot\mathbf{r}')\, \mathbf{J}(\mathbf{r}')\, \mathrm{d}V' \;&=\; \sum_{i,j=1}^{3} f_i \left(\int_R x'_i\, J_j(\mathbf{r}')\, \mathrm{d}V'\right) \mathbf{e}_j = -\sum_{i,j=1}^{3} f_i \left(\int_R x'_j\, J_i(\mathbf{r}')\, \mathrm{d}V'\right) \mathbf{e}_j \\ &=\; \int_R (\mathbf{f}\cdot\mathbf{J}(r'))\, \mathbf{r}'\, \mathrm{d}V' \;. \qquad (3.46)\end{aligned}$$

Combined with the vector triple product identity (A.6), we can thus write

$$\int_R (\mathbf{f}\cdot\mathbf{r}')\, \mathbf{J}(\mathbf{r}')\, \mathrm{d}V' \;=\; -\frac{1}{2}\, \mathbf{f} \times \left[\int_R \mathbf{r}' \times \mathbf{J}(\mathbf{r}')\, \mathrm{d}V'\right] \;. \qquad (3.47)$$

Definition The *magnetic dipole moment* generated by a steady current density $\mathbf{J}$ in R is

$$\mathbf{m} \;\equiv\; \frac{1}{2}\int_{\mathbf{r}'\in R} \mathbf{r}' \times \mathbf{J}(\mathbf{r}')\, \mathrm{d}V' \;. \qquad (3.48)$$

This is analogous to the definition of electric dipole moment we gave earlier, and combining (3.47) for $\mathbf{f} = \mathbf{r}$ with (3.41) we have proven

Proposition 3.4 *The magnetostatic vector potential generated by a steady current density* $\mathbf{J}$ *in a localized region* R *has an expansion with leading order term*

$$\mathbf{A}(\mathbf{r}) \;=\; \frac{\mu_0}{4\pi}\,\frac{\mathbf{m}\times\mathbf{r}}{r^3} + O(1/r^3) \;, \qquad (3.49)$$

as $r \to \infty$*, where* $\mathbf{m}$ *is the magnetic dipole moment* (3.48) *generated by the current density.*

The correction terms $O(1/r^3)$ in (3.49) are called *higher moments* of $\mathbf{J}$, starting with the *quadrupole moment*, and can be computed with more effort.

Definition Analogously to (3.40), we may define the *magnetic dipole vector potential* to be

$$\mathbf{A}(\mathbf{r})_{\text{dipole}} \;\equiv\; \frac{\mu_0}{4\pi}\,\frac{\mathbf{m}\times\mathbf{r}}{r^3} \;. \qquad (3.50)$$

This gives the leading order term in the expansion (3.49). The corresponding magnetic field is

$$\begin{aligned}\mathbf{B}(\mathbf{r})_{\text{dipole}} \;&=\; \nabla\times\mathbf{A}(\mathbf{r})_{\text{dipole}} \;=\; \frac{\mu_0}{4\pi}\,\nabla\times\left(\mathbf{m}\times\frac{\mathbf{r}}{r^3}\right) \;=\; \frac{\mu_0}{4\pi}\left[\mathbf{m}\,\nabla\cdot\left(\frac{\mathbf{r}}{r^3}\right) - \mathbf{m}\cdot\nabla\left(\frac{\mathbf{r}}{r^3}\right)\right] \\ &=\; \frac{\mu_0}{4\pi}\left[-\frac{\mathbf{m}}{r^3} + \frac{3(\mathbf{m}\cdot\mathbf{r})\,\mathbf{r}}{r^5}\right] \;, \qquad (3.51)\end{aligned}$$

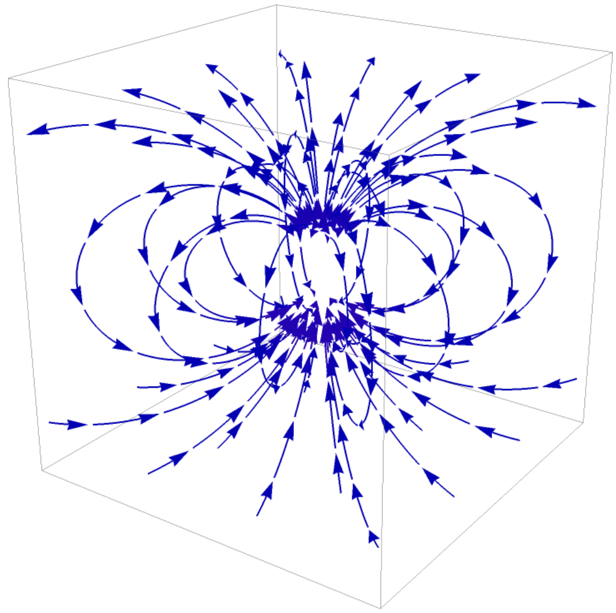


Figure 17: Magnetic field lines for the magnetic dipole (3.51), where we have aligned the magnetic dipole moment $\mathbf{m}$ pointing upwards along the z-axis direction. The magnetic field lines of a bar magnet are similar, where the bar magnet is aligned vertically along the z-axis, with north pole at the top and south pole at the bottom.

where $\mathbf{r} \neq \mathbf{0}$, and on the first line we have used the identity (A.11). Notice here that $\nabla \cdot (\mathbf{r}/r^3) = 0$ for $\mathbf{r} \neq \mathbf{0}$, as in our computation of (1.8). The field lines for $\mathbf{B}(\mathbf{r})_{\text{dipole}}$ are shown in Figure 17.

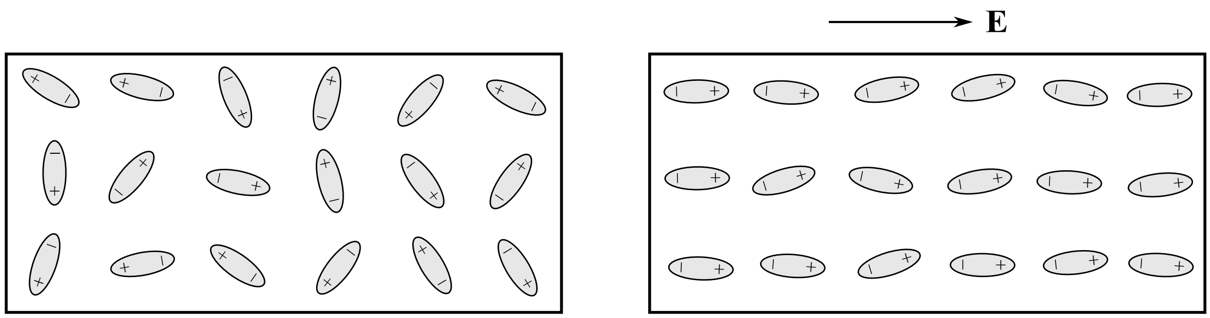
To summarize, we have shown that, far away from *any* localized steady current source, the magnetic field generated will take the form (3.51), to leading order, and thus resemble Figure 17. Notice that in this figure $\mathbf{m}$ points upwards along the z-axis, and the corresponding magnetic field lines come out of the top of the magnetic dipole, which is called the *north pole*, and go into the bottom of the magnetic dipole, which is called the *south pole*.

4 Macroscopic media

4.1 Dielectrics

In order to write down the electrostatic or magnetostatic Maxwell equations, we need to know the charge density ρ or current density $\mathbf{J}$ precisely, in principle everywhere in space. Except for certain idealized situations, such as for point charges in vacuum, this is usually not possible in practice.

For instance, suppose we wish to study the effects of electromagnetism in *water*. In a cubic centimetre of water there are around 10^{22} water molecules. As described in section 3.8, a water molecule is neutral, having total charge zero, but this then behaves like a tiny *electric dipole*: the negative electron and positive proton charges are not uniformy distributed in the molecule, and this generates an electric dipole moment $\mathbf{p}$, with dipole field (3.40). One can view such an electric



(a) Unpolarized. (b) Medium polarized by external $\mathbf{E}$.

Figure 18: The dipoles in a dielectric medium.

dipole as in Figure 18, with a positive end $+$ and a negative end $-$, with $\mathbf{p}$ pointing from the negative end towards the positive end.

A *dielectric medium*, such as water, by definition contains a very large number of these dipoles. With no external electric field applied, the dipoles are aligned somewhat randomly, as in Figure 18a, due to their random thermal motion. Such a configuration is said to be *unpolarized*, with no net macroscopic electric field produced by the dipoles. However, now consider the effect of turning on an *external* electric field $\mathbf{E}$: the electric dipoles will *align* with $\mathbf{E}$, as in Figure 18b. In doing so the large number of aligned dipoles will superpose to generate their *own* macroscopic electric field. In Figure 18b notice that the negative ends of the dipoles are on the left, while the positive ends are on the right, creating an effective surface charge density on the left and right ends of the box. An electric field generated by a positive/negative charge points away from/towards it, respectively, so the electric field generated by the dipoles is in the *opposite* direction to $\mathbf{E}$, *reducing* the overall electric field.

4.2 Electric dipoles

We may describe this more quantitatively by examining electric dipoles in a little more detail. Recall that an electric dipole is realized by starting with two point charges $\pm q$, with separation vector $\mathbf{d}$, and taking the limit in which the distance $d \equiv |\mathbf{d}|$ between the charges tends to zero, the positive charge q is taken to infinity, with the *electric dipole moment* $\mathbf{p} \equiv q\,\mathbf{d}$ held fixed.

If we place such an electric dipole in an external electric field $\mathbf{E}$, what force does it experience? The *total force* on the pair of point charges is (see Figure 19)

$$\begin{aligned}
\mathbf{F} &= q\,\mathbf{E}\left(\mathbf{r}+\tfrac{\mathbf{d}}{2}\right) - q\,\mathbf{E}\left(\mathbf{r}-\tfrac{\mathbf{d}}{2}\right) \\
&= q\,\left[\mathbf{E}(\mathbf{r}) + \left(\tfrac{\mathbf{d}}{2}\cdot\nabla\right)\mathbf{E}(\mathbf{r}) + O(d^2)\right] - q\,\left[\mathbf{E}(\mathbf{r}) - \left(\tfrac{\mathbf{d}}{2}\cdot\nabla\right)\mathbf{E}(\mathbf{r}) + O(d^2)\right] \\
&= q\,\left[(\mathbf{d}\cdot\nabla)\,\mathbf{E}(\mathbf{r}) + O(d^2)\right] \;\longrightarrow\; (\mathbf{p}\cdot\nabla)\,\mathbf{E}\,. \qquad (4.1)
\end{aligned}$$

Here we have placed the charge $+q$ at position vector $\mathbf{r}+\mathbf{d}/2$ and the charge $-q$ at position vector $\mathbf{r}-\mathbf{d}/2$, and then used Taylor's theorem and taken the point dipole limit. This force is *conservative*, with potential energy function V_{dipole} given by

$$\mathbf{F} = -\nabla V_{\mathrm{dipole}}\,, \qquad V_{\mathrm{dipole}} \equiv -\mathbf{p}\cdot\mathbf{E} = \mathbf{p}\cdot\nabla\phi\,. \qquad (4.2)$$

To see this, we may compute

$$-\nabla V_{\mathrm{dipole}} = -\nabla\,[(\mathbf{p}\cdot\nabla)\phi] = (\mathbf{p}\cdot\nabla)(-\nabla\phi) = (\mathbf{p}\cdot\nabla)\,\mathbf{E}\,. \qquad (4.3)$$

Here we have used the fact that $\mathbf{p}$ is constant to pass the gradient ∇ through the directional derivative $\mathbf{p}\cdot\nabla$. The potential energy $V_{\mathrm{dipole}} \equiv -\mathbf{p}\cdot\mathbf{E}$ is minimized when $\mathbf{p}$ points in the same direction as $\mathbf{E}$, and maximized when $\mathbf{p}$ points in the opposite direction to $\mathbf{E}$.

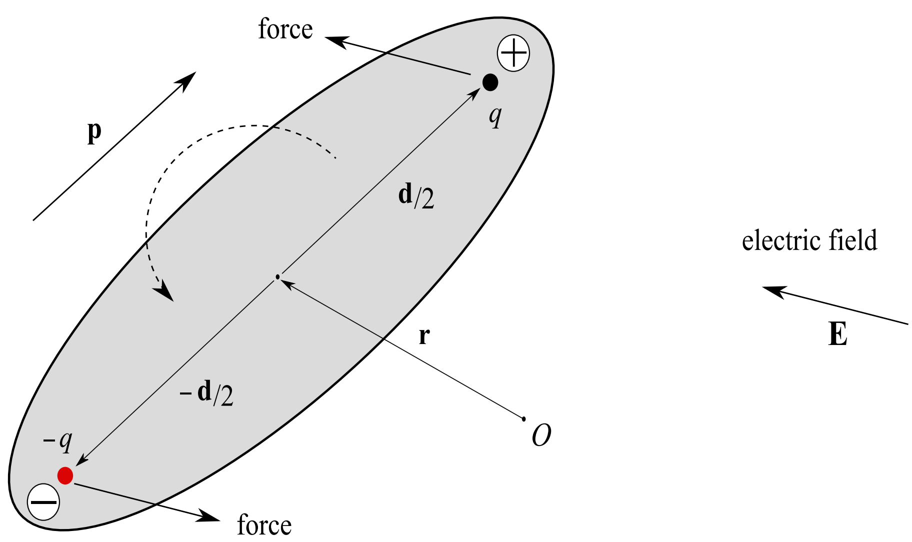


Figure 19: The electrical forces on a dipole consisting of a positive point charge q at position vector $\mathbf{r}+\mathbf{d}/2$, and a point charge $-q$ at position vector $\mathbf{r}-\mathbf{d}/2$.

There is also a *torque* on the dipole, causing it to rotate. The total torque about the point $\mathbf{r}$ is

$$\boldsymbol{\tau} = \frac{\mathbf{d}}{2}\times\left[q\,\mathbf{E}\left(\mathbf{r}+\frac{\mathbf{d}}{2}\right)\right] - \frac{\mathbf{d}}{2}\times\left[-q\,\mathbf{E}\left(\mathbf{r}-\frac{\mathbf{d}}{2}\right)\right] \;\to\; \mathbf{p}\times\mathbf{E}\,. \qquad (4.4)$$

The torque is perpendicular to the plane defined by the electric dipole moment $\mathbf{p}$ and external electric field $\mathbf{E}$, and gives a rotational force about this axis, as shown in Figure 19.[15] In equilibrium the torque on a dipole is by definition zero, but that is the case if and only if $\mathbf{p}$ is aligned with $\mathbf{E}$, so their cross product is zero, precisely as in Figure 18b. (Notice that if $\mathbf{p}$ points in the *opposite* direction to $\mathbf{E}$ this gives an *unstable equilibrium*, being a maximum of the potential energy V_{dipole}).

An electric dipole of course also generates its own electric field. Recall that the electrostatic potential generated by a single electric dipole moment $\mathbf{p}$ at the origin is

$$\phi(\mathbf{r})_{\text{dipole}} = \frac{1}{4\pi\epsilon_0}\frac{\mathbf{p}\cdot\mathbf{r}}{r^3} = -\frac{1}{4\pi\epsilon_0}\,\mathbf{p}\cdot\nabla\left(\frac{1}{r}\right) . \tag{4.5}$$

A dielectric medium by definition consists of a very large number of electric dipoles. By the Principle of Superposition, dipole moments $\mathbf{p}_i$ at position vectors $\mathbf{r}_i$ will generate a potential

$$\phi(\mathbf{r})_{\text{dipoles}} = \frac{1}{4\pi\epsilon_0}\sum_{i=1}^{N}\frac{\mathbf{p}_i\cdot(\mathbf{r}-\mathbf{r}_i)}{|\mathbf{r}-\mathbf{r}_i|^3} \longrightarrow \frac{1}{4\pi\epsilon_0}\int_{\mathbf{r}'\in R}\frac{\mathbf{P}(\mathbf{r}')\cdot(\mathbf{r}-\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|^3}\,\mathrm{d}V' . \tag{4.6}$$

Here we have taken the usual continuum limit on the right hand side, where by definition $\mathbf{P}(\mathbf{r}')\,\delta V'$ is the electric dipole moment in a small volume $\delta V'$, centred at position $\mathbf{r}'$.

Definition The vector field $\mathbf{P}$ is called the *electric polarization density*, describing the distribution of dipoles in the dielectric medium.

On the other hand, as on the right hand side of (4.5) we may then calculate

$$\begin{aligned}
\phi(\mathbf{r})_{\text{dipoles}} &= -\frac{1}{4\pi\epsilon_0}\int_{\mathbf{r}'\in R}\mathbf{P}(\mathbf{r}')\cdot\nabla\left(\frac{1}{|\mathbf{r}-\mathbf{r}'|}\right)\mathrm{d}V' = \frac{1}{4\pi\epsilon_0}\int_{\mathbf{r}'\in R}\mathbf{P}(\mathbf{r}')\cdot\nabla'\left(\frac{1}{|\mathbf{r}-\mathbf{r}'|}\right)\mathrm{d}V' \\
&= \frac{1}{4\pi\epsilon_0}\int_{\mathbf{r}'\in R}\nabla'\cdot\left(\frac{\mathbf{P}(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}\right)\mathrm{d}V' - \frac{1}{4\pi\epsilon_0}\int_{\mathbf{r}'\in R}\frac{\nabla'\cdot\mathbf{P}(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}\,\mathrm{d}V' \\
&= \frac{1}{4\pi\epsilon_0}\int_{\mathbf{r}'\in R}\frac{(-\nabla'\cdot\mathbf{P}(\mathbf{r}'))}{|\mathbf{r}-\mathbf{r}'|}\,\mathrm{d}V' .
\end{aligned} \tag{4.7}$$

Here we have used (3.23) on the first line, and in the last step have used the divergence theorem, with $\mathbf{P}$ by definition being zero on, and outside, the boundary of the region R in which the dipoles are supported. Comparing to (1.28) we may *define*

$$\rho_{\text{bound}}(\mathbf{r}) \equiv -\nabla\cdot\mathbf{P}(\mathbf{r}) . \tag{4.8}$$

This is the *effective* charge density that produces the electrostatic potential (4.7) due to dipoles. It is called $\rho_{\text{bound}}(\mathbf{r})$ as it is effectively generated by charges that are bonded to their overall neutral molecules. Notice that the definition (4.8) looks like Gauss' law (1.16), with $-\mathbf{P}/\epsilon_0$ playing the role of the electric field.

What is Gauss' law in this set-up? We may divide the total charge density ρ into two types: (i) the bound charge density ρ_{bound} in (4.8), that generates an electric field due to the *polarization* of

[15] Recall that an applied torque is equal to the rate of change of angular momentum.

the dielectric medium, and (ii) an "ordinary" charge density, which we call ρ_{free}. The latter consists of any charge *added* to the interior or exterior of the dielectric medium, where we effectively *ignore* the charges that make up the dipoles in the material (those are accounted for in ρ_{bound}). Then Gauss' law (1.16) reads

$$\nabla \cdot \mathbf{E} = \frac{\rho}{\epsilon_0} = \frac{1}{\epsilon_0}\left(\rho_{\mathrm{free}} + \rho_{\mathrm{bound}}\right) = \frac{1}{\epsilon_0}\rho_{\mathrm{free}} - \frac{1}{\epsilon_0}\nabla \cdot \mathbf{P} \ . \tag{4.9}$$

On the other hand, our discussion of Figure 18 led us to conclude that the polarization density $\mathbf{P}$ is everywhere aligned with the electric field $\mathbf{E}$. That is, $\mathbf{P}(\mathbf{r}) \times \mathbf{E}(\mathbf{r}) = \mathbf{0}$, so there is no torque on the dipoles at position $\mathbf{r}$. We thus write

$$\frac{1}{\epsilon_0}\mathbf{P} \equiv \chi_e \, \mathbf{E} \equiv \left(\frac{\epsilon}{\epsilon_0} - 1\right)\mathbf{E} \ . \tag{4.10}$$

Definition ϵ in (4.10) is called the *permittivity* of the dielectric medium, with $\chi_e \equiv (\epsilon/\epsilon_0 - 1)$ called the *electric susceptibility*, which measures the degree of response of the dielectric to an applied electric field.

Here in general ϵ may depend on: (i) position $\mathbf{r}$ if the dielectric medium is not uniform (ϵ may depend on local temperature, pressure, *etc*), (ii) time t if the medium is not static, and (iii) the frequency, magnitude and direction of the electric field. However, in many practical situations ϵ may be treated as approximately constant within the medium. For the *vacuum* $\epsilon/\epsilon_0 = 1$, so that $\mathbf{P} = \mathbf{0}$ and there is no induced polarization, while for other media $\epsilon/\epsilon_0 > 1$, *e.g.* for air $\epsilon/\epsilon_0 \simeq 1.0005$, for water $\epsilon/\epsilon_0 \simeq 1.77$.

Substituting (4.10) into (4.9), we have

$$\frac{1}{\epsilon_0}\rho_{\mathrm{free}} = \nabla \cdot \left(\mathbf{E} + \frac{1}{\epsilon_0}\mathbf{P}\right) = \nabla \cdot \left(\frac{\epsilon}{\epsilon_0}\mathbf{E}\right) \ , \tag{4.11}$$

and the electrostatic Maxwell equations become simply[16]

$$\boxed{\nabla \cdot (\epsilon \, \mathbf{E}) = \rho_{\mathrm{free}} \ , \qquad \nabla \times \mathbf{E} = \mathbf{0} \ .} \tag{4.12}$$

Notice that all we have done is effectively replace $\epsilon_0 \to \epsilon$! When ϵ is constant the resulting Maxwell equations are mathematically identical to the electrostatic Maxwell equations we studied in sections 1 and 2, and may be solved with exactly the same methods. Note here that since $\epsilon/\epsilon_0 > 1$, the electric field $\mathbf{E}$ generated by the free charges ρ_{free} in a dielectric medium is *smaller* than it would have been without the dielectric present, by a factor of ϵ_0/ϵ.

When we have *boundaries* between materials with different values of ϵ, for example, air and water, the following version of Proposition 1.8 applies (we shall need this result in section 6):

Proposition 4.1 *Consider a surface S with (free) surface charge density σ, that is the boundary between two dielectric materials with different permittivities $\epsilon^{\pm}$. Then*

$$\epsilon^+ \, \mathbf{E}^+ \cdot \mathbf{n} - \epsilon^- \, \mathbf{E}^- \cdot \mathbf{n} = \sigma \ , \tag{4.13}$$

[16] The quantity $\mathbf{D} \equiv \epsilon \, \mathbf{E}$ in (4.11) is called the *electric displacement*.

relates the electric fields $\mathbf{E}^{\pm}$ *on the two sides of* S*, with* $\mathbf{n}$ *the unit normal pointing into the "+ side". On the other hand, the components of* $\mathbf{E}$ *tangent to* S *are continuous across* S*.*

Proof The proof is almost identical to that for Proposition 1.8, where instead in the first step we integrate Gauss' law $\nabla \cdot (\epsilon\,\mathbf{E}) = \rho_{\text{free}}$ over the cylindrical region R. The divergence theorem then gives terms $\epsilon^{\pm}\,\mathbf{E}^{\pm}$ integrated through the top and bottom surfaces of this cylinder, respectively. The equation $\nabla \times \mathbf{E} = \mathbf{0}$ is the same as before, and so proof for the components of $\mathbf{E}$ tangent to S is identical to that for Proposition 1.8. ■

4.3 Magnetic dipoles

In section 3 we explained that electric currents generate magnetic fields, via the Biot-Savart law (3.10), and conversely magnetic fields exert a force on electric currents via the magnetic component of the Lorentz force (3.8). In fact these two statements essentially summarize magnetostatics. But you might then ask: where is the electric current that generates the magnetic field of a *permanent magnet*? Similarly, in Figure 15 we mentioned that you might have seen iron filings lining up with the magnetic field lines around a current carrying wire: but *why* do they line up? If the magnetic force in (3.8) is responsible for this, where is the current in an iron filing, or in the needle of a compass that aligns with the Earth's magnetic field?!

*** Remark on spin** Perhaps surprisingly, the answers to these questions involve quantum mechanics in an essential way, even though the effects are macroscopic. In some magnetic materials, the macroscopic magnetic field is indeed produced by the alignment of tiny atomic currents generated by the electrons in the material. However, in materials such as iron it is the alignment of the "spins" of (unpaired) electrons in the atoms that is responsible for producing the magnetic field. If you took the Part A Quantum Theory course, you will have encountered the fact that a single electron has an *intrinsic angular momentum*, called its *spin*. A proper discussion of this quantum mechanical notion is beyond our course here, but we can understand to some extent why "spin" angular momentum might generate a magnetic field, via the following classical argument.

First recall the definition (3.48) of the *magnetic dipole moment*:

$$\mathbf{m} \equiv \frac{1}{2}\int_{\mathbf{r}' \in R} \mathbf{r}' \times \mathbf{J}(\mathbf{r}')\,\mathrm{d}V' \, . \tag{4.14}$$

Here $\mathbf{J}$ is a steady current density supported in a region R, and the magnetic dipole moment determines the leading order behaviour of the vector potential $\mathbf{A}(\mathbf{r})$ in the expansion (3.49). Recalling also that $\mathbf{J} = \rho\,\mathbf{v}$, for a charge density ρ with velocity vector $\mathbf{v}$, notice that the expression (4.14) is very similar to that for *angular momentum*. More precisely, for a mass distribution with density $\rho(\mathbf{r})_{\text{mass}}$ supported in a region R, the angular momentum (about the origin O) is

$$\mathbf{L} = \int_{\mathbf{r}' \in R} \mathbf{r}' \times \rho_{\text{mass}}\,\mathbf{v}\,\mathrm{d}V' \, . \tag{4.15}$$

Here $\rho(\mathbf{r}')_{\text{mass}}\,\delta V'$ is the mass of a small volume $\delta V'$ centred at $\mathbf{r}'$. Let us assume that the charge and mass densities of some matter are proportional to each other, so $\rho = 2\gamma\,\rho_{\text{mass}}$ with γ a constant; for example, this is the case if the matter is all made of the same elementary particles. Then

$$\mathbf{m} \,=\, \frac{1}{2}\int_{\mathbf{r}'\in R} \mathbf{r}'\times\rho\,\mathbf{v}\,\mathrm{d}V' \,=\, \gamma\int_{\mathbf{r}'\in R}\mathbf{r}'\times\rho_{\text{mass}}\,\mathbf{v}\,\mathrm{d}V' \,=\, \gamma\,\mathbf{L}\ . \tag{4.16}$$

The magnetic dipole moment and angular momentum of the matter are hence proportional, with *gyromagnetic ratio* γ satisfying

$$Q \,=\, \int_R \rho\,\mathrm{d}V \,=\, 2\gamma\int_R \rho_{\text{mass}}\,\mathrm{d}V \,=\, 2M\gamma \quad\Rightarrow\quad \gamma \,=\, \frac{Q}{2M}\ , \tag{4.17}$$

where M is the total mass. It is then the *angular momentum* $\mathbf{L}$ of the charge distribution that is effectively generating the magnetic dipole field in (3.51), with $\mathbf{m} = \gamma\,\mathbf{L}$.

We might then crudely imagine an electron as a ball of charge that is spinning about some axis, with the resulting angular momentum of this body being its "spin" angular momentum. The rotating charge gives an electric current, which in turn generates a magnetic dipole moment $\mathbf{m}$. This classical picture is not really a correct description of an electron, but it is nevertheless true that a single electron behaves like a little magnetic dipole, with the magnetic dipole moment aligned with the direction of its spin. In a ferromagnetic material, such as iron, the spins of the electrons can all be aligned, and superposing all the magnetic dipole fields of the electrons leads to the macroscopic magnetic field in your fridge magnets. **[end of remark on spin]**

As in the previous subsection, we can make all of this more quantitative by studying magnetic dipoles. This is relegated to the Additional Material and here we will just state the outcome of these calculations.

We can define the magnetic analogue of the electric polarization density $\mathbf{P}$, called the magnetization density $\mathbf{M}$, which determines the density of magnetic dipole moments. By studying forces action on magnetic dipoles one can derive that $\mathbf{M}$ aligns with $\mathbf{B}$ in media,

$$\mathbf{M} \,\equiv\, \frac{\chi_m}{\mu}\,\mathbf{B} \,\equiv\, \left(\frac{1}{\mu_0} - \frac{1}{\mu}\right)\mathbf{B}\ . \tag{4.18}$$

Definition μ is called the *permeability*, with $\chi_m \equiv (\mu/\mu_0 - 1)$ the *magnetic susceptibility.*

μ is not constant in general, but for uniform materials it is approximately constant. For the vacuum $\mu/\mu_0 = 1$, for air $\mu/\mu_0 \simeq 1.00000037$, for water $\mu/\mu_0 \simeq 0.999992$, while for iron $\mu/\mu_0 \simeq 200{,}000$!

Substituting (4.18) into the magnetic analogue (4.9) (not derived here), we have

$$\mu_0\,\mathbf{J}_{\text{free}} \,=\, \nabla\times(\mathbf{B} - \mu_0\,\mathbf{M}) \,=\, \mu_0\,\nabla\times\left(\frac{1}{\mu}\,\mathbf{B}\right)\ . \tag{4.19}$$

which leads to the effective Maxwell equations

$$\boxed{\nabla\cdot\mathbf{B} \,=\, 0\ , \qquad \nabla\times\left(\frac{1}{\mu}\,\mathbf{B}\right) \,=\, \mathbf{J}_{\text{free}}\ .} \tag{4.20}$$

Notice that the magnetic field generated by $\mathbf{J}_{\text{free}}$ in a magnetic material is μ/μ_0 times the field that would be generated *without* the magnetic material present. This is *e.g.* a little larger for air, since $\chi_m > 0$ (called *paramagnetism*), but a little smaller for water, since $\chi_m < 0$ (called *diamagnetism*).

Definition The quantity $\mathbf{H} \equiv \frac{1}{\mu}\mathbf{B}$ in (4.20) is called the *magnetic field strength.*

Finally, analogously to Proposition 4.1 we have:

Proposition 4.2 *Consider a surface S that is the boundary between two materials with different permeabilities $\mu^{\pm}$. Then the components of $\mathbf{B}$ normal to S are continuous across S, so*

$$\left(\mathbf{B}^{+} - \mathbf{B}^{-}\right) \cdot \mathbf{n} \, = \, 0 \, , \tag{4.21}$$

while the components of $\frac{1}{\mu}\mathbf{B}$ tangent to S are continuous across S, so

$$\left(\frac{1}{\mu^{+}}\mathbf{B}^{+} \, - \, \frac{1}{\mu^{-}}\mathbf{B}^{-}\right) \cdot \mathbf{t} \, = \, 0 \, . \tag{4.22}$$

Here $\mathbf{n}$ is the unit normal to S, pointing into the "+ side", and $\mathbf{t}$ is any tangent vector to S.

Proof Following the proof of Proposition 1.8, (4.21) follows from integrating $\nabla \cdot \mathbf{B} = 0$ over the cylindrical region R, while (4.22) follows from integrating $\nabla \times (\frac{1}{\mu}\mathbf{B}) = \mathbf{0}$ over the surface Σ. ■

5 Electrodynamics and Maxwell's equations

5.1 Maxwell's displacement current

Let's go back to Ampère's law (3.29) in magnetostatics

$$\int_{C=\partial\Sigma} \mathbf{B}\cdot \mathrm{d}\mathbf{r} \;=\; \mu_0 \int_\Sigma \mathbf{J}\cdot \mathrm{d}\mathbf{S}\ . \tag{5.1}$$

Here $C=\partial\Sigma$ is a simple closed curve bounding a surface Σ. Of course, one may use *any* such surface spanning C on the right hand side. If we pick a different surface Σ', with $C=\partial\Sigma'$, then

$$\begin{aligned} 0 \;&=\; \int_\Sigma \mathbf{J}\cdot \mathrm{d}\mathbf{S} - \int_{\Sigma'} \mathbf{J}\cdot \mathrm{d}\mathbf{S} \\ &=\; \int_S \mathbf{J}\cdot \mathrm{d}\mathbf{S}\ . \end{aligned} \tag{5.2}$$

Here S is the *closed* surface obtained by gluing Σ and Σ' together along C. Thus the flux of $\mathbf{J}$ through any closed surface is zero. We may see this in a different way if we assume that $S=\partial R$ bounds a region R, since then

$$\int_S \mathbf{J}\cdot \mathrm{d}\mathbf{S} \;=\; \int_R \nabla\cdot\mathbf{J}\,\mathrm{d}V \;=\; 0\ , \tag{5.3}$$

and in the last step we have used the steady current condition (3.7).

But in general, (3.7) should be replaced by the continuity equation (3.6). The above calculation then changes as follows:

$$\begin{aligned} \int_\Sigma \mathbf{J}\cdot \mathrm{d}\mathbf{S} - \int_{\Sigma'} \mathbf{J}\cdot \mathrm{d}\mathbf{S} \;&=\; \int_S \mathbf{J}\cdot \mathrm{d}\mathbf{S} \;=\; \int_R \nabla\cdot\mathbf{J}\,\mathrm{d}V \;=\; -\int_R \frac{\partial\rho}{\partial t}\,\mathrm{d}V \;=\; -\epsilon_0 \int_R \frac{\partial}{\partial t}\left(\nabla\cdot\mathbf{E}\right)\mathrm{d}V \\ &=\; -\epsilon_0 \int_{S=\partial R} \frac{\partial\mathbf{E}}{\partial t}\cdot \mathrm{d}\mathbf{S} \;=\; -\epsilon_0 \int_\Sigma \frac{\partial\mathbf{E}}{\partial t}\cdot \mathrm{d}\mathbf{S} + \epsilon_0 \int_{\Sigma'} \frac{\partial\mathbf{E}}{\partial t}\cdot \mathrm{d}\mathbf{S}\ . \end{aligned} \tag{5.4}$$

Here in the first line we have used the divergence theorem in the second equality, the continuity equation (3.6) in the third equality, and Gauss' law (1.16) in the final equality. Notice we now regard $\mathbf{E}=\mathbf{E}(\mathbf{r},t)$ as a vector field depending on time. In the second line of (5.4) we have then again used the divergence theorem. We have thus shown that

$$\int_\Sigma \left(\mathbf{J} + \epsilon_0 \frac{\partial\mathbf{E}}{\partial t}\right)\cdot \mathrm{d}\mathbf{S} \;=\; \int_{\Sigma'} \left(\mathbf{J} + \epsilon_0 \frac{\partial\mathbf{E}}{\partial t}\right)\cdot \mathrm{d}\mathbf{S} \tag{5.5}$$

for any two surfaces Σ, Σ' spanning C, and thus suggests replacing Ampère's law (3.28) by

$$\boxed{\nabla\times\mathbf{B} \;=\; \mu_0\left(\mathbf{J} + \epsilon_0 \frac{\partial\mathbf{E}}{\partial t}\right)\ . \qquad (5.6)}$$

This is indeed the correct time-dependent Maxwell equation on the front page. The additional term $\partial\mathbf{E}/\partial t$ is called the *displacement current*, and the above argument is due to Maxwell. It says that a *time-dependent* electric field also produces a magnetic field.

5.2 Faraday's law

The electrostatic equation (1.33) is also modified in the time-dependent case. We can motivate how precisely by the following argument. Consider the electromagnetic field generated by a set of charges all moving with constant velocity $\mathbf{v}$. The charges generate both an $\mathbf{E}$ and a $\mathbf{B}$ field, the latter since the charges are in motion. However, consider instead an observer who is *also* moving at the same constant velocity $\mathbf{v}$. For this observer, the charges are at *rest*, and thus he/she will measure *only* an electric field $\mathbf{E}'$ from the Lorentz force law (3.8) on one of the charges! Indeed, when we wrote down the Lorentz force law (3.8) and Biot-Savart law (3.10), we didn't specify what inertial reference frame we should use to measure the velocities, and this should have worried you at the time!

Since (or assuming) the two observers above must be measuring the same force on a given charge, we conclude that

$$\mathbf{E}' = \mathbf{E} + \mathbf{v} \times \mathbf{B} . \tag{5.7}$$

Now since the field is electrostatic for the moving observer,

$$\begin{aligned} \mathbf{0} &= \nabla \times \mathbf{E}' = \nabla \times \mathbf{E} + \nabla \times (\mathbf{v} \times \mathbf{B}) \\ &= \nabla \times \mathbf{E} + \mathbf{v}\,(\nabla \cdot \mathbf{B}) - (\mathbf{v} \cdot \nabla)\,\mathbf{B} = \nabla \times \mathbf{E} - (\mathbf{v} \cdot \nabla)\,\mathbf{B} . \end{aligned} \tag{5.8}$$

Here in the second line we have used the identity (A.11), and in the last step we have used $\nabla \cdot \mathbf{B} = 0$. Now, for the original observer the charges are all moving with velocity $\mathbf{v}$, so the magnetic field at position $\mathbf{r} + \mathbf{v}\tau$ and time $t + \tau$ is the same as that at position $\mathbf{r}$ and time t:

$$\mathbf{B}(\mathbf{r} + \mathbf{v}\tau, t + \tau) = \mathbf{B}(\mathbf{r}, t) . \tag{5.9}$$

This equation holds for all τ. Dividing through by τ and then taking the limit $\tau \to 0$ leads to the partial differential equation

$$(\mathbf{v} \cdot \nabla)\,\mathbf{B} + \frac{\partial \mathbf{B}}{\partial t} = \mathbf{0} . \tag{5.10}$$

Substituting this into the right hand side of (5.8) then gives

$$\boxed{\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} .} \tag{5.11}$$

This is *Faraday's law*, and is another of Maxwell's equations. The above argument raises issues about what happens in general to our equations when we change to a moving frame. A systematic study of this leads to Einstein's theory of *Special Relativity*, which we shall comment on in the Additional Material.

As usual, the equation (5.11) may be expressed as an integral equation as

Faraday's law *For any simple closed curve* $C = \partial\Sigma$ *bounding a* fixed *surface* Σ

$$\int_{C=\partial\Sigma} \mathbf{E} \cdot \mathrm{d}\mathbf{r} \;=\; -\frac{\mathrm{d}}{\mathrm{d}t}\int_{\Sigma} \mathbf{B} \cdot \mathrm{d}\mathbf{S} \;. \tag{5.12}$$

This says that a *time-dependent* magnetic field produces an electric field. For example, if one moves a bar magnet through a loop of conducting wire C, the resulting electric field from (5.11) induces a current in the wire via the Lorentz force. This is what Faraday did, in fact, in 1831.

Definition The integral $\int_\Sigma \mathbf{B} \cdot \mathrm{d}\mathbf{S}$ is called the *magnetic flux through* Σ.

* The current in the wire then *itself* produces a magnetic field of course, via Ampère's law. However, the signs are such that this magnetic field is in the *opposite* direction to the change in the magnetic field that created it. This is called *Lenz's law*, and a similar effect is what leads to *diamagnetism*, mentioned at the end of section 4. The whole setup may be summarized as follows:

$$\text{changing } \mathbf{B} \;\stackrel{\text{Faraday}}{\longrightarrow}\; \mathbf{E} \;\stackrel{\text{Lorentz}}{\longrightarrow}\; \text{current} \;\stackrel{\text{Ampère}}{\longrightarrow}\; \mathbf{B} \;. \tag{5.13}$$

5.3 Maxwell's equations

We now summarize the full set of Maxwell equations.

There are two scalar equations, namely Gauss' law (1.16) from electrostatics, and the equation (3.17) from magnetostatics that expresses the absence of magnetic monopoles:

$$\nabla \cdot \mathbf{E} \;=\; \frac{\rho}{\epsilon_0} \;, \tag{5.14}$$

$$\nabla \cdot \mathbf{B} \;=\; 0 \;. \tag{5.15}$$

Although we discussed these only in the time-independent case, they are in fact true in general. The are also two vector equations, namely Faraday's law (5.11) and Maxwell's modification (5.6) of Ampère's law (3.28) from magnetostatics:

$$\nabla \times \mathbf{E} \;=\; -\frac{\partial \mathbf{B}}{\partial t} \;, \tag{5.16}$$

$$\nabla \times \mathbf{B} \;=\; \mu_0\left(\mathbf{J} + \epsilon_0\frac{\partial \mathbf{E}}{\partial t}\right) \;. \tag{5.17}$$

Together with the Lorentz force law

$$\mathbf{F} \;=\; q\left(\mathbf{E} + \mathbf{u} \times \mathbf{B}\right) \;, \tag{5.18}$$

which governs the mechanics, this is *all* of electromagnetism. Everything else we have discussed may in fact be derived from these equations.

Maxwell's equations, for given ρ and $\mathbf{J}$, are 8 equations for 6 unknowns. There must therefore be two consistency conditions. To see what these are, we first compute

$$\frac{\partial}{\partial t}\left(\nabla \cdot \mathbf{B}\right) \;=\; -\nabla \cdot \left(\nabla \times \mathbf{E}\right) \;=\; 0 \;, \tag{5.19}$$

where we have used (5.16). This is clearly consistent with (5.15). We get something non-trivial by instead taking the divergence of (5.17), which gives

$$\begin{aligned} 0 &= \nabla \cdot \left(\frac{1}{\mu_0} \nabla \times \mathbf{B} \right) = \nabla \cdot \mathbf{J} + \epsilon_0 \frac{\partial}{\partial t} (\nabla \cdot \mathbf{E}) \\ &= \nabla \cdot \mathbf{J} + \frac{\partial \rho}{\partial t} , \end{aligned} \tag{5.20}$$

where we have used (5.14) in the last step. Thus the continuity equation arises as a *consistency condition* for Maxwell's equations: if ρ and $\mathbf{J}$ do not satisfy (5.20), there is no solution to Maxwell's equations for this choice of charge density and current. Alternatively, we may regard this as saying that Maxwell's equations imply that charge is conserved.

5.4 Electromagnetic potentials and gauge transformations

In the general time-dependent case one can introduce electromagnetic potentials in a similar way to the static cases. We work in a suitable domain in $\mathbb{R}^3$, such as $\mathbb{R}^3$ itself or an open ball therein, as discussed in previous sections. Since $\mathbf{B}$ has zero divergence (5.15), we may again introduce a vector potential

$$\mathbf{B} = \nabla \times \mathbf{A} , \tag{5.21}$$

where now $\mathbf{A} = \mathbf{A}(\mathbf{r}, t)$. It follows from Faraday's law that

$$0 = \nabla \times \mathbf{E} + \frac{\partial \mathbf{B}}{\partial t} = \nabla \times \left(\mathbf{E} + \frac{\partial \mathbf{A}}{\partial t} \right) . \tag{5.22}$$

Thus we may introduce a scalar potential $\phi = \phi(\mathbf{r}, t)$ via

$$\mathbf{E} + \frac{\partial \mathbf{A}}{\partial t} = -\nabla \phi . \tag{5.23}$$

Thus

$$\boxed{\begin{aligned} \mathbf{B} &= \nabla \times \mathbf{A} , & (5.24) \\ \mathbf{E} &= -\nabla \phi - \frac{\partial \mathbf{A}}{\partial t} . & (5.25) \end{aligned}}$$

Note that, by construction, with (5.24) and (5.25) the Maxwell equations (5.15) and (5.16) are automatically satisfied.

Definition Generalizing the discussion in section 3.7, we define the *gauge transformations*

$$\mathbf{A} \to \widehat{\mathbf{A}} \equiv \mathbf{A} + \nabla \psi , \qquad \phi \to \widehat{\phi} \equiv \phi - \frac{\partial \psi}{\partial t} , \tag{5.26}$$

which leave (5.24) and (5.25) invariant.

Again, one may fix this non-uniqueness of $\mathbf{A}$ and ϕ by imposing certain gauge choices. Suppose we have chosen a particular $\mathbf{A}$ and ϕ satisfying (5.24), (5.25), and let $\psi = \psi(\mathbf{r}, t)$ be a solution to the following *wave equation with source*

$$\frac{1}{c^2}\frac{\partial^2 \psi}{\partial t^2} - \nabla^2 \psi \;=\; \frac{1}{c^2}\frac{\partial \phi}{\partial t} + \nabla \cdot \mathbf{A} \;, \tag{5.27}$$

where we have defined

$$c^2 \;\equiv\; \frac{1}{\epsilon_0 \mu_0} \;. \tag{5.28}$$

Compare (5.27) to the analogous time-independent Poisson equation (3.32) in magnetostatics. Then from (5.26) we compute

$$\frac{1}{c^2}\frac{\partial \widehat{\phi}}{\partial t} + \nabla \cdot \widehat{\mathbf{A}} \;=\; \frac{1}{c^2}\frac{\partial \phi}{\partial t} - \frac{1}{c^2}\frac{\partial^2 \psi}{\partial t^2} + \nabla \cdot \mathbf{A} + \nabla^2 \psi \;=\; 0 \;. \tag{5.29}$$

Definition The *Lorenz gauge* (*cf.* (3.33)) for $\mathbf{A}$, ϕ is the condition

$$\frac{1}{c^2}\frac{\partial \phi}{\partial t} + \nabla \cdot \mathbf{A} \;=\; 0 \;. \tag{5.30}$$

In Lorenz gauge Gauss' law (5.14) becomes

$$\frac{\rho}{\epsilon_0} \;=\; \nabla \cdot \mathbf{E} \;=\; \nabla \cdot \left(-\nabla \phi - \frac{\partial \mathbf{A}}{\partial t} \right) \;=\; -\nabla^2 \phi + \frac{1}{c^2}\frac{\partial^2 \phi}{\partial t^2} \;, \tag{5.31}$$

while the Ampère-Maxwell equation (5.17) becomes

$$\nabla \times \mathbf{B} \;=\; \nabla\left(\nabla \cdot \mathbf{A}\right) - \nabla^2 \mathbf{A} \;=\; \mu_0 \left[\mathbf{J} - \epsilon_0 \left(\frac{\partial}{\partial t}\nabla \phi + \frac{\partial^2 \mathbf{A}}{\partial t^2} \right) \right] \;. \tag{5.32}$$

In Lorenz gauge $\nabla\left(\nabla \cdot \mathbf{A}\right) = -\frac{1}{c^2}\frac{\partial}{\partial t}\nabla\phi = -\epsilon_0\mu_0 \frac{\partial}{\partial t}\nabla\phi$, which cancels against the same term on the right hand side of (5.32), giving

$$-\nabla^2 \mathbf{A} + \frac{1}{c^2}\frac{\partial^2 \mathbf{A}}{\partial t^2} \;=\; \mu_0 \, \mathbf{J} \;. \tag{5.33}$$

We may summarize this as follows:

Theorem 5.1 *In Lorenz gauge Maxwell's equations reduce to the wave equations with sources*

$$\Box\, \phi \;=\; -\frac{\rho}{\epsilon_0} \;, \tag{5.34}$$

$$\Box\, \mathbf{A} \;=\; -\mu_0 \, \mathbf{J} \;. \tag{5.35}$$

Here we have defined the d'Alembertian operator (*some references have the opposite overall sign*)

$$\Box \;\equiv\; -\frac{1}{c^2}\frac{\partial^2}{\partial t^2} + \nabla^2 \;. \tag{5.36}$$

It will turn out that the wave speed c, defined in terms of the permittivity and permeability of free space via (5.28), is the speed of light in vacuum, as we discuss in detail in section 6. The Green's function method for solving these wave equations with sources is discussed in section 5.6.

5.5 Electromagnetic energy and Poynting's theorem

Recall that in section 1.7 we derived a formula for the *electrostatic energy density* $\mathcal{E}_{\text{electric}} = \epsilon_0 |\mathbf{E}|^2 / 2$ in terms of the electric field $\mathbf{E}$. The electrostatic energy of a given configuration is the integral of this density over space (1.58). One can motivate the similar formula $\mathcal{E}_{\text{magnetic}} = |\mathbf{B}|^2 / 2\mu_0$ in magnetostatics, although we won't elaborate on this here. This leads to the following:

Definition The *electromagnetic energy density* is

$$\mathcal{E} \equiv \frac{1}{2}\left(\epsilon_0 |\mathbf{E}|^2 + \frac{1}{\mu_0}|\mathbf{B}|^2\right) = \frac{\epsilon_0}{2}\left(|\mathbf{E}|^2 + c^2 |\mathbf{B}|^2\right) . \tag{5.37}$$

We then have

Theorem 5.2 (*Poynting's Theorem*) *The electromagnetic energy density satisfies*

$$\boxed{\frac{\partial \mathcal{E}}{\partial t} + \nabla \cdot \boldsymbol{\mathcal{P}} = -\mathbf{E}\cdot\mathbf{J} ,} \tag{5.38}$$

where we have defined the Poynting vector $\boldsymbol{\mathcal{P}}$ *to be*

$$\boldsymbol{\mathcal{P}} \equiv \frac{1}{\mu_0}\mathbf{E}\times\mathbf{B} . \tag{5.39}$$

Proof Taking the partial derivative of (5.37) with respect to time we compute

$$\begin{aligned}
\frac{\partial \mathcal{E}}{\partial t} &= \epsilon_0\, \mathbf{E}\cdot\frac{\partial \mathbf{E}}{\partial t} + \frac{1}{\mu_0}\mathbf{B}\cdot\frac{\partial \mathbf{B}}{\partial t} \\
&= \frac{1}{\mu_0}\mathbf{E}\cdot(\nabla\times\mathbf{B} - \mu_0\,\mathbf{J}) - \frac{1}{\mu_0}\mathbf{B}\cdot(\nabla\times\mathbf{E}) \\
&= -\nabla\cdot\left(\frac{1}{\mu_0}\mathbf{E}\times\mathbf{B}\right) - \mathbf{E}\cdot\mathbf{J} .
\end{aligned} \tag{5.40}$$

Here after the first step we have used the Maxwell equations (5.17) and (5.16), respectively. The last step uses the identity (A.12). ■

Notice that, in the absence of a source current, $\mathbf{J} = \mathbf{0}$, (5.38) takes the form of a *continuity equation*, analogous to the continuity equation (3.6) that expresses conservation of charge. It is thus natural to interpret (5.38) as a *conservation of energy* equation, and so identify the Poynting vector $\boldsymbol{\mathcal{P}}$ as some kind of rate of energy flow density. One can indeed justify this by examining the above quantities in various physical applications, and we shall look at the particular case of electromagnetic waves in section 6.

Integrating (5.38) over a region R with boundary Σ, using the divergence theorem we obtain

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_R \mathcal{E}\,\mathrm{d}V = -\int_\Sigma \boldsymbol{\mathcal{P}}\cdot\mathrm{d}\mathbf{S} - \int_R \mathbf{E}\cdot\mathbf{J}\,\mathrm{d}V . \tag{5.41}$$

Given our discussion of $\mathcal{E}$, the left hand side is the *rate of increase of energy* in R. The first term on the right hand side is the *rate of energy flow into* the region R. When $\mathbf{J} = \mathbf{0}$, this is precisely analogous to our discussion of charge conservation in section 3.2. The final term on the right hand side of (5.41) is interpreted as (minus) the *rate of work by the field on the sources*. To see this, remember that the force on a charge q moving at velocity $\mathbf{v}$ is $\mathbf{F} = q\,(\mathbf{E} + \mathbf{v} \times \mathbf{B})$. This force does *work* at a rate given by $\mathbf{F} \cdot \mathbf{v} = q\,\mathbf{E} \cdot \mathbf{v}$. Recalling the definition (3.1) of $\mathbf{J} = \rho\,\mathbf{v}$, we see that the force does work on the charge $q = \rho\,\delta V$ in a small volume δV at a rate $\mathbf{F} \cdot \mathbf{v} = \mathbf{E} \cdot \mathbf{J}\,\delta V$. The final term in (5.41) is thus expressing the rate of conversion of electromagnetic energy into mechanical energy, acting on the sources.

5.6 Time-dependent Green's function

The form of the time-dependent Maxwell equations (5.34), (5.35) in Lorenz gauge motivates studying the *wave equation with source*

$$\Box\,\psi \;=\; -4\pi\, f(\mathbf{r}, t) \;, \tag{5.42}$$

for arbitrary source function $f(\mathbf{r}, t)$, where recall $\Box$ is defined in (5.36). Following section 2.2:

Definition A (time-dependent) *Green's function* is a function $G(\mathbf{r}, t\,; \mathbf{r}', t')$ satisfying

$$\Box\, G(\mathbf{r}, t\,; \mathbf{r}', t') \;=\; -4\pi\, \delta(t - t')\, \delta(\mathbf{r} - \mathbf{r}') \;. \tag{5.43}$$

Given such a function, a solution to (5.42) is

$$\psi(\mathbf{r}, t) \;=\; \int_{\mathbf{r}' \in \mathbb{R}^3} \int_{t'=-\infty}^{\infty} G(\mathbf{r}, t\,; \mathbf{r}', t')\, f(\mathbf{r}', t')\, \mathrm{d}V'\, \mathrm{d}t' \;, \tag{5.44}$$

as one sees by applying $\Box$ to the right hand side. As we saw (and also exploited) in section 2, Green's functions are not unique: we may add to G any solution $F = F(\mathbf{r}, t\,; \mathbf{r}', t')$ to the *homogeneous wave equation* $\Box\, F = 0$, and this will be another solution to (5.43). Uniqueness follows after imposing appropriate *boundary conditions*, which are ultimately determined by the precise physical setup.

To gain some insight into the time-dependent problem, let us recall the electrostatics Green's function equation

$$\nabla^2\, G(\mathbf{r}, \mathbf{r}') \;=\; -4\pi\, \delta(\mathbf{r} - \mathbf{r}') \;. \tag{5.45}$$

The unique solution to this equation that is zero "at infinity" in $\mathbb{R}^3$ is $G(\mathbf{r}, \mathbf{r}') = 1/|\mathbf{r} - \mathbf{r}'|$. Physically, this Green's function is $4\pi\epsilon_0$ times the electrostatic potential generated by a unit charge at position $\mathbf{r}'$. Our aim in this subsection is to find the analogous time-dependent solution to (5.43). Physically, from the Maxwell equation (5.34) this should be $4\pi\epsilon_0$ times the electric potential $\phi(\mathbf{r}, t)$ generated by a unit charge that appears at position $\mathbf{r}'$ at an *instant of time* t'. Before this localized disturbance happens at time t' we assume there are no electromagnetic fields excited, and hence the Green's function $G(\mathbf{r}, t\,; \mathbf{r}', t') = 0$ for $t < t'$. Since $\Box$ is a wave operator, we can guess that the disturbance at position $\mathbf{r}'$ and time t' should lead to a *wave* that propagates spherically outwards from the source point $\mathbf{r}'$, at speed c.

To work out the details, we make use of the exponential Fourier expansion/Fourier transform of section 2.4, with complete set of functions (2.56). We already computed the electrostatic Green's function $G(\mathbf{r},\mathbf{r}') = 1/|\mathbf{r}-\mathbf{r}'|$ in this expansion in equation (2.83). Following (2.78), we first write

$$\delta(t-t')\,\delta(\mathbf{r}-\mathbf{r}') \;=\; \frac{1}{(2\pi)^4}\int_{\mathbf{k}\in\mathbb{R}^3}\int_{\omega=-\infty}^{\infty} \mathrm{e}^{-\mathrm{i}\omega(t-t')}\,\mathrm{e}^{\mathrm{i}\mathbf{k}\cdot(\mathbf{r}-\mathbf{r}')}\,\mathrm{d}^3k\,\mathrm{d}\omega\ , \tag{5.46}$$

which expresses the completeness relation for the functions (2.56). Notice that compared to (2.78) there is additional integral over an angular frequency variable ω, which leads to the time dependence. We correspondingly write

$$G(\mathbf{r},t\,;\mathbf{r}',t') = \frac{1}{(2\pi)^2}\int_{\mathbf{k}\in\mathbb{R}^3}\int_{\omega=-\infty}^{\infty} g(\mathbf{k},\omega)\,\mathrm{e}^{-\mathrm{i}\omega(t-t')}\,\mathrm{e}^{\mathrm{i}\mathbf{k}\cdot(\mathbf{r}-\mathbf{r}')}\,\mathrm{d}^3k\,\mathrm{d}\omega\ , \tag{5.47}$$

where we have tacitly assumed that the solution we want depends only on $\mathbf{r}-\mathbf{r}'$ (as for the electrostatic Green's function $1/|\mathbf{r}-\mathbf{r}'|$) and $t-t'$. The function $g(\mathbf{k},\omega)$ is the *Fourier transform* of G. Applying $\square = -\frac{1}{c^2}\frac{\partial^2}{\partial t^2}+\nabla^2$ to (5.47), a similar calculation to (2.80) gives

$$\square\left(\mathrm{e}^{-\mathrm{i}\omega(t-t')}\,\mathrm{e}^{\mathrm{i}\mathbf{k}\cdot(\mathbf{r}-\mathbf{r}')}\right) \;=\; \left(\frac{\omega^2}{c^2}-|\mathbf{k}|^2\right)\mathrm{e}^{-\mathrm{i}\omega(t-t')}\,\mathrm{e}^{\mathrm{i}\mathbf{k}\cdot(\mathbf{r}-\mathbf{r}')}\ . \tag{5.48}$$

Equating $\square\,G$ with -4π times (5.46), we can read off the Fourier coefficient function

$$g(\mathbf{k},\omega) \;=\; \frac{4\pi}{(2\pi)^2}\,\frac{1}{|\mathbf{k}|^2-\omega^2/c^2}\ . \tag{5.49}$$

Substituting this back into (5.47) gives

$$G(\mathbf{r},t\,;\mathbf{r}',t') \;=\; \frac{4\pi}{(2\pi)^4}\int_{\mathbf{k}\in\mathbb{R}^3}\int_{\omega=-\infty}^{\infty}\frac{1}{|\mathbf{k}|^2-\omega^2/c^2}\,\mathrm{e}^{-\mathrm{i}\omega(t-t')}\,\mathrm{e}^{\mathrm{i}\mathbf{k}\cdot(\mathbf{r}-\mathbf{r}')}\,\mathrm{d}^3k\,\mathrm{d}\omega\ . \tag{5.50}$$

This is analogous to the way we derived (2.83), although there is now an additional integral over ω. Looking more closely at the latter integral

$$I \;=\; \int_{\omega=-\infty}^{\infty}\frac{1}{|\mathbf{k}|^2-\omega^2/c^2}\,\mathrm{e}^{-\mathrm{i}\omega(t-t')}\,\mathrm{d}\omega\ , \tag{5.51}$$

we see that there are two simple poles at $\omega=\pm c\,|\mathbf{k}|$. Because of this, as written the integral (5.51) is not actually well-defined, and this is related to the already-mentioned fact that Green's function solutions to (5.43) are not unique.

We may define and then evaluate the integral (5.51) more carefully using a complex contour method. We set

$$\omega \;\equiv\; \mathrm{Re}\,z\ , \qquad a \;\equiv\; -(t-t')\ , \qquad f(z) \;\equiv\; \frac{1}{|\mathbf{k}|^2-z^2/c^2}\ , \tag{5.52}$$

and consider the contour integral

$$I(\Gamma) \;\equiv\; \int_\Gamma f(z)\,\mathrm{e}^{\mathrm{i}az}\,\mathrm{d}z\ . \tag{5.53}$$

This is to be regarded as a *definition* of the ill-defined integral in (5.51). The contour Γ should include the real axis, so as to reproduce the real integral I in (5.51), although as mentioned there are simple poles at $z=\pm c\,|\mathbf{k}|$: these must be avoided in order for the integral to be well-defined. We also recall the following:

Lemma 5.3 (*Jordan's Lemma*) *Let $f(z)$ be a meromorphic function on the complex plane, and suppose that $f(z) \to 0$ as $|z| = R \to \infty$. Let $\gamma_R^{\text{upper}}(\theta) = R\,\mathrm{e}^{\mathrm{i}\theta}$ for $\theta \in [0,\pi]$ be the semi-circular contour of radius R in the upper half plane, and $\gamma_R^{\text{lower}}(\theta) = R\,\mathrm{e}^{\mathrm{i}\theta}$ for $\theta \in [\pi, 2\pi]$ be the semi-circular contour of radius R in the lower half plane. Then*

$$\begin{cases} \lim_{R\to\infty} \int_{\gamma_R^{\text{upper}}} f(z)\,\mathrm{e}^{\mathrm{i}az}\,\mathrm{d}z \;=\; 0\;, & \textit{for } a > 0\;, \\ \lim_{R\to\infty} \int_{\gamma_R^{\text{lower}}} f(z)\,\mathrm{e}^{\mathrm{i}az}\,\mathrm{d}z \;=\; 0\;, & \textit{for } a < 0\;. \end{cases} \tag{5.54}$$

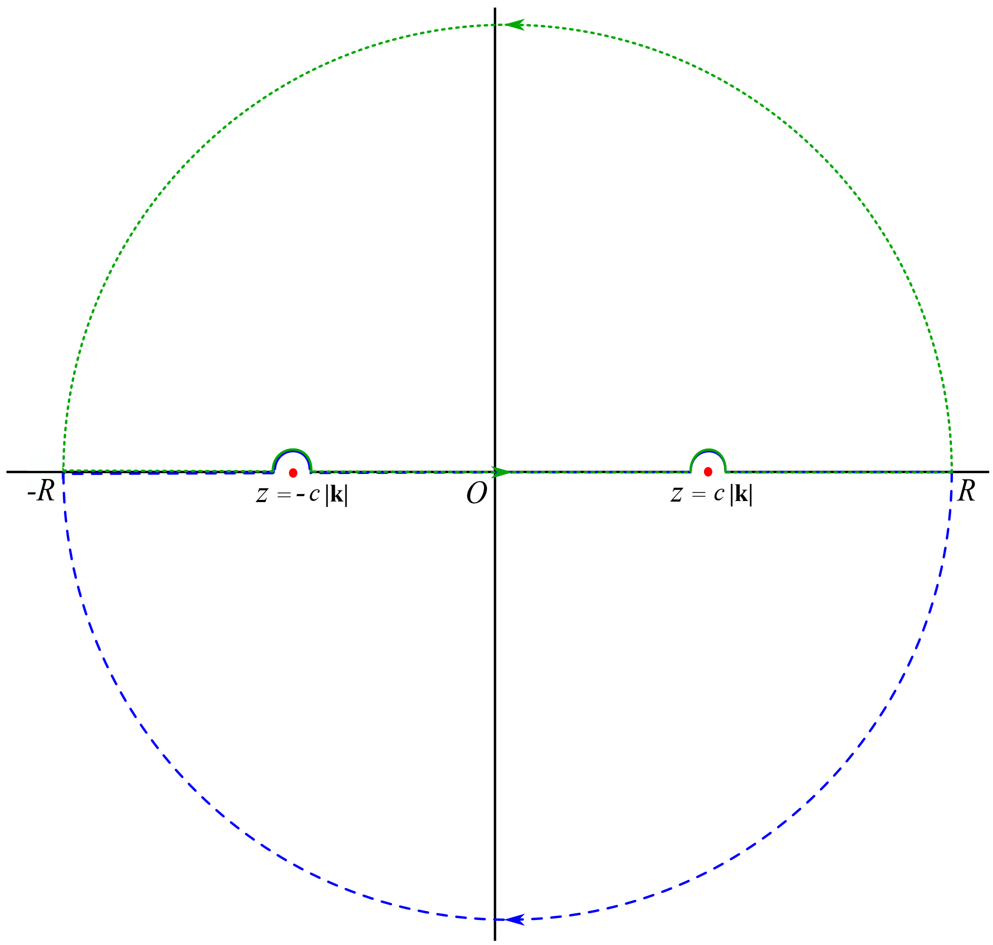


Figure 20: The contour Γ. For $t < t'$ we close in the upper half plane (shown in green/small dashes), while for $t > t'$ we close in the lower half plane (shown in blue/longer dashes). In both cases we indent the contour *above* the simple poles at $z = \pm c\,|\mathbf{k}|$, so that the upper green contour contains no singularities, while the lower blue contour contains the two simple poles.

The correct contour Γ for the Green's function we want is shown in Figure 20. Notice

(i) We choose to indent Γ *above* both simple poles at $z = \pm c\,|\mathbf{k}|$.

(ii) For $a = -(t - t') > 0$, or equivalently $t < t'$, we close the contour in the upper half plane, while for $a = -(t-t') < 0$, or equivalently $t > t'$, we close the contour in the lower half plane.

Point (ii) here allows us to apply Jordan's Lemma 5.3 to conclude that the semi-circular part of the contour does not contribute to the integral $I(\Gamma)$ in the limit $R \to \infty$. The choice of indentation in point (i) then means that the upper contour for $t < t'$ contains no singularities, and the Residue Theorem immediately gives $I(\Gamma) = 0$. With this definition of the ω integral in (5.51), and hence (5.50), we deduce that

$$G(\mathbf{r}, t\,; \mathbf{r}', t') \;=\; 0 \quad \text{for } t < t' \; . \tag{5.55}$$

This was a physical requirement we mentioned earlier: *before* the localized disturbance at time t' we assume there are no electromagnetic fields excited.

Definition Green's functions satisfying (5.55) are called *retarded Green's functions.*

Remark If we had chosen to indent *below* both simple poles, we would instead obtain a Green's function which is zero for $t > t'$ (called an *advanced Green's function*, which is also useful).

Using the lower contour in Figure 20, for $t > t'$ the Residue Theorem instead gives

$$\begin{aligned} I(\Gamma) \;&=\; -2\pi\mathrm{i}\left[\text{Residue}\left(\frac{\mathrm{e}^{-\mathrm{i}(t-t')z}}{|\mathbf{k}|^2 - z^2/c^2}\,,\; z = c\,|\mathbf{k}|\right) + \text{Residue}\left(\frac{\mathrm{e}^{-\mathrm{i}(t-t')z}}{|\mathbf{k}|^2 - z^2/c^2}\,,\; z = -c\,|\mathbf{k}|\right)\right] \\ &=\; 2\pi\mathrm{i}\,c^2\left[\frac{\mathrm{e}^{-\mathrm{i}c\,|\mathbf{k}|(t-t')}}{2c\,|\mathbf{k}|} - \frac{\mathrm{e}^{\mathrm{i}c\,|\mathbf{k}|(t-t')}}{2c\,|\mathbf{k}|}\right] \\ &=\; 2\pi\,\frac{c}{|\mathbf{k}|}\sin\left[c\,|\mathbf{k}|(t-t')\right] \; . \end{aligned} \tag{5.56}$$

Here in the first line note that the overall minus sign is because the lower blue contour Γ runs clockwise in the plane, so that $-\Gamma$ is positively oriented. Inserting (5.56) for the integral I back into Green's function (5.50) hence gives

$$\begin{aligned} G(\mathbf{r}, t\,; \mathbf{r}', t') \;&=\; \frac{4\pi}{(2\pi)^4}\int_{\mathbf{k}\in\mathbb{R}^3}\left\{2\pi\,\frac{c}{|\mathbf{k}|}\sin\left[c\,|\mathbf{k}|(t-t')\right]\right\}\mathrm{e}^{\mathrm{i}\mathbf{k}\cdot(\mathbf{r}-\mathbf{r}')}\,\mathrm{d}^3k \\ &=\; \frac{c}{2\pi^2}\int_{k=0}^{\infty}\frac{1}{k}\cdot k^2\sin\left[ck(t-t')\right]\int_{\theta=0}^{\pi}\int_{\varphi=0}^{2\pi}\mathrm{e}^{\mathrm{i}k|\mathbf{r}-\mathbf{r}'|\cos\theta}\,\sin\theta\,\mathrm{d}k\,\mathrm{d}\theta\,\mathrm{d}\varphi \; . \end{aligned} \tag{5.57}$$

Here in the second line we have introduced spherical polar coordinates for $\mathbf{k} \in \mathbb{R}^3$, with $k \equiv |\mathbf{k}|$, θ being the angle between the vectors $\mathbf{k}$ and $\mathbf{r} - \mathbf{r}'$, and $\mathrm{d}^3k = k^2\,\sin\theta\,\mathrm{d}k\,\mathrm{d}\theta\,\mathrm{d}\varphi$. Integrating first over the angular variables θ, φ immediately gives

$$\begin{aligned} G(\mathbf{r}, t\,; \mathbf{r}', t') \;&=\; \frac{c}{2\pi^2}\int_{k=0}^{\infty}k\sin\left[ck(t-t')\right]\,2\pi\left[-\frac{1}{\mathrm{i}k|\mathbf{r}-\mathbf{r}'|}\mathrm{e}^{\mathrm{i}k|\mathbf{r}-\mathbf{r}'|\cos\theta}\right]_0^{\pi}\mathrm{d}k \\ &=\; \frac{c}{\pi}\frac{1}{|\mathbf{r}-\mathbf{r}'|}\int_{k=0}^{\infty}\sin\left[ck(t-t')\right]\,2\sin\left(k|\mathbf{r}-\mathbf{r}'|\right)\mathrm{d}k \\ &=\; \frac{c}{\pi}\frac{1}{|\mathbf{r}-\mathbf{r}'|}\int_{k=-\infty}^{\infty}\sin\left[ck(t-t')\right]\,\sin\left(k|\mathbf{r}-\mathbf{r}'|\right)\mathrm{d}k \\ &=\; -\frac{c}{4\pi}\frac{1}{|\mathbf{r}-\mathbf{r}'|}\int_{k=-\infty}^{\infty}\left(\mathrm{e}^{\mathrm{i}ck(t-t')} - \mathrm{e}^{-\mathrm{i}ck(t-t')}\right)\left(\mathrm{e}^{\mathrm{i}k|\mathbf{r}-\mathbf{r}'|} - \mathrm{e}^{-\mathrm{i}k|\mathbf{r}-\mathbf{r}'|}\right)\mathrm{d}k \\ &=\; -\frac{c}{2\pi}\frac{1}{|\mathbf{r}-\mathbf{r}'|}\int_{k=-\infty}^{\infty}\left[\mathrm{e}^{\mathrm{i}ck(t-t'+|\mathbf{r}-\mathbf{r}'|/c)} - \mathrm{e}^{\mathrm{i}ck(t-t'-|\mathbf{r}-\mathbf{r}'|/c)}\right]\mathrm{d}k \; . \end{aligned} \tag{5.58}$$

Here in the third line we have used the fact that the integrand is an even function of k, and in the fourth line we have written sine in terms of exponentials. In the final step the four terms one obtains in multiplying out the brackets in the penultimate line are seen to be pairwise equal on replacing $k \mapsto -k$. Finally, recall that the completeness relation (2.60) allows us to identify $\delta(x-x') = \frac{1}{2\pi}\int_{k=-\infty}^{\infty} \mathrm{e}^{\mathrm{i}k(x-x')}\,\mathrm{d}k$. Changing the integration variable $k \equiv \tilde{k}/c$ in the last line of (5.58), and using the completeness relation to integrate over $\tilde{k}$, gives

$$G(\mathbf{r},t\,;\mathbf{r}',t') \;=\; \frac{1}{|\mathbf{r}-\mathbf{r}'|}\left[-\delta\left(t-t'+\frac{|\mathbf{r}-\mathbf{r}'|}{c}\right)+\delta\left(t-t'-\frac{|\mathbf{r}-\mathbf{r}'|}{c}\right)\right] . \tag{5.59}$$

The first Dirac delta function term is zero, since $|\mathbf{r}-\mathbf{r}'|/c \geq 0$ and notice here we are assuming $t > t'$, so that the argument is strictly positive. We have thus proven:

Proposition 5.4 *The retarded Green's function is*

$$G(\mathbf{r},t\,;\mathbf{r}',t') \;=\; \frac{1}{|\mathbf{r}-\mathbf{r}'|}\,\delta\left(t-t'-\frac{|\mathbf{r}-\mathbf{r}'|}{c}\right) . \tag{5.60}$$

Notice here that when $t < t'$ the argument of the Dirac delta function is strictly negative, so that the Green's function is indeed zero. In fact $G(\mathbf{r},t\,;\mathbf{r}',t') = 0$ unless $c(t-t') = |\mathbf{r}-\mathbf{r}'| \geq 0$, which is the equation for a *sphere* of radius $c(t-t')$, centred on the point $\mathbf{r}'$. Indeed, $t-t' = |\mathbf{r}-\mathbf{r}'|/c$ is the time taken for a disturbance at the point $\mathbf{r}'$ to reach the point $\mathbf{r}$, travelling radially outwards with constant speed c. This solution hence has all the properties we were looking for: it gives a wave that originates from a localized electromagnetic disturbance at position $\mathbf{r}'$ and time t', that expands spherically with speed c from this source.

Having obtained the correct Green's function, using (5.44) we can now simply write down the solution to the original wave equation with source (5.42):

$$\begin{aligned}
\psi(\mathbf{r},t) \;&=\; \int_{\mathbf{r}'\in\mathbb{R}^3}\int_{t'=-\infty}^{\infty} G(\mathbf{r},t\,;\mathbf{r}',t')\,f(\mathbf{r}',t')\,\mathrm{d}V'\,\mathrm{d}t' \;, \\
&=\; \int_{\mathbf{r}'\in\mathbb{R}^3}\int_{t'=-\infty}^{\infty} \frac{1}{|\mathbf{r}-\mathbf{r}'|}\,\delta\left(t-t'-\frac{|\mathbf{r}-\mathbf{r}'|}{c}\right) f(\mathbf{r}',t')\,\mathrm{d}V'\,\mathrm{d}t' \;, \\
&=\; \int_{\mathbf{r}'\in\mathbb{R}^3} \frac{1}{|\mathbf{r}-\mathbf{r}'|}\,f\left(\mathbf{r}',t-\frac{|\mathbf{r}-\mathbf{r}'|}{c}\right)\mathrm{d}V' \;,
\end{aligned} \tag{5.61}$$

where in the last step we have integrated over t'. We may hence also write down the electromagnetic potentials that solve the general time-dependent Maxwell equations (5.34), (5.35):

$$\phi(\mathbf{r},t) \;=\; \frac{1}{4\pi\epsilon_0}\int_{\mathbf{r}'\in\mathbb{R}^3} \frac{\rho\left(\mathbf{r}',t-\frac{|\mathbf{r}-\mathbf{r}'|}{c}\right)}{|\mathbf{r}-\mathbf{r}'|}\,\mathrm{d}V' \;, \tag{5.62}$$

$$\mathbf{A}(\mathbf{r},t) \;=\; \frac{\mu_0}{4\pi}\int_{\mathbf{r}'\in\mathbb{R}^3} \frac{\mathbf{J}\left(\mathbf{r}',t-\frac{|\mathbf{r}-\mathbf{r}'|}{c}\right)}{|\mathbf{r}-\mathbf{r}'|}\,\mathrm{d}V' \;, \tag{5.63}$$

Remark One can verify that these potentials satisfy the Lorenz gauge condition (5.30).

Compare (5.62), (5.63) to the static case equations in (3.36). The only difference is the time dependence, where notice that the fields at time t are determined by integrating the charge density ρ and current density $\mathbf{J}$ at the *retarded time*

$$t' = t_{\text{retarded}} \equiv t - \frac{|\mathbf{r} - \mathbf{r}'|}{c} . \tag{5.64}$$

As already remarked, $|\mathbf{r} - \mathbf{r}'|/c$ is the time it takes for the Green's function wave to propagate from the source point $\mathbf{r}'$ to the observation point $\mathbf{r}$. We shall see in section 6 that c is the speed of light in vacuum, and formulas (5.62), (5.63) say there is no instantaneous action at a distance in electromagnetism, but rather there is a delay between a change in the sources and the resulting change in the fields at time t, given by (5.64). We now have the tools to address the following:

5.6.1 Moving point charge

Consider a point charge q moving on an *arbitrary* trajectory $\mathbf{r} = \mathbf{r}_0(t)$, with velocity $\mathbf{v} = \mathrm{d}\mathbf{r}_0(t)/\mathrm{d}t$. If the charge were stationary, we could use Coulomb's law (1.7) to deduce the electric field $\mathbf{E}$ produced. But Coulomb's law led to the electrostatic Maxwell equations, while in a general time-dependent setting we have Faraday's law $\nabla \times \mathbf{E} = -\partial \mathbf{B}/\partial t$. Indeed, because the charge is moving we expect it to produce a magnetic field $\mathbf{B}$ via the Biot-Savart law (3.10). But the latter led to Ampère's law, which in general has the Maxwell displacement current $\partial \mathbf{E}/\partial t$ on the right hand side of (5.6). In fact neither Coulomb's law nor the Biot-Savart law are correct in this setting, because of the complicated way that the fields feed back into each other in Maxwell's equations. But we have solved the *general* time-dependent equations via (5.62), (5.63), and so may use these to solve this problem.

The charge density and current density of the point charge are by definition

$$\rho(\mathbf{r}, t) = q\,\delta[\mathbf{r} - \mathbf{r}_0(t)] , \qquad \mathbf{J}(\mathbf{r}, t) = q\,\mathbf{v}(t)\,\delta[\mathbf{r} - \mathbf{r}_0(t)] , \quad \text{where } \mathbf{v}(t) = \frac{\mathrm{d}\mathbf{r}_0(t)}{\mathrm{d}t} . \tag{5.65}$$

From (5.62) we thus compute

$$\begin{aligned} \phi(\mathbf{r}, t) &= \frac{q}{4\pi\epsilon_0} \int_{\mathbf{r}' \in \mathbb{R}^3} \int_{t'=-\infty}^{\infty} \frac{1}{|\mathbf{r} - \mathbf{r}'|}\, \delta[\mathbf{r}' - \mathbf{r}_0(t')]\, \delta\left[t' - \left(t - |\mathbf{r} - \mathbf{r}'|/c\right)\right] \mathrm{d}V'\, \mathrm{d}t' \\ &= \frac{q}{4\pi\epsilon_0} \int_{t'=-\infty}^{\infty} \frac{1}{|\mathbf{r} - \mathbf{r}_0(t')|}\, \delta[t' - (t - |\mathbf{r} - \mathbf{r}_0(t')|/c)]\, \mathrm{d}t' . \end{aligned} \tag{5.66}$$

Here in the first line we have imposed the retarded time condition (5.64) by (re)inserting a Dirac delta function and integral over t', and in the second line we have integrated over $\mathbf{r}'$. Notice that the Dirac delta function on the second line is a non-trivial function of t', for which we shall need property (iii) in equation (1.18) of Proposition 1.2.

We next define

$$\mathbf{R}(t) \equiv \mathbf{r} - \mathbf{r}_0(t) , \qquad R(t) \equiv |\mathbf{R}(t)| , \tag{5.67}$$

where $\mathbf{R}(t)$ is the position vector of the observation point $\mathbf{r}$, relative to the position vector $\mathbf{r}_0(t)$ of the point charge at time t, and compute

$$\frac{\mathrm{d}R}{\mathrm{d}t} = \frac{\mathrm{d}}{\mathrm{d}t}\sqrt{(\mathbf{r}-\mathbf{r}_0(t))\cdot(\mathbf{r}-\mathbf{r}_0(t))} = -\frac{1}{R}(\mathbf{r}-\mathbf{r}_0(t))\cdot\frac{\mathrm{d}\mathbf{r}_0(t)}{\mathrm{d}t} = -\frac{\mathbf{R}}{R}\cdot\mathbf{v} \ . \tag{5.68}$$

The Dirac delta function on the second line of (5.66) sets $t' = t - R(t')/c$, where $R(t')$ is a known function of t', as the trajectory $\mathbf{r}_0(t')$ is given, and using property (iii) in Proposition 1.2 and (5.68) we obtain

$$\phi(\mathbf{r},t) = \frac{q}{4\pi\epsilon_0}\frac{1}{R(t')}\frac{1}{1-\frac{1}{c}\frac{\mathbf{R}(t')}{R(t')}\cdot\mathbf{v}(t')}\Bigg|_{t' \text{ such that } t'=t-R(t')/c} \ . \tag{5.69}$$

We can see the usual Coulomb field $q/(4\pi\epsilon_0 R)$ generated by a point charge q, but there is an extra term multiplying this, together with the complication of determining the retarded time t', which is determined *implicitly* from $t' = t - R(t')/c$. Notice that $\mathbf{R}(t')/R(t')$ is a unit vector, while $|\mathbf{v}(t')/c| \ll 1$ for speeds small compared to c, so as long as the particle is moving slowly compared to c, the multiplicative correction term to the Coulomb potential in (5.69) is approximately 1.

One can write down a similar general formula for $\mathbf{A}$ using (5.63): Since $\mathbf{J} = \mathbf{v}\rho$, we just get an extra factor of $\mathbf{v}(t')$, concretely

$$\mathbf{A}(\mathbf{r},t) = \frac{\phi(\mathbf{r},t)}{c^2}\mathbf{v}(t') = \frac{q\mu_0}{4\pi}\frac{\mathbf{v}(t')}{R(t')-\mathbf{R}(t')\cdot\mathbf{v}(t')/c}\Bigg|_{t' \text{ such that } t'=t-R(t')/c} \ . \tag{5.70}$$

The equations (5.69) and (5.70) are known as *Liénard–Wiechert potentials.*

Let us now specialize to the case where the particle moves with *constant* velocity $\mathbf{v} = v\,\mathbf{e}_1$ along the x-axis direction, starting at the origin at time $t = 0$, so that $\mathbf{r}_0(t) = vt\,\mathbf{e}_1$. We have $\mathbf{R}(t) = (x - vt)\,\mathbf{e}_1 + y\,\mathbf{e}_2 + z\,\mathbf{e}_3$ and $R(t) = \sqrt{(x-vt)^2 + y^2 + z^2}$, and the retarded time t' solves

$$\begin{aligned} c^2(t-t')^2 &= R(t')^2 = (x-vt')^2 + y^2 + z^2 \\ \Rightarrow \quad \left(1-\frac{v^2}{c^2}\right)t'^2 - 2t'\left(t-\frac{v}{c^2}x\right) + t^2 - \frac{x^2+y^2+z^2}{c^2} &= 0 \ . \end{aligned} \tag{5.71}$$

Introducing

$$\gamma = \gamma(v) \equiv \frac{1}{\sqrt{1-v^2/c^2}} \ , \tag{5.72}$$

the quadratic equation (5.71) in t' has solution

$$t' = \gamma^2\left[t - \frac{v}{c^2}x - \frac{1}{\gamma c}\sqrt{\gamma^2(x-vt)^2+y^2+z^2}\right] \ . \tag{5.73}$$

Here we have chosen the root with $t - t' > 0$. Those who have studied Special Relativity will recognize the Lorentz factor γ making a remarkable appearance, as well as other features of Lorentz transformations! Examining the potential $\phi(\mathbf{r},t)$ in (5.69), some algebra gives

$$\begin{aligned} R(t') - \frac{1}{c}\mathbf{R}(t')\cdot\mathbf{v}(t') &= \sqrt{(x-vt')+y^2+z^2} - \frac{v}{c}(x-vt') = c(t-t') - \frac{v}{c}(x-vt') \\ &= c\left(t - \frac{1}{\gamma^2}t' - \frac{v}{c^2}x\right) = \frac{1}{\gamma}\sqrt{\gamma^2(x-vt)^2+y^2+z^2} \ , \end{aligned} \tag{5.74}$$

where the last step uses (5.73). The potential $\phi(\mathbf{r}, t)$ in (5.69) thus simplifies to

$$\phi(\mathbf{r}, t) = \frac{q\gamma}{4\pi\epsilon_0} \frac{1}{\sqrt{\gamma^2(x - vt)^2 + y^2 + z^2}} . \tag{5.75}$$

This is the Coulomb potential one might naively guess is generated by a point charge q at position $\mathbf{r}_0(t) = vt\,\mathbf{e}_1$, up to the relativistic factors of γ. The vector potential in this case is given by the similar formula

$$\mathbf{A}(\mathbf{r}, t) = \frac{q\gamma}{4\pi\epsilon_0 c^2} \frac{1}{\sqrt{\gamma^2(x - vt)^2 + y^2 + z^2}}\,\mathbf{v} = \frac{\phi(\mathbf{r}, t)}{c^2}\,\mathbf{v} , \tag{5.76}$$

where from (5.28) note $1/(\epsilon_0\, c^2) = \mu_0$.

5.7 Maxwell's equations in macroscopic media

In section 4 we derived the electrostatic Maxwell equations (4.12) and magnetostatic Maxwell equations (4.20) in macroscopic media. One can generalize this discussion to include time dependence, in essentially the same way as we have done for the vacuum Maxwell equations in this section. Rather than repeat the arguments, we here simply present the the final complete set of *macroscopic Maxwell equations* (or *Maxwell equations in matter*):

$$\boxed{\begin{array}{cc} \nabla \cdot (\epsilon\, \mathbf{E}) = \rho_{\text{free}}, & \nabla \times \mathbf{E} = -\dfrac{\partial \mathbf{B}}{\partial t}, \\ \nabla \cdot \mathbf{B} = 0, & \nabla \times \left(\dfrac{1}{\mu}\mathbf{B}\right) = \mathbf{J}_{\text{free}} + \dfrac{\partial(\epsilon\, \mathbf{E})}{\partial t}. \end{array}} \tag{5.77}$$

These are similar in form to the equations originally introduced and studied by Maxwell. Although we have motivated them as an *effective* set of equations, approximately valid in certain media where ϵ, μ characterize the media's electromagnetic properties, notice that setting $\epsilon = \epsilon_0$, $\mu = \mu_0$ gives the *microscopic Maxwell equations* on the front page, which are regarded as fundamental, and exact.

6 Electromagnetic waves

6.1 Source-free equations and electromagnetic waves

We begin by writing down Maxwell's equations in *vacuum*, with no electric charge or current:

$$\nabla \cdot \mathbf{E} = 0 \,, \qquad \nabla \cdot \mathbf{B} = 0 \,, \tag{6.1}$$

$$\nabla \times \mathbf{E} + \frac{\partial \mathbf{B}}{\partial t} = \mathbf{0} \,, \qquad \nabla \times \mathbf{B} - \frac{1}{c^2}\frac{\partial \mathbf{E}}{\partial t} = \mathbf{0} \,, \tag{6.2}$$

where as in (5.28) we have defined

$$c \equiv \sqrt{\frac{1}{\epsilon_0 \mu_0}} \,. \tag{6.3}$$

We have already seen that c is a *speed*; for example, the speed of the spherical wavefront that propagates out from the source point in the Green's function (5.60). The great insight of Maxwell was to realise that this is the *speed of light* in vacuum.

Taking the curl of the first equation in (6.2) we have

$$\begin{aligned} \mathbf{0} &= \nabla(\nabla \cdot \mathbf{E}) - \nabla^2 \mathbf{E} + \nabla \times \frac{\partial \mathbf{B}}{\partial t} = -\nabla^2 \mathbf{E} + \frac{\partial}{\partial t}(\nabla \times \mathbf{B}) \\ &= -\nabla^2 \mathbf{E} + \frac{1}{c^2}\frac{\partial^2 \mathbf{E}}{\partial t^2} \,. \end{aligned} \tag{6.4}$$

Here after the first step we have used the first equation in (6.1), and in the last step we have used the second equation in (6.2). It follows that each component of $\mathbf{E}$ satisfies the *wave equation*

$$\Box\, u = 0 \,, \tag{6.5}$$

where $u = u(\mathbf{r}, t)$, and as in (5.36) the *d'Alembertian operator* is $\Box \equiv \frac{1}{c^2}\frac{\partial^2}{\partial t^2} - \nabla^2$. You can similarly check that $\mathbf{B}$ also satisfies $\Box\, \mathbf{B} = \mathbf{0}$.

The equation (6.5) governs the propagation of waves of speed c in three-dimensional space. It is the natural generalization of the one-dimensional wave equation

$$\frac{1}{c^2}\frac{\partial^2 u}{\partial t^2} - \frac{\partial^2 u}{\partial x^2} = 0 \,, \tag{6.6}$$

which you met in Prelims. Recall that this has particular solutions of the form $u_\pm(x,t) = f(x \mp ct)$, where f is any function which is twice differentiable. In this case, the waves look like the graph of f travelling at constant speed c in the direction of increasing/decreasing x, respectively. The *general* (d'Alembert) solution to (6.6) is $u(x,t) = f(x - ct) + g(x + ct)$, as shown in the Prelims course on Fourier Series and PDEs .

The above generalizes naturally to the three-dimensional equation (6.5), by writing

$$u(\mathbf{r}, t) = f(\mathbf{e} \cdot \mathbf{r} - ct) \,, \tag{6.7}$$

where $\mathbf{e}$ is a fixed *unit vector*, $|\mathbf{e}|^2 = 1$. Indeed, using the chain rule we compute $\nabla^2 u = \mathbf{e} \cdot \mathbf{e}\, f''$, $\partial^2 u/\partial t^2 = c^2 f''$, so that (6.7) solves (6.5) for any twice differentiable function f of one variable.

Definition Solutions to the wave equation (6.5) of the form (6.7) are called *plane-fronted waves*.

The terminology here is justified by noting that at any constant time, u is constant on the planes $\{\mathbf{e}\cdot\mathbf{r} = \text{constant}\}$ orthogonal to $\mathbf{e}$. As time t increases, these plane wavefronts propagate in the direction of $\mathbf{e}$ at speed c. However, unlike the one-dimensional equation, we *cannot* write the general solution to (6.5) as a sum of two plane-fronted waves travelling in opposite directions.

6.2 Monochromatic plane waves

An important special class of plane-fronted waves (6.7) are given by the *complex harmonic waves*

$$u(\mathbf{r},t) = \alpha\,\mathrm{e}^{\mathrm{i}(\mathbf{k}\cdot\mathbf{r}-\omega t)} , \tag{6.8}$$

where α is a complex constant, $\omega > 0$ is the constant *frequency* of the wave, and $\mathbf{k}$ is the constant *wave vector*. To relate to (6.7), note that

$$\mathbf{k}\cdot\mathbf{r} - \omega t = \frac{\omega}{c}\left(\mathbf{e}\cdot\mathbf{r} - ct\right) , \tag{6.9}$$

provided we identify

$$\mathbf{k} = k\,\mathbf{e} , \quad \text{where} \quad k = \frac{\omega}{c} . \tag{6.10}$$

Here $k \equiv |\mathbf{k}|$ is called the *wave number*. Thus (6.8) solves the wave equation (6.5) provided (6.10) holds, which is simply the equation

$$\text{speed } c = \frac{\omega}{k} = \frac{\omega}{2\pi}\cdot\frac{2\pi}{k} = \text{frequency} \times \text{wavelength} . \tag{6.11}$$

The harmonic waves (6.8) are of course complex, although notice that the real and imaginary parts separately solve the wave equation. One can then take linear combinations of these real sine and cosine solutions. Note also that the complex harmonic wave (6.8) is simply the product of exponential Fourier modes (2.56) in each variable x, y, z, t. In fact it is a result of Fourier analysis that *every* solution to the wave equation (6.5) is a linear combination (in general involving an integral) of these harmonic waves, as (6.8) form a complete set of orthonormal functions, (5.46).

Since the components of $\mathbf{E}$ and $\mathbf{B}$ satisfy (6.5), it is natural to look for solutions of the *complex* harmonic wave form

$$\mathbf{E}_{\mathbb{C}}(\mathbf{r},t) \equiv \mathbf{E}_0\,\mathrm{e}^{\mathrm{i}(\mathbf{k}\cdot\mathbf{r}-\omega t)} , \qquad \mathbf{B}_{\mathbb{C}}(\mathbf{r},t) \equiv \mathbf{B}_0\,\mathrm{e}^{\mathrm{i}(\mathbf{k}\cdot\mathbf{r}-\omega t)} , \tag{6.12}$$

where $\mathbf{E}_0$ and $\mathbf{B}_0$ are constant complex vectors. Here and in the following we understand these expressions to mean that we take the *real* part of the complex exponential to obtain the real electromagnetic field, so $\mathbf{E} \equiv \mathrm{Re}\,(\mathbf{E}_{\mathbb{C}})$, $\mathbf{B} \equiv \mathrm{Re}\,(\mathbf{B}_{\mathbb{C}})$. The expressions (6.12) of course satisfy the wave equation, but we must ensure that we satisfy *all* of the Maxwell equations in vacuum (6.1), (6.2). Since $\nabla\cdot\mathbf{E}_{\mathbb{C}} = \mathbf{E}_0\cdot\nabla\,\mathrm{e}^{\mathrm{i}(\mathbf{k}\cdot\mathbf{r}-\omega t)} = \mathrm{i}\,\mathbf{k}\cdot\mathbf{E}_0\,\mathrm{e}^{\mathrm{i}(\mathbf{k}\cdot\mathbf{r}-\omega t)}$, the two equations in (6.1) immediately give

$$\mathbf{k}\cdot\mathbf{E}_0 = 0 = \mathbf{k}\cdot\mathbf{B}_0 . \tag{6.13}$$

The first equation in (6.2) reads

$$\mathbf{0} \;=\; \nabla\times\mathbf{E}_{\mathbb{C}} + \frac{\partial\mathbf{B}_{\mathbb{C}}}{\partial t} \;=\; (\mathrm{i}\mathbf{k}\times\mathbf{E}_0 - \mathrm{i}\omega\,\mathbf{B}_0)\,\mathrm{e}^{\mathrm{i}(\mathbf{k}\cdot\mathbf{r}-\omega t)}\;, \tag{6.14}$$

allowing us to read off

$$\mathbf{B}_0 \;=\; \frac{1}{\omega}\,\mathbf{k}\times\mathbf{E}_0 \;=\; \frac{1}{c}\,\mathbf{e}\times\mathbf{E}_0\;, \tag{6.15}$$

where in the second equality we have used (6.10). One can then verify that the second equation in (6.2) is automatically satisfied, using $\mathbf{k}\times\left(\frac{1}{\omega}\,\mathbf{k}\times\mathbf{E}_0\right)+\frac{\omega}{c^2}\,\mathbf{E}_0 = \frac{1}{\omega}\,(\mathbf{k}\cdot\mathbf{E}_0)\,\mathbf{k}-\frac{k^2}{\omega}\,\mathbf{E}_0+\frac{\omega}{c^2}\,\mathbf{E}_0=\mathbf{0}$, where the first equality uses the vector triple product (A.6), and in the second we have used (6.13) and (6.10). To summarize, we have shown

Proposition 6.1 *The* monochromatic electromagnetic plane wave*, given by*

$$\mathbf{E}_{\mathbb{C}}(\mathbf{r},t) \;=\; \mathbf{E}_0\,\mathrm{e}^{\mathrm{i}(\mathbf{k}\cdot\mathbf{r}-\omega t)}\;, \qquad \mathbf{B}_{\mathbb{C}}(\mathbf{r},t) \;=\; \frac{1}{\omega}\,\mathbf{k}\times\mathbf{E}_{\mathbb{C}}(\mathbf{r},t)\;, \tag{6.16}$$

solves the vacuum Maxwell equations (6.1)*,* (6.2)*, provided* $|\mathbf{k}|=\omega/c$*, and* $\mathbf{k}\cdot\mathbf{E}_0=0$*.*

Notice that the solution is specified by the angular frequency ω, the direction of propagation $\mathbf{e}\equiv\mathbf{k}/|\mathbf{k}|$, and with the electric field direction specified by a constant vector $\mathbf{E}_0$ that is orthogonal to the direction of propagation. In fact both $\mathbf{E}$ and $\mathbf{B}$ are orthogonal to the direction of propagation (6.13), which is known as a *transverse wave*. One can contrast this with, *e.g.* sound waves, which are *longitudinal* waves, with the wave motion aligned with the direction of propagation. $\mathbf{E}$ and $\mathbf{B}$ are also orthogonal to each other. Fourier analysis implies that the general vacuum solution is a combination of these monochromatic plane waves.

6.3 Polarization

So far we have written complex solutions to the vacuum Maxwell equations, but as mentioned the actual electric and magnetic fields are given by the real (or imaginary) parts of the monochromatic plane waves in (6.16). Focusing on the electric field, we thus have

$$\mathbf{E}(\mathbf{r},t) \;=\; \mathrm{Re}\,(\mathbf{E}_{\mathbb{C}}(\mathbf{r},t)) \;=\; \boldsymbol{\alpha}\,\cos(\mathbf{k}\cdot\mathbf{r}-\omega t) - \boldsymbol{\beta}\,\sin(\mathbf{k}\cdot\mathbf{r}-\omega t)\;, \tag{6.17}$$

where we have defined $\boldsymbol{\alpha}\equiv\mathrm{Re}\,\mathbf{E}_0$, $\boldsymbol{\beta}\equiv\mathrm{Im}\,\mathbf{E}_0$, which are real vectors, orthogonal to the direction of propagation $\mathbf{k}$. The magnetic field is

$$\mathbf{B}(\mathbf{r},t) \;=\; \frac{1}{\omega}\,\mathbf{k}\times\mathbf{E}(\mathbf{r},t) \;=\; \frac{1}{\omega}\,\mathbf{k}\times\boldsymbol{\alpha}\,\cos(\mathbf{k}\cdot\mathbf{r}-\omega t) - \frac{1}{\omega}\,\mathbf{k}\times\boldsymbol{\beta}\,\sin(\mathbf{k}\cdot\mathbf{r}-\omega t)\;. \tag{6.18}$$

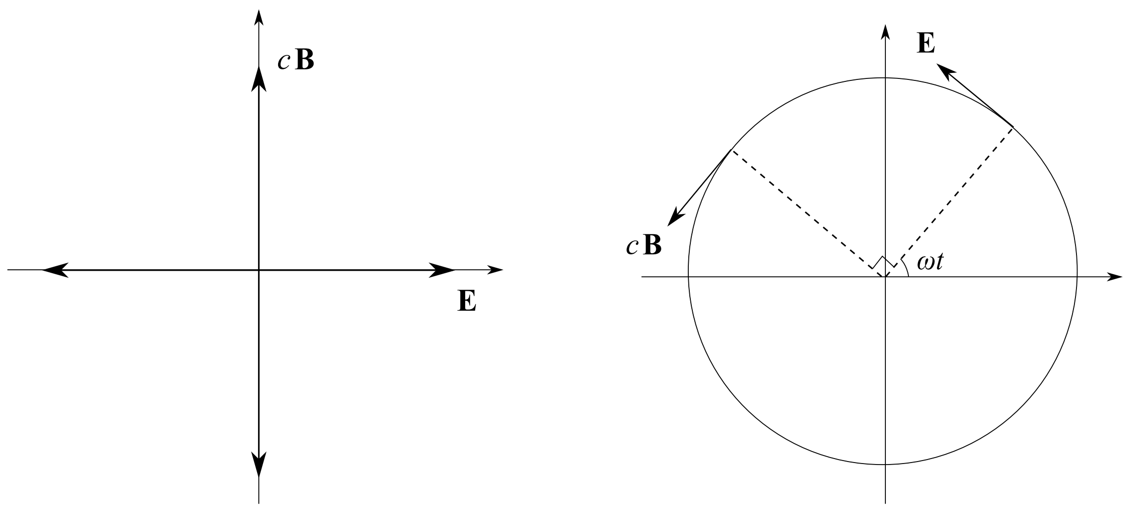
If we fix a particular point in space, say the origin $\mathbf{r}=\mathbf{0}$, then the electric field (6.17) is

$$\mathbf{E}(0,t) \;=\; \boldsymbol{\alpha}\,\cos\omega t \,+\, \boldsymbol{\beta}\,\sin\omega t\;. \tag{6.19}$$

As t varies, this sweeps out an *ellipse* in the plane spanned by $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ (similar remarks apply to the $\mathbf{B}$-field in (6.18)). As a simple example, taking $\boldsymbol{\alpha}=\alpha\,\mathbf{e}_1$, $\boldsymbol{\beta}=\beta\,\mathbf{e}_2$ where $\mathbf{e}_1$, $\mathbf{e}_2$ are orthonormal vectors, so that $\mathbf{e} \equiv \mathbf{k}/|\mathbf{k}| = \mathbf{e}_3$ is the direction of propagation, then the 1 and 2 components of $\mathbf{E}$ in (6.19) are $E_1 = \alpha \cos \omega t$, $E_2 = \beta \sin \omega t$, and thus

$$\frac{E_1^2}{\alpha^2} + \frac{E_2^2}{\beta^2} = 1 \; . \tag{6.20}$$

This is an ellipse, with semi-major(minor) axis length α, semi-minor(major) axis length β, centred on the origin.



(a) Linear polarization.

(b) Circular polarization (right-handed).

Figure 21: Polarizations of monochromatic electromagnetic plane waves, viewed from the direction of propagation.

There are two special choices of $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$, which have names:

(i) **Linear polarization**: If $\boldsymbol{\alpha}$ is proportional to $\boldsymbol{\beta}$, so that the ellipse degenerates to a *line*, then the monochromatic plane wave is said to be *linearly polarized*. In this case, $\mathbf{E}$ and $\mathbf{B}$ oscillate in two fixed orthogonal directions – see Figure 21a. For example, taking $\boldsymbol{\beta} = \mathbf{0}$ the electric and magnetic fields at the origin are

$$\mathbf{E} = \alpha \, \mathbf{e}_1 \cos \omega t \; , \qquad \mathbf{B} = \frac{1}{c} \alpha \, \mathbf{e}_2 \cos \omega t \; . \tag{6.21}$$

(ii) **Circular polarization**: If $\boldsymbol{\alpha} \cdot \boldsymbol{\beta} = 0$ (as in the example (6.20) above) and also $|\boldsymbol{\alpha}| = |\boldsymbol{\beta}|$, so that the ellipse is a *circle*, then the monochromatic plane wave is said to be *circularly polarized*. In this case, $\mathbf{E}$ and $\mathbf{B}$ rotate at constant angular velocity about the direction of propagation – see Figure 21b. A circularly polarized wave is said to be *right-handed* or *left-handed*, depending on whether $|\boldsymbol{\alpha}|^2 \, \mathbf{e} = \boldsymbol{\alpha} \times \boldsymbol{\beta}$ or $|\boldsymbol{\alpha}|^2 \, \mathbf{e} = -\boldsymbol{\alpha} \times \boldsymbol{\beta}$, respectively. With your thumb aligned with the direction of propagation $\mathbf{e}$, the direction of the electric and magnetic fields is given by the curl of your fingers on your right and left hands, respectively.

In general, electromagnetic waves are combinations of waves with different polarizations.

Finally, notice the formula (6.15) gives

$$\mathbf{B} \;=\; \frac{1}{c}\,\mathbf{e}\times\mathbf{E} \quad\Longrightarrow\quad c|\mathbf{B}| \;=\; |\mathbf{E}|\ , \tag{6.22}$$

and thus the *energy density* (5.37) of a monochromatic plane wave is

$$\mathcal{E} \;\equiv\; \frac{\epsilon_0}{2}\left(|\mathbf{E}|^2 + c^2|\mathbf{B}|^2\right) \;=\; \epsilon_0\,|\mathbf{E}|^2\ . \tag{6.23}$$

The Poynting vector (5.39) is

$$\boldsymbol{\mathcal{P}} \;\equiv\; \frac{1}{\mu_0}\,\mathbf{E}\times\mathbf{B} \;=\; \frac{1}{\mu_0}\mathbf{E}\times\left(\frac{1}{c}\,\mathbf{e}\times\mathbf{E}\right) \;=\; \frac{1}{\mu_0 c}|\mathbf{E}|^2\,\mathbf{e} \;=\; c\,\mathcal{E}\,\mathbf{e}\ , \tag{6.24}$$

which is in the direction of propagation of the wave, with magnitude $|\boldsymbol{\mathcal{P}}| = c\,\mathcal{E}$. Thus electromagnetic waves *carry energy*, a fact which anyone who has made a mircowave pot noodle can confirm.

6.4 Reflection and refraction

Maxwell postulated that the electromagnetic waves we have been discussing describe *light*. If that's the case, then all observed (classical) properties of light must be a consequence of Maxwell's equations. In this section we prove the *laws of optics*. Here a ray of light that hits the boundary between two transparent materials, such as air and water, is divided into a *reflected ray* and a *refracted* ray. These obey (see Figure 22):

(i) **Law of reflection**: the reflected ray lies in the plane of incidence, with the angle of incidence θ equal to the angle of reflection θ''.

(ii) **Law of refraction**: the refracted ray lies in the plane of incidence, with the angle of incidence θ and angle of refraction θ' related by *Snell's law*

$$n\,\sin\theta \;=\; n'\,\sin\theta'\ , \tag{6.25}$$

where the materials have a *refractive index* n, n', respectively.

For example, $n_{\mathrm{air}} \simeq 1$, while $n_{\mathrm{water}} \simeq 1.3$. In fact we will show that

$$n \;=\; \sqrt{\frac{\epsilon\mu}{\epsilon_0\mu_0}}\ , \tag{6.26}$$

where ϵ, μ are the permittivity and permeability of the medium. Snell's law (6.25) rearranges to

$$\sin\theta' \;=\; \frac{n}{n'}\,\sin\theta\ , \tag{6.27}$$

so that when moving from a material with a lower n to a higher n' (such as from air to water), the angle of refraction θ' is *smaller* than the angle of incidence θ, causing the light to bend towards the normal direction to the boundary (as in Figure 22). On the other hand, moving from a material with a higher n to a lower n' (such as from water to air), there is a *critical angle of incidence*

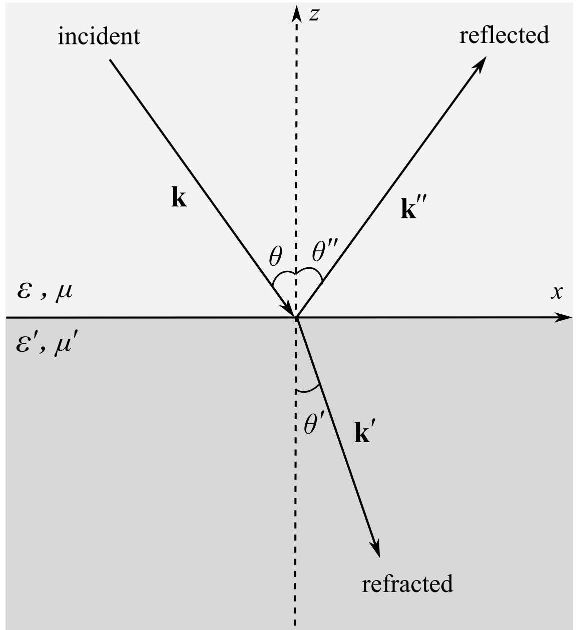


Figure 22: A light ray striking the boundary between two materials. The incident light ray, refracted light ray and reflected light ray travel in the directions $\mathbf{k}$, $\mathbf{k}'$, $\mathbf{k}''$, respectively, with angle of incidence θ, angle of refraction θ', and angle of reflection θ''.

$\theta_c \equiv \arcsin{(n'/n)}$, where for $\theta > \theta_c$ the right hand side of equation (6.27) is larger than 1, and hence there is no solution for θ'. There is then no refracted light ray, a phenomenon called *total internal reflection*.

To derive these laws from Maxwell's equations, we model the incident light ray as a monochromatic plane wave, travelling in the direction $\mathbf{k}$, with electric field

$$\mathbf{E}_{\text{incident}} \;=\; \mathbf{E}_0\, \mathrm{e}^{\mathrm{i}(\mathbf{k}\cdot\mathbf{r}-\omega t)}\;. \tag{6.28}$$

The corresponding magnetic field is $\mathbf{B} = \frac{1}{\omega}\,\mathbf{k}\times\mathbf{E}$, where it is convenient to work with complex plane waves and drop the subscript $\mathbb{C}$ from $\mathbf{E}_{\mathbb{C}}$ and $\mathbf{B}_{\mathbb{C}}$. Without loss of generality, we take the plane that divides the two materials to be the (x,y)-plane $\{z=0\}$, and take

$$\mathbf{e} \;\equiv\; \frac{\mathbf{k}}{|\mathbf{k}|} \;=\; (\,\sin\theta\,,\,0\,,\,-\cos\theta\,)\;, \tag{6.29}$$

so that the *plane of incidence*, spanned by the normal $\mathbf{e}_3$ to $\{z=0\}$ and incident direction $\mathbf{k}$, is the (x,z)-plane $\{y=0\}$, as in Figure 22. Notice then that the electric field (6.28) is independent of the y coordinate. Necessarily $\mathbf{k}\cdot\mathbf{E}_0 = 0$, and we take

$$\mathbf{E}_0 \;=\; E_0\,\mathbf{e}_2\;, \tag{6.30}$$

so that $\mathbf{E}_{\text{incident}}$ points into the page in Figure 22.

Similarly including the reflected and refracted waves, we may write

$$\mathbf{E} \;=\; \begin{cases} \mathbf{E}_{\text{incident}}+\mathbf{E}_{\text{reflected}} \;=\; \mathbf{E}_0\,\mathrm{e}^{\mathrm{i}(\mathbf{k}\cdot\mathbf{r}-\omega t)}+\mathbf{E}_0''\,\mathrm{e}^{\mathrm{i}(\mathbf{k}''\cdot\mathbf{r}-\omega'' t)}\;, & z>0\;, \\ \mathbf{E}_{\text{refracted}} \;=\; \mathbf{E}_0'\,\mathrm{e}^{\mathrm{i}(\mathbf{k}'\cdot\mathbf{r}-\omega' t)}\;, & z<0\;. \end{cases} \tag{6.31}$$

The Maxwell equations in a general macroscopic medium are (5.77), and simply involve replacing $\epsilon_0 \to \epsilon$, $\mu_0 \to \mu$. It follows from (6.3) that in a general macroscopic medium the wave speed propagation is not the speed of light c in vacuum, but rather

$$v = \sqrt{\frac{1}{\epsilon\mu}} = \sqrt{\frac{\epsilon_0\mu_0}{\epsilon\mu}}\, c\,, \qquad v' = \sqrt{\frac{1}{\epsilon'\mu'}} = \sqrt{\frac{\epsilon_0\mu_0}{\epsilon'\mu'}}\, c\,. \tag{6.32}$$

Thus (6.10) becomes

$$|\mathbf{k}| = \frac{\omega}{v}\,, \qquad |\mathbf{k}''| = \frac{\omega''}{v}\,, \qquad |\mathbf{k}'| = \frac{\omega'}{v'}\,. \tag{6.33}$$

Given the incident ray, it remains to then determine the directions of $\mathbf{k}'$, $\mathbf{k}''$, which will give the laws of refraction and reflection, and also ω', ω'', $\mathbf{E}_0'$, $\mathbf{E}_0''$. This comes from analysing the *boundary conditions* at $\{z=0\}$.

Proposition 4.1 and Proposition 4.2 say[17]:

(i) The components of $\mathbf{E}$ and $\frac{1}{\mu}\mathbf{B}$ *tangent* to $\{z=0\}$ are continuous across this surface.

(ii) The components of $\epsilon\,\mathbf{E}$ and $\mathbf{B}$ *normal* to $\{z=0\}$ are continuous across this surface.

Focusing first on the electric field, the two tangent directions are $\mathbf{e}_1$ and $\mathbf{e}_2$, so we may write down

$$\begin{aligned}
(\mathbf{E}_{\text{incident}} + \mathbf{E}_{\text{reflected}})\cdot\mathbf{e}_i\Big|_{z=0} &= \mathbf{E}_{\text{refracted}}\cdot\mathbf{e}_i\Big|_{z=0}\,, \qquad i = 1,2\,,\\
\epsilon\,(\mathbf{E}_{\text{incident}} + \mathbf{E}_{\text{reflected}})\cdot\mathbf{e}_3\Big|_{z=0} &= \epsilon'\,\mathbf{E}_{\text{refracted}}\cdot\mathbf{e}_3\Big|_{z=0}\,.
\end{aligned} \tag{6.34}$$

More explicitly, these read

$$\begin{aligned}
(\mathbf{E}_0''\cdot\mathbf{e}_1)\,\mathrm{e}^{\mathrm{i}(\mathbf{k}''\cdot(x,y,0)-\omega'' t)} &= (\mathbf{E}_0'\cdot\mathbf{e}_1)\,\mathrm{e}^{\mathrm{i}(\mathbf{k}'\cdot(x,y,0)-\omega' t)}\,,\\
E_0\,\mathrm{e}^{\mathrm{i}(|\mathbf{k}|\,x\sin\theta-\omega t)} + (\mathbf{E}_0''\cdot\mathbf{e}_2)\,\mathrm{e}^{\mathrm{i}(\mathbf{k}''\cdot(x,y,0)-\omega'' t)} &= (\mathbf{E}_0'\cdot\mathbf{e}_2)\,\mathrm{e}^{\mathrm{i}(\mathbf{k}'\cdot(x,y,0)-\omega' t)}\,,\\
\epsilon\,(\mathbf{E}_0''\cdot\mathbf{e}_3)\,\mathrm{e}^{\mathrm{i}(\mathbf{k}''\cdot(x,y,0)-\omega'' t)} &= \epsilon'\,(\mathbf{E}_0'\cdot\mathbf{e}_3)\,\mathrm{e}^{\mathrm{i}(\mathbf{k}'\cdot(x,y,0)-\omega' t)}\,,
\end{aligned} \tag{6.35}$$

where we have used the form of $\mathbf{E}_0$ and $\mathbf{k}$ in (6.30) and (6.29), respectively. These equations hold for *all* x, y and t, which are clearly quite strong conditions! In particular, notice that the exponential Fourier modes $\mathrm{e}^{\mathrm{i}kx}$ are *linearly independent* functions of x, for different k. So for example the functional equation $\alpha\,\mathrm{e}^{\mathrm{i}kx} + \gamma = \beta\,\mathrm{e}^{\mathrm{i}k'x}$, with coefficients $\alpha, \beta, \gamma \neq 0$, implies $k = k' = 0$, $\alpha + \gamma = \beta$.

Unpacking (6.35) is a little fiddly. We look first at the middle equation. Notice we cannot have both $(\mathbf{E}_0''\cdot\mathbf{e}_2)$ and $(\mathbf{E}_0'\cdot\mathbf{e}_2)$ equal to zero, otherwise $E_0 = 0$ and there is no incident ray. Suppose these coefficients are *both non-zero*.[18] Then from the remark above about linear independence, looking at the t dependence of the middle equation in (6.35) immediately gives

$$\omega = \omega' = \omega''\,, \tag{6.36}$$

[17]The alert reader will notice these were derived for *statics*. However, one can verify that the additional time-dependent terms in the surface integral over the rectangular surface Σ do not contribute on taking $\varepsilon \to 0$.

[18]If instead say $(\mathbf{E}_0''\cdot\mathbf{e}_2) \neq 0$ but $(\mathbf{E}_0'\cdot\mathbf{e}_2) = 0$, then if there is a refracted ray one of $(\mathbf{E}_0'\cdot\mathbf{e}_1)$ or $(\mathbf{E}_0'\cdot\mathbf{e}_3)$ must be non-zero. Then just add a multiple of the first or last equation in (6.35) to the middle equation, respectively, and the reasoning below proceeds in the same way.

so that reflected and refracted frequencies are the same as the incident frequency. Since the term proportional to E_0 is independent of y, and non-zero, looking at the y dependence implies that the other two terms in the middle equation of (6.35) are also independent of y. That is,

$$\mathbf{e}_2 \cdot \mathbf{k}' \;=\; \mathbf{e}_2 \cdot \mathbf{k}'' \;=\; 0\ . \tag{6.37}$$

But this says that the reflected and refracted rays lie in the plane of incidence, namely the (x,z)-plane. As in Figure 22, we may hence write

$$\frac{\mathbf{k}'}{|\mathbf{k}'|} \;=\; (\,\sin\theta'\,,\,0\,,\,-\cos\theta'\,)\ , \qquad \frac{\mathbf{k}''}{|\mathbf{k}''|} \;=\; (\,\sin\theta''\,,\,0\,,\,\cos\theta''\,)\ . \tag{6.38}$$

The x dependence in the middle equation of (6.35) then gives

$$|\mathbf{k}|\,\sin\theta \;=\; |\mathbf{k}'|\,\sin\theta' \;=\; |\mathbf{k}''|\,\sin\theta''\ . \tag{6.39}$$

But substituting (6.36) into (6.33), this implies $|\mathbf{k}| = |\mathbf{k}''|$ and $v|\mathbf{k}| = v'|\mathbf{k}'|$ and hence

$$\theta \;=\; \theta''\ , \qquad \frac{1}{v}\,\sin\theta \;=\; \frac{1}{v'}\,\sin\theta'\ . \tag{6.40}$$

We have thus proven the law of reflection, and the law of refraction (6.25), with the dimensionless refractive index given by (6.26), which using (6.32) may also be written as $n = c/v$.

To finish solving the problem we must also impose the boundary conditions on the **B**-field at $\{z = 0\}$, although in the following we just outline the steps. These boundary conditions read

$$\begin{aligned} \frac{1}{\mu}\,\left(\mathbf{k}\times\mathbf{E}_0 + \mathbf{k}''\times\mathbf{E}_0''\right)\cdot\mathbf{e}_i \;&=\; \frac{1}{\mu'}\,\left(\mathbf{k}'\times\mathbf{E}_0'\right)\cdot\mathbf{e}_i\ , \qquad i \;=\; 1,2\ , \\ \left(\mathbf{k}\times\mathbf{E}_0 + \mathbf{k}''\times\mathbf{E}_0''\right)\cdot\mathbf{e}_3 \;&=\; \left(\mathbf{k}'\times\mathbf{E}_0'\right)\cdot\mathbf{e}_3\ . \end{aligned} \tag{6.41}$$

One can verify that (6.35), (6.41) are solved by taking

$$\mathbf{E}_0' \;=\; E_0'\,\mathbf{e}_2\ , \qquad \mathbf{E}_0'' \;=\; E_0''\,\mathbf{e}_2\ , \tag{6.42}$$

so that the incident, reflected and refracted electromagnetic fields all have the same polarization (linear, in the y-axis direction), where equation (6.35) then imposes only

$$E_0 + E_0'' \;=\; E_0'\ , \tag{6.43}$$

while after a little work one checks that (6.41) imposes only

$$\sqrt{\frac{\epsilon}{\mu}}\,\left(E_0 - E_0''\right)\cos\theta \;=\; \sqrt{\frac{\epsilon'}{\mu'}}\,E_0'\,\cos\theta'\ , \tag{6.44}$$

where we have used (6.32), (6.33). Assuming $\mu = \mu'$, which is approximately the case for air and water, one can solve (6.43), (6.44) using (6.40) to find

$$E_0' \;=\; \frac{2\cos\theta\,\sin\theta'}{\sin(\theta+\theta')}\,E_0\ , \qquad E_0'' \;=\; -\frac{\sin(\theta-\theta')}{\sin(\theta+\theta')}\,E_0\ . \tag{6.45}$$

We have thus determined completely the reflected and refracted waves, in terms of the incident wave, in the case when the wave is polarized orthogonal to the incidence plane.

For polarization parallel to the incidence plane,

$$\begin{aligned} &\mathbf{E}_0 = E_0 \begin{pmatrix} \cos\theta \\ 0 \\ \sin\theta \end{pmatrix} \quad \mathbf{E}'_0 = E'_0 \begin{pmatrix} \cos\theta' \\ 0 \\ \sin\theta' \end{pmatrix} \quad \mathbf{E}''_0 = E''_0 \begin{pmatrix} -\cos\theta \\ 0 \\ \sin\theta \end{pmatrix} \\ &\mathbf{B}_0 = \frac{kE_0}{\omega} \begin{pmatrix} 0 \\ -1 \\ 0 \end{pmatrix} \quad \mathbf{B}'_0 = \frac{k'E'_0}{\omega} \begin{pmatrix} 0 \\ -1 \\ 0 \end{pmatrix} \quad \mathbf{B}''_0 = \frac{kE''_0}{\omega} \begin{pmatrix} 0 \\ -1 \\ 0 \end{pmatrix} \end{aligned} \tag{6.46}$$

solving the boundary conditions similarly as above leads to the solution (for $\mu = \mu'$)

$$E'_0 = 2E_0 \frac{\sin\theta' \cos\theta}{\sin(\theta+\theta')\cos(\theta-\theta')}, \qquad E''_0 = 2E_0 \frac{\sin(\theta-\theta')\cos(\theta+\theta')}{\sin(\theta+\theta')\cos(\theta-\theta')}. \tag{6.47}$$

Note that when $\theta + \theta' = \pi/2$, then $E''_0 = 0$, that is there is no reflection! This angle θ is called *Brewster angle*. It can be used to produce linearly polarized light: shine light with arbitrary polarizations under the Brewster angle θ on the material; then the reflected beam is linearly polarized (orthogonally to the plane of incidence).

A Vector calculus

The following is a summary of some results from the Prelims Multivariable Calculus course. As in the main text, all functions and vector fields are assumed to be sufficiently well-behaved in order for formulae to make sense. For example, one might take everything to be *smooth* (partial derivatives to all orders exist). Similar remarks apply to (the parametrizations of) curves and surfaces in $\mathbb{R}^3$.

A.1 Vectors in $\mathbb{R}^3$

We work in $\mathbb{R}^3$, or a domain therein, in Cartesian coordinates. If $\mathbf{e}_1 = (1,0,0)$, $\mathbf{e}_2 = (0,1,0)$, $\mathbf{e}_3 = (0,0,1)$ denote the standard orthonormal basis vectors, then a position vector is

$$\mathbf{r} = \sum_{i=1}^{3} x_i \, \mathbf{e}_i \; , \tag{A.1}$$

where $x_1 = x$, $x_2 = y$, $x_3 = z$ are the Cartesian coordinates in this basis. We denote the Euclidean length of $\mathbf{r}$ by

$$|\mathbf{r}| = r = \sqrt{x_1^2 + x_2^2 + x_3^2} \; , \tag{A.2}$$

so that $\hat{\mathbf{r}} \equiv \mathbf{r}/r$ is a unit vector for $\mathbf{r} \neq \mathbf{0}$. A vector field $\mathbf{f} = \mathbf{f}(\mathbf{r})$ may be written in this basis as

$$\mathbf{f}(\mathbf{r}) = \sum_{i=1}^{3} f_i(\mathbf{r}) \, \mathbf{e}_i \; . \tag{A.3}$$

The *scalar product* of two vectors $\mathbf{a}$, $\mathbf{b}$ is denoted by

$$\mathbf{a} \cdot \mathbf{b} \equiv \sum_{i=1}^{3} a_i b_i \; , \tag{A.4}$$

while their *vector cross product* is the vector

$$\mathbf{a} \times \mathbf{b} \equiv (a_2 b_3 - a_3 b_2) \, \mathbf{e}_1 + (a_3 b_1 - a_1 b_3) \, \mathbf{e}_2 + (a_1 b_2 - a_2 b_1) \, \mathbf{e}_3 \; . \tag{A.5}$$

The *vector triple product identity* reads

$$\mathbf{a} \times (\mathbf{b} \times \mathbf{c}) = (\mathbf{a} \cdot \mathbf{c}) \, \mathbf{b} - (\mathbf{a} \cdot \mathbf{b}) \, \mathbf{c} \; , \tag{A.6}$$

which holds for all vectors $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$.

A.2 Vector operators

The *gradient* of a function $\psi = \psi(\mathbf{r})$ is the vector field

$$\mathbf{grad}\, \psi = \nabla \psi \equiv \sum_{i=1}^{3} \frac{\partial \psi}{\partial x_i} \, \mathbf{e}_i \; . \tag{A.7}$$

The *divergence* of a vector field $\mathbf{f} = \mathbf{f}(\mathbf{r})$ is the function (scalar field)

$$\mathrm{div}\ \mathbf{f} \ = \ \nabla \cdot \mathbf{f} \ \equiv \ \sum_{i=1}^{3} \mathbf{e}_i \cdot \frac{\partial \mathbf{f}}{\partial x_i} \ = \ \sum_{i=1}^{3} \frac{\partial f_i}{\partial x_i} \ , \tag{A.8}$$

while the *curl* is the vector field

$$\begin{aligned} \mathbf{curl}\, \mathbf{f} \ &= \ \nabla \times \mathbf{f} \ \equiv \ \sum_{i=1}^{3} \mathbf{e}_i \times \frac{\partial \mathbf{f}}{\partial x_i} \\ &= \ \left(\frac{\partial f_3}{\partial x_2} - \frac{\partial f_2}{\partial x_3} \right) \mathbf{e}_1 + \left(\frac{\partial f_1}{\partial x_3} - \frac{\partial f_3}{\partial x_1} \right) \mathbf{e}_2 + \left(\frac{\partial f_2}{\partial x_1} - \frac{\partial f_1}{\partial x_2} \right) \mathbf{e}_3 \ . \end{aligned} \tag{A.9}$$

Two important identities are

$$\nabla \times (\nabla \psi) \ = \ \mathbf{0} \ , \qquad \nabla \cdot (\nabla \times \mathbf{f}) \ = \ 0 \ , \tag{A.10}$$

or in words: the curl of a gradient is zero, and the divergence of a curl is zero. Two more identities we shall need are

$$\nabla \times (\mathbf{a} \times \mathbf{b}) \ = \ \mathbf{a}\,(\nabla \cdot \mathbf{b}) \ - \ \mathbf{b}\,(\nabla \cdot \mathbf{a}) \ + \ (\mathbf{b} \cdot \nabla)\, \mathbf{a} \ - \ (\mathbf{a} \cdot \nabla)\, \mathbf{b} \ , \tag{A.11}$$

$$\nabla \cdot (\mathbf{a} \times \mathbf{b}) \ = \ \mathbf{b} \cdot (\nabla \times \mathbf{a}) \ - \ \mathbf{a} \cdot (\nabla \times \mathbf{b}) \ . \tag{A.12}$$

The second order operator ∇^2, defined by

$$\nabla^2 \psi \ \equiv \ \nabla \cdot (\nabla \psi) \ = \ \sum_{i=1}^{3} \frac{\partial^2 \psi}{\partial x_i^2} \ , \tag{A.13}$$

is called the *Laplacian*. We shall also use the identity

$$\nabla \times (\nabla \times \mathbf{f}) \ = \ \nabla\, (\nabla \cdot \mathbf{f}) \ - \ \nabla^2\, \mathbf{f} \ . \tag{A.14}$$

Notice from the definitions (A.2) and (A.7) that

$$\nabla r \ = \ \sum_{i=1}^{3} \frac{\partial r}{\partial x_i}\, \mathbf{e}_i \ = \ \sum_{i=1}^{3} \frac{x_i}{r}\, \mathbf{e}_i \ = \ \frac{\mathbf{r}}{r} \ = \ \hat{\mathbf{r}} \ , \tag{A.15}$$

and, by translating $\mathbf{r} \to \mathbf{r} - \mathbf{r}'$, more generally $\nabla\, |\mathbf{r} - \mathbf{r}'| = (\mathbf{r} - \mathbf{r}')/|\mathbf{r} - \mathbf{r}'|$.

A.3 Integral theorems

Definition (Line integral) Let C be a curve in $\mathbb{R}^3$, parametrized by $\mathbf{r} : [t_0, t_1] \to \mathbb{R}^3$, or $\mathbf{r}(t)$ for short. Then the *line integral* of a scalar field ψ and vector field $\mathbf{f}$ along C are respectively

$$\int_C \psi\, \mathrm{d}s \ \equiv \ \int_{t_0}^{t_1} \psi(\mathbf{r}(t)) \left| \frac{\mathrm{d}\mathbf{r}(t)}{\mathrm{d}t} \right| \mathrm{d}t \ , \qquad \int_C \mathbf{f} \cdot \mathrm{d}\mathbf{r} \ \equiv \ \int_{t_0}^{t_1} \mathbf{f}(\mathbf{r}(t)) \cdot \frac{\mathrm{d}\mathbf{r}(t)}{\mathrm{d}t}\, \mathrm{d}t \ . \tag{A.16}$$

Note that $\boldsymbol{\tau}(t) \equiv \mathrm{d}\mathbf{r}/\mathrm{d}t$ is the *tangent vector* to the curve – a vector field defined on C. The values of these integrals are independent of the choice of *oriented* parametrization (proof uses the chain rule).

A curve is *simple* if $\mathbf{r} : [t_0, t_1] \to \mathbb{R}^3$ is injective (then C is non-self-intersecting), and is *closed* if $\mathbf{r}(t_0) = \mathbf{r}(t_1)$ (then C forms a loop).

Definition (Surface integral) Let Σ be a surface in $\mathbb{R}^3$, parametrized by $\mathbf{r}(u,v)$, with $(u,v) \in D \subseteq \mathbb{R}^2$. The *unit normal* $\mathbf{n}$ to the surface is

$$\mathbf{n} \equiv \frac{\mathbf{t}_u \times \mathbf{t}_v}{|\mathbf{t}_u \times \mathbf{t}_v|} , \tag{A.17}$$

where

$$\mathbf{t}_u \equiv \frac{\partial \mathbf{r}}{\partial u}, \qquad \mathbf{t}_v \equiv \frac{\partial \mathbf{r}}{\partial v} \tag{A.18}$$

are two *tangent vectors* to the surface. These are all vector fields defined on Σ. The *surface integral* of a function ψ over Σ is

$$\int_\Sigma \psi \, \mathrm{d}S \equiv \iint_D \psi(\mathbf{r}(u,v)) \left| \frac{\partial \mathbf{r}}{\partial u} \times \frac{\partial \mathbf{r}}{\partial v} \right| \mathrm{d}u \, \mathrm{d}v . \tag{A.19}$$

The *sign* of $\mathbf{n}$ in (A.17) is not in general independent of the choice of parametrization. Typically, the whole of a surface cannot be parametrized by a single domain D; rather, one needs to cover Σ with several parametrizations using domains $D_I \subseteq \mathbb{R}^2$, where I labels the domain. The surface integral (A.19) is then defined in the obvious way, as a sum of integrals over $D_I \subseteq \mathbb{R}^2$. However, in doing this it might not be possible to define a continuous $\mathbf{n}$ over the whole of Σ (an example being a Möbius strip).

Definition (Orientations) A surface Σ is *orientable* if there is a choice of continuous unit normal vector field $\mathbf{n}$ on Σ. If an orientable Σ has boundary $\partial\Sigma$, a simple closed curve, then the normal $\mathbf{n}$ induces an orientation of $\partial\Sigma$: we require that $\boldsymbol{\tau} \times \mathbf{n}$ points *away* from Σ, where $\boldsymbol{\tau}$ denotes the oriented tangent vector to $\partial\Sigma$ – see Figure 23.

The point here is that the choice of direction $\boldsymbol{\tau}$ along the curve $\partial\Sigma$ in turn fixes the choice of sign when integrating over $\partial\Sigma$.

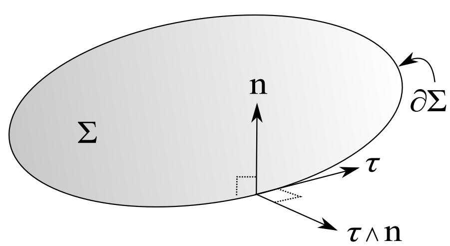


Figure 23: Surface Σ with unit normal $\mathbf{n}$, and boundary $\partial\Sigma$ with oriented tangent vector $\boldsymbol{\tau}$. The direction of $\boldsymbol{\tau}$ is fixed by requiring $\boldsymbol{\tau} \times \mathbf{n}$ to point away from Σ. (Right hand rule for the cross product: $\boldsymbol{\tau}$ points along the index finger, $\mathbf{n}$ along the middle finger, and $\boldsymbol{\tau} \times \mathbf{n}$ along the thumb.)

With this in hand, we may now state

Theorem A.1 (*Stokes*) *Let Σ be an orientable surface in $\mathbb{R}^3$, with unit normal vector $\mathbf{n}$ and boundary curve $\partial\Sigma$. If $\mathbf{f}$ is a vector field then*

$$\int_\Sigma (\nabla \times \mathbf{f}) \cdot \mathbf{n} \, \mathrm{d}S = \int_{\partial\Sigma} \mathbf{f} \cdot \mathrm{d}\mathbf{r} . \tag{A.20}$$

Definition (Volume integral) The integral of a function ψ in a (bounded) region R in $\mathbb{R}^3$ is

$$\int_R \psi \,\mathrm{d}V \;\equiv\; \iiint_R \psi(\mathbf{r}) \,\mathrm{d}x_1 \,\mathrm{d}x_2 \,\mathrm{d}x_3 \ . \tag{A.21}$$

Theorem A.2 (*Divergence*) *Let R be a bounded region in $\mathbb{R}^3$ with boundary surface ∂R. If $\mathbf{f}$ is a vector field then*

$$\int_R \nabla \cdot \mathbf{f} \,\mathrm{d}V \;=\; \int_{\partial R} \mathbf{f} \cdot \mathbf{n} \,\mathrm{d}S \ , \tag{A.22}$$

where $\mathbf{n}$ is the outward unit normal vector to ∂R.

Note that the surface Σ in Stokes' theorem has a boundary $\partial\Sigma$, whereas the surface ∂R in the divergence theorem does not (it is itself the boundary of the region R).

Finally, we will need the following result proved in the Prelims Multivariable Calculus course:

Lemma A.3 *If f is a continuous function such that*

$$\int_R f \,\mathrm{d}V \;=\; 0 \ , \tag{A.23}$$

for all *bounded regions R, then $f \equiv 0$.*

The proof of this is standard: if f is non-zero at a point $\mathbf{r}_0 \in \mathbb{R}^3$, say $f(\mathbf{r}_0) > 0$ is positive, then by continuity f is also positive in a small neighbourhood around $\mathbf{r}_0$. But then the intergal of f over that neighbourhood will be positive, contradicting (A.23).