Electromagnetic Theory AMA3001

Lecture Notes 2017/2018

Gleb Gribakin
Department of Applied Mathematics
and Theoretical Physics
Queen's University Belfast

May 15, 2018

Contents

Preface

These lecture notes are an evolution of the set of notes by Professor James Walters who taught this module at Queen's for many years. In delivering Electromagnetic Theory, Professor Walters followed two principal sources:

1. C. A. Coulson and T. J. M. Boyd, *Electricity*, 2nd ed. (Longman, London, 1979).
2. J. R. Reitz, F. J. Milford, R. W. Christy, *Foundations of electromagnetic theory*, 4th ed. (Addison-Wesley, Reading, 1993).

My own understanding of Electromagnetism has been formed by a number of other books, and I am happy to recommend these to the reader:

3. E. M. Purcell, *Electricity and Magnetism*, 2nd ed. (McGraw-Hill, New York, 1985).
4. R. P. Feynman, *The Feynman lectures on physics, Volume 2* (BasicBooks, New York, 2011).
   Read online: http://www.feynmanlectures.caltech.edu
5. L. D. Landau and E. M. Lifshitz, *The Classical Theory of Fields*, 4th ed. (Pergamon, Oxford, 1975).

While delivering the module for the first time in 2015–2016, I consulted these titles:

6. J. D. Jackson, *Classical Electrodynamics*, 3rd ed. (Wiley, New York, 1999).
7. A. Sommerfeld, *Electrodynamics* (Academic Press, New York, 1952).

I am also happy to mention another book that was recommended to me by a colleague who taught this subject at Oakland University in the US:

8. G. L. Pollack and D. R. Stump, *Electromagnetism* (Addison Wesley, San Francisco, 2002).

Finally, I am grateful to my wife Anna, who proofread most of the notes, spotting numerous typos and helping me to identify and improve many instances of obscure or akward phrasing. Naturally, all the remaining defects and deficiencies are my sole responsibility.

0 Introduction: electromagnetism in one lecture

0.1 Maxwell's equations

Charge is a fundamental property of matter.

Electric and magnetic fields ($\boldsymbol{E}$ and $\boldsymbol{B}$) are produced by stationary or moving charges, but can also be sustained in vacuum in the absence of charges as electromagnetic waves (see below).

Force acting on a particle with charge q (Lorentz's force) is

$$\boldsymbol{F} = q\boldsymbol{E} + \frac{q}{c}\boldsymbol{v} \times \boldsymbol{B}, \tag{0.1}$$

where $\boldsymbol{v}$ is the velocity of the particle and c is the speed of light. In this section we use CGS system, in which E and B are measured in the same units. This is quite natural, since the electric and magnetic fields transform into each other if we consider them in different inertial frames of reference. Also, in these units the magnitudes of the electric and magnetic fields in an electromagnetic wave are equal. Equation (0.1) then shows that for slow-moving particles ($v \ll c$) the electric force (first term) is much greater than the magnetic force (second term).

Maxwell found that all the information about electric and magnetic phenomena could be presented concisely as[1]

$$\nabla \cdot \boldsymbol{E} = 4\pi\rho, \tag{0.2}$$

$$\nabla \cdot \boldsymbol{B} = 0, \tag{0.3}$$

$$\nabla \times \boldsymbol{B} = \frac{4\pi}{c}\boldsymbol{j} + \frac{1}{c}\frac{\partial \boldsymbol{E}}{\partial t}, \tag{0.4}$$

$$\nabla \times \boldsymbol{E} = -\frac{1}{c}\frac{\partial \boldsymbol{B}}{\partial t}, \tag{0.5}$$

where ρ is the electric charge density and $\boldsymbol{j}$ is the current density[2].

Equation (0.2) is the differential form of *Coulomb's law*. It shows that the electric field originates on the charges. It "diverges" from the regions in space that have charge, as $\nabla \cdot \boldsymbol{E} \equiv \text{div}\boldsymbol{E}$, is the divergence of $\boldsymbol{E}$. By analogy, Eq. (0.3) implies that there are *no magnetic charges* (or monopoles).

Equation (0.4) without the second term on the right-hand side (which was added by Maxwell, see below) shows that magnetic fields are caused by currents, and is equivalent to the *Biot-Savart-Laplace-Ampere law*.

[1] The equations were cast in this form by Oliver Heaviside (1850–1925), who developed vector calculus (with J. W. Gibbs), while Maxwell originally wrote them in components. The equations shown here apply in the absence of dielectric or magnetic materials.

[2] Recall also the vector differentiation (nabla) operator $\nabla = \boldsymbol{i}\frac{\partial}{\partial x} + \boldsymbol{j}\frac{\partial}{\partial y} + \boldsymbol{k}\frac{\partial}{\partial x}$.

Equation (0.5) is *Faraday's law* of induction. It shows that a time-dependent magnetic field gives rise to an electric field. This field is different from that generated by charges in that its work over a closed contour is not zero.

If $\nabla\cdot$ is applied to Eq. (0.4) (without the second term), one obtains[3]

$$\nabla \cdot \boldsymbol{j} = 0.$$

This equation is in general incorrect. It works for stationary current, but does not describe the situation when charges flow ("diverge") out of a certain region of space, causing a decrease in the charge density there. This means that Eq. (0.4) with only the first term on the right-hand side is incomplete.

The correct form of *charge conservation* is

$$\nabla \cdot \boldsymbol{j} = -\frac{\partial \rho}{\partial t}. \tag{0.6}$$

It shows that, as expected, a decreasing charge density cause the emergence of a current out of this region of space.

To make Eq. (0.4) compatible with (0.6), Maxwell added the second term on the right-hand side of (0.4). It involves the time-derivative of the electric field, and is known as the *displacement current*.

Applying $\nabla\times$ to Eq. (0.5) and using Eqs. (0.2) and (0.4) in vacuum where $\rho = 0$ and $\boldsymbol{j} = 0$, gives[4]

$$\nabla^2 \boldsymbol{E} - \frac{1}{c^2}\frac{\partial^2 \boldsymbol{E}}{\partial^2 t} = 0. \tag{0.7}$$

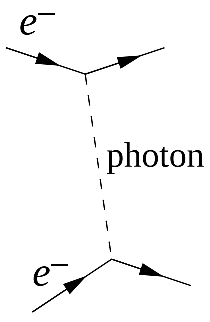
This is the *wave equation*. It shows that electric fields accompanied by magnetic fields can propagate in vacuum as *electromagnetic waves*, travelling at the speed of light. This mathematical discovery immediately told Maxwell that light, whose nature was not known until then, is an electromagnetic wave!

[3]We make use of the identity $\nabla \cdot (\nabla \times \boldsymbol{B}) = (\nabla \times \nabla) \cdot \boldsymbol{B} = 0$.
[4]We also use the identity $\nabla \times (\nabla \times \boldsymbol{E}) = \nabla(\nabla \cdot \boldsymbol{E}) - \nabla^2 \boldsymbol{E}$.

1 Electrostatics

Electrostatics is the study of electric fields created by stationary charges.

1.1 Electric charge



Electric charge is a property of matter. It can be positive or negative. The charge characterises the ability of particles to participate in the electromagnetic interaction, or, at the elementary particle level, the ability to emit or absorb photons[5].

The usual notation for the charge is q or Q.

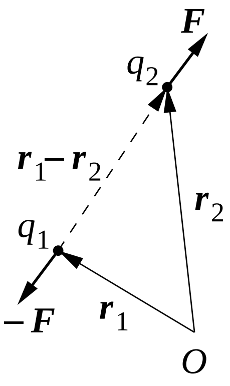
The unit of charge in SI is coulomb (C). It is defined in terms of the units of current (ampere, A) and time (s), as

$$1\ \mathrm{C} = 1\ \mathrm{A} \times 1\ \mathrm{s}.$$

The smallest observable amount of charge, or the elementary charge, is the charge of the electron $-e$, or that of the proton[6] e, where

$$e \approx 1.6 \times 10^{-19}\ \mathrm{C}.$$

1.2 Coulomb's law



Coulomb's law states that the force between charges is proportional to their magnitudes and inversely proportional to the distance squared.

Mathematically, the force on charge q_2 with position vector $\boldsymbol{r}_2$ due to charge q_1 at $\boldsymbol{r}_1$ is (in SI units)

$$\boldsymbol{F} = \frac{q_1 q_2}{4\pi\varepsilon_0} \frac{\boldsymbol{r}_2 - \boldsymbol{r}_1}{|\boldsymbol{r}_2 - \boldsymbol{r}_1|^3}, \tag{1.8}$$

where

$$\varepsilon_0 = 8.854 \times 10^{-12}\ \mathrm{N^{-1}C^2m^{-2}}, \tag{1.9}$$

which gives $1/(4\pi\varepsilon_0) = 9 \times 10^9\ \mathrm{NC^{-2}m^2}$.

For $q_1 q_2 > 0$ (charges of the same sign) the force is repulsive, and for $q_1 q_2 < 0$ (charges of opposite sign) it is attractive.

The force acting on q_1 is obtained from Eq. (1.8) by interchanging indices 1 and 2, and is just the negative of $\boldsymbol{F}$, in accordance with Newton's 3rd law.

[5] The Feynman diagram on the left shows the interaction between two electrons (solid lines) via exchange of a photon (dashed line).

[6] The proton p is not an elementary particle. It consists of three quarks, which have fractional charges: $p = uud$, where u is the "up" quark with the charge $\frac{2}{3}e$ and d is the "down" quark with the charge $-\frac{1}{3}e$. The neutron n also consists of three quarks, $n = udd$. However, free quarks are never observed. They are always *confined* within protons or neutrons, or other particles, in sets of three or two (as quark-antiquark pairs). This ensures that observed charges are always integer multiples of e.

1.3 Electric field

According to Coulomb's law, the force on charge q at point $\boldsymbol{r}$ due to charge q_1 at point $\boldsymbol{r}_1$ is

$$\boldsymbol{F} = \frac{qq_1}{4\pi\varepsilon_0} \frac{\boldsymbol{r} - \boldsymbol{r}_1}{|\boldsymbol{r} - \boldsymbol{r}_1|^3}. \tag{1.10}$$

It is convenient mathematically and correct physically to write this force as

$$\boldsymbol{F} = q\boldsymbol{E}(\boldsymbol{r}), \tag{1.11}$$

where

$$\boldsymbol{E}(\boldsymbol{r}) = \frac{q_1}{4\pi\varepsilon_0} \frac{\boldsymbol{r} - \boldsymbol{r}_1}{|\boldsymbol{r} - \boldsymbol{r}_1|^3}, \tag{1.12}$$

is the *electric field* that charge q_1 creates at point $\boldsymbol{r}$.

The physical reason for writing the force as (1.11) is that the charge q does not "know" that there is charge q_1 somewhere in the distance, but is affected by something acting on it locally at $\boldsymbol{r}$. This "something" is the electric field. (If q_1 is moved, q does not "feel" this change instantaneously[7].)

Looking at Eq. (1.11), we can state the following:

Definition: electric field is the force acting on a unit positive charge.

The force on q from a system of N charges q_i at $\boldsymbol{r}_i$ $(i = 1, \ldots, N)$ is

$$\boldsymbol{F} = \sum_{i=1}^{N} \frac{qq_i}{4\pi\varepsilon_0} \frac{\boldsymbol{r} - \boldsymbol{r}_i}{|\boldsymbol{r} - \boldsymbol{r}_i|^3}, \tag{1.13}$$

which assumes that the charges q_i act on q independently. Their electric field then is

$$\boldsymbol{E}(\boldsymbol{r}) = \sum_{i=1}^{N} \frac{q_i}{4\pi\varepsilon_0} \frac{\boldsymbol{r} - \boldsymbol{r}_i}{|\boldsymbol{r} - \boldsymbol{r}_i|^3}. \tag{1.14}$$

Equation (1.14) is a manifestation of the *principle of superposition*, which states that the electric field due to a system of charges is equal to the sum of the fields created by each of these charges[8].

Electric fields can be visualised by drawing a number of field lines (or lines of force) to which vector $\boldsymbol{E}$ is tangential at every point. These lines must begin and end on charges or at infinity. The density of field lines characterises the strength of the electric field. Some examples are shown in Fig. 1.

---

[7] According to Einstein's theory of relativity, no signal can propagate faster than light, so the charge q at $\boldsymbol{r}$ will not "learn" about this change until $|\boldsymbol{r} - \boldsymbol{r}_1|/c$ time later.

[8] The strong interaction between protons and neutrons (collectively known as nucleons) in a nucleus does not obey the superposition principle. These particles consist of quarks which interact with each other by means of gluons. This interaction is described by Quantum Chromodynamics and is strongly nonlinear, since gluons can emit gluons and interact with each other. Although the basic structure of this theory has been known since late 1960's, reasonably accurate calculations of nucleons' masses, have only become feasible now [see, e.g., Sz. Borsanyi *et al.*, *Ab initio calculation of the neutron-proton mass difference*, Science, Vol. 347, Issue 6229, pp. 1452–1455 (2015)].

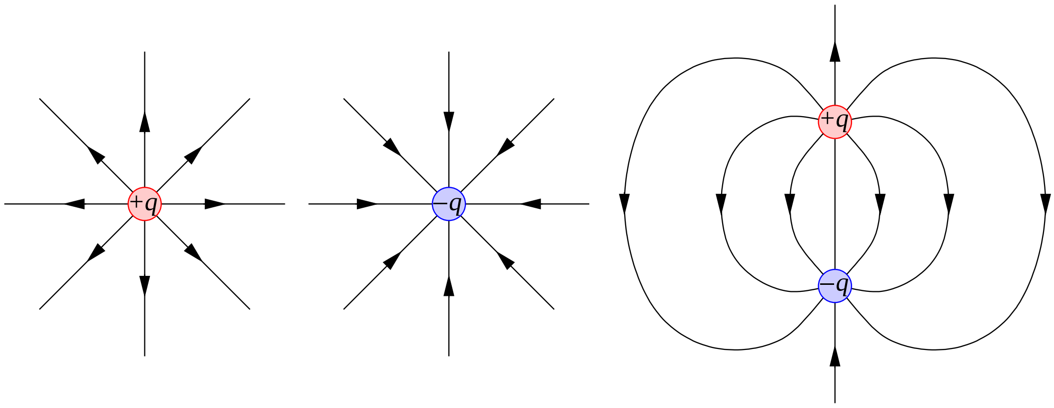


Figure 1: Field lines for the positive charge $+q$, negative charge $-q$, and the system of two charges $+1$ and $-q$.

1.4 Electrostatic potential

Let us compute the gradient of the function $1/r$. This can be done directly using Cartesian coordinates,

$$\nabla\frac{1}{r} = \nabla\frac{1}{\sqrt{x^2+y^2+z^2}} = \ldots.$$

However, it is instructive to do this in a slightly different way.

Since the function $1/r$ depends on x, y and z through r, we can use chain rule:

$$\nabla\frac{1}{r} \equiv \frac{d}{d\boldsymbol{r}}\frac{1}{r} = \frac{d}{dr}\left(\frac{1}{r}\right)\frac{dr}{d\boldsymbol{r}} = -\frac{1}{r^2}\nabla r,$$

where we wrote ∇ formally as the "vector derivative" $d/d\boldsymbol{r}$. The last derivative is

$$\begin{aligned}\nabla r &= \nabla\sqrt{x^2+y^2+z^2}\\ &= \boldsymbol{i}\frac{2x}{2\sqrt{x^2+y^2+z^2}} + \boldsymbol{j}\frac{2y}{2\sqrt{x^2+y^2+z^2}} + \boldsymbol{k}\frac{2z}{2\sqrt{x^2+y^2+z^2}}\\ &= \frac{x\boldsymbol{i}+y\boldsymbol{j}+z\boldsymbol{k}}{\sqrt{x^2+y^2+z^2}}\\ &= \frac{\boldsymbol{r}}{r},\end{aligned}$$

which is a unit vector in the direction of $\boldsymbol{r}$. Hence, we have[9]

$$\nabla\frac{1}{r} = -\frac{\boldsymbol{r}}{r^3}. \tag{1.15}$$

Shifting the origin to an arbitrary point r_i gives

$$\nabla\frac{1}{|\boldsymbol{r}-\boldsymbol{r}_i|} = -\frac{\boldsymbol{r}-\boldsymbol{r}_i}{|\boldsymbol{r}-\boldsymbol{r}_i|^3}. \tag{1.16}$$

[9] In a similar way, the gradient of any function $f(r)$ that depends on the distance from the origin is $\nabla f(r) = f'(r)\nabla r = f'(r)\boldsymbol{r}/r$.

Comparing this identity with Eq. (1.14), we can see that the electric field can be written as

$$\boldsymbol{E}(\boldsymbol{r}) = -\nabla\phi(\boldsymbol{r}), \tag{1.17}$$

where

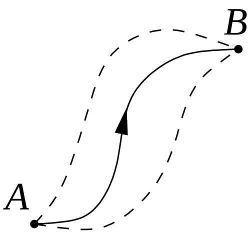
$$\phi(\boldsymbol{r}) = \sum_{i=1}^{N} \frac{q_i}{4\pi\varepsilon_0} \frac{1}{|\boldsymbol{r} - \boldsymbol{r}_i|}. \tag{1.18}$$

is the *electrostatic potential* of N point charges q_i at $\boldsymbol{r}_i$.

Unlike the electric field, which is a vector, the electrostatic potential is a scalar. When finding the electric field for a system of charges, it is often easier to determine its electrostatic potential first [e.g., by means of Eq. (1.18)], and then find the electric field from Eq. (1.17).

The SI units of the potential is volt (V), which is equal to JC^{-1} (J is joule, the SI unit of energy or work).

Equation (1.17) shows that the unit of the electric field is Vm^{-1}. Alternatively it be written as NC^{-1}, as seen from Eq. (1.11).



Let us now consider work by the electric field on moving charge q from point A to point B along some path. From the definition of work, we have

$$W = \int_A^B \boldsymbol{F} \cdot d\boldsymbol{r},$$

and using Eq. (1.11) and then Eq. (1.17), we obtain [10]

$$W = \int_A^B q\boldsymbol{E} \cdot d\boldsymbol{r} = -q \int_A^B \nabla\phi \cdot d\boldsymbol{r} = -q \int_A^B d\phi = -q[\phi(\boldsymbol{r}_B) - \phi(\boldsymbol{r}_A)]. \tag{1.19}$$

The quantity in brackets is the *potential difference* between points B and A.

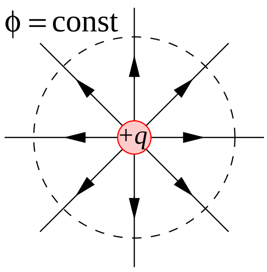
Equation (1.19) shows that the work by the electric field does not depend on the shape of the path, but only on the initial and final points. Forces or fields for which this is true are called *conservative*. Hence, the electric field $\boldsymbol{E}(\boldsymbol{r})$ is conservative.

The condition that $\boldsymbol{E}$ is conservative can be written in a compact mathematical form. We know that

$$\nabla \times \nabla\psi = 0,$$

for *any* scalar function ψ. From Eq. (1.17), we then have

$$\nabla \times \boldsymbol{E} = 0. \tag{1.20}$$



Electric fields can be visualised by showing the *equipotential surfaces* on which

$$\phi(\boldsymbol{r}) = \text{const}. \tag{1.21}$$

Field lines are perpendicular to equipotential surfaces, since $d\phi = \nabla\phi \cdot d\boldsymbol{r} = -\boldsymbol{E} \cdot d\boldsymbol{r} = 0$ means $\boldsymbol{E} \perp d\boldsymbol{r}$, where $d\boldsymbol{r}$ lies on the equipotential serface.

[10] Recall that the differential of ϕ (i.e., the change of ϕ related to the displacement $d\boldsymbol{r} = dx\boldsymbol{i} + dy\boldsymbol{j} + dz\boldsymbol{k}$) is

$$d\phi = \frac{\partial\phi}{\partial x}dx + \frac{\partial\phi}{\partial y}dy + \frac{\partial\phi}{\partial z}dz = \nabla\phi \cdot d\boldsymbol{r}.$$

1.5 Volume and surface distributions of charge

When charge is distributed continuously in space, it is characterised by the *volume charge density* ρ (charge per unit volume). The charge of a volume element dV is then given by $dq = \rho dV$.

When charge is distributed continuously over a surface, it is characterised by the *surface charge density* σ (charge per unit area)[11]. The charge of a surface element dS is then found as $dq = \sigma dS$.

In accordance with Eq. (1.12), the electric field due to charge $dq' = \rho(\boldsymbol{r}')dV'$ within a small volume element dV' located at $\boldsymbol{r}'$, is

$$\frac{1}{4\pi\varepsilon_0}\frac{\rho(\boldsymbol{r}')dV'(\boldsymbol{r}-\boldsymbol{r}')}{|\boldsymbol{r}-\boldsymbol{r}'|^3}.$$

By the superposition principle, the total electric field is found as

$$\boldsymbol{E}(\boldsymbol{r}) = \frac{1}{4\pi\varepsilon_0}\int_V \frac{\rho(\boldsymbol{r}')(\boldsymbol{r}-\boldsymbol{r}')}{|\boldsymbol{r}-\boldsymbol{r}'|^3}dV' + \frac{1}{4\pi\varepsilon_0}\int_S \frac{\sigma(\boldsymbol{r}')(\boldsymbol{r}-\boldsymbol{r}')}{|\boldsymbol{r}-\boldsymbol{r}'|^3}dS', \tag{1.22}$$

where the first integral is over all volume charges, and where we have also added the contribution of the surface charges (second term). The corresponding electrostatic potential is given by

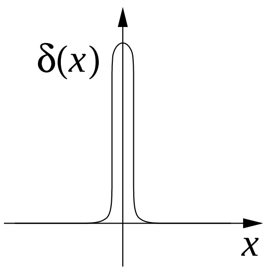
$$\phi(\boldsymbol{r}) = \frac{1}{4\pi\varepsilon_0}\int_V \frac{\rho(\boldsymbol{r}')}{|\boldsymbol{r}-\boldsymbol{r}'|}dV' + \frac{1}{4\pi\varepsilon_0}\int_S \frac{\sigma(\boldsymbol{r}')}{|\boldsymbol{r}-\boldsymbol{r}'|}dS'. \tag{1.23}$$

The charge density can also describe point charges. Thus, the density of point charge q at $\boldsymbol{r}'$ can be written as

$$\rho(\boldsymbol{r}) = q\delta(\boldsymbol{r}-\boldsymbol{r}'), \tag{1.24}$$

where $\delta(\boldsymbol{r}-\boldsymbol{r}')$ is the *Dirac delta function* defined by

$$\delta(\boldsymbol{r}-\boldsymbol{r}') = \begin{cases} 0 & \text{for } \boldsymbol{r} \neq \boldsymbol{r}' \\ +\infty & \text{for } \boldsymbol{r} = \boldsymbol{r}' \end{cases}, \tag{1.25}$$



and

$$\int \delta(\boldsymbol{r}-\boldsymbol{r}')dV = 1. \tag{1.26}$$

Note that although the above integral is over the whole space, it is only an arbitrarily small vicinity of the point $\boldsymbol{r} = \boldsymbol{r}'$ that contributes to it, since the delta function vanishes everywhere else[12].

From Eqs. (1.25) and (1.26), for any continuous function $f(\boldsymbol{r})$,

$$\int_V f(\boldsymbol{r})\delta(\boldsymbol{r}-\boldsymbol{r}')dV = \begin{cases} f(\boldsymbol{r}'), & \text{if } \boldsymbol{r}' \text{ is inside } V \\ 0, & \text{if } \boldsymbol{r}' \text{ is outside } V \end{cases}. \tag{1.27}$$

[11] In some problems we also consider charges distributed on a line, which are characterised by linear charge density λ.

[12] The graph on the left shows the function $\delta(x)$ in one dimension as a spike which should be infinitely tall and infinitely narrow, with a unit area underneath: $\int \delta(x)dx = 1$. The three-dimensional delta function is $\delta(\boldsymbol{r}) = \delta(x)\delta(y)\delta(z)$.

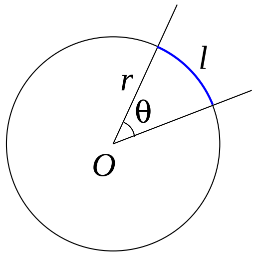
The first line above is obtained by realising that the only point that contributes to the intergral is that where $\boldsymbol{r} = \boldsymbol{r}'$, since $\delta(\boldsymbol{r} - \boldsymbol{r}') = 0$ elsewhere. The function $f(\boldsymbol{r})$ in the integrand can then be evaluated at $\boldsymbol{r} = \boldsymbol{r}'$ this point and, as a constant, taken outside the integral, which then reduces to (1.26).

For a system of N point charges q_i at $\boldsymbol{r}_i$, the charge density is

$$\rho(\boldsymbol{r}) = \sum_{i=1}^{N} q_i \delta(\boldsymbol{r} - \boldsymbol{r}_i). \tag{1.28}$$

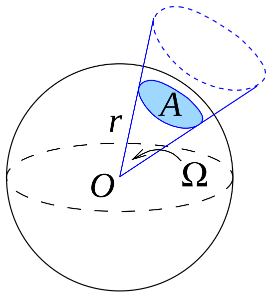
It is a simple exercise to verify that substitution of this in (1.22) gives (1.14).

1.6 Solid angle



The planar angle is defined as the ratio of the length l of an arc subtended by the angle, to the radius of the circle r, i.e., $\theta = l/r$ (in radians). In particular, the full angle is 2π radians.

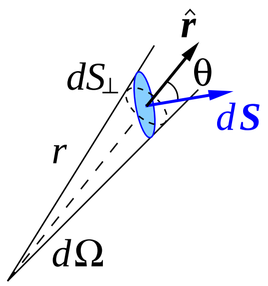
Similarly, in three dimensions the magnitude of a *solid angle* is given by the ratio of the area A that it subtends on the surface of a sphere to the square of its radius r: $\Omega = A/r^2$. The unit of the solid angle is steradian, and the full angle is $4\pi r^2/r^2 = 4\pi$ steradians.



For the element of solid angle $d\Omega$ (i.e., an arbitrarily small solid angle) we have

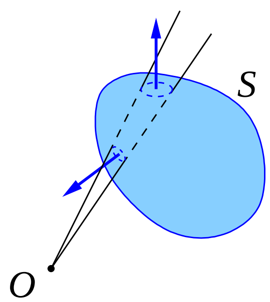
$$d\Omega = \frac{dS_\perp}{r^2} = \frac{dS\cos\theta}{r^2} = \frac{\hat{\boldsymbol{r}} \cdot d\boldsymbol{S}}{r^2} = \frac{\boldsymbol{r} \cdot d\boldsymbol{S}}{r^3}. \tag{1.29}$$

Here $dS_\perp$ is the cross sectional area of the solid angle perpendicular to the radius, θ is the angle between the unit vector $\hat{\boldsymbol{r}} = \boldsymbol{r}/r$ and $d\boldsymbol{S}$, where dS is the element of a surface subtended by the solid angle, and $d\boldsymbol{S} = \boldsymbol{n} dS$, $\boldsymbol{n}$ being a unit normal to the surface.



The magnitude of a finite solid angle subtended by surface S is given by the integral over this surface

$$\Omega = \int_S \frac{\boldsymbol{r} \cdot d\boldsymbol{S}}{r^3}. \tag{1.30}$$



For a *closed* surface, and taking the direction of $d\boldsymbol{S}$ as the outward normal, we have

$$\Omega = \oint_S \frac{\boldsymbol{r} \cdot d\boldsymbol{S}}{r^3} = \begin{cases} 4\pi, & \text{if the origin } O \text{ is inside } S, \\ 0, & \text{if the origin } O \text{ is outside } S. \end{cases} \tag{1.31}$$

Indeed, when the surface enloses the origin, the above integral gives the full solid angle. When the origin is outside the surface, there are contributions from parts of the surface facing the origin and those facing away from it that are equal in magnitude but opposite in sign. They cancel each other, and the resulting integral is zero.

1.7 Gauss's Law

The flux of the electric field $\boldsymbol{E}$ across surface S is defined as [13]

$$\int_S \boldsymbol{E} \cdot d\boldsymbol{S}. \tag{1.32}$$

Let us consider the flux of the electric field due to N point charges q_i with positions $\boldsymbol{r}_i$, Eq. (1.14), through a *closed* surface S:

$$\oint_S \boldsymbol{E} \cdot d\boldsymbol{S} = \oint_S \sum_{i=1}^{N} \frac{q_i}{4\pi\varepsilon_0} \frac{\boldsymbol{r}-\boldsymbol{r}_i}{|\boldsymbol{r}-\boldsymbol{r}_i|^3} \cdot d\boldsymbol{S} = \frac{1}{4\pi\varepsilon_0} \sum_{i=1}^{N} q_i \oint_S \frac{(\boldsymbol{r}-\boldsymbol{r}_i)\cdot d\boldsymbol{S}}{|\boldsymbol{r}-\boldsymbol{r}_i|^3}. \tag{1.33}$$

According to Eq. (1.31), the integral on the right-hand side of (1.33) is equal to 4π if $\boldsymbol{r}_i$ is inside S, and is equal to zero if $\boldsymbol{r}_i$ is outside S.

Hence, only the charges that are enclosed by S contribute to the sum over i, and we have

$$\oint \boldsymbol{E} \cdot d\boldsymbol{S} = \frac{Q}{\varepsilon_0}, \tag{1.34}$$

where Q is the total charge enclosed by S.

Equation (1.34) is *Gauss's law*: the flux of the electric field out of any closed surface is equal to the total charge enclosed by this surface, divided by ε_0.

Gauss's law can be used to determine the electric field in cases where it possesses some symmetry due to the symmetry of the charge distribution.

Example 1. Electric field of a spherically symmetric charge distribution.

A spherically symmetric charge density $\rho(r)$ depends on the distance from the origin r, but not on the direction of $\boldsymbol{r}$. The corresponding electric field is in the radial direction, $\boldsymbol{E}(\boldsymbol{r}) = E(r)\hat{\boldsymbol{r}}$, and its magnitude depends on r only.

Choose the Gaussian surface S as a sphere of radius r (on which $d\boldsymbol{S} = \hat{\boldsymbol{r}}dS$), we obtain the flux as

$$\oint_S \boldsymbol{E} \cdot d\boldsymbol{S} = \oint_S E(r)\hat{\boldsymbol{r}} \cdot \hat{\boldsymbol{r}} dS = E(r) \oint_S dS = E(r) 4\pi r^2, \tag{1.35}$$

where we used the fact that $E(r)$ is constant on this sphere.

The charge inside the sphere is found by integration over the volume V of this sphere,

$$Q = \int_V \rho(r')dV' = \int_0^r \rho(r') 4\pi r'^2 dr', \tag{1.36}$$

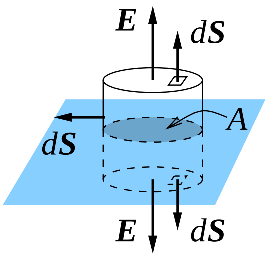
where $dV' = 4\pi r'^2 dr'$ is the volume of a spherical shell of radius r' and thickness dr'.

[13] "In the case of fluxes, we have to take the integral, over a surface, of the flux through every element of the surface. The result of this operation is called the surface integral of the flux. It represents the quantity which passes through the surface." (J. C. Maxwell, *Treatise on Electricity and Magnetism*, 1873). For example, if we consider the motion of a fluid with velocity $\boldsymbol{v}$, the intergal $\int_S \boldsymbol{v} \cdot d\boldsymbol{S}$ will give the volume of fluid that flows across S in unit time.

Substituting the flux (1.35) and the charge (1.36) into Gauss's law (1.34), we obtain

$$E(r) = \frac{1}{4\pi\varepsilon_0 r^2} \int_0^r \rho(r') 4\pi r'^2 dr'. \tag{1.37}$$

Example 2. Electric field of a uniformly charged plane.

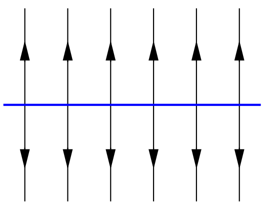


Consider an infinite plane with surface charge density σ. By symmetry, the electric field produced by this charge distribution must be perpendicular to the plane, and its magnitude can depend only on the distance from the plane.

We choose the Gaussian surface as a right cylinder, placed symmetrically with respect to the plane, with flat surfaces parallel to it. On each of the flat surfaces of area A, the electric field is constant and perpendicular to it (i.e., parallel to $d\boldsymbol{S}$), and the corresponding flux is EA. The flux across the curved surface is zero, since $\boldsymbol{E}$ and $d\boldsymbol{S}$ are perpendicular to each other (so that $\boldsymbol{E} \cdot d\boldsymbol{S} = 0$). Hence, the total flux is $2EA$.

The charge enclosed by the cylinder is σA, and from (1.34), we have

$$2EA = \frac{\sigma A}{\varepsilon_0},$$



which gives

$$E = \frac{\sigma}{2\varepsilon_0}. \tag{1.38}$$

We see that on either side of the plane, the electric field is uniform (i.e., $\boldsymbol{E} = \text{const}$). It is directed away from the plane for $\sigma > 0$, and towards the plane for $\sigma < 0$.

Let us now derive the differential form of Gauss's law. Assuming the charge is distributed with density ρ, the charge Q in Eq. (1.34) is given by

$$Q = \int_V \rho dV,$$

where V is the volume bounded by S. Using Gauss's theorem[14], the left-hand side of Eq. (1.34) is transformed into a volume integral. Thus, we obtain

$$\int_V \nabla \cdot \boldsymbol{E}\, dV = \frac{1}{\varepsilon_0} \int_V \rho dV,$$

or

$$\int_V \left(\nabla \cdot \boldsymbol{E} - \frac{\rho}{\varepsilon_0} \right) dV = 0.$$

Since this is true for *any* volume V, the expression in brackets must vanish, which gives *Gauss's law in differential form*:

$$\nabla \cdot \boldsymbol{E} = \rho/\varepsilon_0. \tag{1.39}$$

[14]Gauss's theorem is one of a family of formulae that relates an integral of the derivative of a function over an interval, an area, or a volume, to the contribution of the function on the boundary of this interval, area of volume. The simplest of such formulae is the "fundamental theorem of calculus", $\int_a^b (dF/dx)dx = F(b) - F(a)$. Gauss's (or divergence) theorem states that for a vector field $\boldsymbol{a}$, the volume integral of its divergence, $\text{div}\, \boldsymbol{a} \equiv \nabla \cdot \boldsymbol{a}$, is equal to the integral of $\boldsymbol{a}$ over the surface S bounding V: $\int_V \nabla \cdot \boldsymbol{a}\, dV = \oint_S \boldsymbol{a} \cdot d\boldsymbol{S}$.

Though very different from Eq. (1.12) in appearence, it is completely equivalent to Coulomb's law that states that the field of a charge is proportional to its magnitude and inversely proportional to the squared distance from it.

Substituting (1.17) into (1.39) gives *Poisson's equation* for the electrostatic potential[15],

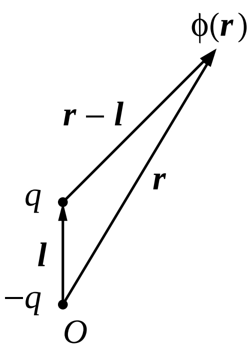
$$\nabla^2 \phi = -\rho/\varepsilon_0. \tag{1.40}$$

In vacuum where $\rho = 0$, this becomes *Laplace's equation*:

$$\nabla^2 \phi = 0. \tag{1.41}$$

Although this equation looks very simple, its solutions (known as harmonic functions) are not trivial. They have interesting properties and we will examine some of them in Ch. 3.

1.8 Dipoles



Let us determine the electrostatic potential of two charges, q and $-q$, separated by distance l. This system is known as the electric dipole[16]. If the negative charge is at the origin and the positive charge has position $\boldsymbol{l}$, the potential is

$$\phi(\boldsymbol{r}) = \frac{q}{4\pi\varepsilon_0|\boldsymbol{r}-\boldsymbol{l}|} - \frac{q}{4\pi\varepsilon_0 r} = \frac{q}{4\pi\varepsilon_0}\left(\frac{1}{|\boldsymbol{r}-\boldsymbol{l}|} - \frac{1}{r}\right). \tag{1.42}$$

The size of the dipole is usually regarded as small compared with other distances, so we are interested in the potential $\phi(\boldsymbol{r})$ for $r \gg l$. Hence, we expand the first term in brackets in (1.42) in Taylor series to first order in $\boldsymbol{l}$[17]:

$$\frac{1}{|\boldsymbol{r}-\boldsymbol{l}|} = \frac{1}{r} + \nabla\frac{1}{r}\cdot(-\boldsymbol{l}) + \ldots \simeq \frac{1}{r} + \frac{\boldsymbol{r}}{r^3}\cdot\boldsymbol{l},$$

where we used (1.15) and neglected terms with powers of $\boldsymbol{l}$ higher than unity. Substituting this into (1.42) gives the electrostatic potential of the dipole,

$$\phi(\boldsymbol{r}) = \frac{\boldsymbol{p}\cdot\boldsymbol{r}}{4\pi\varepsilon_0 r^3}, \tag{1.43}$$

where

$$\boldsymbol{p} = q\boldsymbol{l}, \tag{1.44}$$

is the *electric dipole moment* of the system. Note that vector $\boldsymbol{l}$ is directed from the negative to the positive charge.

---

[15] The scalar product of nabla operators is the *Laplacian*,

$$\nabla\cdot\nabla \equiv \nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}.$$

[16] This is a model of a *polar* molecule, such as NaCl or HCl, in which one of the atoms (Na or H) is positively charged, while the other one (Cl) is negatively charged.

[17] Expanding a scalar function in Taylor series gives $f(\boldsymbol{r}+\boldsymbol{a}) = f(\boldsymbol{r}) + \nabla f\cdot\boldsymbol{a} + \ldots$.

The corresponding electric field is found from Eq. (1.17),

$$\begin{aligned} \boldsymbol{E}(\boldsymbol{r}) = -\nabla\phi(\boldsymbol{r}) &= -\frac{1}{4\pi\varepsilon_0}\nabla\left[\frac{1}{r^3}\boldsymbol{p}\cdot\boldsymbol{r}\right] \\ &= -\frac{1}{4\pi\varepsilon_0}\left[\nabla\left(\frac{1}{r^3}\right)\boldsymbol{p}\cdot\boldsymbol{r} + \frac{1}{r^3}\nabla(\boldsymbol{p}\cdot\boldsymbol{r})\right] \\ &= -\frac{1}{4\pi\varepsilon_0}\left[-\frac{3}{r^4}\frac{\boldsymbol{r}}{r}\boldsymbol{p}\cdot\boldsymbol{r} + \frac{1}{r^3}\boldsymbol{p}\right], \end{aligned}$$

where we used product rule and the fact that $\nabla(\boldsymbol{p}\cdot\boldsymbol{r}) = \boldsymbol{p}$, which is easily verified in Cartesian coordinates (see also[9]). Hence, the field of the dipole is

$$\boldsymbol{E}(\boldsymbol{r}) = \frac{3(\boldsymbol{p}\cdot\boldsymbol{r})\boldsymbol{r} - \boldsymbol{p}r^2}{4\pi\varepsilon_0 r^5}. \tag{1.45}$$

The dipole moment can also be defined for a system of N charges q_i with positions $\boldsymbol{r}_i$, as

$$\boldsymbol{p} = \sum_{i=1}^{N} q_i\boldsymbol{r}_i. \tag{1.46}$$

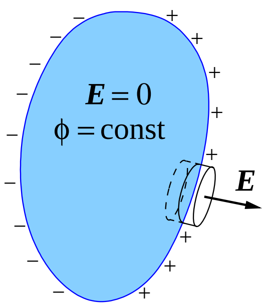
For the charges q and $-q$ separated by $\boldsymbol{l}$ this definition agrees with (1.44). It is also easy to check that the dipole moment (1.46) does not depend on the position of the origin if the total charge of the system is zero, i.e., $\sum_{i=1}^{N} q_i = 0$.

1.9 Conductors

Depending on their response to external electric fields, all materials can be divided into two classes, conductors and dielectrics (or insulators)[18].

A *conductor* is a material that contains many "free" charges, i.e., electrons that can move freely inside the material.

By contrast, the charges in dielectrics are "bound" (see Ch. 2).



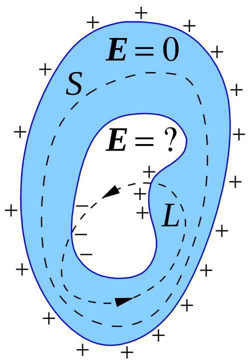
When a conductor is placed in an electric field, the charges will move until the field inside it becomes zero. In practice, this happens in a small fraction of a second.

As a result, in electrostatics, we have the following.

1. The field inside a conductor is zero, $\boldsymbol{E} = 0$.
2. By Eq. (1.39), the charge density inside the conductor is zero, $\rho = 0$, and the charges can only be found on its surface.
3. By Eq. (1.17), the potential of the conductor is constant, $\phi = \text{const}$.
4. The electric field immediately outside the conductor is perpendicular to its surface (since this surface is an equipotential surface, see end of Sec. 1.4), and is given by
$$E = \sigma/\varepsilon_0, \tag{1.47}$$
where σ is the surface charge density.

[18]There are also semiconductors. They behave as insulators at very low temperatures, but at raised temperatures they can conduct electricity, though not as well as metals.

To prove the latter point, we consider a small cylinder with flat surfaces parallel to the surface of the conductor and apply Gauss's law (1.34). The only contribution to the flux comes from the flat surface outside the conductor, and is given by EA, where A is the area of the flat surface. The charge inside the cylinder is σA, which yields (1.47).

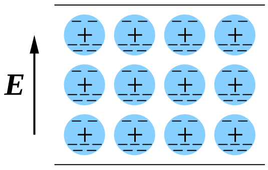


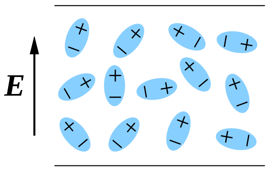
We can show that if a conductor has a cavity inside, the electric field inside it must be zero. First, the total charge on the inner surface is zero (by Gauss's law applied to a surface S that lies entirely within the conductor and encloses the cavity). Hence, the inner surface can only have equal amounts of positive and negative charges on it. If such charges were present, then a certain electric field would be present in the cavity (its field lines starting on positive charges and ending on negative charges). This means that the integral along the loop L that follows one of the field lines inside the cavity and closes inside the body of the conductor would be nonzero, $\oint_L \boldsymbol{E} \cdot d\boldsymbol{r} \neq 0$ (since the field inside the conductor is zero). However, the electric field is conservative, and such integral must always be zero. Hence, there can be no field inside the cavity and no charges on the inside surface[19].

[19]This principle lies behind the idea of electromagnetic shielding which was discovered by Michael Faraday in 1836. A metal enclosure or a metal mesh cage (known as *Faraday cage*) placed around a piece of equipment ensures that it is unaffected by external electric fields.

2 Dielectrics

2.1 Polarisation





When a dielectric is placed in an electric field, the positive and negative charges inside it separate. This effect is known as *polarisation*.

Polarisation occures when the negatively charged atomic electrons are displaced relative to the positively charged nuclei. In dielectrics which consist of molecular dipoles, polarisation can also occur as the dipoles develop a preferential orientation due to the external field[20].

Quantitatively, the *polarisation vector* is the dipole moment per unit volume. It can be written as

$$\boldsymbol{P} = n\boldsymbol{p}_m, \tag{2.1}$$

where n is the number density of atomic or molecular (mean) dipoles $\boldsymbol{p}_m$.

The polarisation increases with the electric field, and for *isotropic* dielectrics we write

$$\boldsymbol{P} = \varepsilon_0 \chi(E)\boldsymbol{E}, \tag{2.2}$$

where $\chi(E)$ is the *susceptibility*[21].

For many materials the vector $\boldsymbol{P}$ is proportional to $\boldsymbol{E}$ (if the field is not too strong), so one can assume that $\chi =$ const. Such dielectrics are called *linear*.

The SI units of the dipole moment are Cm, and the units of polarisation are C/m^2, i.e., the same as the units of $\varepsilon_0 E$ [see, e.g., Eq. (1.12) or (1.38)]. This shows that susceptibility χ defined by Eq. (2.2) is dimensionless.

2.2 Polarisation charge densities

Let us find the electrostatic potential produced by a polarised piece of dielectric of volume V.

A small volume element dV' of the dielectric at point $\boldsymbol{r}'$ has the dipole momement $\boldsymbol{P}(\boldsymbol{r}')dV'$. According to Eq. (1.43), its contribution to the potential at point $\boldsymbol{r}$ is

$$d\phi = \frac{\boldsymbol{P}(\boldsymbol{r}') \cdot (\boldsymbol{r} - \boldsymbol{r}')dV'}{4\pi\varepsilon_0 |\boldsymbol{r} - \boldsymbol{r}'|^3}.$$

[20]Without the external field the dipoles are randomly oriented. Thermal motion prevents complete orientation of the dipoles by the field. Molecular orientation takes time (e.g., about 10^{-11} s for water molecules at room temperature). A time-dependent, oscillating external field keeps reorienting the molecular dipoles. Heat generated in this process is dissipated in the medium. This is the principle behind heating of food (which contains large amounts of water) in a microwave oven.

[21]In anistropic dielectrics the charges are displaced more easily along some directions than along others, and the connection between the components of $\boldsymbol{P}$ and $\boldsymbol{E}$ is given by $P_i = \varepsilon_0 \sum_{j=1}^{3} \chi_{ij} E_j$, where χ_{ij} is the susceptibility tensor (represented by a 3×3 matrix).

The total potential of the dielectric is obtained by integrating $d\phi$ over its volume,

$$\phi(\boldsymbol{r}) = \frac{1}{4\pi\varepsilon_0}\int_V \frac{\boldsymbol{P}(\boldsymbol{r}')\cdot(\boldsymbol{r}-\boldsymbol{r}')}{|\boldsymbol{r}-\boldsymbol{r}'|^3}dV'. \tag{2.3}$$

This expression can be re-written using the relation

$$\nabla'\frac{1}{|\boldsymbol{r}-\boldsymbol{r}'|} = \frac{\boldsymbol{r}-\boldsymbol{r}'}{|\boldsymbol{r}-\boldsymbol{r}'|^3} \tag{2.4}$$

where the operator ∇' acts on $\boldsymbol{r}'$ [cf. Eq. (1.16)], which gives

$$\phi(\boldsymbol{r}) = \frac{1}{4\pi\varepsilon_0}\int_V \boldsymbol{P}(\boldsymbol{r}')\cdot\nabla'\frac{1}{|\boldsymbol{r}-\boldsymbol{r}'|}dV'.$$

We now use the vector identity ("product rule")

$$\nabla\cdot(\boldsymbol{P}f) = \boldsymbol{P}\cdot\nabla f + f\nabla\cdot\boldsymbol{P}, \tag{2.5}$$

in which f is any scalar function, and obtain

$$\phi(\boldsymbol{r}) = \frac{1}{4\pi\varepsilon_0}\int_V \nabla'\cdot\left[\boldsymbol{P}(\boldsymbol{r}')\frac{1}{|\boldsymbol{r}-\boldsymbol{r}'|}\right]dV' - \frac{1}{4\pi\varepsilon_0}\int_V \frac{1}{|\boldsymbol{r}-\boldsymbol{r}'|}\nabla'\cdot\boldsymbol{P}(\boldsymbol{r}')dV'.$$

Using Gauss's theorem we transform the first integral into the integral over the surface S of the dielectric, which gives

$$\phi(\boldsymbol{r}) = -\frac{1}{4\pi\varepsilon_0}\int_V \frac{\nabla'\cdot\boldsymbol{P}(\boldsymbol{r}')}{|\boldsymbol{r}-\boldsymbol{r}'|}dV' + \frac{1}{4\pi\varepsilon_0}\oint_S \frac{\boldsymbol{P}(\boldsymbol{r}')\cdot d\boldsymbol{S}'}{|\boldsymbol{r}-\boldsymbol{r}'|}. \tag{2.6}$$

Comparison with Eq. (1.23) shows that

$$\rho_p = -\nabla\cdot\boldsymbol{P}, \tag{2.7}$$

is the volume density of polarisation charges, and

$$\sigma_p = \boldsymbol{P}\cdot\boldsymbol{n}, \tag{2.8}$$

is their surface density, where $\boldsymbol{n}$ is the outward unit normal to the surface of the dielectric [as $d\boldsymbol{S}' = \boldsymbol{n}dS'$ in (2.6)].

2.3 Electric displacement

In the presence of dielectrics, Eq. (1.39) can be written as

$$\nabla\cdot\boldsymbol{E} = \frac{1}{\varepsilon_0}(\rho+\rho_p),$$

where ρ now denotes the density of *free charges* and ρ_p is the density of polarisation charges. Substituting (2.7) and rearranging, we obtain

$$\nabla\cdot\boldsymbol{D} = \rho, \tag{2.9}$$

where

$$\boldsymbol{D} = \varepsilon_0\boldsymbol{E} + \boldsymbol{P} \tag{2.10}$$

is the *electric displacement* vector.

Equation (2.9) replaces Eq. (1.39) in the presence of dialectrics. For an isotropic medium,

$$\boldsymbol{D} = \varepsilon_0(1+\chi)\boldsymbol{E} = \varepsilon\boldsymbol{E}, \tag{2.11}$$

where $\varepsilon = (1+\chi)\varepsilon_0$ is *permittivity*, and the ratio

$$\kappa \equiv \frac{\varepsilon}{\varepsilon_0} = 1+\chi, \tag{2.12}$$

is the *dielectric constant* (also known as the relative permittivity).

Using (1.17), we have from Eq. (2.11),

$$\boldsymbol{D} = -\varepsilon\nabla\phi, \tag{2.13}$$

and (2.9) yields

$$\nabla\cdot(\varepsilon\nabla\phi) = -\rho. \tag{2.14}$$

For a homogeneous dielectric ($\varepsilon = \text{const}$) we have Poisson's equation

$$\nabla^2\phi = -\rho/\varepsilon. \tag{2.15}$$

Note that inside a conductor we have $\boldsymbol{E} = 0$ and $\boldsymbol{P} = 0$, hence,

$$\boldsymbol{D} = 0. \tag{2.16}$$

2.4 Gauss's law

Integrating Eq. (2.9) over volume V bounded by surface S, and using Gauss's theorem, we obtain

$$\oint_S \boldsymbol{D}\cdot d\boldsymbol{S} = Q_f, \tag{2.17}$$

where $Q_f = \int_V \rho dV$ is the total *free* charge enclosed by S.

Equation (2.17) is the form of *Gauss's law* used in the presence of dielectrics.

2.5 Electric displacement outside a conducting surface

Consider a conductor embedded in a dielectric and apply Gauss's law (2.17) to a cylindrical surface with faces parallel to the surface (cf. diagram in Sec. 1.9). Similarly to Eq. (1.47), we find the normal component of the electric displacement near the conductor's surface,

$$D_n = \sigma, \tag{2.18}$$

where σ is the surface density of free charge on the conductor.

The electric field outside a conducting surface is normal to it,

$$\boldsymbol{E} = E\boldsymbol{n}. \tag{2.19}$$

Using (2.11) for an isotropic dielectric,

$$\boldsymbol{D} = \varepsilon E\boldsymbol{n}, \tag{2.20}$$

we obtain

$$E = \sigma/\varepsilon, \tag{2.21}$$

which replaces Eq. (1.47) when a dielectric is present.

2.6 Boundary conditions

At the boundary between two dielectrics, the tangential components

$$E_{1t} - E_{2t} = 0, \tag{2.22}$$

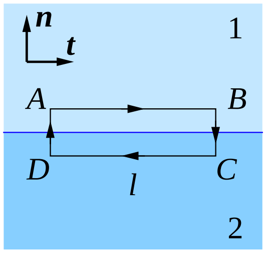
and the normal components

$$D_{1n} - D_{2n} = \sigma, \tag{2.23}$$

where the subscripts 1 and 2 refer to quantities in dielectrics 1 and 2, respectively, the unit normal $\boldsymbol{n}$ is from dielectric 2 into 1, and σ is the surface density of free charge on the boundary.

Proof. To prove (2.22), we use the fact that the electric field is conservative, and hence that

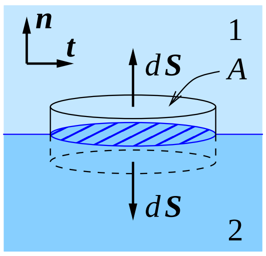
$$\oint_L \boldsymbol{E} \cdot d\boldsymbol{r} = 0 \tag{2.24}$$



for any closed loop L.[22] We choose L as a rectangle $ABCD$ with long sides AB and CD parallel to the interface, as shown in the diagram. We also set $AB = CD \equiv l \gg BC = DA$, which ensures that the contribution of BC and DA to the integral can be neglected. At the same time we assume that l is sufficiently small, so that the electric field remains approximately constant on AB and CD. On AB we have $d\boldsymbol{r} = \boldsymbol{t}dr$, and on CD $d\boldsymbol{r} = -\boldsymbol{t}dr$, where $\boldsymbol{t}$ is a unit vector tangential to the surface (see diagram). Hence, we obtain

$$\begin{aligned}\oint_L \boldsymbol{E} \cdot d\boldsymbol{r} &= \int_A^B \boldsymbol{E} \cdot d\boldsymbol{r} - \int_A^B \boldsymbol{E} \cdot d\boldsymbol{r} = \int_A^B \boldsymbol{E} \cdot \boldsymbol{t}dr - \int_A^B \boldsymbol{E} \cdot \boldsymbol{t}dr \\ &= E_{1t} \int_A^B dr - E_{2t} \int_C^D dr = (E_{1t} - E_{2t})l = 0,\end{aligned}$$

which proves (2.22).



Equation (2.23) is proved by using Gauss's law (2.17) and choosing the surface S as a cylinder with flat faces parallel to the interface (see diagram). We assume that the height of the cylinder is small, so that the contribution of the flux across the curved surface can be neglected. In this case only the flux across the two flat surfaces (each of area A) must be included, and we have

$$\begin{aligned}\oint_S \boldsymbol{D} \cdot d\boldsymbol{S} &= \int_{\text{top}} \boldsymbol{D} \cdot d\boldsymbol{S} + \int_{\text{bottom}} \boldsymbol{D} \cdot d\boldsymbol{S} \\ &= \int_{\text{top}} \boldsymbol{D} \cdot \boldsymbol{n}dS + \int_{\text{bottom}} \boldsymbol{D} \cdot (-\boldsymbol{n})dS = (D_{1n} - D_{2n})A = \sigma A,\end{aligned}$$

where the right-hand side is the free charge enclosed by the cylinder. Dividing both sides of the last equality by A gives (2.23).

The potential ϕ is continuous across the interface. This follows from (1.19), since the integral $\int_A^B \boldsymbol{E} \cdot d\boldsymbol{r}$ vanishes if we consider the points A and B in dielectrics 1 and 2, respectively, but arbitrarily close to each other.

[22] This follows from Eq. (1.19) if we consider a path for which points A and B coincide.

2.7 Capacitors

A capacitor (or condenser) consists of two conductors, called plates, separated by vacuum or dielectric. It is *charged* by placing a positive charge Q on one of them and the charge $-Q$ on the other.

The electric field in this system is proportional to Q, and so is the potential difference $\Delta\phi \equiv V$ between the plates. Hence, we introduce the *capacitance*,

$$C = Q/V, \tag{2.25}$$

depends only on the shape and position of the plates and the dielectrics used.

The SI unit of capacitance is *farad* (F), 1 F = 1 CV^{-1}.

Example 1: Isolated sphere.

Let us determine the capacitance of a conducting sphere of radius a surrounded by vacuum. The second plate of such capacitor is assumed to be at infinity.

If the sphere carries charge Q, the electric field outside it is given by[23]

$$\boldsymbol{E} = \frac{Q}{4\pi\varepsilon_0 r^2}\hat{\boldsymbol{r}}.$$

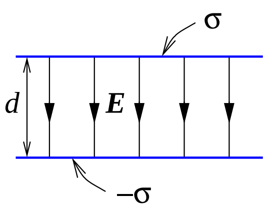
The potential difference between the sphere and a point at infinity is

$$V = \int_a^\infty \boldsymbol{E}\cdot d\boldsymbol{r} = \int_a^\infty \frac{Q}{4\pi\varepsilon_0 r^2}dr = \frac{Q}{4\pi\varepsilon_0 a},$$

which gives[24]

$$C = \frac{Q}{V} = 4\pi\varepsilon_0 a. \tag{2.26}$$

Example 2: Parallel plate capacitor.



A parallel plate capacitor consists of two parallel planar conductors separated by distance d. The electric field in this system is found as a superposition of the fields of two uniformly charged planes (Example 2 in Sec. 1.7) with surface charge densities σ and $-\sigma$. The fields outside the plates cancel, and the fields between them add to give

$$E = \sigma/\varepsilon_0. \tag{2.27}$$

The potential difference between the plates is

$$V = \int_L \boldsymbol{E}\cdot d\boldsymbol{r} = Ed = \frac{\sigma d}{\varepsilon_0}, \tag{2.28}$$

where the path L connects any point on the positively charged plate with a point on the negatively charged plate, e.g., is a straight line perpendicular to the plates. If the total charge on the positively charged plate is Q, then

[23]This result immediatly follows from the spherical symmetry of the system and application of Gauss's law (1.34) to a spherical Gaussian surface of radius r.

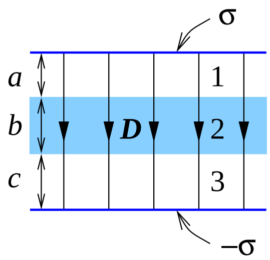
[24]Note that in CGS the capacitance of an isolated sphere is equal to its radius, $C = a$.

$\sigma = Q/A$, where A is the area of the plate. Substituting this into (2.28), we find the capacitance

$$C = \frac{Q}{V} = \frac{\varepsilon_0 A}{d}. \tag{2.29}$$

This result is valid for capacitors with plates whose linear dimensions are much greater than the spacing between them, which allows one to neglect the edge effects (e.g., that Eq. (2.27) does not hold near the edges of the plates).

Example 3: Parallel plate capacitor with dielectric.



Consider a parallel plate capacitor in which a layer of dielectric of thickness b and dielectric constant κ is inserted parallel to the plates and separated from them by vacuum gaps of thickness a and c (see diagram).

If we neglect the edge effects, $\boldsymbol{D} = 0$ outside the plates, and

$$D = \sigma \tag{2.30}$$

between the plates, where $\sigma = Q/A$ and A is the area of each plate[25]. From (2.11) the electric field in the gaps (denoted 1 and 3) is

$$E_1 = E_3 = D/\varepsilon_0 = \sigma/\varepsilon_0, \tag{2.31}$$

and the field in the dielectric (denoted 2) is

$$E_2 = D/\varepsilon_0\kappa = \sigma/\varepsilon_0\kappa. \tag{2.32}$$

The potential difference between plates is

$$V = E_1 a + E_2 b + E_3 c = \frac{\sigma}{\varepsilon_0}(a + b/\kappa + c) = \frac{Q}{\varepsilon_0 A}(a + b/\kappa + c), \tag{2.33}$$

so that

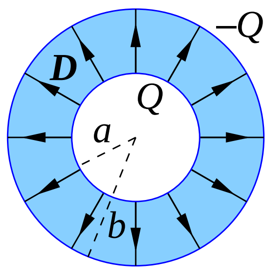
$$C = \frac{Q}{V} = \frac{\varepsilon_0 A}{a + b/\kappa + c}. \tag{2.34}$$

Note that for $b = 0$ (no dielectric) and $a + c = d$, the above result coincides with Eq. (2.29). However, if we set $a = c = 0$ and $b = d$ in Eq. (2.34), we obtain

$$C = \frac{\kappa\varepsilon_0 A}{d} = \frac{\varepsilon A}{d}. \tag{2.35}$$

which shows that the capacitance of a capacitor filled with dielectric is κ times greater than that with vacuum between the plates.

Example 4: Concentric spherical capacitor.



A spherical capacitor consists of two conducting spheres of radii a and b ($a < b$) with dielectric of constant κ in between (diagram shows the cross section).

We assume that the inner sphere carries charge Q and the outer one $-Q$. Due to the symmetry of the system, the fields $\boldsymbol{E}$ and $\boldsymbol{D}$ are radial, i.e., $\boldsymbol{E} = E(r)\hat{\boldsymbol{r}}$ and $\boldsymbol{D} = D(r)\hat{\boldsymbol{r}}$. Applying Gauss's law (2.17) to a sphere of radius r, we find for $a < r < b$,

$$D(r) = \frac{Q}{4\pi r^2}, \tag{2.36}$$

[25] This result can be obtained using the symmetry of the system and Gauss's law (2.17).

and from (2.11),

$$E(r) = \frac{Q}{4\pi\varepsilon_0\kappa r^2}, \tag{2.37}$$

while the fields inside the inner sphere and outside the outer sphere are zero.

Using $\boldsymbol{E} = -\nabla\phi = -(\partial\phi/\partial r)\hat{\boldsymbol{r}}$ (i.e., keeping only the radial component of the gradient in spherical coordinates, since $\boldsymbol{E}$ is radial), we have

$$\frac{d\phi}{dr} = -\frac{Q}{4\pi\varepsilon_0\kappa r^2},$$

so that

$$\phi(r) = \frac{Q}{4\pi\varepsilon_0\kappa r} + \text{const.} \tag{2.38}$$

This gives the potential difference[26]

$$V = \phi(a) - \phi(b) = \frac{Q}{4\pi\varepsilon_0\kappa}\left(\frac{1}{a} - \frac{1}{b}\right),$$

and capacitance

$$C = \frac{Q}{V} = \frac{4\pi\varepsilon_0\kappa ab}{b-a}. \tag{2.39}$$

Note that if we let $b \to \infty$ and set $\kappa = 1$, the above formula reproduces Eq. (2.26), as expected. If on the other hand, we consider a thin spherical capacitor for which $b \approx a$ and $b - a \equiv d \ll a, b$, then Eq. (2.39) gives

$$C = \frac{4\pi\varepsilon_0\kappa a^2}{b-a} = \frac{\varepsilon_0\kappa A}{d},$$

where $A = 4\pi a^2$ is the surface area of the sphere. This result is in agreement with the capacitance of a parallel plate capacitor filled with dielectric.

[26]This could also be obtained from $V = \int_a^b \boldsymbol{E}\cdot d\boldsymbol{r} = \int_a^b E(r)dr$ using Eq. (2.37).

3 Solutions of electrostatic problems – Potential theory

3.1 Properties of the electrostatic potential

In this section we summarise the properties of the electrostatic potential derived in chapters 1 and 2.

According to Eqs. (2.14) and (2.15), for linear isotropic media, we have

$$\nabla \cdot (\varepsilon \nabla \phi) = -\rho, \tag{3.1}$$

and if $\varepsilon = \text{const}$, then

$$\nabla^2 \phi = -\rho/\varepsilon, \tag{3.2}$$

where ρ is the volume density of free charges.

For $\rho = 0$ the potential satisfies Laplace's equation

$$\nabla^2 \phi = 0. \tag{3.3}$$

Additionally:

(i) On a conductor

$$\phi = \text{const}. \tag{3.4}$$

(ii) The surface charge density on the surface of a conductor is [by Eqs. (1.17) and (2.21)],

$$\sigma = -\varepsilon \frac{\partial \phi}{\partial n}, \tag{3.5}$$

where $\partial\phi/\partial n$ is the derivative along the direction of the outward normal ($\partial\phi/\partial n = \boldsymbol{n} \cdot \nabla\phi$), taken at the conductor's surface.

Hence, the total charge on the conductor is

$$Q = -\oint_S \varepsilon \frac{\partial \phi}{\partial n} dS, \tag{3.6}$$

where the integral is over the surface of the conductor.

(iii) For a finite system of charges

$$\phi \to 0 \quad \text{at infinity}, \tag{3.7}$$

see, e.g., Eq. (1.18) for a system of point charges. Note that the potential decreases as $1/r$ if the total charge of the system is nonzero, and faster if the system is neutral [cf. Eq. (1.43) for the potential of the dipole].

(iv) If there is a uniform field $\boldsymbol{E}_0$ in the z direction at infinity, then

$$\phi \simeq -E_0 z \quad \text{as} \quad z \to \pm\infty. \tag{3.8}$$

Indeed, for this potential we have $\boldsymbol{E} = -\nabla(-E_0 z) = E_0 \boldsymbol{k} \equiv \boldsymbol{E_0}$.

(v) Potential ϕ has no singularities, except at point charges,

$$\phi \simeq \frac{q}{4\pi\varepsilon}\frac{1}{r} \quad \text{as} \quad r \to 0, \tag{3.9}$$

where r is the distance from a point charge q.

(vi) At the boundary between two media the potential is continuous,

$$\phi_1 = \phi_2, \tag{3.10}$$

where ϕ_1 and ϕ_2 are the potentials in the two media (see end of Sec. 2.6). Also, from (2.23), using (2.13), we have

$$\varepsilon_2 \frac{\partial \phi_2}{\partial n} = \varepsilon_1 \frac{\partial \phi_1}{\partial n} + \sigma, \tag{3.11}$$

where the direction of $\boldsymbol{n}$ is from 2 into 1, and σ is the surface density of free charges at the boundary.

3.2 Uniqueness theorem

Consider a system which consists of a set of conductors, each carrying a given charge, another set of conductors, each kept at a given potential, and a given volume distribution of charge, in the presence of given linear isotropic dielectrics. Then there cannot be more than one potential function ϕ for this system. This statement is known as the *uniqueness theorem*.

The uniqueness theorem means that if, for a given system, we have found a potential that satisfies the equations and conditions listed in Sec. 3.1, this potential gives the solution to the problem.

Proof. Suppose there are two potentials, ϕ_1 and ϕ_2, which satisfy Eq. (3.1), i.e.,

$$\nabla \cdot (\varepsilon \nabla \phi_1) = -\rho \quad \text{and} \quad \nabla \cdot (\varepsilon \nabla \phi_2) = -\rho, \tag{3.12}$$

as well as the conditions (i)–(iii) and (v) from Sec. 3.1.

Let

$$\phi = \phi_1 - \phi_2. \tag{3.13}$$

From (3.12), we have

$$\nabla \cdot (\varepsilon \nabla \phi) = 0. \tag{3.14}$$

Since ϕ_1 and ϕ_2 satisfy identical conditions [given by Eq. (3.4) or by Eq. (3.6), as well as (3.7) and (3.9)], we find

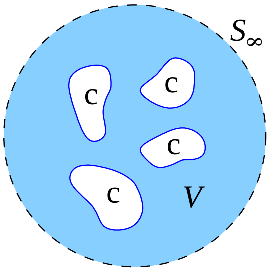
(i) $\phi = 0$ on the surfaces of all conductors with fixed potentials;

(ii) on any conductor with a given charge,

$$\oint_S \varepsilon \frac{\partial \phi}{\partial n} dS = 0, \tag{3.15}$$

(iii) $\phi \to 0$ as $1/r$ or faster at infinity;

(iv) ϕ has no singularities.



Consider the integral

$$\int_V \varepsilon \nabla\phi \cdot \nabla\phi \, dV, \tag{3.16}$$

over the volume outside the conductors (c) and bounded by a *sphere at infinity* S_∞ (which is a sphere whose radius can be made arbitrarily large). From

$$\nabla \cdot (\phi\varepsilon\nabla\phi) = \varepsilon\nabla\phi \cdot \nabla\phi + \phi\nabla \cdot (\varepsilon\nabla\phi) = \varepsilon\nabla\phi \cdot \nabla\phi, \tag{3.17}$$

where we used (3.14), we have (using Gauss's theorem):

$$\begin{aligned} \int_V \varepsilon\nabla\phi \cdot \nabla\phi \, dV = \int_V \nabla \cdot (\phi\varepsilon\nabla\phi) dV = \oint_S \phi\varepsilon\nabla\phi \cdot d\boldsymbol{S} = \oint_S \phi\varepsilon\frac{\partial\phi}{\partial n} dS \\ = \sum_i \oint_{S_i} \phi\varepsilon\frac{\partial\phi}{\partial n} dS + \oint_{S_\infty} \phi\varepsilon\frac{\partial\phi}{\partial n} dS, \end{aligned} \tag{3.18}$$

where the integral is over the surfaces of all conductors (S_i) and S_∞ that bound V, and the direction of the normal $\boldsymbol{n}$ is out of volume V.

Note that for the contribution of the surface of each conductor S_i to (3.18),

$$\oint_{S_i} \phi\varepsilon\frac{\partial\phi}{\partial n} dS = \phi \oint_{S_i} \varepsilon\frac{\partial\phi}{\partial n} dS,$$

since $\phi = \text{const}$ there, and all such contributions vanish because of (i) or (ii).

The contribution of the sphere at infinity to (3.18) vanishes because of (iii). Indeed, for $\phi \sim 1/r$, its derivative behaves as $\nabla\phi \sim 1/r^2$, and the integrand decreases as $1/r^3$ with the radius r of the sphere, while the surface area of S_∞ increases only as $\sim r^2$. Hence, in the limit $r \to \infty$ this integral is zero.

Therefore,

$$\int_V \varepsilon(\nabla\phi)^2 dV = 0, \tag{3.19}$$

and since $\varepsilon > 0$,

$$\nabla\phi = 0$$

everywhere, so that

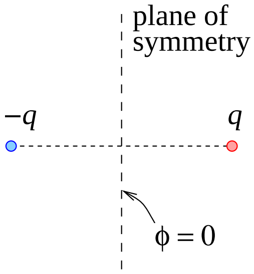
$$\phi = \phi_1 - \phi_2 = \text{const}.$$

Since $\phi \to 0$ at infinity, this constant must be zero, which means that

$$\phi_1 = \phi_2. \tag{3.20}$$

3.3 Method of images

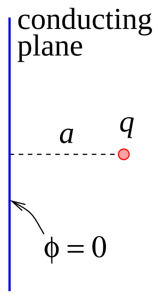
Images in a plane.



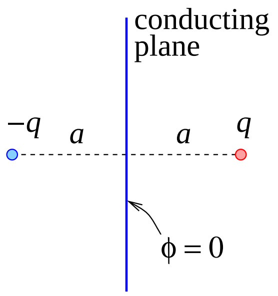
The electrostatic potential of two charges q and $-q$ is given by

$$\phi = \frac{1}{4\pi\varepsilon_0}\left(\frac{q}{r_1} - \frac{q}{r_2}\right), \tag{3.21}$$

where r_1 and r_2 are the distances from the two charges. For any point that lies in the plane of symmetry of the system (i.e., the plane that bisects the line between q and $-q$ and is perpendiculr to it), $r_1 = r_2$, which gives $\phi = 0$.



Consider the problem of finding the electric field for a system that consists of a point charge $q > 0$ at a distance a from an infinite conducting plane held at zero potential. The presence of the positive charge will cause accumulation of negative charges in the plane near the point closest to charge q. The corresponding surface charge density σ will be such that the condition $\phi = 0$ is satisfied everywhere in the plane. However, what is this charge distribution?



Looking at Eq. (3.21), it is easy to see that this potential provides the solution for the electrostatic problem in the half-space where the positive charge is, if we assume that the charge $-q$ is placed symmetrically to q on the other side of the plane. Indeed, (3.21) satisfies the condition $\phi = 0$ on the plane and has the correct behaviour near the point charge q [see (i) and (v) in Sec. 3.1].

This means that in the presence of q, the charges in the plane distribute themselves in such a way that their potential coincides with the potential of charge $-q$ placed at a distance a behind the plane. This fictitious charge is known as the *image charge*[27].

If we place the origin in the plane and choose the z axis perpendicular to it and through the charge q, the potential (3.21) will be given by

$$\phi(\boldsymbol{r}) = \frac{q}{4\pi\varepsilon_0}\left(\frac{1}{\sqrt{\rho^2 + (a-z)^2}} - \frac{1}{\sqrt{\rho^2 + (a+z)^2}}\right), \tag{3.22}$$

where ρ is the distance from the z axis[28].

The surface charge density on the plane can be found from (3.22) using Eq. (3.5), where the direction of $\boldsymbol{n}$ is along the z axis. This gives

$$\sigma(\rho) = -\,\varepsilon_0 \frac{\partial\phi}{\partial z}\bigg|_{z=0} = -\frac{aq}{2\pi(a^2+\rho^2)^{3/2}}. \tag{3.23}$$

As expected, the density of surface charges on the conducting plane is negative (for $q > 0$) and is largest for $\rho = 0$, at the foot of the perpendicular from q[29].

The electric field in the system coincides with the field of two point charges, with the negative (image) charge contribution being due to the charges in the plane. Hence, the force acting on charge q is towards the plane and given by

$$F = \frac{q^2}{4\pi\varepsilon_0(2a)^2} = \frac{q^2}{16\pi\varepsilon_0 a^2}. \tag{3.24}$$

Images with spheres and cylinders.

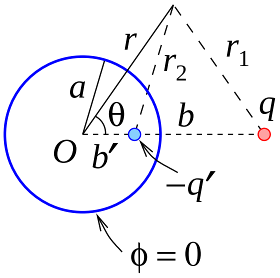
Consider a system which consists of charge q at a distance b from the centre of a conducting sphere of radius a, held at zero potential (with $b > a$). We

[27] Equation (3.21) also provides the solution to the problem of a negative charge outside an infinite conducting plane, in which case it is the positive charge q that plays the role of the image charge.

[28] In cylindrical coordinates ρ, ψ and z, the Cartesian components $\boldsymbol{r} = (\rho\cos\psi, \rho\sin\psi, z)$, $\boldsymbol{a} = (0,0,a)$ (for the position of charge q), with $r_1 = |\boldsymbol{r} - \boldsymbol{a}|$ and $r_2 = |\boldsymbol{r} + \boldsymbol{a}|$.

[29] It is easy to show that the total charge on the plane is $-q$, e.g., using integration in plane polar coordinates: $\int \sigma dS = \int_0^{2\pi}\int_0^{\infty} \sigma(\rho)\rho\, d\rho d\psi = -q$.

will show that by placing an image charge $-q'$ inside the sphere at a distance b' from its centre, we can make $\phi = 0$ on the surface of the sphere.



The potential of the real and image charges is

$$\phi(\boldsymbol{r}) = \frac{1}{4\pi\varepsilon_0}\left(\frac{q}{r_1} - \frac{q'}{r_2}\right), \tag{3.25}$$

where r_1 and r_2 are the distances from the charges q and $-q'$, respectively. If vector $\boldsymbol{r}$ forms angle θ with the direction towards q, we have

$$\phi(\boldsymbol{r}) = \frac{1}{4\pi\varepsilon_0}\left(\frac{q}{\sqrt{r^2+b^2-2rb\cos\theta}} - \frac{q'}{\sqrt{r^2+b'^2-2rb'\cos\theta}}\right). \tag{3.26}$$

To set this potential to zero on the surface of the sphere ($r = a$), we require

$$\frac{q}{\sqrt{a^2+b^2-2ab\cos\theta}} = \frac{q'}{\sqrt{a^2+b'^2-2ab'\cos\theta}}, \tag{3.27}$$

or equivalently,

$$\sqrt{\frac{a^2+b^2}{q^2} - \frac{2ab}{q^2}\cos\theta} = \sqrt{\frac{a^2+b'^2}{q'^2} - \frac{2ab'}{q'^2}\cos\theta}.$$

For this to be satisfied for all angles θ, we must have

$$\frac{a^2+b^2}{q^2} = \frac{a^2+b'^2}{q'^2} \quad \text{and} \quad \frac{ab}{q^2} = \frac{ab'}{q'^2}.$$

Solving these equations simultaneously, we find[30]

$$q' = \frac{aq}{b}, \quad b' = \frac{a^2}{b}, \tag{3.28}$$

for the magnitude and position of the image charge[31].

Substitution of (3.28) into Eq. (3.28) gives the potential outside the sphere as

$$\phi = \frac{1}{4\pi\varepsilon_0}\left(\frac{q}{\sqrt{r^2+b^2-2rb\cos\theta}} - \frac{q}{\sqrt{(rb/a)^2+a^2-2rb\cos\theta}}\right). \tag{3.29}$$

[30] A simple way to see this is to introduce the position vectors $\boldsymbol{b}$ and $\boldsymbol{b}'$ of the charges q and $-q'$, and vector $\boldsymbol{a}$ for an arbitrary point on the sphere. Condition (3.27) then reads

$$\frac{q}{|\boldsymbol{a}-\boldsymbol{b}|} = \frac{q'}{|\boldsymbol{b}'-\boldsymbol{a}|} \quad \text{or} \quad \left|\frac{\boldsymbol{a}}{q} - \frac{\boldsymbol{b}}{q}\right| = \left|\frac{\boldsymbol{b}'}{q'} - \frac{\boldsymbol{a}}{q'}\right|.$$

Since the angles between the vectors $\boldsymbol{a}/q$ and $\boldsymbol{b}/q$, and $\boldsymbol{b}'/q'$ and $\boldsymbol{a}/q'$ are the same, the above equation will hold if their lengths are pairwise equal, i.e., $a/q = b'/q'$ and $b/q = a/q'$, which gives (3.28).

[31] If charge q is just outside the sphere, i.e., b is only slighly greater than a, $b = a(1+\xi)$, where $\xi \ll 1$, then from Eq. (3.28) $q' = q/(1+\xi) \simeq q$ and $b' = a/(1+\xi) \simeq a(1-\xi)$. [Recall the binomial expansion $(1+\xi)^{-1} = 1 - \xi + \xi^2 - \xi^3 + \ldots$.] We see that in this case the magnitude of the image charge is equal to that of charge q, and that its distance from the surface of the sphere is the same as that of charge q (both equal to ξa), as in the case of an image charge in a plane.

Using this, we find the charge density on the sphere from Eq. (3.5):

$$\sigma = -\varepsilon_0 \frac{\partial \phi}{\partial n} = -\varepsilon_0 \left.\frac{\partial \phi}{\partial r}\right|_{r=a} = -\frac{q(b^2 - a^2)}{4\pi a(a^2 + b^2 - 2ab\cos\theta)^{3/2}}. \tag{3.30}$$

The total charge on the sphere $\int \sigma dS = \int_0^{2\pi} \int_0^{\pi} \sigma a^2 \sin\theta d\theta d\psi = -qa/b = -q'$, as it should be. (Apply Gauss's law to a surface enclosing the sphere.)

If the potential of the sphere is nonzero, we can place a second image charge q'' at the centre of the sphere, making the potential

$$\phi = \frac{1}{4\pi\varepsilon_0}\left(\frac{q}{\sqrt{r^2 + b^2 - 2rb\cos\theta}} - \frac{q}{\sqrt{(rb/a)^2 + a^2 - 2rb\cos\theta}} + \frac{q''}{r}\right). \tag{3.31}$$

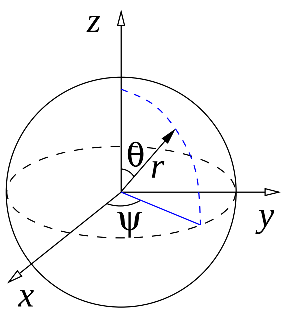
The total image charge enclosed by the sphere is now $-q' + q''$. By setting $q'' = Q + aq/b$, we obtain the potential for a conducting sphere of radius a, carrying charge Q, and a point charge q at a distance b from its centre.

A similar system of image charges can be constructed for a uniformly charged line parallel to a conducting cylinder.

Combining images.

By combining two or more sets of images, more complicated problems can be solved, e.g., that of a dipole near a conducting plane, whose image is a dipole.

3.4 Solution of Laplace's equation in spherical polar coordinates



Electrostatics problems for linear isotropic media with constant permittivity and no free volume charges give rise to Laplace's equation (3.3). Some of them can be solved by considering Laplace's equation in spherical polar coordinates,

$$\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2 \frac{\partial \phi}{\partial r}\right) + \frac{1}{r^2 \sin\theta}\frac{\partial}{\partial \theta}\left(\sin\theta \frac{\partial \phi}{\partial \theta}\right) + \frac{1}{r^2 \sin^2\theta}\frac{\partial^2 \phi}{\partial \psi^2} = 0. \tag{3.32}$$

If the system is symmetric about the z axis, i.e., $\phi = \phi(r, \theta)$, it becomes

$$\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2 \frac{\partial \phi}{\partial r}\right) + \frac{1}{r^2 \sin\theta}\frac{\partial}{\partial \theta}\left(\sin\theta \frac{\partial \phi}{\partial \theta}\right) = 0. \tag{3.33}$$

Particular solutions of this equation can be found by variable separation, seeking solution in the form

$$\phi(r, \theta) = R(r)\Theta(\theta). \tag{3.34}$$

Substituting this into Eq. (3.33) and dividing it through by $R(r)\Theta(\theta)$ gives

$$\underbrace{\frac{1}{R(r)}\frac{1}{r^2}\frac{d}{dr}\left(r^2 \frac{dR}{dr}\right)}_{\text{depends only on } r} + \frac{1}{r^2}\underbrace{\frac{1}{\Theta(\theta)}\frac{1}{\sin\theta}\frac{d}{d\theta}\left(\sin\theta \frac{d\Theta}{d\theta}\right)}_{\text{depends only on } \theta} = 0. \tag{3.35}$$

For this equation to hold for all r and θ, the term that depends on θ only must be equal to a constant. Denoting this constant $-\lambda$, we obtain the equation for $\Theta(\theta)$,

$$\frac{1}{\sin\theta}\frac{d}{d\theta}\left(\sin\theta \frac{d\Theta}{d\theta}\right) + \lambda\Theta(\theta) = 0. \tag{3.36}$$

One can show that this equation has solutions that are finite on the interval $0 \leq \theta \leq \pi$ only for $\lambda = n(n+1)$, where $n = 0, 1, 2, \ldots$. Written in the standartd form, these solutions are known as *Legendre polynomials* $P_n(\cos\theta)$[32]. Explicitly, the first three of them are

$$P_0(\cos\theta) = 1, \tag{3.37}$$

$$P_1(\cos\theta) = \cos\theta, \tag{3.38}$$

$$P_2(\cos\theta) = \frac{1}{2}(3\cos^2\theta - 1). \tag{3.39}$$

Legendre polynomials with even n are even functions, and those with odd n are odd functions of $\cos\theta$. They have the following orthogonality property,

$$\int_0^\pi P_n(\cos\theta)P_m(\cos\theta)\sin\theta d\theta = \frac{2\delta_{nm}}{2n+1}, \tag{3.40}$$

where $\delta_{nm} = 1$ for $n = m$ and 0 for $n \neq m$ (Kronecker delta). In particular,

$$\int_0^\pi P_n(\cos\theta)\sin\theta d\theta = 0 \quad \text{for} \quad n \neq 0. \tag{3.41}$$

For $\theta = 0$ or π,

$$P_n(\cos 0) = P_n(1) = 1, \tag{3.42}$$

$$P_n(\cos\pi) = P_n(-1) = (-1)^n. \tag{3.43}$$

Substituting $-\lambda = -n(n+1)$ in place of the θ-dependent term in Eq. (3.35) yields the equation for $R(r)$:

$$\frac{d}{dr}\left(r^2\frac{dR}{dr}\right) - n(n+1)R(r) = 0. \tag{3.44}$$

We seek its solution in the form $R(r) = r^s$, which gives a quadratic equation

$$s(s+1) = n(n+1),$$

with solutions $s = n$ and $s = -n-1$, corresponding to $R(r) = r^n$ and $R(r) = r^{-n-1}$. Combining the radial and angular parts we see that Laplace's equation (3.33) possesses independent solutions of the form

$$r^n P_n(\cos\theta) \quad \text{and} \quad r^{-n-1}P_n(\cos\theta) \quad (n = 0, 1, 2, \ldots).$$

The most general solution of equation (3.33) is obtained by taking a linear combination of these with arbitrary coefficients A_n and B_n:

$$\phi = \sum_{n=0}^{\infty}\left(A_n r^n + \frac{B_n}{r^{n+1}}\right)P_n(\cos\theta). \tag{3.45}$$

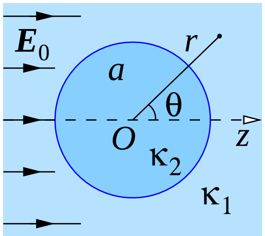
[32]By variable substitution $x = \cos\theta$, Eq. (3.36) is cast in the form

$$\frac{d}{dx}\left[(1-x^2)\frac{d\Theta}{dx}\right] + \lambda\Theta(x) = 0.$$

Seeking its solution in the form $\Theta(x) = \sum_{k=0}^{\infty} a_k x^k$ gives the recurrency relation $a_{k+2} = a_k[k(k+1) - \lambda]/[(k+1)(k+2)]$, which leads to polynomial solutions (finite on $-1 \leq x \leq 1$) only if $\lambda = n(n+1)$, where n is a non-negative integer (the degree of the polynomial).

Dielectric sphere in a uniform field.

Consider a dielectric sphere of radius a and constant κ_2 in a medium with dielectric constant κ_1 in a uniform electric field $\boldsymbol{E}_0$ along the z axis.



Let us choose the z axis in the direction of $\boldsymbol{E}_0$ through the centre of the sphere. The system is symmetric about the z axis (i.e., axially symmetric), and the potential inside and outside the sphere satisfies Eq. (3.33). Hence, we can write the potential ϕ_1 in dielectric 1 and ϕ_2 in dielectric 2 in the form (3.45):

$$\phi_1 = \sum_{n=0}^{\infty} \left(A_n^1 r^n + \frac{B_n^1}{r^{n+1}} \right) P_n(\cos\theta) \qquad (r > a), \tag{3.46}$$

$$\phi_2 = \sum_{n=0}^{\infty} \left(A_n^2 r^n + \frac{B_n^2}{r^{n+1}} \right) P_n(\cos\theta) \qquad (r < a), \tag{3.47}$$

where A_n^1, B_n^1, A_n^2 and B_n^2 are some coefficients.

From the boundary condition at infinity ($\phi_1 \simeq -E_0 z = -E_0 r\cos\theta$) we obtain

$$A_1^1 = -E_0, \qquad A_n^1 = 0 \quad \text{for} \quad n \geq 2,$$

since there no terms in (3.46) may increase faster than r at large distances. The potential (3.47) should have no singularity at the origin ($r = 0$), so

$$B_n^2 = 0 \quad \text{for all} \quad n.$$

Hence, we have

$$\phi_1 = A_0^1 - E_0 r\cos\theta + \sum_{n=0}^{\infty} \frac{B_n^1}{r^{n+1}} P_n(\cos\theta), \tag{3.48}$$

$$\phi_2 = \sum_{n=0}^{\infty} A_n^2 r^n P_n(\cos\theta). \tag{3.49}$$

The boundary conditions on the sphere $r = a$ are (3.10) and (3.11) (with $\sigma = 0$),

$$\phi_1 = \phi_2 \quad \text{and} \quad \kappa_1 \frac{\partial \phi_1}{\partial r}\bigg|_{r=a} = \kappa_2 \frac{\partial \phi_2}{\partial r}\bigg|_{r=a}.$$

Substituting ϕ_1 and ϕ_2 from Eqs.(3.48) and (3.49) into these equation, we have

$$A_0^1 - E_0 a\cos\theta + \sum_{n=0}^{\infty} \frac{B_n^1}{a^{n+1}} P_n(\cos\theta) = \sum_{n=0}^{\infty} A_n^2 a^n P_n(\cos\theta),$$

$$\kappa_1 \left[-E_0\cos\theta - \sum_{n=0}^{\infty} \frac{(n+1)B_n^1}{a^{n+2}} P_n(\cos\theta) \right] = \kappa_2 \left[\sum_{n=0}^{\infty} n A_n^2 a^{n-1} P_n(\cos\theta) \right].$$

For these equation to hold for all θ, the coefficients that multiply each of the Legendre polynomials, including P_0 and P_1 [recall Eqs. (3.37) and (3.38)] on the left- and right-hand sides must be equal. Hence, for $n = 0$, we have

$$A_0^1 + \frac{B_0^1}{a} = A_0^2 \quad \text{and} \quad -\kappa_1 \frac{B_0^1}{a^2} = 0,$$

which gives

$$B_0^1 = 0, \qquad A_0^1 = A_0^2 \equiv A_0. \tag{3.50}$$

For $n = 1$, we have

$$-E_0 a + \frac{B_1^1}{a^2} = A_1^2 a, \tag{3.51}$$

$$\kappa_1 \left(-E_0 - \frac{2B_1^1}{a^3} \right) = \kappa_2 A_1^2 \tag{3.52}$$

and solving these simultaneously[33], we find

$$B_1^1 = \frac{\kappa_2 - \kappa_1}{2\kappa_1 + \kappa_2} E_0 a^3, \qquad A_1^2 = -\frac{3\kappa_1}{2\kappa_1 + \kappa_2} E_0. \tag{3.53}$$

Finally, for $n \geq 2$, we have

$$\frac{B_n^1}{a^{n+1}} = A_n^2 a^n \quad \text{and} \quad -\kappa_1 \frac{(n+1)B_n^1}{a^{n+2}} = \kappa_2 n A_n^2 a^{n-1},$$

which gives[34]

$$A_n^2 = 0, \quad B_n^1 = 0 \quad \text{for all} \quad n \geq 2. \tag{3.54}$$

Substituting all of these coefficients into Eqs (3.48) and (3.49) gives the potential outisde and inside the sphere,

$$\phi = A_0 - E_0 r \cos\theta + \frac{\kappa_2 - \kappa_1}{\kappa_2 + 2\kappa_1} E_0 a^3 \frac{\cos\theta}{r^2}, \quad (r > a) \tag{3.55}$$

$$\phi = A_0 - \frac{3\kappa_1}{\kappa_2 + 2\kappa_1} E_0 r \cos\theta, \quad (r < a) \tag{3.56}$$

where A_0 can be set to zero[35].

Note that inside the sphere,

$$\phi = -\frac{3\kappa_1 E_0}{\kappa_2 + 2\kappa_1} E_0 z,$$

the field is uniform,

$$\boldsymbol{E} = -\nabla\phi = \frac{3\kappa_1}{\kappa_2 + 2\kappa_1} E_0 \boldsymbol{k}. \tag{3.57}$$

For $\kappa_2 > \kappa_1$ the field is reduced compared to E_0, for $\kappa_2 < \kappa_1$ it is enhanced.

The potential outside the sphere (3.55) is the sum of the potential of the uniform external field $\boldsymbol{E}_0$ and the potential of the dipole (last term) due to polarisation of the sphere. Comparison with Eq. (1.43) shows that the dipole moment induced on the sphere is

$$p = 4\pi\varepsilon_0 \frac{\kappa_2 - \kappa_1}{\kappa_2 + 2\kappa_1} a^3 E_0, \tag{3.58}$$

[33] The quickest way to do this is probably by dividing Eq. (3.51) by a and Eq. (3.52) by κ_2 and subtracting them to find B_1^1, and then substituting it into (3.51) to find A_1^2.

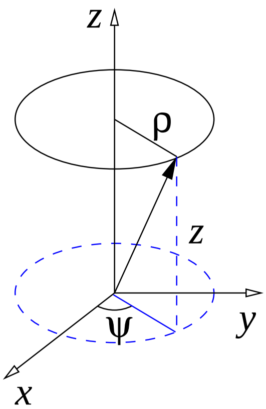
[34] According to the first equation, B_n^1 and A_n^2 have the same sign, but according to the second one, they have opposite signs, which is only possible if both coefficients are zero.

[35] Adding an arbitrary constant to ϕ does not change the electric field $\boldsymbol{E} = -\nabla\phi$.

and is in the direction of $\boldsymbol{E}_0$ (for $\kappa_2 > \kappa_1$).

Letting $\kappa_2 \to \infty$ in Eqs. (3.55) and (3.56) gives the solution for a conducting (e.g., metallic) sphere. Indeed, in this limit the potential inside the sphere [Eq. (3.56)] is constant and the field inside [Eq. (3.57)] vanishes. From Eq. (3.58), the dipole moment of the metallic sphere is $\boldsymbol{p} = 4\pi\varepsilon_0 a^3 \boldsymbol{E}_0$. It is proportional to the external electric field[36].

3.5 Solution of Laplace's equation in cylindrical polar coordinates



In cylindrical coordinates (ρ, ψ, z), Laplace's equation is

$$\frac{1}{\rho}\frac{\partial}{\partial \rho}\left(\rho \frac{\partial \phi}{\partial \rho}\right) + \frac{1}{\rho^2}\frac{\partial^2 \phi}{\partial \psi^2} + \frac{\partial^2 \phi}{\partial z^2} = 0. \tag{3.59}$$

For systems with translational symmetry along the z axis, the potential ϕ does not depend on z, and Laplace's equation reads

$$\frac{1}{\rho}\frac{\partial}{\partial \rho}\left(\rho \frac{\partial \phi}{\partial \rho}\right) + \frac{1}{\rho^2}\frac{\partial^2 \phi}{\partial \psi^2} = 0. \tag{3.60}$$

This equation possesses independent solutions[37]

$$\ln \rho, \quad \rho^n \sin n\psi, \quad \rho^n \cos n\psi \quad (n \in \mathbb{Z}). \tag{3.61}$$

The most general solution of equation (3.60) is then

$$\phi = C \ln \rho + \sum_{n=-\infty}^{\infty} (A_n \cos n\psi + B_n \sin n\psi)\rho^n, \tag{3.62}$$

where C, A_n and B_n are arbitrary coefficients.

Equation (3.62) can be used, e.g., to solve the problem of a dielectric cylinder in the uniform electric field perpendicular to its axis.

---

[36]In general, this relation is written as $\boldsymbol{p} = \alpha \boldsymbol{E}_0$, where the coefficient α is known as the *dipole polarisability*. The dipole polarisability of a metallic sphere is proportional to the cube of radius $\alpha = 4\pi\varepsilon_0 a^3$ (in SI) or, in CGS units, equals it: $\alpha = a^3$. Interestingly, this relation holds approximately for atoms, which have electrons in them, but do not really look like metallic spheres (and must be described quantum-mechanically). If we adopt CGS but use atomic units (a.u.) of length, the mean radius of the hydrogen atom is 1.5 and polarisability is 4.5 a.u., while for caesium (Cs), the largest atom in the periodic table, the mean radius is 6.3 and polarisability is 400 a.u. (atomic units cubed still denoted as a.u.).

[37]To derive this, use variable separation. Seek solution of Eq. (3.60) in the form $\phi(\rho, \psi) = R(\rho)\Phi(\psi)$. Following the same steps as in Sec. 3.4, one obtains the equation for $\Phi(\psi)$,

$$\Phi'' + \lambda \Phi(\psi) = 0,$$

with solutions $\sin(\sqrt{\lambda}\psi)$ and $\cos(\sqrt{\lambda}\psi)$, which are periodic with period 2π only if $\lambda = n^2$, where n is an integer. The coresponding radial equation is

$$\rho \frac{d}{d\rho}\left(\rho \frac{dR}{d\rho}\right) + n^2 R(\rho) = 0.$$

For $n = 0$ its two independent solutions are 1 and $\ln \rho$, while for $n \neq 0$ these are ρ^n and ρ^{-n} (found by seeking solution in the form $R(\rho) = \rho^s$). Note that the unit solution $R(\rho) = 1$ (for $n = 0$) is included in Eq. (3.61), since $\rho^n \cos n\psi = 1$ for $n = 0$.

4 Electrostatic energy

4.1 Energy of a charge distribution

Definition: Electrostatic energy of a charge distribution is the work required to assemble the distribution from separated charges initially at rest at infinity.

For a general distribution of free charges with volume density ρ and surface density σ, this energy is given by

$$U = \frac{1}{2}\int_V \rho(\boldsymbol{r})\phi(\boldsymbol{r})dV + \frac{1}{2}\int_S \sigma(\boldsymbol{r})\phi(\boldsymbol{r})dS. \tag{4.1}$$

Proof. The work by an electric field on moving charge q between points A and B is [see Eq. (1.19)]

$$W = -q[\phi(\boldsymbol{r}_B) - \phi(\boldsymbol{r}_A)].$$

When we consider work by external forces *against* the electric field, the sign of this expression must be reversed. If point A is at infinity where $\phi(\boldsymbol{r}_A) = 0$, and point B has position $\boldsymbol{r}$, the work required to move charge dq to this point is

$$dq\phi(\boldsymbol{r}). \tag{4.2}$$

Let us assume that the charge distributions which appear in Eq. (4.1) are built gradually, with the densities increasing from 0 to their final values as

$$\alpha\rho(\boldsymbol{r}) \quad \text{and} \quad \alpha\sigma(\boldsymbol{r}),$$

as the parameter α increases from 0 to 1. For a given value of α the electrostatic potential is $\alpha\phi(\boldsymbol{r})$, since the potential depends linearly on the charge density [see Eq. (2.15)].

When α increases by a small amount $d\alpha$, the extra charge in the volume dV at point $\boldsymbol{r}$ is

$$dq = d\alpha\rho(\boldsymbol{r})dV$$

with a similar increase $d\alpha\sigma dS$ of the charge on dS. Hence, by Eq. (4.2), the work required to produce such increase in the whole space is

$$dU = \int_V d\alpha\rho(\boldsymbol{r})\alpha\phi(\boldsymbol{r})dV + \int_S d\alpha\sigma(\boldsymbol{r})\alpha\phi(\boldsymbol{r})dS.$$

The total work is the sum of dU for all α from 0 to 1, i.e., the integral

$$\begin{aligned} U = \int dU &= \int_0^1 \left[\int_V d\alpha\rho(\boldsymbol{r})\alpha\phi(\boldsymbol{r})dV + \int_S d\alpha\sigma(\boldsymbol{r})\alpha\phi(\boldsymbol{r})dS\right] \\ &= \int_0^1 \alpha d\alpha \left[\int_V d\rho(\boldsymbol{r})\phi(\boldsymbol{r})dV + \int_S \sigma(\boldsymbol{r})\phi(\boldsymbol{r})dS\right], \end{aligned} \tag{4.3}$$

which gives Eq. (4.1), since $\int_0^1 \alpha d\alpha = \frac{1}{2}$.

If the only free charges in the system are on conducting surfaces S_i, Eq. (4.1) becomes

$$U = \frac{1}{2}\sum_i Q_i\phi_i, \tag{4.4}$$

where Q_i and ϕ_i are the charge and potential of conductor i.

We can write a similar expression for a system of N point charges q_i at positions $\boldsymbol{r}_i$ in vacuum, assuming that point charges are very small metalic spheres:

$$U = \frac{1}{2}\sum_{i=1}^{N} q_i\phi(\boldsymbol{r}_i) = \frac{1}{2}\sum_{i=1}^{N} q_i \frac{1}{4\pi\varepsilon_0}\sum_{j=1}^{N}\frac{q_j}{|\boldsymbol{r}_i - \boldsymbol{r}_j|}.$$

[see Eq. (1.18)]. However, for point charges the term with $j = i$ in the last expression gives a zero in the denominator. This contribution corresponds to the interaction of a charge q_i with itself. Such contributions do not depend on the actual positions of the charges, so it would be meaningful not to include them in the total electrostatic energy[38]. Hence, the energy of the system of point charges in vacuum is given by

$$U = \frac{1}{2}\sum_{i=1}^{N} q_i \frac{1}{4\pi\varepsilon_0}\sum_{\substack{j=1\\ j\neq i}}^{N}\frac{q_j}{|\boldsymbol{r}_i - \boldsymbol{r}_j|} = \frac{1}{2}\sum_{\substack{i,j=1\\ j\neq i}}^{N}\frac{q_i q_j}{4\pi\varepsilon_0|\boldsymbol{r}_i - \boldsymbol{r}_j|} \tag{4.5}$$

Note that the last sum is a double sum over all i and j from 1 to N, excluding $j = i$. In this sum each pair of charges is included twice, and the factor $\frac{1}{2}$ removes this "double counting". Alternatively, one can write the energy of the system of point charges as a sum over all *distinct* pairs of charges:

$$U = \sum_{i<j}^{N}\frac{q_i q_j}{4\pi\varepsilon_0|\boldsymbol{r}_i - \boldsymbol{r}_j|}. \tag{4.6}$$

<u>Example.</u> By Eq. (4.4), the electrostatic energy of a capacitor with charges Q and $-Q$ on the plates, which have potentials ϕ_1 and ϕ_2, respectively, is

$$\begin{aligned} U &= \frac{1}{2}(Q\phi_1 - Q\phi_2) = \frac{1}{2}Q(\phi_1 - \phi_2) \\ &= \frac{1}{2}QV = \frac{1}{2}CV^2 = \frac{Q^2}{2C}, \end{aligned} \tag{4.7}$$

where $V = \phi_1 - \phi_2$ is the potential difference between the plates, and Eq. (2.25) was used.

[38] In doing so, we neglect the mathematically infinite amount of energy required to put together each of the point charges. For moving charges, the problem of self-interaction becomes significantly more difficult, and there is no consistent solution to it, see, e.g., R. P. Feynman, *The Feynman lectures on physics, Vol. 2*, Ch. 28 (Ref. [4] on page 5). When charged particles, such as electrons, are accelerated, they emit electromagnetic radiation (see Ch. 11), which means that charges lose energy and should experience a radiation resistance force, which can only be due to the action of the electromagnetic field on the electron. So, it is not possible to totally neglect the self-interation in electrodynamics. The problem persists when electrons are described quantum-mechanically, or using the relativisitic quantum theory known as Quantum Electrodynamics (QED), which describes the interactions of electrons, their antiparticles (positrons) and photons. There is a rigorous mathematical procedure of dealing with infinite contributions in QED, called renormalisation. It takes account of the radiative corrections (i.e., self-interaction) and allows one to predict all measurable quantities, except the mass and charge of the electron itself!

4.2 Energy density of the electrostatic field

Consider a finite distribution of charge in which all the surface charges σ are on conducting surfaces S_c and ρ is the volume density of free charges[39]. Its electrostatic energy is

$$U = \frac{1}{2}\int_V \rho(\boldsymbol{r})\phi(\boldsymbol{r})dV + \frac{1}{2}\int_{S_c} \sigma(\boldsymbol{r})\phi(\boldsymbol{r})dS, \qquad (4.8)$$

where V is the volume outside the conductors enclosed in a sphere of an arbitrarily large radius (sphere at infinity S_∞).

Equation (4.8), as well as Eqs. (4.4) and (4.6), suggest that the electrostatic energy "belongs" to the interacting charges. We will now show that Eq. (4.8) can be cast in a completely different mathematical form, which will make it clear that energy is actually "owned" by the field![40]

From Eq. (2.9), we can replace the volume charge density in Eq. (4.8) by $\nabla\cdot\boldsymbol{D}$,

$$U = \frac{1}{2}\int_V \nabla\cdot\boldsymbol{D}\phi dV + \frac{1}{2}\int_{S_c}\sigma\phi dS,$$

where we have also dropped the argument $\boldsymbol{r}$ for brevity. From the product rule formula, $\nabla\cdot(\phi\boldsymbol{D}) = \nabla\phi\cdot\boldsymbol{D} + \phi\nabla\cdot\boldsymbol{D}$, we have

$$\phi\nabla\cdot\boldsymbol{D} = \nabla\cdot(\phi\boldsymbol{D}) - \nabla\phi\cdot\boldsymbol{D},$$

which gives

$$U = \frac{1}{2}\int_V \nabla\cdot(\phi\boldsymbol{D})dV - \frac{1}{2}\int_V \nabla\phi\cdot\boldsymbol{D}dV + \frac{1}{2}\int_{S_c}\sigma\phi dS.$$

Substituting $\nabla\phi = -\boldsymbol{E}$ in the middle term [see Eq. (1.17)], and using Gauss's theorem to transform the first term into a surface integral, we obtain

$$U = \frac{1}{2}\int_V \boldsymbol{E}\cdot\boldsymbol{D}dV + \frac{1}{2}\oint_{S_\infty}\phi\boldsymbol{D}\cdot d\boldsymbol{S} + \frac{1}{2}\oint_{S_c}\phi\boldsymbol{D}\cdot d\boldsymbol{S} + \frac{1}{2}\int_{S_c}\sigma\phi dS. \qquad (4.9)$$

Here we used the fact that V is bounded by S_∞ and the surfaces of the conductors (cf. diagram in Sec. 3.2).

Since the potential ϕ and the field D decrease as $1/r$ and $1/r^2$ (or faster), respectively, the integral over S_∞ (whose area grows as r^2) vanishes for $r\to\infty$.

[39] True free surface charges can only be found on outer surfaces of conductors. Mathematically, such charges are distributed in an infinitely thin layer, since the field inside the conductor is zero. Any other distributions of "surface" charges are thin layers of volume charges with the volume density that peaks sharply near the surface.

[40] In his fundamental paper, *A Dynamical Theory of the Electromagnetic Field* [Phil. Trans. R. Soc. Lond. **155**, 459 (1865)], James Clerk Maxwell, who performed a similar derivation, writes about the electromagnetic energy: "Where does it reside? On the old theories it resides in the electrified bodies, conducting circuits, and magnets, in the form of an unknown quantity called potential energy, or the power of producing certain effects at a distance. On our theory it resides in the electromagnetic field, in the space surrounding the electrified and magnetic bodies, as well as in those bodies themselves, and is in two different forms, which may be described without hypothesis as magnetic polarization and electric polarization". The magnetic part of the energy will be derived in Sec. 8.5.

In the second last term in Eq. (4.9), $d\boldsymbol{S}$ has the direction of the outward normal for V, so that $d\boldsymbol{S} = -\boldsymbol{n}dS$, where $\boldsymbol{n}$ is the outward normal for the conductors. Hence,

$$\boldsymbol{D} \cdot d\boldsymbol{S} = -\boldsymbol{D} \cdot \boldsymbol{n}dS = -D_n dS = -\sigma dS,$$

where σ is the surface charge density on the conductor [see Eq. (2.18)], and the two last terms in Eq. (4.9) cancel.

Therefore we find the electrostatic energy as

$$U = \frac{1}{2}\int_V \boldsymbol{E} \cdot \boldsymbol{D} dV, \tag{4.10}$$

where V can be extended to the whole space (since $\boldsymbol{E} = 0$ inside conductors). Equation (4.10) shows that each volume element dV contributes to the energy, as long as the electric field is nonzero there. Hence, the energy is stored in the field locally, rather than in the long-range interaction between the charges. The quantity

$$\frac{1}{2}\boldsymbol{E} \cdot \boldsymbol{D} \tag{4.11}$$

is the *energy density* of the electrostatic field.

Example. Consider a parallel-plate capacitor whose plates carry free charges with surface densities σ and $-\sigma$ (see Sec. 2.7). Inside the capacitor $D = \sigma$ and $E = \sigma/\varepsilon$, and the energy density obtained from Eq. (4.11),

$$\frac{1}{2}\frac{\sigma^2}{\varepsilon},$$

is constant. Multiplying it by the volume of the capacitor Ad, where d is the spacing between the plates and A is the area of each plate, we find

$$U = \frac{1}{2}\frac{\sigma^2}{\varepsilon}Ad = \frac{1}{2}\frac{(Q/A)^2}{\varepsilon}Ad = \frac{Q^2}{2(\varepsilon A/d)} = \frac{Q^2}{2C},$$

where $C = \varepsilon A/d$ is the capacitance [cf. Eq. (2.35)], the same result as (4.7).

4.3 Forces

Consider an *isolated* system composed of a number of parts (point charges, conductors, dielectrics). Let one of the parts move through a small displacement $d\boldsymbol{r}$ under the influence of the electric force $\boldsymbol{F}$ acting on it. The work by $\boldsymbol{F}$,

$$dW = \boldsymbol{F} \cdot d\boldsymbol{r} \tag{4.12}$$

is related to the change in the electrostatic potential energy[41],

$$dW = -dU, \tag{4.13}$$

so that

$$-dU = F_x dx + F_y dy + F_z dz, \tag{4.14}$$

[41] The minus sign in Eq. (4.13) is necessary, because when the field does a certain amount of work, its own energy decreases.

and

$$F_x = -\left(\frac{\partial U}{\partial x}\right)_Q, \quad F_y = -\left(\frac{\partial U}{\partial y}\right)_Q, \quad F_z = -\left(\frac{\partial U}{\partial z}\right)_Q, \tag{4.15}$$

where the subscript Q indicates that the charges that create the field remain constant (because the system is isolated).

In practice, electric fields are often created by conductors that are maintained at constant potentials, as they are connected to batteries. In this case the work done by the system on moving one of the parts through $d\boldsymbol{r}$ is

$$dW = dW_b - dU, \tag{4.16}$$

where dW_b is the work done by the batteries.

For a system of charged conductors, from (4.4), and given that $\phi_j = \text{const}$,

$$dU = \frac{1}{2}\sum_j \phi_j dQ_j. \tag{4.17}$$

The work supplied by the batteries is required to move each of the charges dQ_j from zero potential to ϕ_j, so, by Eq. (4.2),

$$dW_b = \sum_j \phi_j dQ_j = 2dU. \tag{4.18}$$

Hence, from (4.16) and (4.18),

$$dW = 2dU - dU = dU,$$

and the forces are

$$F_x = \left(\frac{\partial U}{\partial x}\right)_\phi, \quad F_y = \left(\frac{\partial U}{\partial y}\right)_\phi, \quad F_z = \left(\frac{\partial U}{\partial z}\right)_\phi, \tag{4.19}$$

where the subscript ϕ indicates that the potentials are maintained constant.

Example 1. For a parallel-plate capacitor with vacuum between the plates, the capacitance is $C = \varepsilon_0 A/z$ where A is the area of each of the plates and z is the distance between them [see Eq. (2.29)]. Using $U = Q^2/2C$ from (4.7), we find that for $Q = \text{const}$, the force acting on either of the plates is

$$F_z = -\frac{\partial U}{\partial z} = -\frac{Q^2}{2\varepsilon_0 A}. \tag{4.20}$$

The minus sign shows that the force is attractive, which was to be expected, given that the plates carry opposite charges.

To see clearer the meaning of (4.20), replace Q/A by the surface charge density $\sigma = Q/A$, and recall the field of a uniformly charged plane $E_1 = \sigma/2\varepsilon_0$ [Eq. (1.38)]. This gives

$$F_z = -\frac{1}{2}\frac{\sigma}{\varepsilon_0}Q = -QE_1,$$

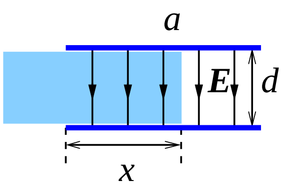
i.e., the force on charge $-Q$ in the field E_1 of the positively charged plate.

Using $U = CV^2/2$ from (4.7), we find that for the constant potentials case ($V = \text{const}$), the force is

$$F_z = \frac{\partial U}{\partial z} = -\frac{V^2}{2}\frac{\varepsilon_0 A}{z^2} = -\frac{Q^2}{2\varepsilon_0 A},$$

where we used $CV = Q$, which is the same answer as for $Q = \text{const}$.

Example 2. A parallel-plate capacitor with rectangular plates of sides a and b and distance d between them, has a slab of dielectric with permittivity ε and thickness d inserted through distance x along the side a. What is the force acting on the dielectric?



According to Eq. (4.11), the energy density is

$$\frac{1}{2}ED \quad \text{(in vacuum)}, \quad \frac{1}{2}E_d D_d \quad \text{(in dielectric)},$$

where E and $D = \varepsilon_0 E$ are the electric field and displacement in the vacuum part, and E_d and $D_d = \varepsilon E$ are those in the dielectric. The fields are uniform and in the same direction, and the energy density is constant within each part.

The electrostatic energy is found from equation (4.10), which in our case means multiplying the energy density by the volume of the corresponding part and adding the two contributions:

$$U = \frac{1}{2}[ED(a-x)bd + E_d D_d xbd],$$

Note that the electric field is the same in both parts, since the potential difference between the plates, $V = Ed$, is constant across the plates. Hence, the energy is

$$U = \frac{1}{2}[\varepsilon_0 E^2(a-x) + \varepsilon E^2 x]bd = \frac{1}{2}\underbrace{\frac{[\varepsilon_0(a-x)+\varepsilon x]b}{d}}_{C}V^2,$$

where the factor marked by the underbrace is the capacitance [cf. Eq. (4.7)].

Therefore, the force acting on the dielectric if the potentials are kept constant ($V = \text{const}$), is

$$F_x = \frac{\partial U}{\partial x} = \frac{(\varepsilon - \varepsilon_0)b}{2d}V^2,$$

and since $\varepsilon > \varepsilon_0$, the dielectric is drawn inside by this force.

One can show that for $Q = \text{const}$ the force from equation (4.15) is the same.

5 Steady currents

5.1 Electric current

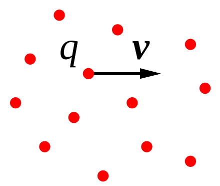
Current is the flow of electric charge.

Definition. The current I is the amount of charge transported across a given surface in unit time,

$$I = \frac{dQ}{dt}. \tag{5.1}$$

Units: ampere[42] (A), $1\ \text{A} = 1\ \text{C}\,\text{s}^{-1}$.

5.2 Current density



If there are n charge carriers per unit volume, with charge q each, moving with velocity $\boldsymbol{v}$, the *current density* is

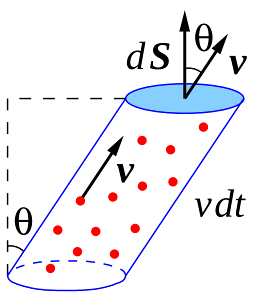
$$\boldsymbol{j} = nq\boldsymbol{v}. \tag{5.2}$$

The charge that flows through the surface element $d\boldsymbol{S}$ in time dt is

$$dQ = \boldsymbol{j} \cdot d\boldsymbol{S}dt, \tag{5.3}$$

hence, the current through $d\boldsymbol{S}$ is

$$dI = \boldsymbol{j} \cdot d\boldsymbol{S}. \tag{5.4}$$



To derive Eq. (5.3), note that for the charges to pass through $d\boldsymbol{S}$ in time dt, they should be within the distance vdt of the surface element, as measured along the line parallel to $\boldsymbol{v}$. Hence, these charges must be within the oblique cylinder, whose volume is $dSvdt\cos\theta$, where θ is the angle between $d\boldsymbol{S}$ and $\boldsymbol{v}$. The amount of charge within the cylinder is

$$dQ = qndSvdt\cos\theta = nq\boldsymbol{v} \cdot d\boldsymbol{S}dt = \boldsymbol{j} \cdot d\boldsymbol{S}dt,$$

where the definition of the current density (5.2) was used.

By Eq. (5.4), the current through a finite surface S is

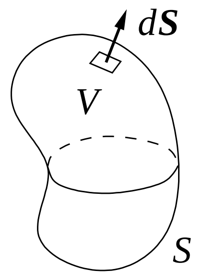
$$I = \int_S \boldsymbol{j} \cdot d\boldsymbol{S}. \tag{5.5}$$

If there are several types of charge carriers with number densities n_i, charges q_i and velocities $\boldsymbol{v}_i$, then

$$\boldsymbol{j} = \sum_i n_i q_i \boldsymbol{v}_i. \tag{5.6}$$

[42]Of the two related units, coulomb and ampere ($1\ \text{C} = 1\ \text{A} \times 1\ \text{s}$), the ampere is among the seven SI base units. "The ampere is that constant current which, if maintained in two straight parallel conductors of infinite length, of negligible circular cross-section, and placed 1 metre apart in vacuum, would produce between these conductors a force equal to 2×10^{-7} newton per metre of length. It follows that the magnetic constant, μ_0, also known as the permeability of free space, is exactly $4\pi \times 10^{-7}$ henries per metre, $\mu_0 = 4\pi \times 10^{-7}$ H/m." *SI Brochure: The International System of Units (SI)* [8th edition, 2006; updated in 2014]

5.3 Equation of continuity



Consider a charge distribution with volume density $\rho(\boldsymbol{r}, t)$, and an arbitrary volume V bounded by surface S. The charge inside it is given by

$$Q = \int_V \rho(\boldsymbol{r}, t) dV.$$

Its time derivative,

$$\frac{dQ}{dt} = \frac{d}{dt} \int_V \rho(\boldsymbol{r}, t) dV = \int_V \frac{\partial \rho(\boldsymbol{r}, t)}{\partial t} dV,$$

is equal to the current that flows into V, or the negative of the current that flows out of V across surface S:

$$\int_V \frac{\partial \rho(\boldsymbol{r}, t)}{\partial t} dV = -\oint_S \boldsymbol{j} \cdot d\boldsymbol{S}. \tag{5.7}$$

(Since $d\boldsymbol{S}$ is in the direction of the outer normal, the current out of V is regarded as positive.) Using Gauss's theorem on the right-hand side of Eq. (5.7),

$$\int_V \frac{\partial \rho(\boldsymbol{r}, t)}{\partial t} dV = -\int_V \nabla \cdot \boldsymbol{j} dV,$$

and rearranging, we have

$$\int_V \left[\frac{\partial \rho(\boldsymbol{r}, t)}{\partial t} + \nabla \cdot \boldsymbol{j} \right] dV = 0.$$

Since this is true for any volume V, we obtain the *equation of continuity*

$$\frac{\partial \rho}{\partial t} + \nabla \cdot \boldsymbol{j} = 0. \tag{5.8}$$

This equation is a mathematical form of charge conservation.

5.4 Ohm's law

It is found experimentally that for many substances at a given temperature, the current density is proportional to the electric field,

$$\boldsymbol{j} = \sigma \boldsymbol{E}. \tag{5.9}$$

This is *Ohm's Law*, and σ is the *conductivity*. Its reciprocal,

$$\rho = 1/\sigma, \tag{5.10}$$

is the *resistivity*, so that

$$\boldsymbol{E} = \rho \boldsymbol{j}.$$

The units of resistivity are found from $\rho = E/j$ as

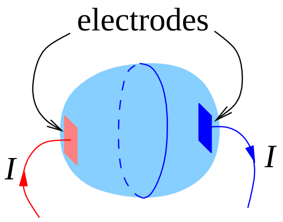
$$\frac{\text{volt (metre)}^{-1}}{\text{ampere (metre)}^{-2}} = \left(\frac{\text{volt}}{\text{ampere}} \right) \times \text{metre}.$$

In SI,

$$1 \text{ ohm} = \frac{1 \text{ volt}}{1 \text{ ampere}} \equiv 1\ \Omega,$$

so the unit of resistivity is $\Omega\,\text{m}$, and the unit is conductivity is $\Omega^{-1}\text{m}^{-1}$.

5.5 Steady currents in continuous media



Steady conditions mean that there is no variation with time. In particular, the charge density is independent of time: $\rho(\boldsymbol{r}, t) \to \rho(\boldsymbol{r})$.

Places where current enters or leaves a medium are known as *electrodes*. The current flowing from an electrode is given by

$$I = \int_S \boldsymbol{j} \cdot d\boldsymbol{S}, \tag{5.11}$$

where S is the surface surrounding the electrode.

The current flow satisfies the following equations:

$$\boldsymbol{E} = -\nabla\phi, \tag{5.12}$$

Ohm's law,

$$\boldsymbol{j} = \sigma\boldsymbol{E}, \tag{5.13}$$

and the continuity equation (5.8) with $\partial\rho/\partial t = 0$, so that

$$\nabla \cdot \boldsymbol{j} = 0. \tag{5.14}$$

Substituting (5.12) into (5.13), and (5.13) into (5.14), we find

$$\nabla \cdot (\sigma\nabla\phi) = 0, \tag{5.15}$$

or, for a uniform medium ($\sigma = \text{const}$),

$$\nabla^2\phi = 0,$$

i.e., Laplace's equation.

It must be solved subject to the following boundary conditions:

(i) On the electrodes,

$$\phi = \text{const},$$

which means that $\boldsymbol{E}$ and $\boldsymbol{j}$ are normal to the electrode surface. (Electric fields are perpendicular to equipotential surfaces, see end of Sec. 1.4.)

(ii) The current leaving the electrode is given by equation (5.11).

(iii) On the boundary between two media with different conductivities,

$$j_{1n} = j_{2n}, \tag{5.16}$$

for the normal components of the current density[43]. Consequently,

$$\sigma_1 E_{1n} = \sigma_2 E_{2n},$$

while the tangential component of the electric field is continuous (Sec. 2.6),

$$E_{1t} = E_{2t}.$$

[43] This can be proved by considering $\oint_S \boldsymbol{j} \cdot d\boldsymbol{S}$ for a cylindrical surface S whose flat faces are parallel to the boundary between media 1 and 2, and much larger than the curved surface (cf. Sec. 2.6). For steady currents this intergral must be zero, so the contributions of the two flat faces cancel, which yileds Eq. (5.16).

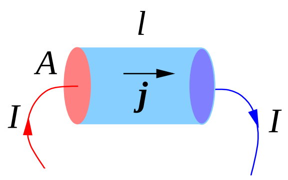
5.6 Resistance

For a conductor of any shape, with current I entering at one electrode and leaving at another, the potential difference between the electrodes being V, the *resistance* of the conductor is

$$R = \frac{V}{I}. \tag{5.17}$$

It is supposed to be independent of V and I. Sometimes (5.17) is referred to as Ohm's law.

In SI the resistance is measured in ohms (Ω) (see Sec. 5.4).



Example. Consider a straight homogeneous wire of constant cross section A and length l, with constant $\boldsymbol{j}$ along the wire. This satisfies all the boundary conditions, and the magnitude of the current density is (using $\phi = -Ez$, $E = (\phi_1 - \phi_2)/l$),

$$j = \sigma(\phi_1 - \phi_2)/l.$$

hence, the current is

$$I = jA = \sigma A(\phi_1 - \phi_2)/l = \sigma AV/l,$$

and the resistance of the wire is

$$R = \frac{l}{\sigma A} = \frac{\rho l}{A}.$$

5.7 Joule heating

The work done by the electric field when moving charge Q through a potential difference V is

$$QV,$$

and the corresponding power (i.e., work per unit time) is

$$IV.$$

Using (5.17), it can be written in different forms,

$$IV = I^2 R = \frac{V^2}{R}. \tag{5.18}$$

When this power is dissipated in the material, it experiences *Joule heating*.

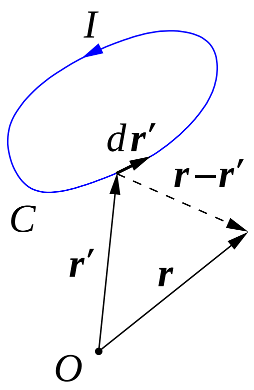
6 Magnetic field of steady currents

6.1 Magnetic field

6.1.1 The origin of magnetism.

Magnetism is an effect produced by the flow of electric charges. There are no magnetic charges (or *magnetic monopoles*[44]).

6.1.2 Magnetic field produced by a current.



Consider a circuit C carrying current I. The contribution of a small segment $d\boldsymbol{r}'$ of the circuit located at point $\boldsymbol{r}'$, to the *magnetic field* $\boldsymbol{B}$ at point $\boldsymbol{r}$ is

$$d\boldsymbol{B} = \frac{\mu_0}{4\pi} I \frac{d\boldsymbol{r}' \times (\boldsymbol{r} - \boldsymbol{r}')}{|\boldsymbol{r} - \boldsymbol{r}'|^3}, \tag{6.1}$$

where $d\boldsymbol{r}'$ is in the direction of the current. This mathematical form was found by generalising experimental data. It resembles Coulomb's law in that the magnitude of the magnetic field descreases as the inverse square of the distance. However, the direction of the field is neither away nor towards its source, but perpendicular to both $d\boldsymbol{r}'$ and $\boldsymbol{r} - \boldsymbol{r}'$[45].

The magnetic field of the entire circuit is obtained by adding the contributions of all the segments,

$$\boldsymbol{B}(\boldsymbol{r}) = \frac{\mu_0}{4\pi} I \int_C \frac{d\boldsymbol{r}' \times (\boldsymbol{r} - \boldsymbol{r}')}{|\boldsymbol{r} - \boldsymbol{r}'|^3}. \tag{6.2}$$

Equation (6.2) is *Biot-Savart's law*, and the constant μ_0 is

$$\mu_0 = 4\pi \times 10^{-7}\ \mathrm{N\,s^2C^{-2}}. \tag{6.3}$$

The unit of magnetic field in SI is tesla (T)[46]. The CGS unit is gauss (G),

$$1\ \mathrm{T} = 10^4\ \mathrm{G}. \tag{6.4}$$

The Earth's magnetic field near its surface is ∼1 G (0.25–0.65 G, depending on location).

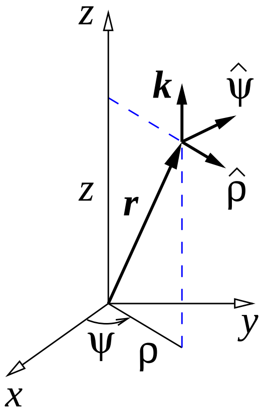
[44] So far, magnetic monopoles have not been discovered, and they may not exist. However, this hypothetical particle can have important implications for particle physics. For example, Dirac showed that the existence of a magnetic monopole would explain quantisation of electric charge. Magnetic monopoles can also catalyse the decay of protons, as predicted by some elementary particle theories. (Current experiments set the lower limit for the proton lifetime at 2×10^{29} years.)

[45] The vector product of vectors $\boldsymbol{a}$ and $\boldsymbol{b}$ is a vector $\boldsymbol{c} = \boldsymbol{a} \times \boldsymbol{b}$ such that (i) it is perpendicular to both $\boldsymbol{a}$ and $\boldsymbol{b}$ (i.e., perpendicular to the plane that contains $\boldsymbol{a}$ and $\boldsymbol{b}$), (ii) $c = ab \sin\theta$, where θ is the angle between $\boldsymbol{a}$ and $\boldsymbol{b}$, and (iii) from the top of $\boldsymbol{c}$ the rotation from $\boldsymbol{a}$ to $\boldsymbol{b}$ is in the positive direction (i.e., anticlockwise).

[46] 1 T is a strong field. Such fields are produced by neodimium magnets, the strongest permanent magnets commercially available. Their structure is $Nd_2Fe_{14}B$ and their strength is due to lined-up electron spins (elementary "magnets") in the neodimium (Nd) atoms.

Example: magnetic field of an infinite straight wire carrying current I.

Let us assume that the wire is along the z axis. The position vector $\boldsymbol{r}' = z'\boldsymbol{k}$ is the point on the wire with the z-coordinate z'. It takes values in the range $-\infty < z' < \infty$, and $d\boldsymbol{r}' = dz'\boldsymbol{k}$.



Biot-Savart's law (6.2) gives

$$\boldsymbol{B}(\boldsymbol{r}) = \frac{\mu_0}{4\pi} I \int_{-\infty}^{\infty} \frac{dz'\boldsymbol{k} \times (\boldsymbol{r} - z'\boldsymbol{k})}{|\boldsymbol{r} - z'\boldsymbol{k}|^3}. \tag{6.5}$$

The system has axial symmetry about the z axis, so it is convenient to use cylindrical coordinates in which $\boldsymbol{r} = \rho\hat{\boldsymbol{\rho}} + z\boldsymbol{k}$, where $\hat{\boldsymbol{\rho}}$ is the unit vector associated with coordinate ρ. In this case

$$\boldsymbol{k} \times (\boldsymbol{r} - z'\boldsymbol{k}) = \boldsymbol{k} \times (\rho\hat{\boldsymbol{\rho}} + z\boldsymbol{k} - z'\boldsymbol{k}) = \rho\boldsymbol{k} \times \hat{\boldsymbol{\rho}} = \rho\hat{\boldsymbol{\psi}},$$

where $\hat{\boldsymbol{\psi}}$ is the unit vector associated with coordinate ψ and perpendicular to both $\boldsymbol{k}$ and $\hat{\boldsymbol{\rho}}$. The distance in the denominator of (6.5) is

$$|\boldsymbol{r} - z'\boldsymbol{k}| = |\rho\hat{\boldsymbol{\rho}} + z\boldsymbol{k} - z'\boldsymbol{k}| = \sqrt{\rho^2 + (z - z')^2},$$

and we obtain

$$\boldsymbol{B}(\boldsymbol{r}) = \frac{\mu_0}{4\pi} I\rho\hat{\boldsymbol{\psi}} \int_{-\infty}^{\infty} \frac{dz'}{[\rho^2 + (z' - z)^2]^{3/2}}.$$

Changing the integration variable from z' to $\xi = z' - z$, we have

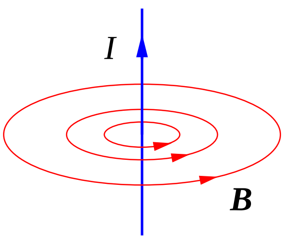
$$\boldsymbol{B}(\boldsymbol{r}) = \frac{\mu_0}{4\pi} I\rho\hat{\boldsymbol{\psi}} \int_{-\infty}^{\infty} \frac{d\xi}{(\rho^2 + \xi^2)^{3/2}}.$$

Using the variable substitution $\xi = \rho \tan\alpha$,

$$d\xi = \frac{\rho d\alpha}{\cos^2\alpha}, \qquad \rho^2 + \xi^2 = \rho^2(1 + \tan^2\alpha) = \frac{\rho^2}{\cos^2\alpha},$$

with the new integration limits $\alpha = -\pi/2$ and $\pi/2$, we have

$$\boldsymbol{B}(\boldsymbol{r}) = \frac{\mu_0}{4\pi} \frac{I\hat{\boldsymbol{\psi}}}{\rho} \int_{-\pi/2}^{\pi/2} \cos\alpha d\alpha,$$



which gives

$$\boldsymbol{B} = \frac{\mu_0 I}{2\pi\rho} \hat{\boldsymbol{\psi}}. \tag{6.6}$$

As expected from the symmetry of the system, the magnitude of the magnetic field depends only on the distance from the wire. Vector $\boldsymbol{B}$ is tangential to the circles centred on the wire, which represent the magnetic field lines.

6.1.3 The Lorentz force.

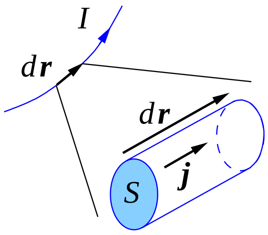
The force on a charge q moving with velocity $\boldsymbol{v}$ in electric field $\boldsymbol{E}$ and magnetic field $\boldsymbol{B}$ is the *Lorentz force*,

$$\boldsymbol{F} = q\boldsymbol{E} + q\boldsymbol{v} \times \boldsymbol{B}. \tag{6.7}$$

The first term in Eq. (6.7) is the familiar electric force (Sec. 1.3). The second term is the force due to the magentic field. It acts only on moving charges, and is perpendicular to both the particle's velocity and the magnetic field.

6.1.4 Force on a current.

Let us determine the force that acts on a circuit carrying current I in the magnetic field $\boldsymbol{B}$.



Consider a small segment $d\boldsymbol{r}$ of the wire carrying current I. When enlarged, this segment looks as a cylinder. By Eq. (5.4), the current I is related to the current density j as

$$I = jS, \tag{6.8}$$

where S is the perpendicular cross section area of the wire. Multiplying Eq. (6.8) by $d\boldsymbol{r}$ and taking into account the fact that $d\boldsymbol{r}$ and $\boldsymbol{j}$ are in the same direction (so that $jd\boldsymbol{r} = \boldsymbol{j}|d\boldsymbol{r}|$), we have

$$Id\boldsymbol{r} = \boldsymbol{j}dV, \tag{6.9}$$

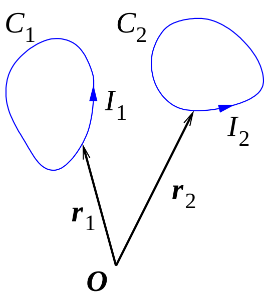
where $dV = |d\boldsymbol{r}|S$ is the volume of the segment of the wire.

The segment contains ndV charge carriers, where n is their number density. The charge carries have charge q each and move with velocity $\boldsymbol{v}$ (the current density $\boldsymbol{j} = nq\boldsymbol{v}$, see Sec. 5.2). Hence, the magnetic force acting on the segment is

$$d\boldsymbol{F} = q\boldsymbol{v} \times \boldsymbol{B} ndV = \boldsymbol{j} \times \boldsymbol{B} dV = Id\boldsymbol{r} \times \boldsymbol{B}, \tag{6.10}$$

and the total force acting on the circuit C is

$$\boldsymbol{F} = \int_C Id\boldsymbol{r} \times \boldsymbol{B}(\boldsymbol{r}). \tag{6.11}$$



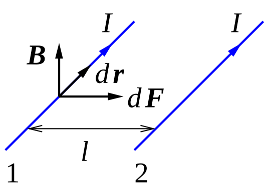
Example 1. Consider two circuits, C_1 and C_2, carrying currents I_1 and I_2, respectively. The magnetic field created by circuit C_2 at point $\boldsymbol{r}_1$ is, according to Eq. (6.2),

$$\boldsymbol{B}(\boldsymbol{r}_1) = \frac{\mu_0}{4\pi} I_2 \oint_{C_2} \frac{d\boldsymbol{r}_2 \times (\boldsymbol{r}_1 - \boldsymbol{r}_2)}{|\boldsymbol{r}_1 - \boldsymbol{r}_2|^3}.$$

Using (6.11) we find the force acting on circuit C_1 as

$$\boldsymbol{F}_1 = \frac{\mu_0}{4\pi} I_1 I_2 \oint_{C_1} \oint_{C_2} \frac{d\boldsymbol{r}_1 \times [d\boldsymbol{r}_2 \times (\boldsymbol{r}_1 - \boldsymbol{r}_2)]}{|\boldsymbol{r}_1 - \boldsymbol{r}_2|^3}. \tag{6.12}$$

The force $\boldsymbol{F}_2$ acting on C_2 is obtained from Eq. (6.12) by interchanging indices 1 and 2. Of course, $\boldsymbol{F}_2 = -\boldsymbol{F}_1$, in accordance with Newton's 3rd law[47].



Example 2. Let us find the force between two parallel currents I separared by distance l. The magnetic field created by wire 2 at the position of wire 1 is $B = \mu_0 I/(2\pi l)$ [see Eq. (6.6)], perpendicular to the plane containing the currents. The force acting on segment $d\boldsymbol{r}$ of wire 1 is $d\boldsymbol{F} = Id\boldsymbol{r} \times \boldsymbol{B}$, towards wire 2. Hence, the two wires with parallel currents attract each other with the force

$$F = \frac{\mu_0 I^2}{2\pi l}$$

per unit length of the wire. For $I = 1$ A and $l = 1$ m, $F = 2 \times 10^{-7}$ N/m.

[47] To show this, apply the *"bac–cab"* rule to the numerator of the integrand, and use the identity

$$\oint_{C_2} d\boldsymbol{r}_2 \cdot \frac{(\boldsymbol{r}_2 - \boldsymbol{r}_1)}{|\boldsymbol{r}_2 - \boldsymbol{r}_1|^3} = -\oint_{C_2} \nabla_2 \frac{1}{|\boldsymbol{r}_2 - \boldsymbol{r}_1|} \cdot d\boldsymbol{r}_2 = 0,$$

which holds since $\int_A^B \nabla f \cdot d\boldsymbol{r} = f(\boldsymbol{r}_B) - f(\boldsymbol{r}_A)$ and $\oint \nabla f \cdot d\boldsymbol{r} = 0$ for any scalar function.

6.2 Laws of magnetostatics

For a volume distribution of current with current density $\boldsymbol{j}$, we use Eq. (6.9) and replace $Id\boldsymbol{r}'$ by $\boldsymbol{j}(\boldsymbol{r}')dV'$ in Biot-Savart's law (6.2), obtaining the field as the volume integral

$$\boldsymbol{B}(\boldsymbol{r}) = \frac{\mu_0}{4\pi}\int_V \frac{\boldsymbol{j}(\boldsymbol{r}') \times (\boldsymbol{r}-\boldsymbol{r}')}{|\boldsymbol{r}-\boldsymbol{r}'|^3}\,dV'. \tag{6.13}$$

We will now show that this leads to the following key equations:

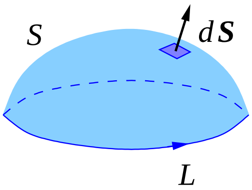
$$\nabla\cdot\boldsymbol{B} = 0, \tag{6.14}$$

and

$$\nabla\times\boldsymbol{B} = \mu_0\boldsymbol{j}. \tag{6.15}$$

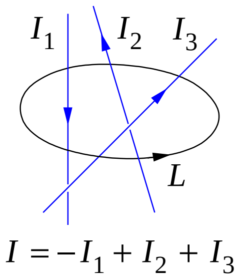
Equation (6.14) means the absence of magnetic charges [cf. Eq. (1.39) for the electric field], and it also holds in the nonstatic case.

Equation (6.15) shows how the magnetic field is created by currents. It can be cast in the integral form by considering a line integral of $\boldsymbol{B}$ over a closed loop L and using Stokes's theorem:



$$\oint_L \boldsymbol{B}\cdot d\boldsymbol{r} = \int_S \nabla\times\boldsymbol{B}\cdot d\boldsymbol{S},$$

where S is any surface bounded by L[48]. Substituting $\nabla\times\boldsymbol{B}$ from Eq. (6.15) and using Eq. (5.5), we obtain



$$\oint_L \boldsymbol{B}\cdot d\boldsymbol{r} = \mu_0 I, \tag{6.16}$$

where I is the total current through the closed curve L, determined using the right-hand corkscrew sign convention. Equation (6.16) is *Ampere's law.*

Derivation of Eqs. (6.14) and (6.15).

Applying $\nabla\cdot$ to Eq. (6.13), we have

$$\begin{aligned}
\nabla\cdot\boldsymbol{B} &= \frac{\mu_0}{4\pi}\int_V \nabla\cdot\left[\frac{\boldsymbol{j}(\boldsymbol{r}')\times(\boldsymbol{r}-\boldsymbol{r}')}{|\boldsymbol{r}-\boldsymbol{r}'|^3}\right]dV' \\
&= -\frac{\mu_0}{4\pi}\int_V \nabla\cdot\left[\boldsymbol{j}(\boldsymbol{r}')\times\nabla\frac{1}{|\boldsymbol{r}-\boldsymbol{r}'|}\right]dV' \\
&= \frac{\mu_0}{4\pi}\int_V \boldsymbol{j}(\boldsymbol{r}')\cdot\left[\underbrace{\nabla\times\nabla}_{=0}\frac{1}{|\boldsymbol{r}-\boldsymbol{r}'|}\right]dV' \\
&= 0,
\end{aligned}$$

where we used Eq. (1.16) to obtain the 2nd line, and $\boldsymbol{a}\cdot(\boldsymbol{b}\times\boldsymbol{c}) = -\boldsymbol{b}\cdot(\boldsymbol{a}\times\boldsymbol{c})$ in line three [∇ acts on functions of $\boldsymbol{r}$, so $\boldsymbol{j}(\boldsymbol{r}')$ can be treated as a constant].

[48] The direction of $d\boldsymbol{S}$ is related to the direction in which L is traversed by the right-hand corkscrew rule, i.e., if the corkscrew is rotated in the direction in which L is traversed, it will move in the direction of $d\boldsymbol{S}$. Alternatively, one can say that when viewed from the tip of $d\boldsymbol{S}$, the loop is traversed in the positive (anticlockwise) direction.

Applying $\nabla\times$ to Eq. (6.13), we have

$$\begin{aligned}\nabla\times\boldsymbol{B} &= -\frac{\mu_0}{4\pi}\int_V \nabla\times\left[\boldsymbol{j}(\boldsymbol{r}')\times\nabla\frac{1}{|\boldsymbol{r}-\boldsymbol{r}'|}\right]dV' \\ &= -\frac{\mu_0}{4\pi}\int_V \left[\boldsymbol{j}(\boldsymbol{r}')\underbrace{(\nabla\cdot\nabla)}_{=\nabla^2}\frac{1}{|\boldsymbol{r}-\boldsymbol{r}'|} - (\boldsymbol{j}(\boldsymbol{r}')\cdot\nabla)\nabla\frac{1}{|\boldsymbol{r}-\boldsymbol{r}'|}\right]dV', \quad (6.17)\end{aligned}$$

where we used $\boldsymbol{a}\times(\boldsymbol{b}\times\boldsymbol{c}) = \boldsymbol{b}(\boldsymbol{a}\cdot\boldsymbol{c})-\boldsymbol{c}(\boldsymbol{a}\cdot\boldsymbol{b}) = \boldsymbol{b}(\boldsymbol{a}\cdot\boldsymbol{c})-(\boldsymbol{a}\cdot\boldsymbol{b})\boldsymbol{c}$, treating $\boldsymbol{j}(\boldsymbol{r}')$ as a constant and keeping the function of $\boldsymbol{r}$ on the right of the ∇ operators.

Recalling that the potential of a point charge q at $\boldsymbol{r}'$

$$\phi = \frac{q}{4\pi\varepsilon_0}\frac{1}{|\boldsymbol{r}-\boldsymbol{r}'|}$$

satisfies Poisson's equation $\nabla^2\phi = -\rho/\varepsilon_0$ with $\rho = q\delta(\boldsymbol{r}-\boldsymbol{r}')$ [see Eqs. (1.40) and (1.24)], we have

$$\nabla^2\frac{1}{|\boldsymbol{r}-\boldsymbol{r}'|} = -4\pi\delta(\boldsymbol{r}-\boldsymbol{r}'). \quad (6.18)$$

Hence, the first term in square brackets in Eq. (6.17) gives,

$$\mu_0\int_V \boldsymbol{j}(\boldsymbol{r}')\delta(\boldsymbol{r}-\boldsymbol{r}')dV' = \mu_0\boldsymbol{j}(\boldsymbol{r}),$$

[using Eq. (1.27)], which is the right-hand side of Eq. (6.15).

It remains to show that the 2nd term in Eq. (6.17) is zero. Using the identity

$$\boldsymbol{j}(\boldsymbol{r}')\times\left(\underbrace{\nabla\times\nabla}_{=0}\frac{1}{|\boldsymbol{r}-\boldsymbol{r}'|}\right) = \nabla\left(\boldsymbol{j}(\boldsymbol{r}')\cdot\nabla\frac{1}{|\boldsymbol{r}-\boldsymbol{r}'|}\right) - (\boldsymbol{j}(\boldsymbol{r}')\cdot\nabla)\nabla\frac{1}{|\boldsymbol{r}-\boldsymbol{r}'|},$$

in which left-hand side is zero, we re-write the 2nd term in Eq. (6.17) as

$$\frac{\mu_0}{4\pi}\nabla\int_V \boldsymbol{j}(\boldsymbol{r}')\cdot\nabla\frac{1}{|\boldsymbol{r}-\boldsymbol{r}'|}dV' = -\frac{\mu_0}{4\pi}\nabla\int_V \boldsymbol{j}(\boldsymbol{r}')\cdot\nabla'\frac{1}{|\boldsymbol{r}-\boldsymbol{r}'|}dV', \quad (6.19)$$

where ∇' acts on $\boldsymbol{r}'$. Making use of the product rule

$$\nabla'\cdot\left(\boldsymbol{j}(\boldsymbol{r}')\frac{1}{|\boldsymbol{r}-\boldsymbol{r}'|}\right) = \nabla'\cdot\boldsymbol{j}(\boldsymbol{r}')\frac{1}{|\boldsymbol{r}-\boldsymbol{r}'|} + \boldsymbol{j}(\boldsymbol{r}')\cdot\nabla'\frac{1}{|\boldsymbol{r}-\boldsymbol{r}'|},$$

we transform the integral on the right-hand-side (6.19) into

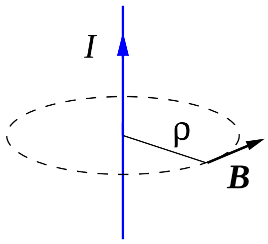
$$\int_V \nabla'\cdot\left(\boldsymbol{j}(\boldsymbol{r}')\frac{1}{|\boldsymbol{r}-\boldsymbol{r}'|}\right)dV' - \int_V \nabla'\cdot\boldsymbol{j}(\boldsymbol{r}')\frac{1}{|\boldsymbol{r}-\boldsymbol{r}'|}dV'.$$

The second term above vanishes, since $\nabla\cdot\boldsymbol{j} = 0$ for steady currents [see Eq. (5.14)]. By Gauss's theorem, the first term transforms into the integral

$$\oint_S \frac{\boldsymbol{j}(\boldsymbol{r}')}{|\boldsymbol{r}-\boldsymbol{r}'|}\cdot d\boldsymbol{S}'$$

over the surface bounding V. Assuming that the currents exist only in a finite range of space completely enclosed by S, this integral also vanishes.

Example. Similar to Gauss's law in electrostatics (Sec. 1.7), Ampere's law (6.16) allows one to determine the magnetic field in cases where the system possesses some symmetry.



For a straight wire carrying current I, the magnetic field is tangential to the circles centred on the wire (since the $\boldsymbol{B}$-field lines must be closed), and its magnitude depends only on the distance ρ from the wire. Choosing the loop L in Eq. (6.16) as a circle of radius ρ centred on the wire, we obtain

$$\oint_L \boldsymbol{B}\cdot d\boldsymbol{r} = B(\rho)\oint_L |d\boldsymbol{r}| = B(\rho)2\pi\rho = \mu_0 I,$$

which gives

$$B(\rho) = \frac{\mu_0 I}{2\pi\rho},$$

in agreement with Eq. (6.6).

6.3 Magnetic flux

Definition: the magnetic flux Φ through a surface S is

$$\Phi = \int_S \boldsymbol{B}\cdot d\boldsymbol{S}. \tag{6.20}$$

For a closed surface, using Gauss's theorem and equation (6.14), we have

$$\oint_S \boldsymbol{B}\cdot d\boldsymbol{S} = 0. \tag{6.21}$$

The unit of magnetic flux is the weber (Wb): $1\ \text{Wb} = 1\ \text{T}\,\text{m}^2$.

6.4 Magnetic scalar potential

In the region of space with no currents, $\boldsymbol{j} = 0$, equation (6.15) gives

$$\nabla\times\boldsymbol{B} = 0.$$

Hence, the magnetic field can be represented by the gradient of a scalar field[49]. The *magnetic scalar potential* ϕ_m is defined by

$$\boldsymbol{B} = -\mu_0\nabla\phi_m. \tag{6.22}$$

Using (6.14), we see that ϕ_m satisfies Laplace's equation,

$$\nabla^2\phi_m = 0.$$

Thus, we can use the methods for solving Laplace's equation developed in Secs. 3.4 and 3.5 (with suitable boundary conditions) to find magnetic fields.

[49] Compare this with the condition that the electrostatic field is conservative, Eq. (1.20), which is related to the possibility to represent it as $\boldsymbol{E} = -\nabla\phi$.

6.5 Magnetic vector potential

The field $\boldsymbol{B}$ that satisfies equation (6.14) can be represented by

$$\boldsymbol{B} = \nabla \times \boldsymbol{A}, \tag{6.23}$$

since $\nabla \cdot \boldsymbol{B} = \nabla \cdot (\nabla \times \boldsymbol{A}) = (\nabla \times \nabla)\boldsymbol{A} = 0$ for any $\boldsymbol{A}$.

The vector function $\boldsymbol{A}$ is the *magnetic vector potential*.

The vector potential $\boldsymbol{A}$ is not unique. Adding the gradient of any scalar field χ to $\boldsymbol{A}$ leaves $\boldsymbol{B}$ unchanged. Indeed, replacing $\boldsymbol{A} \to \boldsymbol{A} + \nabla\chi$ in (6.23) gives

$$\boldsymbol{B} = \nabla \times (\boldsymbol{A} + \nabla\chi) = \nabla \times \boldsymbol{A},$$

since $\nabla \times \nabla\chi = 0$.

This allows one to impose extra conditions on $\boldsymbol{A}$, e.g.,

$$\nabla \cdot \boldsymbol{A} = 0, \tag{6.24}$$

which is a convenient choice in magnetostatics.

Proof. Let $\boldsymbol{A}'$ be a vector potential such that $\nabla \cdot \boldsymbol{A}' = f \neq 0$. We construct a new vector potential as

$$\boldsymbol{A} = \boldsymbol{A}' + \nabla\chi,$$

and impose (6.24),

$$\nabla \cdot \boldsymbol{A} = \nabla \cdot \boldsymbol{A}' + \nabla \cdot \nabla\chi = f + \nabla^2\chi = 0,$$

which gives

$$\nabla^2\chi = -f. \tag{6.25}$$

This equation is similar to Poisson's equation (1.40) in electrostatics,

$$\nabla^2\phi = -\rho/\varepsilon_0, \tag{6.26}$$

whose solution is given by Eq. (1.23) (in the absence of surface charges),

$$\phi(\boldsymbol{r}) = \frac{1}{4\pi\varepsilon_0}\int_V \frac{\rho(\boldsymbol{r}')}{|\boldsymbol{r} - \boldsymbol{r}'|}dV'. \tag{6.27}$$

Hence, Eq. (6.25) has a solution

$$\chi(\boldsymbol{r}) = \frac{1}{4\pi}\int_V \frac{f(\boldsymbol{r}')}{|\boldsymbol{r} - \boldsymbol{r}'|}dV',$$

and the condition (6.24) can always be satisfied.

Let us now find the equation that relates $\boldsymbol{A}$ to the current density. Substituting (6.23) into (6.15), we have

$$\nabla \times (\nabla \times \boldsymbol{A}) = \mu_0\boldsymbol{j},$$

or

$$\nabla(\nabla \cdot \boldsymbol{A}) - \nabla^2\boldsymbol{A} = \mu_0\boldsymbol{j}.$$

If $\boldsymbol{A}$ is chosen so that (6.24) holds, this gives

$$\nabla^2 \boldsymbol{A} = -\mu_0 \boldsymbol{j}. \tag{6.28}$$

Comparing Eq. (6.28) with Possion's equation (6.26), we notice that it repre-sentes "three copies" of the latter, one for each of the components x, y and z. Hence, its solution can be constructed in the same way as Eq. (6.27),

$$\boldsymbol{A}(\boldsymbol{r}) = \frac{\mu_0}{4\pi} \int_V \frac{\boldsymbol{j}(\boldsymbol{r}')}{|\boldsymbol{r} - \boldsymbol{r}'|} dV'. \tag{6.29}$$

It is easy to show that the vector potential from Eq. (6.29) satisfies (6.24), and that substitution of (6.29) into Eq. (6.23) leads to Biot-Savart's law (6.13).

For a circuit C carrying current I, the vector potential is given by

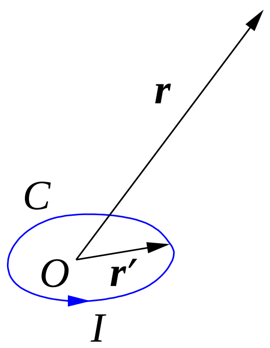
$$\boldsymbol{A}(\boldsymbol{r}) = \frac{\mu_0}{4\pi} I \int_C \frac{d\boldsymbol{r}'}{|\boldsymbol{r} - \boldsymbol{r}'|}, \tag{6.30}$$

which is obtained from Eq. (6.29) using (6.9). Substitution of (6.30) into (6.23) gives Biot-Savart's law in the form of Eq. (6.2).

The formulae for the vector potenial (6.29) and (6.30) are simpler than the corresponding expressions for the magnetic field, Eqs. (6.13) and (6.2). This is why it is often easier to find $\boldsymbol{A}$ first and then determine $\boldsymbol{B}$ from Eq. (6.23).

6.6 Field at a large distance from a current loop – magnetic dipole

Let us use equation (6.30) to find the magnetic field of a current loop C at distances r much greater than the size of the loop.



Expanding to first order in $\boldsymbol{r}'$ for $r' \ll r$ (cf. Sec. 1.8),

$$\frac{1}{|\boldsymbol{r} - \boldsymbol{r}'|} \simeq \frac{1}{r} - \nabla \frac{1}{r} \cdot \boldsymbol{r}' = \frac{1}{r} + \frac{\boldsymbol{r} \cdot \boldsymbol{r}'}{r^3},$$

and substituting into (6.30), we find

$$\boldsymbol{A}(\boldsymbol{r}) = \frac{\mu_0 I}{4\pi} \frac{1}{r} \int_C d\boldsymbol{r}' + \frac{\mu_0 I}{4\pi} \frac{1}{r^3} \int_C (\boldsymbol{r} \cdot \boldsymbol{r}') d\boldsymbol{r}'. \tag{6.31}$$

The first term above vanishes for the closed loop C. To transform the second term, we use two identities,

$$\boldsymbol{r} \times (\boldsymbol{r}' \times d\boldsymbol{r}') = \boldsymbol{r}'(\boldsymbol{r} \cdot d\boldsymbol{r}') - d\boldsymbol{r}'(\boldsymbol{r} \cdot \boldsymbol{r}') \tag{6.32}$$

and

$$d[(\boldsymbol{r} \cdot \boldsymbol{r}')\boldsymbol{r}'] = (\boldsymbol{r} \cdot d\boldsymbol{r}')\boldsymbol{r}' + (\boldsymbol{r} \cdot \boldsymbol{r}') d\boldsymbol{r}', \tag{6.33}$$

where $\boldsymbol{r}$ is regarded as a constant vector. Subtracting (6.32) from (6.33), we find

$$(\boldsymbol{r} \cdot \boldsymbol{r}') d\boldsymbol{r}' = \frac{1}{2} d[(\boldsymbol{r} \cdot \boldsymbol{r}')\boldsymbol{r}'] + \frac{1}{2} (\boldsymbol{r}' \times d\boldsymbol{r}') \times \boldsymbol{r}.$$

Substituting into (6.31), we obtain

$$\boldsymbol{A}(\boldsymbol{r}) = \frac{\mu_0 I}{4\pi}\frac{1}{r^3}\left\{\frac{1}{2}\int_C d[(\boldsymbol{r}\cdot\boldsymbol{r}')\boldsymbol{r}'] + \frac{1}{2}\int_C (\boldsymbol{r}'\times d\boldsymbol{r}')\times\boldsymbol{r}\right\}.$$

The first term on the right-hand side gives zero[50], while the second yields

$$\boldsymbol{A}(\boldsymbol{r}) = \frac{\mu_0}{4\pi}\frac{\boldsymbol{m}\times\boldsymbol{r}}{r^3}, \tag{6.34}$$

where

$$\boldsymbol{m} = \frac{I}{2}\int_C \boldsymbol{r}'\times d\boldsymbol{r}' \tag{6.35}$$

is the *magnetic dipole moment* of the circuit C.

Substituting (6.34) into (6.23), we find the magnetic field,

$$\boldsymbol{B}(\boldsymbol{r}) = \frac{\mu_0}{4\pi}\nabla\times\frac{\boldsymbol{m}\times\boldsymbol{r}}{r^3} = -\frac{\mu_0}{4\pi}\nabla\left(\frac{\boldsymbol{m}\cdot\boldsymbol{r}}{r^3}\right) = \frac{\mu_0}{4\pi r^3}\left[\frac{3(\boldsymbol{m}\cdot\boldsymbol{r})\boldsymbol{r}}{r^2} - \boldsymbol{m}\right], \tag{6.36}$$

where the last step is identical to the derivation of Eq. (1.45), and the magnetic field of the dipole is similar to the electric field of the point dipole (1.45).

To prove the first step in Eq. (6.36), transform the double vector product using the "bac–cab" rule, keeping in mind that $\boldsymbol{m}$ is a constant vector,

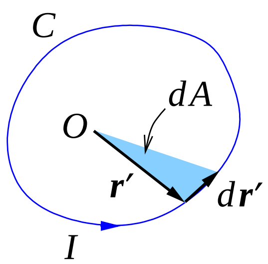
$$\nabla\times\frac{\boldsymbol{m}\times\boldsymbol{r}}{r^3} = -\nabla\times\left(\boldsymbol{m}\times\nabla\frac{1}{r}\right) = -\boldsymbol{m}\left(\nabla\cdot\nabla\frac{1}{r}\right) + (\boldsymbol{m}\cdot\nabla)\nabla\frac{1}{r}.$$

The first term on the right-hand side vanishes since $\nabla^2(1/r) = 0$ for $r \neq 0$. The second term can be transformed with the help of another "bac–cab" identity

$$\boldsymbol{m}\times\Big(\underbrace{\nabla\times\nabla}_{=0}\frac{1}{r}\Big) = \nabla\left(\boldsymbol{m}\cdot\nabla\frac{1}{r}\right) - (\boldsymbol{m}\cdot\nabla)\nabla\frac{1}{r},$$

which shows that

$$\nabla\times\frac{\boldsymbol{m}\times\boldsymbol{r}}{r^3} = \nabla\left(\boldsymbol{m}\cdot\nabla\frac{1}{r}\right) = -\nabla\left(\frac{\boldsymbol{m}\cdot\boldsymbol{r}}{r^3}\right).$$



Geometrically,

$$\frac{1}{2}\int_C \boldsymbol{r}'\times d\boldsymbol{r}'$$

is the "vector area" of the current loop. If the loop is planar, the magnitude of this integral is equal to the area of the loop A, and its direction is perpendicular to plane of the loop, given by the right-hand corkscrew rule relative to the current. (As seen from the diagram, $\frac{1}{2}\boldsymbol{r}'\times d\boldsymbol{r}' = dA\,\boldsymbol{n}$.) In this case

$$\boldsymbol{m} = IA\boldsymbol{n},$$

where $\boldsymbol{n}$ is the unit vector perpendicular to the loop.

Comparison of the second last expression in (6.36) with (6.22) shows the field of the magnetic dipole corresponds to the magnetic scalar potential

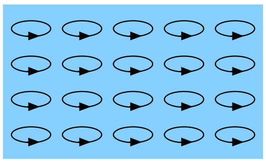
$$\phi_m = \frac{\boldsymbol{m}\cdot\boldsymbol{r}}{4\pi r^3}.$$

---

[50] For a linear integral of a complete differential, we have $\int_A^B d[\boldsymbol{a}(\boldsymbol{r})] = \boldsymbol{a}(\boldsymbol{r}_B) - \boldsymbol{a}(\boldsymbol{r}_A)$. When such an integral is taken over a closed loop, the initial and final points coincide and the integral vanishes.

7 Magnetic properties of matter

7.1 Magnetisation



The moving electrical charges in atoms and molecules behave as little current loops, characterised by their magnetic dipole moments $\boldsymbol{m}$. The *magnetisation* $\boldsymbol{M}$ of a material is the magnetic moment per unit volume. The magnetic moment of volume dV is then

$$\boldsymbol{M}\,dV.$$

From equation (6.34), the vector potential produced by a volume V of magnetic material is

$$\boldsymbol{A}(\boldsymbol{r}) = \frac{\mu_0}{4\pi}\int_V \frac{\boldsymbol{M}(\boldsymbol{r}')\times(\boldsymbol{r}-\boldsymbol{r}')}{|\boldsymbol{r}-\boldsymbol{r}'|^3}\,dV' \tag{7.1}$$

$$= \frac{\mu_0}{4\pi}\int_V \boldsymbol{M}(\boldsymbol{r}')\times\nabla'\frac{1}{|\boldsymbol{r}-\boldsymbol{r}'|}\,dV'$$

$$= \frac{\mu_0}{4\pi}\left[\int_V \frac{\nabla'\times\boldsymbol{M}(\boldsymbol{r}')}{|\boldsymbol{r}-\boldsymbol{r}'|}\,dV' - \int_V \nabla'\times\left(\frac{\boldsymbol{M}(\boldsymbol{r}')}{|\boldsymbol{r}-\boldsymbol{r}'|}\right)dV'\right], \tag{7.2}$$

where we used Eq. (2.4) and a rearrangement of the product rule $\nabla\times(\boldsymbol{a}f) = \nabla\times\boldsymbol{a}f - \boldsymbol{a}\times\nabla f$, to obtain the last line.

Transforming the second term in (7.2) with the help of the identity[51]

$$\int_V \nabla\times\boldsymbol{a}\,dV = -\oint_S \boldsymbol{a}\times d\boldsymbol{S}, \tag{7.3}$$

we find

$$\boldsymbol{A}(\boldsymbol{r}) = \frac{\mu_0}{4\pi}\int_V \frac{\nabla'\times\boldsymbol{M}(\boldsymbol{r}')}{|\boldsymbol{r}-\boldsymbol{r}'|}\,dV' + \frac{\mu_0}{4\pi}\oint_S \frac{\boldsymbol{M}(\boldsymbol{r}')\times d\boldsymbol{S}'}{|\boldsymbol{r}-\boldsymbol{r}'|}. \tag{7.4}$$

Comparing the first term with equation (6.29), we see that magnetisation $\boldsymbol{M}$ results in the volume current density

$$\boldsymbol{j}_m = \nabla\times\boldsymbol{M}, \tag{7.5}$$

known as the *magnetisation current density*. It is similar to the polarisation charge density in the case of dielectrics, Sec. 2.2.

The second term in (7.4) is due to the magnetisation *surface current density*

$$\boldsymbol{J}_m = \boldsymbol{M}\times\boldsymbol{n}, \tag{7.6}$$

where $\boldsymbol{n}$ is the outward unit normal ($\boldsymbol{M}\times d\boldsymbol{S} = \boldsymbol{M}\times\boldsymbol{n}dS$). The surface current density is the current per unit length perpendicular to it. Its contribution can be absorbed into $\boldsymbol{j}_m$ if we smear the surface to a finite thickness.

[51]Proof. Considering the dot product of a constant vector $\boldsymbol{t}$ with the left-hand side of (7.3), and using properties of the triple scalar product and Gauss's theorem, we have

$$\boldsymbol{t}\cdot\int_V \nabla\times\boldsymbol{a}\,dV = \int_V \nabla\cdot(\boldsymbol{t}\times\boldsymbol{a})\,dV = \oint_S(\boldsymbol{a}\times\boldsymbol{t})\cdot d\boldsymbol{S} = -\boldsymbol{t}\cdot\oint_S \boldsymbol{a}\times d\boldsymbol{S},$$

where S is the surface that bounds V, which proves (7.3), since $\boldsymbol{t}$ is arbitrary.

7.2 Equations of the magnetic field

In the presence of magnetic materials the total current density is

$$\boldsymbol{j} + \boldsymbol{j}_m,$$

where $\boldsymbol{j}$ is the free current density. In this case Eq. (6.15) takes the form

$$\nabla \times \boldsymbol{B} = \mu_0(\boldsymbol{j} + \boldsymbol{j}_m). \tag{7.7}$$

Note that Eq. (6.15) was derived on the assumption of steady currents, i.e., $\nabla \cdot \boldsymbol{j} = 0$, which is also satisfied by $\boldsymbol{j}_m$, since

$$\nabla \cdot \boldsymbol{j}_m = \nabla \cdot (\nabla \times \boldsymbol{M}) = 0.$$

Using (7.5) in Eq. (7.7), we obtain

$$\nabla \times \boldsymbol{B} = \mu_0 \boldsymbol{j} + \mu_0 \nabla \times \boldsymbol{M}.$$

After rearrangement, this gives

$$\nabla \times \boldsymbol{H} = \boldsymbol{j}, \tag{7.8}$$

where

$$\boldsymbol{H} = \frac{\boldsymbol{B}}{\mu_0} - \boldsymbol{M}, \tag{7.9}$$

is the *magnetic field intensity*[52], and

$$\boldsymbol{B} = \mu_0(\boldsymbol{H} + \boldsymbol{M}). \tag{7.10}$$

Equations (7.8) and (7.10) replace (6.15) when magnetic materials are present.

Using Stokes's theorem, we obtain the integral form of (7.8) as

$$\oint_L \boldsymbol{H} \cdot d\boldsymbol{r} = I, \tag{7.11}$$

where I is the free current through loop L, which replaces Ampere's law (6.16).

7.3 Magnetic susceptibility and permeability

For a large class of materials (isotropic and linear),

$$\boldsymbol{M} = \chi_m \boldsymbol{H}, \tag{7.12}$$

where χ_m is the *magnetic susceptibility*.

If $\chi_m < 0$ the material is called *diamagnetic*, and if $\chi_m > 0$ – *paramagnetic*[53].

[52] This vector is analogous to the displacement $\boldsymbol{D}$ of the electric field. Note, however, how a minus sign in Eq. (2.7) led to the plus sign in the definition of $\boldsymbol{D}$ (2.10), while (7.5) leads to the minus sign in the definition of $\boldsymbol{H}$ (7.9).

[53] A material is diamagnetic if it has no "elementary current loops" (i.e., atomic or molecular magnetic moments) in the absence of the external magnetic field. Application of such field gives rise to magnetisation which opposes the field, which is a consequence of Lenz's rule [see Sec. 8.2]. In paramagnetic materials the magnetic dipole moments exist even in the absence of the field, but they are randomly oriented (i.e., no net magnetisation). When the field is switched on, they orient preferentially along the field, so that $\boldsymbol{M}$ and $\boldsymbol{H}$ are in the same direction. In both cases typical values are $\chi_m \sim 10^{-5}$.

From Eqs. (7.10) and (7.12) we have

$$\boldsymbol{B} = \mu \boldsymbol{H}, \tag{7.13}$$

where

$$\mu = \mu_0(1 + \chi_m) \tag{7.14}$$

is the *permeability*. In vacuum $\chi_m = 0$, which is why μ_0 [see Eq. (6.3)] is known as the vacuum permeability, or permeability of free space.

Similar to Eq. (2.12), the ratio

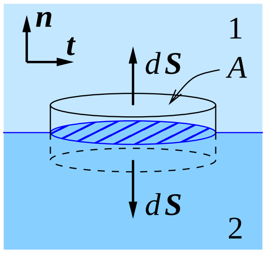
$$\kappa_m \equiv \frac{\mu}{\mu_0} = 1 + \chi_m, \tag{7.15}$$

is the *relative permeability*.

Another class of materials (*ferromagnetics*) is characterised by much larger values of magnetisation which can be nonzero in the absence of external magnetic fields. Relations (7.12) and (7.13) do not hold for them, and the magnetisation depends not only on $\boldsymbol{H}$, but also on the "history" of the sample.

7.4 Boundary conditions

The boundary conditions for the magnetic field are derived in a way similar to the boundary conditions at the interface of two dielectrics (Sec. 2.6).



Consider the boundary between two magnetic materials 1 and 2, with unit normal vector $\boldsymbol{n}$ (from 2 into 1) and unit tangential vector $\boldsymbol{t}$. The boundary condition for the $\boldsymbol{B}$ field is

$$B_{1n} = B_{2n}. \tag{7.16}$$

It is obtained from $\nabla \cdot \boldsymbol{B} = 0$ in the integral form (6.21), applied to a cylinder with flat faces parallel to the boundary and negligibly small curved surface [cf. derivation of Eq. (2.23)].

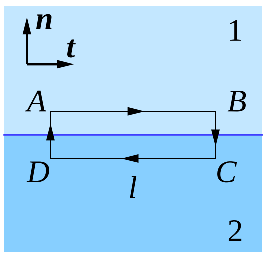
Using Eq. (7.10), the boundary condition (7.16) can be written as

$$H_{1n} + M_{1n} = H_{2n} + M_{2n}, \tag{7.17}$$

or, for two linear, isotropic magnetics which satisfy (7.13),

$$\mu_1 H_{1n} = \mu_2 H_{2n}, \tag{7.18}$$

where μ_1 and μ_2 are the permeabilities of the two materials.



The boundary condition for $\boldsymbol{H}$ is derived by applying Eq. (7.11) to a rectangular loop with long sides parallel to the interface. If surface currents are present, the current through the loop is

$$I = \boldsymbol{J} \cdot (\boldsymbol{n} \times \boldsymbol{t}) l,$$

where $\boldsymbol{J}$ is the surface current density, l is the length of AB and $\boldsymbol{n} \times \boldsymbol{t}$ is the unit normal to the loop, linked by the right-hand corkcscrew rule with the direction in which $ABCD$ is traversed. Equation (7.11) then yields

$$(\boldsymbol{H}_1 - \boldsymbol{H}_2) \cdot \boldsymbol{t} = (\boldsymbol{J} \times \boldsymbol{n}) \cdot \boldsymbol{t}, \tag{7.19}$$

which can also be written in alternative forms[54] [cf. derivation of Eq. (2.22)].

In the absence of surface currents ($\boldsymbol{J} = 0$), we have

$$H_{1t} = H_{2t}. \tag{7.20}$$

7.5 Permanent magnets

A permanent magnet is a piece of material that retains magnetisation in the absence of free currents, i.e., for $\boldsymbol{j} = 0$. Equation (7.8) then gives

$$\nabla \times \boldsymbol{H} = 0,$$

which implies that $\boldsymbol{H}$ can be described through a scalar potential ϕ_m,

$$\boldsymbol{H} = -\nabla \phi_m, \tag{7.21}$$

[cf. Eqs. (1.17) and (1.20) for the electric field and electrostatic potential].

From Eqs. (7.9) and (6.14), we have

$$\nabla \cdot \boldsymbol{H} = -\nabla \cdot \boldsymbol{M}. \tag{7.22}$$

Equations (7.21) and (7.22) are similar to Eqs. (1.17) and (1.39), which means that

$$\rho_m = -\nabla \cdot \boldsymbol{M} \tag{7.23}$$

can be considered as the source "magnetic charge" density for the field $\boldsymbol{H}$.

Then, by analogy with equations (1.23) and (2.6), we can write

$$\phi_m(\boldsymbol{r}) = -\frac{1}{4\pi} \int_V \frac{\nabla' \cdot \boldsymbol{M}(\boldsymbol{r}')}{|\boldsymbol{r} - \boldsymbol{r}'|} dV' + \frac{1}{4\pi} \oint_S \frac{\boldsymbol{M}(\boldsymbol{r}') \cdot d\boldsymbol{S}'}{|\boldsymbol{r} - \boldsymbol{r}'|}, \tag{7.24}$$

where V is the volume of the magnet, and S is its surface. The second term in (7.24) is due to the surface density

$$\sigma_m = \boldsymbol{M} \cdot \boldsymbol{n} \tag{7.25}$$

of "magnetic charges", where $\boldsymbol{n}$ is the unit vector out of S.

Note that if the magnet is uniformly magnetised ($\boldsymbol{M} = \text{const}$), then only the second term in Eq. (7.24) contributes to the magnetic scalar potential.

7.6 Potential problems

We consider problems involving magnetic materials in which there are no free currents, i.e., $\boldsymbol{j} = 0$. In this case,

$$\nabla \times \boldsymbol{H} = 0 \tag{7.26}$$

and

$$\boldsymbol{H} = -\nabla \phi_m. \tag{7.27}$$

[54] Equation (7.19) can be written as $(\boldsymbol{H}_1 - \boldsymbol{H}_2) \cdot \boldsymbol{t} = \boldsymbol{J} \cdot (\boldsymbol{n} \times \boldsymbol{t})$ or $\boldsymbol{n} \times (\boldsymbol{H}_1 - \boldsymbol{H}_2) = \boldsymbol{J}$.

In addition, if the material is linear and homogeneous, $\nabla \cdot \boldsymbol{B} = 0$ gives

$$\nabla \cdot \boldsymbol{H} = 0. \tag{7.28}$$

This equation also holds if the magnetisation is uniform, i.e., $\boldsymbol{M} = \text{const}$ [which is obtained by applying $\nabla \cdot$ to Eq. (7.10) and using $\nabla \cdot \boldsymbol{M} = 0$].

By equation (7.28), the magnetic potential satisfies Laplace's equation,

$$\nabla^2 \phi_m = 0. \tag{7.29}$$

This equation must be solved with boundary conditions (7.16) and (7.20), which can be done using methods from Ch. 3.

In terms of ϕ_m, the normal component of $\boldsymbol{H}$ (7.27) is $-\partial\phi_m/\partial n$. The boundary condition (7.16) can then be written for linear, isotropic materials, as

$$\mu_1 \frac{\partial \phi_1}{\partial n} = \mu_2 \frac{\partial \phi_2}{\partial n}, \tag{7.30}$$

or, more generally, as

$$-\frac{\partial \phi_1}{\partial n} + M_{1n} = -\frac{\partial \phi_2}{\partial n} + M_{2n}, \tag{7.31}$$

where ϕ_1 and ϕ_2 are the magnetic scalar potentials in materials 1 and 2, $\boldsymbol{M}_1$ and $\boldsymbol{M}_2$ being the corresponding magnetisations.

The second boundary condition (7.20) is for the tangential derivatives,

$$\frac{\partial \phi_1}{\partial t} = \frac{\partial \phi_2}{\partial t}. \tag{7.32}$$

It can be integrated along the boundary (i.e., in the $\boldsymbol{t}$ direction) and replaced by the requirement that the magnetic scalar potential is continuous across the boundary,

$$\phi_1 = \phi_2. \tag{7.33}$$

8 Electromagnetic induction

8.1 Electromotive force

The *electromotive force* (emf) in a circuit C is defined as the work by the "motive" force $\boldsymbol{F}$ to move a unit positive charge around the circuit:

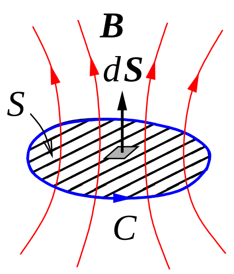
$$\mathcal{E} = \oint_C \boldsymbol{F} \cdot d\boldsymbol{r}. \tag{8.1}$$

This force can be due to an electric field, or chemical forces in a battery, or other forces. The concept of emf may be generalised by considering the integral in equation (8.1) along any closed curve in space.

Ohm's law (5.9) can be generalised to forces other than the electric field, to read

$$\boldsymbol{j} = \sigma \boldsymbol{F}. \tag{8.2}$$

8.2 The law of induction



According to Faraday, when the magnetic flux (Sec. 6.3)

$$\Phi = \int_S \boldsymbol{B} \cdot d\boldsymbol{S} \tag{8.3}$$

through a circuit changes, there is an associated *induced* emf $\mathcal{E}$ in the circuit,

$$\mathcal{E} = -\frac{d\Phi}{dt}. \tag{8.4}$$

This emf is due to an electric field $\boldsymbol{E}$ produced by the changing magnetic field,

$$\mathcal{E} = \oint_C \boldsymbol{E} \cdot d\boldsymbol{r}. \tag{8.5}$$

The direction of $d\boldsymbol{S}$ (normal to the surface S bounded by the circuit C) in (8.3) and the direction of $d\boldsymbol{r}$, in which the circuit is traversed in (8.5), are related by the right-hand corkscrew rule.

If C is a conducting circuit, then the emf (8.5) produces a current, which causes a magnetic field that *opposes* the change in the magnetic flux that caused it. This is *Lenz's law*, manifested by the minus sign in equation (8.4).

From equations (8.3), (8.4) and (8.5), we find

$$\begin{aligned}\oint_C \boldsymbol{E} \cdot d\boldsymbol{r} &= -\frac{d}{dt}\int_S \boldsymbol{B} \cdot d\boldsymbol{S} \\ &= -\int_S \frac{\partial \boldsymbol{B}}{\partial t} \cdot d\boldsymbol{S}.\end{aligned}$$

Using Stokes's theorem on the left-hand side, we have

$$\int_S \nabla \times \boldsymbol{E} \cdot d\boldsymbol{S} = -\int_S \frac{\partial \boldsymbol{B}}{\partial t} \cdot d\boldsymbol{S}.$$

Since this is true for *any* surface S, we conclude that

$$\nabla \times \boldsymbol{E} = -\frac{\partial \boldsymbol{B}}{\partial t}, \tag{8.6}$$

which is the differential form of Faraday's law.

Comparing equation (8.6) with (1.20), we see that the electric field produced by the time-dependent magnetic field is no longer conservative.

8.3 Mutual inductance and self-inductance

Consider two circuits, C_1 and C_2, carrying currents I_1 and I_2. The magnetic field $\boldsymbol{B}_1$ produced by C_1 is proportional to I_1, and the flux of this magnetic field through C_2 is also proportional to I_1. Denoting this flux by Φ_2, we can write

$$\Phi_2 = L_{21} I_1 \tag{8.7}$$

where L_{21} is the *mutual inductance* between circuits 1 and 2. This means that if current I_1 changes in time, this will cause an emf in circuit 2, given by

$$\mathcal{E}_2 = -L_{21} \frac{dI_1}{dt}. \tag{8.8}$$

Similarly, the flux of the magnetic field $\boldsymbol{B}_2$ created by the current I_2 through the circuit C_1 is

$$\Phi_1 = L_{12} I_2, \tag{8.9}$$

and the corresponding emf in circuit 1 is

$$\mathcal{E}_1 = -L_{12} \frac{dI_2}{dt}. \tag{8.10}$$

Let us show that

$$L_{12} = L_{21}. \tag{8.11}$$

Using the definition of the vector potential (6.23) and Stokes's theorem, we can write the flux in C_2 due to the current in circuit I_1 as

$$\Phi_2 = \int_{S_2} \boldsymbol{B}_1 \cdot d\boldsymbol{S}_2 = \int_{S_2} \nabla \times \boldsymbol{A}_1 \cdot d\boldsymbol{S}_2 = \oint_{C_2} \boldsymbol{A}_1(\boldsymbol{r}_2) \cdot d\boldsymbol{r}_2, \tag{8.12}$$

where $\boldsymbol{A}_1$ is the vector potential created by current I_1. Using equation (6.30) for the vector potential of a current, we have

$$\Phi_2 = \oint_{C_2} \frac{\mu_0}{4\pi} I_1 \oint_{C_1} \frac{d\boldsymbol{r}_1}{|\boldsymbol{r}_2 - \boldsymbol{r}_1|} \cdot d\boldsymbol{r}_2.$$

Comparing with (8.7), we obtain the formula for the mutual inductance,

$$L_{21} = \frac{\mu_0}{4\pi} \oint_{C_2} \oint_{C_1} \frac{d\boldsymbol{r}_1 \cdot d\boldsymbol{r}_2}{|\boldsymbol{r}_2 - \boldsymbol{r}_1|}, \tag{8.13}$$

which is symmetric with respect to interchanging indices 1 and 2, hence (8.11).

Similar to equations (8.8) and (8.10), one can also consider the emf in a circuit caused by the change in *its own* current I,

$$\mathcal{E} = -L\frac{dI}{dt}, \tag{8.14}$$

with the corresponding expression for the magnetic flux

$$\Phi = LI. \tag{8.15}$$

The coefficient L here is the *self-inductance* of the circuit.

Note that the self-inductance cannot be found from equation (8.13) by making the contours C_1 and C_2 identical, since the integrand would then contain a singularity for $|\boldsymbol{r}_2 - \boldsymbol{r}_1| \to 0$, and the integral would diverge logarithmically. A correct calculation must take into account the finite thickness of the wire that makes up the circuit.

The SI unit of inductance is the henry (H), 1 H = 1 Wb A^{-1}.

8.4 Magnetic energy

Consider a set of N current carrying circuits in the presence of magnetic media. Let I_k be the current in kth circuit and Φ_k the magnetic flux through it. Let the currents be changed to $I_k + dI_k$ over a small time interval dt, with resulting change $d\Phi_k$ in the fluxes.

The change of flux through kth circuit causes an induced emf $\mathcal{E}_k$ [see (8.4)]. The work done *against* it (to sustain the current) is

$$dW_k = -\mathcal{E}_k dQ_k = \frac{d\Phi_k}{dt} dQ_k,$$

where $dQ_k = I_k dt$ is the charge transported through the circuit. Hence,

$$dW_k = I_k d\Phi_k.$$

The total work done for all the circuits is

$$dW = \sum_{k=1}^{N} I_k d\Phi_k. \tag{8.16}$$

This is the work done in changing the magnetic field produced by the currents.

Let the currents be raised from zero to their final values I_k by changing a parameter α from zero to unity, so that the currents are αI_k. Assuming that the magnetic materials are linear, the magnetic field will also increase linearly with α. In this case the fluxes will be $\alpha\Phi_k$, with Φ_k being their final values. The work done in changing from α to $\alpha + d\alpha$ is

$$dW = \sum_{k=1}^{N} (\alpha I_k) d(\alpha\Phi_k) = \sum_{k=1}^{N} I_k \Phi_k \alpha d\alpha.$$

The total work is therefore

$$W = \int_0^1 \left(\sum_{k=1}^{N} I_k \Phi_k \right) \alpha d\alpha,$$

which gives

$$W = \frac{1}{2} \sum_{k=1}^{N} I_k \Phi_k. \tag{8.17}$$

This is the work done in establishing the magnetic field produced by the currents, i.e., the *magnetic energy* [cf. Sec. 4.1 and equations (4.1) and (4.4)].

Using the inductances, the total flux through kth circuit may be written as

$$\Phi_k = \sum_{l=1}^{N} L_{kl} I_l, \tag{8.18}$$

where $L_{kk} \equiv L_k$ is the self-inductance of circuit k. The magnetic energy (8.17) then takes the form

$$W = \frac{1}{2} \sum_{k=1}^{N} \sum_{l=1}^{N} L_{kl} I_k I_l. \tag{8.19}$$

8.5 Energy density of the magnetic field

Similarly to equation (8.12), the total flux through circuit k can be written as

$$\Phi_k = \int_{S_k} \boldsymbol{B} \cdot d\boldsymbol{S} = \int_{S_k} \nabla \times \boldsymbol{A} \cdot d\boldsymbol{S} = \oint_{C_k} \boldsymbol{A} \cdot d\boldsymbol{r}.$$

Hence, equation (8.17) becomes

$$W = \frac{1}{2} \sum_{k=1}^{N} I_k \oint_{C_k} \boldsymbol{A} \cdot d\boldsymbol{r}.$$

For a volume distribution of currents, we replace $I_k d\boldsymbol{r}$ by $\boldsymbol{j} dV$ and integrate over the volume V containing the currents, instead of summing over k and integrating over C_k. This gives

$$W = \frac{1}{2} \int_V \boldsymbol{A} \cdot \boldsymbol{j} \, dV. \tag{8.20}$$

Using Ampere's law (7.8), we can write this as

$$W = \frac{1}{2} \int_V \boldsymbol{A} \cdot (\nabla \times \boldsymbol{H}) dV,$$

and using the identity $\nabla \cdot (\boldsymbol{A} \times \boldsymbol{H}) = \boldsymbol{H} \cdot (\nabla \times \boldsymbol{A}) - \boldsymbol{A} \cdot (\nabla \times \boldsymbol{H})$, we obtain

$$\begin{aligned} W &= \frac{1}{2} \int_V \boldsymbol{H} \cdot (\nabla \times \boldsymbol{A}) \, dV - \frac{1}{2} \int_V \nabla \cdot (\boldsymbol{A} \times \boldsymbol{H}) \, dV \\ &= \frac{1}{2} \int_V \boldsymbol{H} \cdot \boldsymbol{B} \, dV - \frac{1}{2} \oint_S (\boldsymbol{A} \times \boldsymbol{H}) \cdot d\boldsymbol{S}, \end{aligned} \tag{8.21}$$

where in the last step we used Gauss's theorem to transform the second term, and S is the surface that bounds V.

For currents contained in a finite volume, the potential and magnetic field decrease as

$$A \sim \frac{1}{r} \quad \text{and} \quad H \sim \frac{1}{r^2},$$

or faster at large distances r. This means that if we extend the integration volume to the whole space ($r \to \infty$), the surface integral in (8.21) vanishes (since the area of S is proportional to r^2). The magnetic energy is then given by

$$W = \frac{1}{2} \int_{\text{all space}} \boldsymbol{H} \cdot \boldsymbol{B} \, dV. \tag{8.22}$$

This shows that the quantity

$$\frac{1}{2} \boldsymbol{H} \cdot \boldsymbol{B} \tag{8.23}$$

is the *energy density* of the magnetic field.

8.6 Forces

Consider an *isolated* system which consists of N circuits in which currents I_k are maintained ($k = 1, \ldots, N$), in the presence of linear magnetic media. Suppose one of the components of the system is displaced by $d\boldsymbol{r}$. The work by the force $\boldsymbol{F}$ acting on this component, is

$$\boldsymbol{F} \cdot d\boldsymbol{r} = -dW + dW_b, \tag{8.24}$$

where dW is the change in the magnetic energy and dW_b is the work by the batteries on maintaining the currents. From equation (8.17),

$$dW = \frac{1}{2} \sum_k I_k d\Phi_k.$$

The change in the flux in circuit k causes the emf $\mathcal{E}_k = -d\Phi_k/dt$, and the work by batteries on moving charges *against* these emf is

$$dW_b = \sum_k (-\mathcal{E}_k) dQ_k = \sum_k \frac{d\Phi_k}{dt} dQ_k = \sum_k d\Phi_k I_k = 2dW.$$

Hence, the work (8.24) is

$$\boldsymbol{F} \cdot d\boldsymbol{r} = dW,$$

and the force is

$$F_x = \left(\frac{\partial W}{\partial x} \right)_I, \quad F_y = \left(\frac{\partial W}{\partial y} \right)_I, \quad F_z = \left(\frac{\partial W}{\partial z} \right)_I, \tag{8.25}$$

where the subscript I indicates that the currents are maintained constant.

9 Maxwell's equations

9.1 Summary of field equations to date

From Sec. 2.3,

$$\nabla \cdot \boldsymbol{D} = \rho, \tag{9.1}$$

is an expression of Coulomb's law in electrostatics (charges at rest).

From Sec. 6.2,

$$\nabla \cdot \boldsymbol{B} = 0, \tag{9.2}$$

means there are no magnetic charges; magnetic fields are produced by currents.

From Sec. 8.2, we have the law of induction,

$$\nabla \times \boldsymbol{E} = -\frac{\partial \boldsymbol{B}}{\partial t}. \tag{9.3}$$

From Sec. 7.2,

$$\nabla \times \boldsymbol{H} = \boldsymbol{j}, \tag{9.4}$$

follows from Biot-Savart's law for steady currents ($\nabla \cdot \boldsymbol{j} = 0$).

What happens in a general situation when the charges are free to move and fields vary in time? Do equations (9.1)–(9.4) still apply?

9.2 Displacement current

One of the fundamental laws in nature is conservation of charge. This is expressed by the continuity equation (5.8),

$$\nabla \cdot \boldsymbol{j} + \frac{\partial \rho}{\partial t} = 0. \tag{9.5}$$

However, from Eq. (9.4) we have

$$\nabla \cdot \boldsymbol{j} = \nabla \cdot (\nabla \times \boldsymbol{H}) = 0,$$

which is only true for steady currents. In general $\partial \rho / \partial t \neq 0$ in Eq. (9.5), which means that Eq. (9.4) must be modified.

Maxwell realised that the correct form of Eq. (9.4) is obtained by adding the derivative $\partial \boldsymbol{D}/\partial t$ to the right-hand side[55], giving

$$\nabla \times \boldsymbol{H} = \boldsymbol{j} + \frac{\partial \boldsymbol{D}}{\partial t}. \tag{9.6}$$

He called this term the *displacement current*. Taking $\nabla \cdot$ of Eq. (9.6) we obtain

$$0 = \nabla \cdot \boldsymbol{j} + \nabla \cdot \frac{\partial \boldsymbol{D}}{\partial t}.$$

[55]In his own words, "The variations of the electrical displacement must be added to the currents to get the total motion of electricity" [Phil. Trans. R. Soc. Lond. **155**, 459 (1865)].

Changing the order of differentiation with respect to time and coordinates gives

$$0 = \nabla \cdot \boldsymbol{j} + \frac{\partial}{\partial t} \nabla \cdot \boldsymbol{D},$$

and substituting $\nabla \cdot \boldsymbol{D}$ from Eq. (9.1) gives the continuity equation (9.5).

9.3 Maxwell's equations

Following Maxwell's work, it has been established that the dynamics of the electric and magnetic fields in the general case is governed by the four *Maxwell equations*,

$$\nabla \cdot \boldsymbol{D} = \rho, \tag{9.7a}$$

$$\nabla \cdot \boldsymbol{B} = 0, \tag{9.7b}$$

$$\nabla \times \boldsymbol{E} = -\frac{\partial \boldsymbol{B}}{\partial t}, \tag{9.7c}$$

$$\nabla \times \boldsymbol{H} = \boldsymbol{j} + \frac{\partial \boldsymbol{D}}{\partial t}. \tag{9.7d}$$

The electric fields $\boldsymbol{D}$ and $\boldsymbol{E}$ and magnetic fields $\boldsymbol{H}$ and $\boldsymbol{B}$ related by

$$\boldsymbol{D} = \varepsilon_0 \boldsymbol{E} + \boldsymbol{P}, \tag{9.8a}$$

$$\boldsymbol{B} = \mu_0 (\boldsymbol{H} + \boldsymbol{M}), \tag{9.8b}$$

and it is assumed that the media providing the polarisation $\boldsymbol{P}$ and magnetisation $\boldsymbol{M}$ are at rest.

For the system of equations (9.7a)–(9.7d), (9.8a) and (9.8b) to be complete, we need to add relations between $\boldsymbol{E}$ and $\boldsymbol{P}$, and between $\boldsymbol{H}$ and $\boldsymbol{M}$. For example, if the media are linear and isotropic, we have $\boldsymbol{P} = \varepsilon_0 \chi \boldsymbol{E}$ and $\boldsymbol{M} = \chi_m \boldsymbol{H}$, or equivalently, $\boldsymbol{D} = \varepsilon \boldsymbol{E}$ and $\boldsymbol{B} = \mu \boldsymbol{H}$ (see Secs. 2.3 and 7.3).

The electric and magnetic fields determine the force experienced by a charge q, which is given by Lorentz's formula (Sec. 6.1.3),

$$\boldsymbol{F} = q\boldsymbol{E} + q\boldsymbol{v} \times \boldsymbol{B}. \tag{9.9}$$

9.4 Electromagnetic energy and Poynting vector

It is natural to assume that the total energy density of the electromagnetic field is the sum of the electric and magnetic contributions (4.11) and (8.23),

$$w = \frac{1}{2}(\boldsymbol{E} \cdot \boldsymbol{D} + \boldsymbol{H} \cdot \boldsymbol{B}). \tag{9.10}$$

To verify that this definition gives meaningful results, let us consider the total electromagnetic energy within a given volume V

$$\int_V w dV = \frac{1}{2} \int_V (\boldsymbol{E} \cdot \boldsymbol{D} + \boldsymbol{H} \cdot \boldsymbol{B}) dV,$$

and determine its rate of change. Assuming linear and isotropic media, we have

$$\begin{aligned}\frac{\partial}{\partial t}\int_V w dV &= \frac{1}{2}\frac{\partial}{\partial t}\int_V \left(\frac{1}{\varepsilon}\boldsymbol{D}\cdot\boldsymbol{D} + \frac{1}{\mu}\boldsymbol{B}\cdot\boldsymbol{B}\right) dV \\ &= \int_V \left(\frac{1}{\varepsilon}\boldsymbol{D}\cdot\frac{\partial \boldsymbol{D}}{\partial t} + \frac{1}{\mu}\boldsymbol{B}\cdot\frac{\partial \boldsymbol{B}}{\partial t}\right) dV \\ &= \int_V [\boldsymbol{E}\cdot(\nabla\times\boldsymbol{H} - \boldsymbol{j}) - \boldsymbol{H}\cdot(\nabla\times\boldsymbol{E})] dV, \end{aligned} \tag{9.11}$$

where we used Maxwell's equations (9.7c) and (9.7d). Using the identity

$$\nabla\cdot(\boldsymbol{E}\times\boldsymbol{H}) = \boldsymbol{H}\cdot(\nabla\times\boldsymbol{E}) - \boldsymbol{E}\cdot(\nabla\times\boldsymbol{H}),$$

we transform (9.11) into

$$\frac{\partial}{\partial t}\int_V w dV = -\int_V \nabla\cdot(\boldsymbol{E}\times\boldsymbol{H}) dV - \int_V \boldsymbol{j}\cdot\boldsymbol{E} dV.$$

Using Gauss's theorem we change the first term on the right-hand side into the surface integral, and obtain

$$\frac{\partial}{\partial t}\int_V w dV = -\oint_S (\boldsymbol{E}\times\boldsymbol{H})\cdot d\boldsymbol{S} - \int_V \boldsymbol{j}\cdot\boldsymbol{E}\, dV, \tag{9.12}$$

where S is the surface that bounds volume V.

The first term on the right-hand side of equation (9.12) describes the loss of energy in volume V due to the flow of electromagnetic energy across the surface S. The corresponding energy flux density

$$\boldsymbol{S} = \boldsymbol{E}\times\boldsymbol{H} \tag{9.13}$$

is known as the *Poynting vector*[56].

The second term in (9.12) describes *Joule heating*, i.e., the loss of energy due to dissipation into heat. Indeed, recall that the current density can be written as $\boldsymbol{j} = qn\boldsymbol{v}$, where q and n are the charge and number density of current carriers, and $\boldsymbol{v}$ is their velocity (see Sec. 5.2). Lorentz's force acting on each charge carrier is given by equation (9.9). The work by this force is

$$\boldsymbol{F}\cdot d\boldsymbol{r} = (q\boldsymbol{E} + q\boldsymbol{v}\times\boldsymbol{B})\cdot\boldsymbol{v}dt = q\boldsymbol{E}\cdot\boldsymbol{v}dt,$$

and the work per unit volume per unit time (i.e., power density) is given by

$$qn\boldsymbol{v}\cdot\boldsymbol{E} = \boldsymbol{j}\cdot\boldsymbol{E}.$$

This can be compared with the Joule heating in conductors (Sec. 5.7).

Thus we see that equation (9.12) describes energy conservation for the electromagnetic field.

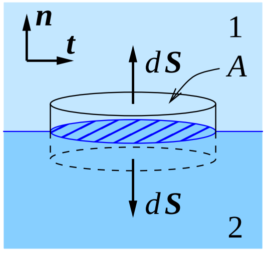
[56]Note that the first term on the right-hand side of Eq. (9.12) can now be written as $\oint \boldsymbol{S}\cdot d\boldsymbol{S}$, where $\boldsymbol{S}$ stands for the Poynting vector while $d\boldsymbol{S}$ is the usual surface element. The two $\boldsymbol{S}$ should not be confused!

9.5 Boundary conditions

Consider the boundary between two materials, 1 and 2, with the normal vector $\boldsymbol{n}$ from 2 into 1 and tangential vector $\boldsymbol{t}$. The boundary conditions for the electric and magnetic fields are the same as obtained in Secs. 2.6 and 7.4:

$$E_{1t} - E_{2t} = 0, \tag{9.14a}$$
$$D_{1n} - D_{2n} = \sigma, \tag{9.14b}$$
$$B_{1n} - B_{2n} = 0, \tag{9.14c}$$
$$H_{1t} - H_{2t} = \boldsymbol{J} \cdot (\boldsymbol{n} \times \boldsymbol{t}), \tag{9.14d}$$

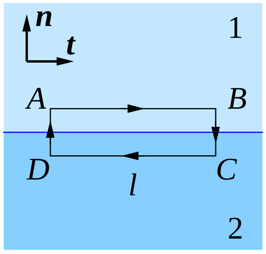
where subscripts 1 and 2 label quantities in materials 1 and 2, σ is the surface charge density on the boundary, and $\boldsymbol{J}$ is the surface current density.



The boundary conditions (9.14b) and (9.14c) are derived from Maxwell's equations (9.7a) and (9.7b), respectively. In integral form, these equations read

$$\oint_S \boldsymbol{D} \cdot d\boldsymbol{S} = Q, \qquad \oint_S \boldsymbol{B} = 0,$$

where Q is the free charge enclosed by S. We choose S as a small cylinder with bases parallel to the surface and of infinitesimal height, so that the flux across the curved surface can be neglected (see Sec. 2.6 and 7.4 for details).



Boundary conditions (9.14a) and (9.14d) are derived from Maxwell's equations (9.7c) and (9.7d). In integral form, they read

$$\oint_L \boldsymbol{E} \cdot d\boldsymbol{r} = -\frac{d}{dt} \int_S \boldsymbol{B} \cdot d\boldsymbol{S}, \qquad \oint_L \boldsymbol{H} \cdot d\boldsymbol{r} = I + \frac{d}{dt} \int_S \boldsymbol{D} \cdot d\boldsymbol{S},$$

where S is any surface bounded by the closed curve L, and I is the current across S (or through L). We choose L as a rectangle with long sides parallel to the surface and vanishingly small short sides. In this case the flux of $\boldsymbol{B}$ or $\boldsymbol{D}$ across S is negligible and the static result holds (Sec. 2.6 and 7.4).

9.6 Electromagnetic potentials

The equation $\nabla \cdot \boldsymbol{B} = 0$ is satisfied by both static and time-dependent magnetic fields. Hence, as in Sec. 6.5, the magnetic induction $\boldsymbol{B}$ can be described by the vector potential $\boldsymbol{A}$,

$$\boldsymbol{B} = \nabla \times \boldsymbol{A}. \tag{9.15}$$

Using this in Maxwell's equation (9.7c), changing the order of the spatial and time derivatives and rearranging, we obtain

$$\nabla \times \left(\boldsymbol{E} + \frac{\partial \boldsymbol{A}}{\partial t} \right) = 0.$$

Any vector whose curl is identically zero can be written as the gradient of a scalar function, hence,

$$\boldsymbol{E} + \frac{\partial \boldsymbol{A}}{\partial t} = -\nabla \phi.$$

This means that the electric field is given by

$$\boldsymbol{E} = -\nabla\phi - \frac{\partial \boldsymbol{A}}{\partial t}, \tag{9.16}$$

where ϕ is the *scalar potential*. For time-independent fields $\boldsymbol{E} = -\nabla\phi$, and ϕ is the familiar electrostatic potential (Sec. 1.4).

As we saw in Sec. 6.5, the vector potential is not unique. Replacing

$$\boldsymbol{A} \to \boldsymbol{A}' = \boldsymbol{A} + \nabla\chi, \tag{9.17}$$

does not change the magnetic field:

$$\boldsymbol{B} = \nabla \times \boldsymbol{A}' = \nabla \times (\boldsymbol{A} + \nabla\chi) = \nabla \times \boldsymbol{A} + \underbrace{\nabla \times \nabla}_{=0}\chi = \nabla \times \boldsymbol{A}.$$

On the other hand, substituting

$$\boldsymbol{A} = \boldsymbol{A}' - \nabla\chi$$

into the right-hand side of equation (9.16) gives

$$\begin{aligned}\boldsymbol{E} &= -\nabla\phi - \frac{\partial}{\partial t}(\boldsymbol{A}' - \nabla\chi) \\ &= -\nabla\phi + \frac{\partial\nabla\chi}{\partial t} - \frac{\partial \boldsymbol{A}'}{\partial t} \\ &= -\nabla\left(\phi - \frac{\partial\chi}{\partial t}\right) - \frac{\partial \boldsymbol{A}'}{\partial t}.\end{aligned}$$

We see that if, simultaneously with (9.17), we change the scalar potential,

$$\phi \to \phi' = \phi - \frac{\partial\chi}{\partial t}, \tag{9.18}$$

the electric field

$$\boldsymbol{E} = -\nabla\phi' - \frac{\partial \boldsymbol{A}'}{\partial t}$$

will be the same as (9.16).

The transformation of the vector and scalar potentials (9.17) and (9.18) is known as a *gauge transformation*. The property of the electric and magnetic fields to remain unchanged upon such transformations is called *gauge invariance*. A particular choice of ϕ and $\boldsymbol{A}$ is often referred to as a *gauge*[57].

The vector and scalar potentials are introduced in a way that Maxwell's equations (9.7b) and (9.7c) are automatically satisfied. The equations for the potentials are then obtained by substituting the electric and magnetic fields, Eqs. (9.15) and (9.16), in the other two Maxwell's equations (9.7a) and (9.7d).

[57] In both Classical and Quantum Electrodynamics the freedom to choose a gauge offers mathematical convenience, by simplifying equations and helping to solve particular problems. In quantum field theories, such as those of electroweak and strong interactions, and in the Standard Model, gauge invariance becomes the leading principle for introducing the interaction between elementary particles in a consistent and effective way.

Assuming linear, isotropic media ($\boldsymbol{D} = \varepsilon \boldsymbol{E}$ and $\boldsymbol{H} = \boldsymbol{B}/\mu$), we obtain from (9.7d):

$$\nabla \times (\nabla \times \boldsymbol{A}) = \mu \boldsymbol{j} - \varepsilon\mu \frac{\partial}{\partial t}\left(\nabla\phi + \frac{\partial \boldsymbol{A}}{\partial t}\right). \tag{9.19}$$

Using the identity

$$\nabla \times (\nabla \times \boldsymbol{A}) = \nabla(\nabla \cdot \boldsymbol{A}) - \nabla^2 \boldsymbol{A}, \tag{9.20}$$

we obtain from (9.19),

$$\nabla^2 \boldsymbol{A} - \varepsilon\mu \frac{\partial^2 \boldsymbol{A}}{\partial t^2} - \nabla\left(\nabla \cdot \boldsymbol{A} + \varepsilon\mu \frac{\partial \phi}{\partial t}\right) = -\mu \boldsymbol{j}. \tag{9.21}$$

From Maxwell's equation (9.7a) we have

$$-\varepsilon \nabla \cdot \left(\nabla\phi + \frac{\partial \boldsymbol{A}}{\partial t}\right) = \rho,$$

which can be written as

$$\nabla^2 \phi - \varepsilon\mu \frac{\partial^2 \phi}{\partial t^2} + \frac{\partial}{\partial t}\left(\nabla \cdot \boldsymbol{A} + \varepsilon\mu \frac{\partial \phi}{\partial t}\right) = -\frac{\rho}{\varepsilon}. \tag{9.22}$$

One can show (see below) that the potentials $\boldsymbol{A}$ and ϕ can always be chosen in such a way that they satisfy the condition

$$\nabla \cdot \boldsymbol{A} + \varepsilon\mu \frac{\partial \phi}{\partial t} = 0. \tag{9.23}$$

For this choice of the potentials (known as *Lorenz gauge*[58]), equations (9.21) and (9.22) assume a particularly simple form:

$$\nabla^2 \boldsymbol{A} - \varepsilon\mu \frac{\partial^2 \boldsymbol{A}}{\partial t^2} = -\mu \boldsymbol{j}, \tag{9.24a}$$

$$\nabla^2 \phi - \varepsilon\mu \frac{\partial^2 \phi}{\partial t^2} = -\frac{\rho}{\varepsilon}. \tag{9.24b}$$

For $\boldsymbol{j} = 0$ and $\rho = 0$ these are the *wave equations*. Their solutions describe waves propagating with velocity $v = 1/\sqrt{\varepsilon\mu}$ (see Ch. 10).

Let us now show that by performing a suitable gauge transformation, the potentials can always be made to satisfy the Lorenz condition (9.23). Suppose the initial choice of the potentials $\boldsymbol{A}'$ and ϕ' is such that

$$\nabla \cdot \boldsymbol{A}' + \varepsilon\mu \frac{\partial \phi'}{\partial t} = f \neq 0. \tag{9.25}$$

Using the transformations (9.17) and (9.18), we obtain

$$\nabla \cdot \boldsymbol{A} + \varepsilon\mu \frac{\partial \phi}{\partial t} + \nabla^2 \chi - \varepsilon\mu \frac{\partial^2 \chi}{\partial t^2} = f.$$

By choosing the function χ that satisfies the equation

$$\nabla^2 \chi - \varepsilon\mu \frac{\partial^2 \chi}{\partial t^2} = f, \tag{9.26}$$

(which can always be done, see Ch. 11.1), we ensure that the potentials $\boldsymbol{A}$ and ϕ do satisfy Lorenz's condition (9.23).

[58]This is named after Ludvig Lorenz (1829–1891), a Danish physicist who proposed it, not to be confused with Hendrik Antoon Lorentz (1953–1928), a Dutch physicist whose name is associated with the Lorentz force and Lorentz transformations of special relativity.

10 Electromagnetic waves

10.1 Wave equation

Taking the curl of 3rd Maxwell's equation (9.7c), we have

$$\nabla \times (\nabla \times \boldsymbol{E}) = -\frac{\partial}{\partial t}\nabla \times \boldsymbol{B}.$$

Assuming linear, isotropic and homogeneous media ($\boldsymbol{B} = \mu\boldsymbol{H}$ and $\boldsymbol{D} = \varepsilon\boldsymbol{E}$ with constant ε and μ), and using vector identity (9.20) on the left-hand side and 4th Maxwell's equation (9.7d) on the right-hand side, we obtain

$$\nabla(\nabla \cdot \boldsymbol{E}) - \nabla^2\boldsymbol{E} = -\frac{\partial}{\partial t}\left(\mu\boldsymbol{j} + \mu\varepsilon\frac{\partial \boldsymbol{E}}{\partial t}\right).$$

Finally, making use of Ohm's law ($\boldsymbol{j} = \sigma\boldsymbol{E}$), expressing $\nabla \cdot \boldsymbol{E}$ from 1st Maxwell's equation (9.7a), and rearranging, we find the equation for the electric field,

$$\nabla^2\boldsymbol{E} - \varepsilon\mu\frac{\partial^2 \boldsymbol{E}}{\partial t^2} - \mu\sigma\frac{\partial \boldsymbol{E}}{\partial t} = \frac{1}{\varepsilon}\nabla\rho. \tag{10.1}$$

Similarly, taking curl of 4th Maxwell's equation (9.7d),

$$\nabla \times (\nabla \times \boldsymbol{B}) = \mu\nabla \times \boldsymbol{j} + \varepsilon\mu\frac{\partial}{\partial t}\nabla \times \boldsymbol{E},$$

then making use of Ohm's law and equations (9.7b) and (9.7c),

$$-\nabla^2\boldsymbol{B} = \mu\sigma\nabla \times \boldsymbol{E} - \varepsilon\mu\frac{\partial^2 \boldsymbol{B}}{\partial t^2},$$

and finally using (9.7c) again, gives the equation for the magnetic field,

$$\nabla^2\boldsymbol{B} - \varepsilon\mu\frac{\partial^2 \boldsymbol{B}}{\partial t^2} - \mu\sigma\frac{\partial \boldsymbol{B}}{\partial t} = 0. \tag{10.2}$$

For a charge-free ($\rho = 0$), nonconducting ($\sigma = 0$) medium, equations (10.1) and (10.2) become

$$\nabla^2\boldsymbol{E} - \varepsilon\mu\frac{\partial^2 \boldsymbol{E}}{\partial t^2} = 0, \tag{10.3a}$$

$$\nabla^2\boldsymbol{B} - \varepsilon\mu\frac{\partial^2 \boldsymbol{B}}{\partial t^2} = 0. \tag{10.3b}$$

These equations have the form of the standard *wave equation*,

$$\nabla^2 u - \frac{1}{v^2}\frac{\partial^2 u}{\partial t^2} = 0, \tag{10.4}$$

whose solutions describe waves propagating with speed v (see Sec. 10.2). Comparing (10.4) with (10.3a) and (10.3b), we see that for *electromagnetic waves*

$$v = \frac{1}{\sqrt{\varepsilon\mu}}. \tag{10.5}$$

In particular, in vacuum this speed (denoted c) is given by

$$c = \frac{1}{\sqrt{\varepsilon_0\mu_0}} = \left(8.854 \times 10^{-12} \times 4\pi \times 10^{-7}\right)^{-1/2} = 2.998 \times 10^8 \ \mathrm{m\,s^{-1}}. \tag{10.6}$$

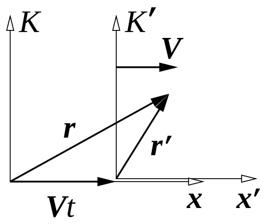
This is in fact the speed of light. When this was first noticed by Maxwell, it immediately led him to conclude[59] that light is an electromagnetic wave!

The ratio

$$n = \frac{c}{v} = \frac{\sqrt{\varepsilon\mu}}{\sqrt{\varepsilon_0\mu_0}} = \sqrt{\kappa\kappa_m}, \tag{10.7}$$

where κ is the dielectric constant and κ_m is the relative permeability, is called the *refractive index* of the medium.

Understanding of the electromagnetic nature of light posed a serious question that was only resolved by Einstein. In classical mechanics, if we consider an inertial reference frame K' which moves with constant velocity $\boldsymbol{V}$ relative to an inertial frame K, the position vectors of a point in the two frames are related by[60]



$$\boldsymbol{r} = \boldsymbol{r}' + \boldsymbol{V}t. \tag{10.8}$$

Differentiating (10.8) with respect to time gives the classical law of addition of velocities,

$$\boldsymbol{v} = \boldsymbol{v}' + \boldsymbol{V}, \tag{10.9}$$

where $\boldsymbol{v}$ and $\boldsymbol{v}'$ are the velocities of the point in frames K and K', respectively.

According to (10.9), electromagnetic waves, and light in particular, will propagate with different speeds in different frames of reference. Mathematically, this also means that if we change variables in the wave equation, e.g., in one dimension,

$$\frac{\partial^2 u}{\partial x^2} - \frac{1}{c^2}\frac{\partial^2 u}{\partial t^2} = 0, \tag{10.10}$$

using

$$x' = x - Vt, \qquad t' = t, \tag{10.11}$$

the form of the wave equation will change. This would also be true for Maxwell's equations (9.7a)–(9.7d), i.e., they would not look so nice any more. This means that *physics* itself would be different in different inertial frames!

One way out of this was to pronounce that there is a substance (ether) that mediates the propagation of electromagnetic waves, and, hence, a special frame

[59] "This velocity is so nearly that of light, that it seems we have strong reason to conclude that light itself (including radiant heat, and other radiations if any) is an electromagnetic disturbance in the form of waves propagated through the electromagnetic field according to electromagnetic laws." [J. Clerk Maxwell, *A Dynamical Theory of the Electromagnetic Field*, Phil. Trans. R. Soc. Lond. **155**, 459–512 (1865)].

[60] Time is measured in a way that at $t = 0$ the origins of the two frames coincide.

of reference in which Maxwell's equations hold. Another point of view, argued by Einstein[61], was that all inertial frames are equivalent, and that the wave equation and Maxwell's equations that lead to it, should look the same in all of them. This, however, requires one to replace the intuitive relation (10.11) and the notion of absolute time, by *Lorentz's transformations*,

$$x' = \frac{x - Vt}{\sqrt{1 - \dfrac{V^2}{c^2}}}, \qquad t' = \frac{t - \dfrac{V}{c^2}x}{\sqrt{1 - \dfrac{V^2}{c^2}}}. \tag{10.12}$$

Exercise: Verify that the change of variables from x, t to x', t', described by (10.12) leaves the wave equation (10.10) unchanged.

10.2 Plane waves

Consider a uniform, charge-free ($\rho = 0$) nonconducting ($\boldsymbol{j} = 0$) medium. In this case we can choose a gauge in which

$$\phi = 0 \quad \text{and} \quad \nabla \cdot \boldsymbol{A} = 0. \tag{10.13}$$

For the scalar potential to vanish, we perform a gauge transformation (9.18) with $\chi = \int \phi dt$. For $\nabla \cdot \boldsymbol{A} = 0$, from $\nabla \cdot \boldsymbol{E} = 0$, we notice [using (9.16)] that

$$\frac{\partial}{\partial t} \nabla \cdot \boldsymbol{A} = 0,$$

which means that $\nabla \cdot \boldsymbol{A}$ does not depend on time. We can then perform a second gauge transformation (9.17) with a time-independent χ, which does not change ϕ and makes $\nabla \cdot \boldsymbol{A}' = 0$. (For this we need χ to satisfy $\nabla^2 \chi = -\nabla \cdot \boldsymbol{A}$.)

For (10.13), Lorenz's condition (9.23) is also fulfilled. Hence, the potential satisfies the equation

$$\nabla^2 \boldsymbol{A} - \varepsilon\mu \frac{\partial^2 \boldsymbol{A}}{\partial t^2} = 0, \tag{10.14}$$

and the fields are given by

$$\boldsymbol{E} = -\frac{\partial \boldsymbol{A}}{\partial t}, \qquad \boldsymbol{B} = \nabla \times \boldsymbol{A}. \tag{10.15}$$

Let us consider a particular type of electromagnetic waves in which the fields depend only on one spatial coordinate, say, on x. Such waves are known as *plane waves*. The field equations in this case take the form of a wave equation in one dimension,

$$\frac{\partial^2 u}{\partial x^2} - \frac{1}{c^2}\frac{\partial^2 u}{\partial t^2} = 0, \tag{10.16}$$

where $c = 1/\sqrt{\varepsilon\mu}$ here is the speed of light in the medium. To solve this equation, we re-write it as

$$\left(\frac{\partial}{\partial x} - \frac{1}{c}\frac{\partial}{\partial t}\right)\left(\frac{\partial}{\partial x} + \frac{1}{c}\frac{\partial}{\partial t}\right) u = 0, \tag{10.17}$$

[61] Albert Einstein put forward his Special Theory of Relativity in 1905.

and introduce new variables

$$\xi = x - ct, \quad \eta = x + ct. \tag{10.18}$$

The derivatives transform as

$$\frac{\partial u}{\partial x} = \frac{\partial u}{\partial \xi}\frac{\partial \xi}{\partial x} + \frac{\partial u}{\partial \eta}\frac{\partial \eta}{\partial x} = \frac{\partial u}{\partial \xi} + \frac{\partial u}{\partial \eta},$$

and

$$\frac{\partial u}{\partial t} = \frac{\partial u}{\partial \xi}\frac{\partial \xi}{\partial t} + \frac{\partial u}{\partial \eta}\frac{\partial \eta}{\partial t} = c\frac{\partial u}{\partial \eta} - c\frac{\partial u}{\partial \xi},$$

so that

$$\frac{\partial}{\partial x} - \frac{1}{c}\frac{\partial}{\partial t} = 2\frac{\partial}{\partial \xi}, \qquad \frac{\partial}{\partial x} + \frac{1}{c}\frac{\partial}{\partial t} = 2\frac{\partial}{\partial \eta}.$$

Equation (10.17) then becomes

$$\frac{\partial^2 u}{\partial \xi \partial \eta} = 0. \tag{10.19}$$

Integrating it with respect to ξ and η we find the solution $u = f(\xi) + g(\eta)$, i.e.,

$$u = f(x - ct) + g(x + ct), \tag{10.20}$$

where f and g are arbitrary functions[62].

To understand the meaning of the solution (10.20), consider the case $g = 0$, so that $u = f(x - ct)$. At $t = 0$ this solution has a shape given by $u = f(x)$. The solution for $t > 0$ is the same function shifted by ct along the x axis. We see that according to this solution the values of the electromagnetic field are propagated along the x axis with a constant velocity c. Similarly, the second term in (10.20) represents a wave that propagates in the negative x direction.

Since the fields in a plane wave depend only on x and t, the condition $\nabla \cdot \boldsymbol{A} = 0$ gives

$$\frac{\partial A_x}{\partial x} = 0.$$

When substituted into the wave equation, this gives

$$\frac{\partial^2 A_x}{\partial t^2} = 0 \quad \Longrightarrow \quad \frac{\partial A_x}{\partial t} = \text{const}.$$

Since $\partial \boldsymbol{A}/\partial t$ determines the electric field, the latter condition allows for the presence of a constant electric field in the x direction. Such field has no relation to the electromagnetic wave and we can set $A_x = 0$.

The vector potential $\boldsymbol{A}$ of the plane wave can thus be chosen perpendicular to the x axis, i.e., the direction of propagation of the wave. Since $\boldsymbol{A} = \boldsymbol{A}(\xi) = \boldsymbol{A}(x - ct)$ for the wave propagating in the positive x direction, we have

$$\boldsymbol{E} = -\frac{\partial \boldsymbol{A}}{\partial t} = c\boldsymbol{A}', \tag{10.21}$$

[62] This is known as d'Alembert's solution.

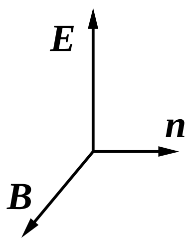
where the prime denotes differentiation with respect to ξ. The magnetic field is

$$\boldsymbol{B} = \nabla \times \boldsymbol{A} = \begin{vmatrix} \boldsymbol{i} & \boldsymbol{j} & \boldsymbol{k} \\ \dfrac{\partial}{\partial x} & \dfrac{\partial}{\partial y} & \dfrac{\partial}{\partial z} \\ 0 & A_y & A_z \end{vmatrix} = -\boldsymbol{j}\frac{\partial A_z}{\partial x} + \boldsymbol{k}\frac{\partial A_y}{\partial x} = \boldsymbol{i} \times \frac{\partial}{\partial x}(A_y\boldsymbol{j} + A_z\boldsymbol{k}),$$

so that

$$\boldsymbol{B} = \boldsymbol{n} \times \boldsymbol{A}', \tag{10.22}$$

where $\boldsymbol{n}$ is the unit vector in the direction of propagation of the wave[63].



From (10.21) and (10.22) we see that the electric and magnetic fields are perpendicular to the direction of propagation of the wave (i.e., the wave is *transverse*), and are related by

$$\boldsymbol{B} = \frac{1}{c}\boldsymbol{n} \times \boldsymbol{E}. \tag{10.23}$$

For their magnitudes, we have

$$B = E/c = \sqrt{\varepsilon\mu}E. \tag{10.24}$$

The Poynting vector of the plane wave is

$$\boldsymbol{S} = \boldsymbol{E} \times \boldsymbol{H} = \frac{\boldsymbol{E} \times \boldsymbol{B}}{\mu} = \frac{EB\boldsymbol{n}}{\mu} = \sqrt{\frac{\varepsilon}{\mu}}E^2\boldsymbol{n}, \tag{10.25}$$

and the energy density is

$$w = \frac{1}{2}(\boldsymbol{E}\cdot\boldsymbol{D} + \boldsymbol{B}\cdot\boldsymbol{H}) = \frac{1}{2}\left(\varepsilon E^2 + \frac{1}{\mu}B^2\right) = \frac{1}{2}(\varepsilon E^2 + \varepsilon E^2) = \varepsilon E^2. \tag{10.26}$$

These quantities are related by

$$\boldsymbol{S} = cw\boldsymbol{n}, \tag{10.27}$$

as it should be for a wave travelling with velocity c.

10.3 Monochromatic plane waves

If the electromagnetic field of a wave has a simple periodic dependence on time, the wave is called *monochromatic*. All quantities that describe such wave depend on time as $\cos\omega t$ or $\sin\omega t$, or, in general, $\cos(\omega t + \alpha)$, where ω is the *angular frequency* (often simply called frequency).

[63] A quicker way of finding $\boldsymbol{B}$ makes use of the chain rule. Since ∇ is the vector operator of "differentiation with respect to $\boldsymbol{r}$", and $\boldsymbol{A}$ depends of $\boldsymbol{r}$ through ξ, we have:

$$\boldsymbol{B} = \nabla \times \boldsymbol{A} = \nabla(\xi) \times d\boldsymbol{A}/d\xi = \nabla(x - ct) \times \boldsymbol{A}' = \boldsymbol{i} \times \boldsymbol{A}' = \boldsymbol{n} \times \boldsymbol{A}'.$$

Note that the order of the vector (operator) ∇ and $\boldsymbol{A}$ in the cross product is preserved.

For a plane wave in the positive x direction, the field depends on

$$x - ct = \frac{c}{\omega}\left(\frac{\omega}{c}x - \omega t\right).$$

Introducing the *wave number*,

$$k \equiv \frac{\omega}{c}, \tag{10.28}$$

we see that the fields in a monochromatic plane wave depend on $kx - \omega t$.

In general, if the wave travels in the direction of unit vector $\boldsymbol{n}$, we replace x by $\boldsymbol{n} \cdot \boldsymbol{r}$ (projection of $\boldsymbol{r}$ onto the direction of $\boldsymbol{n}$), and $kx - \omega t$ is changed to $\boldsymbol{k} \cdot \boldsymbol{r} - \omega t$, where

$$\boldsymbol{k} = k\boldsymbol{n} \tag{10.29}$$

is the *wave vector*.

Since Maxwell's equations and other relations satisfied by the potentials and fields are linear, it is convenient to use complex exponents, rather than sines or cosines, for studying monochromatic plane waves, e.g., $\boldsymbol{A} = \boldsymbol{A}_0 e^{i(\boldsymbol{k}\cdot\boldsymbol{r}-\omega t)}$. The physical potentials and fields are then given by the real parts of the complex fields,

$$\boldsymbol{A} = \mathrm{Re}\left[\boldsymbol{A}_0 e^{i(\boldsymbol{k}\cdot\boldsymbol{r}-\omega t)}\right], \tag{10.30}$$

$$\boldsymbol{E} = \mathrm{Re}\left[\boldsymbol{E}_0 e^{i(\boldsymbol{k}\cdot\boldsymbol{r}-\omega t)}\right], \tag{10.31}$$

$$\boldsymbol{B} = \mathrm{Re}\left[\boldsymbol{B}_0 e^{i(\boldsymbol{k}\cdot\boldsymbol{r}-\omega t)}\right]. \tag{10.32}$$

where $\boldsymbol{A}_0$, $\boldsymbol{E}_0$ and $\boldsymbol{B}_0$ are complex amplitudes[64].

For example, the electric field of a wave satisfies the wave equation (10.3a),

$$\nabla^2 \boldsymbol{E} - \frac{1}{c^2}\frac{\partial^2 \boldsymbol{E}}{\partial t^2} = 0, \tag{10.33}$$

where $c = 1/\sqrt{\varepsilon\mu}$. Using $\boldsymbol{E} = \boldsymbol{E}_0 e^{i(\boldsymbol{k}\cdot\boldsymbol{r}-\omega t)}$, we find

$$\frac{\partial^2 \boldsymbol{E}}{\partial t^2} = (-i\omega)^2 \boldsymbol{E}_0 e^{i(\boldsymbol{k}\cdot\boldsymbol{r}-\omega t)} = -\omega^2 \boldsymbol{E}_0 e^{i(\boldsymbol{k}\cdot\boldsymbol{r}-\omega t)}.$$

To find $\nabla^2 \boldsymbol{E}$, note that for a constant vector $\boldsymbol{E}_0$, we have

$$\nabla^2 \boldsymbol{E}_0 e^{i(\boldsymbol{k}\cdot\boldsymbol{r}-\omega t)} = \boldsymbol{E}_0 \nabla^2 e^{i(\boldsymbol{k}\cdot\boldsymbol{r}-\omega t)} = \boldsymbol{E}_0 (\nabla \cdot \nabla) e^{i(\boldsymbol{k}\cdot\boldsymbol{r}-\omega t)}.$$

The last expression is found using the chain rule,

$$\nabla e^{i(\boldsymbol{k}\cdot\boldsymbol{r}-\omega t)} = e^{i(\boldsymbol{k}\cdot\boldsymbol{r}-\omega t)} \nabla[i(\boldsymbol{k}\cdot\boldsymbol{r} - \omega t)] = i\boldsymbol{k} e^{i(\boldsymbol{k}\cdot\boldsymbol{r}-\omega t)},$$

where we used $\nabla(\boldsymbol{k}\cdot\boldsymbol{r}) = \boldsymbol{k}$, followed by

$$\nabla \cdot \left[i\boldsymbol{k} e^{i(\boldsymbol{k}\cdot\boldsymbol{r}-\omega t)}\right] = i\boldsymbol{k} \cdot \nabla e^{i(\boldsymbol{k}\cdot\boldsymbol{r}-\omega t)} = (i\boldsymbol{k})^2 e^{i(\boldsymbol{k}\cdot\boldsymbol{r}-\omega t)},$$

so that

$$\nabla^2 \boldsymbol{E}_0 e^{i(\boldsymbol{k}\cdot\boldsymbol{r}-\omega t)} = -k^2 \boldsymbol{E}_0 e^{i(\boldsymbol{k}\cdot\boldsymbol{r}-\omega t)}.$$

[64] If the equations are satisfied by complex potentials or fields, both their real and imaginary parts also satisfy them.

The wave equation (10.33) then gives

$$\left(-k^2 + \frac{\omega^2}{c^2}\right) \boldsymbol{E}_0 e^{i(\boldsymbol{k}\cdot\boldsymbol{r}-\omega t)} = 0,$$

which yields

$$k^2 = \frac{\omega^2}{c^2} \quad \text{or} \quad k = \frac{\omega}{c},$$

in agreement with (10.28).

Similarly, from Maxwell's third equation (9.7c),

$$\nabla \times \boldsymbol{E} = -\frac{\partial \boldsymbol{B}}{\partial t},$$

we find

$$\begin{aligned} \nabla \times \boldsymbol{E}_0 e^{i(\boldsymbol{k}\cdot\boldsymbol{r}-\omega t)} &= -\frac{\partial}{\partial t} \boldsymbol{B}_0 e^{i(\boldsymbol{k}\cdot\boldsymbol{r}-\omega t)}, \\ \nabla \left[e^{i(\boldsymbol{k}\cdot\boldsymbol{r}-\omega t)}\right] \times \boldsymbol{E}_0 &= i\omega \boldsymbol{B}_0 e^{i(\boldsymbol{k}\cdot\boldsymbol{r}-\omega t)}, \\ \boldsymbol{k} \times \boldsymbol{E}_0 e^{i(\boldsymbol{k}\cdot\boldsymbol{r}-\omega t)} &= \omega \boldsymbol{B}_0 e^{i(\boldsymbol{k}\cdot\boldsymbol{r}-\omega t)}. \end{aligned} \tag{10.34}$$

Hence we obtain the relation between the electric and magnetic fields in the wave

$$\boldsymbol{B} = \frac{\boldsymbol{k}}{\omega} \times \boldsymbol{E} = \frac{1}{c} \boldsymbol{n} \times \boldsymbol{E}, \tag{10.35}$$

seen earlier in (10.23), or, cancelling the exponents in (10.34), the relation between their amplitudes,

$$\boldsymbol{B}_0 = \frac{\boldsymbol{k}}{\omega} \times \boldsymbol{E}_0 = \frac{1}{c} \boldsymbol{n} \times \boldsymbol{E}_0 = \sqrt{\varepsilon\mu}\, \boldsymbol{n} \times \boldsymbol{E}_0. \tag{10.36}$$

Exercise: Using $\nabla \cdot \boldsymbol{E} = 0$ show that a monochromatic plane wave is transverse.

Note that in using complex fields, caution must be exercised when *multiplying* them, e.g., when calculating the field energy density (9.10) or the Poynting vector (9.13)

$$\boldsymbol{S} = \boldsymbol{E} \times \boldsymbol{H}.$$

If $\boldsymbol{E}$ and $\boldsymbol{H} = \boldsymbol{B}/\mu$ are complex, we must use

$$\boldsymbol{S} = \operatorname{Re} \boldsymbol{E} \times \operatorname{Re} \boldsymbol{H},$$

which gives[65]

$$\begin{aligned} \boldsymbol{S} &= \frac{1}{2}\left[\boldsymbol{E}_0 e^{i(\boldsymbol{k}\cdot\boldsymbol{r}-\omega t)} + \boldsymbol{E}_0^* e^{-i(\boldsymbol{k}\cdot\boldsymbol{r}-\omega t)}\right] \times \frac{1}{2}\left[\boldsymbol{H}_0 e^{i(\boldsymbol{k}\cdot\boldsymbol{r}-\omega t)} + \boldsymbol{H}_0^* e^{-i(\boldsymbol{k}\cdot\boldsymbol{r}-\omega t)}\right] \\ &= \frac{1}{4}\left[\boldsymbol{E}_0 \times \boldsymbol{H}_0^* + \boldsymbol{E}_0^* \times \boldsymbol{H}_0 + \boldsymbol{E}_0 \times \boldsymbol{H}_0 e^{2i(\boldsymbol{k}\cdot\boldsymbol{r}-\omega t)} + \boldsymbol{E}_0^* \times \boldsymbol{H}_0^* e^{-2i(\boldsymbol{k}\cdot\boldsymbol{r}-\omega t)}\right]. \end{aligned}$$

[65] Recall that for a complex number z, $\operatorname{Re} z = \frac{1}{2}(z+z^*)$, where z^* is its complex conjugate.

If we average the Poynting vector over time (i.e., over one period of the field), the terms containing $e^{2i(\boldsymbol{k}\cdot\boldsymbol{r}-\omega t)}$ and $e^{-2i(\boldsymbol{k}\cdot\boldsymbol{r}-\omega t)}$ give zero[66], and we find

$$\langle \boldsymbol{S} \rangle = \frac{1}{2}\mathrm{Re}\,(\boldsymbol{E}_0 \times \boldsymbol{H}_0^*) = \frac{1}{2}\mathrm{Re}\,(\boldsymbol{E} \times \boldsymbol{H}^*). \tag{10.37}$$

Similarly, the time-averaged energy density is given by $\langle w \rangle = \frac{1}{4}\,\mathrm{Re}\,(\boldsymbol{E}_0 \cdot \boldsymbol{D}_0^* + \boldsymbol{B}_0 \cdot \boldsymbol{H}_0^*) = \frac{1}{4}\,\mathrm{Re}\,(\boldsymbol{E} \cdot \boldsymbol{D}^* + \boldsymbol{B} \cdot \boldsymbol{H}^*)$.

Let us now find how the direction of the electric field $\boldsymbol{E}$ in a monochromatic plane wave depends on time. [The magnetic field $\boldsymbol{B}$ of the wave is perpendicular to $\boldsymbol{E}$ at every point, see (10.35).]

For a complex amplitude $\boldsymbol{E}_0$, its square is also generally complex,

$$\boldsymbol{E}_0^2 = \left|\boldsymbol{E}_0^2\right| e^{-2i\alpha},$$

where we denoted the argument of this complex number by -2α. We can then write

$$\boldsymbol{E}_0 = \boldsymbol{b}e^{-i\alpha},$$

where vector $\boldsymbol{b}$ may be complex, but $\boldsymbol{b}^2$ is real. Separating the real and imaginary parts, we can write

$$\boldsymbol{b} = \boldsymbol{b}_1 + i\boldsymbol{b}_2, \tag{10.38}$$

where $\boldsymbol{b}_1$ and $\boldsymbol{b}_2$ are real vectors. They are also orthogonal, since

$$\boldsymbol{b}^2 = (\boldsymbol{b}_1 + i\boldsymbol{b}_2)^2 = \boldsymbol{b}_1^2 + 2i\underbrace{\boldsymbol{b}_1 \cdot \boldsymbol{b}_2}_{=0} + \boldsymbol{b}_2^2$$

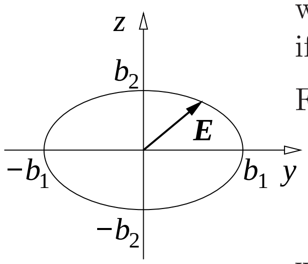
must be real. Assuming that the wave propagates in the x direction, let us choose the y axis in the direction of $\boldsymbol{b}_1$, with vector $\boldsymbol{b}_2$ parallel to the z axis. From

$$\boldsymbol{E} = \mathrm{Re}\,\left[\boldsymbol{E}_0 e^{i(\boldsymbol{k}\cdot\boldsymbol{r}-\omega t)}\right] = \mathrm{Re}\,\left[\boldsymbol{b}e^{i(\boldsymbol{k}\cdot\boldsymbol{r}-\omega t-\alpha)}\right]$$

we then have

$$E_y(t) = b_1 \cos(\omega t - \boldsymbol{k}\cdot\boldsymbol{r} + \alpha), \tag{10.39a}$$

$$E_z(t) = \pm b_2 \sin(\omega t - \boldsymbol{k}\cdot\boldsymbol{r} + \alpha), \tag{10.39b}$$



where in the last equation we use + if $\boldsymbol{b}_2$ is in the positive z direction, and − if it is in the negative z direction.

From equations (10.39a) and (10.39b) we see that

$$\frac{E_y^2}{b_1^2} + \frac{E_z^2}{b_2^2} = 1, \tag{10.40}$$

which means that the electric field vector describes an ellipse[67] in the y-z plane. When viewed from the positive x direction (i.e., with the wave travelling

---

[66]The average value of a quantity over one period is given by $\langle \dots \rangle = \frac{1}{T}\int_0^T \dots dt$, where $T = 2\pi/\omega$ is the period.

[67]The equation of the ellipse in Cartesian coordinates, $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$, where a and b are its semiaxes.

towards the observer), the electric field vector rotates anticlockwise if the sign in (10.39b) is $+$, and clockwise if it is $-$. Such wave is called *elliptically polarised*.

If $b_1 = b_2$, the ellipse (10.40) reduces to a circle. In this case the wave is *circularly polarised*. The ratio of y and z components of the complex amplitude $\boldsymbol{E}_0$ is

$$\frac{E_{0z}}{E_{0y}} = \pm i,$$

corresponding to the $\pm$ sign in (10.39b), with the top and bottom signs describing *right* and *left* circular polarisation, respectively[68].

If $b_1 = 0$ or $b_2 = 0$, the wave is *linearly polarised*.

10.4 Reflection and refraction at a dielectric boundary

Consider a monochromatic plane wave with wave vector $\boldsymbol{k}_1$ and frequency ω, incident on the interface between dielectric media 1 and 2 with permittivities ε_1 and ε_2 ($\mu_1 = \mu_2 = \mu_0$). In this case the solution that satisfies all the boundary conditions can be written as the sum of the incident and *reflected* waves in medium 1, and a transmitted (*refracted*) wave in medium 2.

Let the wave vector of the reflected wave be $\boldsymbol{k}_1'$ and that of the transmitted wave $\boldsymbol{k}_2$. The corresponding wave numbers are related to that of the incident wave by

$$k_1' = k_1 = \frac{\omega}{c} = \sqrt{\varepsilon_1 \mu_0}\, \omega, \tag{10.41}$$

$$k_2 = \sqrt{\varepsilon_2 \mu_0}\, \omega = \sqrt{\frac{\varepsilon_2}{\varepsilon_1}} k_1, \tag{10.42}$$

assuming that the frequencies of the three waves are the same (see below).

At the interface, the electromagnetic fields must satisfy the boundary conditions (9.14a)–(9.14d) which, in the absence of charges and currents, read

$$E_{1t} = E_{2t}, \tag{10.43a}$$

$$D_{1n} = D_{2n}, \tag{10.43b}$$

$$B_{1n} = B_{2n}, \tag{10.43c}$$

$$H_{1t} = H_{2t}, \tag{10.43d}$$

Directing the unit normal vector $\boldsymbol{n}$ from medium 2 into 1, and unit vector $\boldsymbol{t}$ parallel to the interface, we have from (10.43a):

$$\underbrace{\boldsymbol{E}_0 e^{i(\boldsymbol{k}_1\cdot\boldsymbol{r}-\omega t)}}_{\text{incident}} \cdot \boldsymbol{t} + \underbrace{\boldsymbol{E}_0' e^{i(\boldsymbol{k}_1'\cdot\boldsymbol{r}-\omega t)}}_{\text{reflected}} \cdot \boldsymbol{t} = \underbrace{\boldsymbol{E}_0'' e^{i(\boldsymbol{k}_2\cdot\boldsymbol{r}-\omega t)}}_{\text{refracted}} \cdot \boldsymbol{t}, \tag{10.44}$$

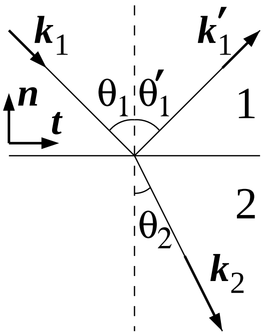
[68] The rotation of the electric field in these two cases corresponds to that of a right-hand or left-hand screw pointing along the x axis. This convention is adopted in theoretical physics and electrical engineering. It corresponds to right-circularly polarised photons carrying angular momentum with a positive projection onto their direction of motion, and left circularly polarised photons – negative. Note that the opposite convention is sometimes used, especially in optics.

where $\boldsymbol{E}_0$, $\boldsymbol{E}_0'$ and $\boldsymbol{E}_0''$ are the (complex) amplitudes of the incident, reflected and transmitted waves.

Since the boundary condition (10.44) must hold for all points $\boldsymbol{r}$ in the interface plane, we must have

$$\boldsymbol{k}_1 \cdot \boldsymbol{r} = \boldsymbol{k}_1' \cdot \boldsymbol{r} = \boldsymbol{k}_2 \cdot \boldsymbol{r},$$

which means that the components of the three wave vectors parallel to the plane are equal[69]. The three vectors thus lie in the same plane perpendicular to the interface, known as the *plane of incidence.*



Let the angles between $\boldsymbol{k}_1$, $\boldsymbol{k}_1'$ and $\boldsymbol{k}_2$ and the normal to the interface be θ_1, θ_1' and θ_2, respectively (see diagram). From the condition for the parallel components, we have

$$k_1 \sin\theta_1 = k_1' \sin\theta_1',$$

and since $k_1 = k_1'$, we have

$$\theta_1 = \theta_1',$$

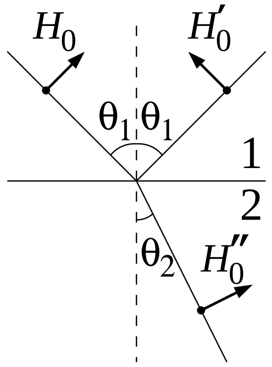
which shows that the angle of incidence equals the angle of reflection. Also,

$$k_1 \sin\theta_1 = k_2 \sin\theta_2,$$

and using (10.42), we find

$$\frac{\sin\theta_1}{\sin\theta_2} = \sqrt{\frac{\varepsilon_2}{\varepsilon_1}} = \frac{n_2}{n_1}, \qquad (10.45)$$

where n_1 and n_2 are the refractive indices [see (10.7)] of media 1 and 2, respectively. Equation (10.45) is known as *Snell's law.*



Consider the incident wave linearly polarised in the direction perpendicular to the plane of incidence. In this case the electric field amplitudes are parallel to the interface (into the page, on the diagram), with the magnetic field amplitudes H_0, H_0' and H_0'' as shown. From boundary condition (10.44), we have

$$E_0 + E_0' = E_0''. \qquad (10.46)$$

Applying the boundary condition (10.43d), we find[70]

$$H_0 \cos\theta_1 - H_0' \cos\theta_1 = H_0'' \cos\theta_2,$$

and using

$$H = \frac{1}{\mu_0} B = \frac{1}{\mu_0 c} E = \sqrt{\frac{\varepsilon}{\mu_0}} E,$$

we have

$$\sqrt{\varepsilon_1}(E_0 \cos\theta_1 - E_0' \cos\theta_1) = \sqrt{\varepsilon_2} E_0'' \cos\theta_2. \qquad (10.47)$$

[69] We set the frequencies of the incident, reflected and transmitted waves equal from the outset. This is necessary for (10.44) to be satisfied at all times t.

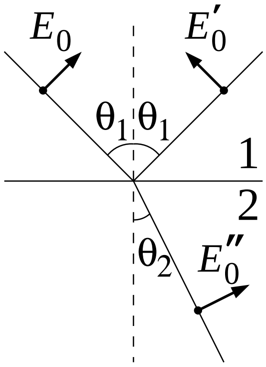
[70] Of the other boundary conditions, (10.43b) is satisfied trivially, as $D_n = 0$ for this polarisation, and (10.43c) simply reproduces Snell's law (10.45).

Using $\sqrt{\varepsilon_2/\varepsilon_1} = \sin\theta_1/\sin\theta_2$ and solving equations (10.46) and (10.47) simultaneously, we find the amplitude of the reflected wave

$$E_0' = \frac{\sin(\theta_2 - \theta_1)}{\sin(\theta_2 + \theta_1)} E_0, \tag{10.48}$$

and the amplitude of the transmitted wave

$$E_0'' = \frac{2\sin\theta_2\cos\theta_1}{\sin(\theta_2 + \theta_1)} E_0. \tag{10.49}$$



Let us now consider the incident wave linearly polarised in the plane of incidence. The corresponding electric field amplitudes are shown on the diagram, with the magnetic field amplitudes perpendicular to the plane of incidence (and to the page plane, towards us in the diagram). Boundary condition (10.43d) gives

$$H_0 + H_0' = H_0'',$$

or, equivalently,

$$\sqrt{\varepsilon_1}(E_0 + E_0') = \sqrt{\varepsilon_2}E_0''. \tag{10.50}$$

From boundary condition (10.43a), we have

$$E_0\cos\theta_1 - E_0'\cos\theta_1 = E_0''\cos\theta_2. \tag{10.51}$$

Solving (10.50) and (10.51) simultaneously, we find the amplitude of the reflected wave[71]

$$E_0' = \frac{\tan(\theta_1 - \theta_2)}{\tan(\theta_1 + \theta_2)} E_0, \tag{10.52}$$

and the amplitude of the transmitted wave

$$E_0'' = \frac{2\sin\theta_2\cos\theta_1}{\sin(\theta_1 + \theta_2)\cos(\theta_1 - \theta_2)} E_0. \tag{10.53}$$

Equations (10.48), (10.49) and (10.52), (10.53) are known as *Fresnel equations*. They point to a number of important phenomena.

For polarisation parallel to the incident plane, the reflected wave amplitude (10.52) vanishes if $\theta_1 + \theta_2 = \pi/2$. (since $\tan(\theta_1+\theta_2) \to \infty$). The corresponding angle of incidence is known as *Brewster's angle*. When unpolarised light (such as solar light, with equal amounts of parallel and perpendicular polarisation) is incident at such angle, the reflected light becomes totally polarised in the direction perpendicular to the plane of incidence[72]. Note that when a wave is incident at Brewster's angle, the wave vectors of the transmitted and reflected waves are at the right angle to each other. The particles in medium 2 are driven by the electromagnetic wave in the direction perpendicular to $\boldsymbol{k}_2$, i.e., parallel

[71] This is helped by the identity $\dfrac{\sin\theta_1\cos\theta_1 - \sin\theta_2\cos\theta_2}{\sin\theta_1\cos\theta_1 + \sin\theta_2\cos\theta_2} = \dfrac{\tan(\theta_1-\theta_2)}{\tan(\theta_1+\theta_2)}$.

[72] For other angles the reflected light contains both polarisations, but is still partially polarised in the perpendicular direction. By wearing Polaroid® glasses in which the lenses only let vertically polarised light through, reflections and glare from horizontal surfaces are strongly reduced.

to $\boldsymbol{k}'_1$ (for in-plane polarisation). As shown in Sec. 11.2, electric charges that oscillate in a line, do not radiate along this line. This provides a physical explanation as to why for parallel polarisation the amplitude of the reflected wave is zero at Brewster's angle.

Note also that for $n_2 < n_1$, equation (10.45) gives $\sin\theta_2 > \sin\theta_1$ and $\theta_2 > \theta_1$. For the angle of incidence such that

$$\sin\theta_1 = \frac{n_2}{n_1}, \tag{10.54}$$

we find $\theta_2 = \pi/2$. In this case the absolute values of the reflected wave amplitudes for either polarisation, (10.48) or (10.52), become equal to the incident wave amplitudes. This means that the energy flux in the reflected wave is equal to that of the incident wave, and no energy is transmitted into medium 2. The angle (10.54) is known as the angle of *total internal reflection*. For this and greater angles there is no transmitted wave and the interface between the two media acts as a perfect mirror[73].

Another effect described by Fresnel equations is that for $n_2 > n_1$ (i.e., $\theta_2 < \theta_1$), the reflected wave amplitude from equation (10.48) has the opposite sign to E_0. This means that the wave acquires an extra phase of π upon reflection from a more optically dense medium. [The same is true for the parallel polarisation; although equation (10.52) gives the same sign, the change of the phase is taken care through the explicit choice of the directions of $\boldsymbol{E}_0$ and $\boldsymbol{E}'_0$.] This effect is important when studying interference in thin films. In particular, when the thickness of a film decreases towards zero, the waves reflected off the front and back surfaces of the film cancel out, so that an infinitely thin film does not reflect at all.

Finally, we consider reflection at normal incidence ($\theta_1 \to 0$, or $\theta_1 \ll 1$). In this case $\sin\theta_1 \simeq \theta_1$ and $\sin\theta_2 \simeq \theta_2$, and equation (10.45) gives

$$\frac{\theta_1}{\theta_2} \simeq \frac{n_2}{n_1}.$$

The ratio of the energy flux in the reflected wave to that in the incident wave, known as the *reflection coefficient*, is found from either (10.48) or (10.52), as

$$R = \left|\frac{E'_0}{E_0}\right|^2 = \left(\frac{n_1 - n_2}{n_1 + n_2}\right)^2. \tag{10.55}$$

Taking the air-glass interface as an example, $n_1 = 1$, $n_2 = 1.5$, we find

$$R = \left(\frac{1 - 1.5}{1 + 1.5}\right)^2 = \frac{1}{25} = 0.04.$$

This means that a double-glazed windows reflects about 16% of the incident light, no matter how clean the glass in the window is!

[73]This phenomenon is behind the use of optical fibres in communications. They allow one to transmit light through very large distances with minimal losses. Also, *Porro prisms* employ total internal reflection. Using them in pairs allowed binoculars to be made more compact, by increasing the optical path length between the objective and eyepiece lenses without making the size of the instrument bigger.

11 Electromagnetic radiation

11.1 Inhomogeneous wave equation

The electromagnetic potentials in Lorenz gauge satisfy equations (9.24a) and (9.24b):

$$\nabla^2 \boldsymbol{A} - \frac{1}{c^2}\frac{\partial^2 \boldsymbol{A}}{\partial t^2} = -\mu \boldsymbol{j}, \tag{11.1a}$$

$$\nabla^2 \phi - \frac{1}{c^2}\frac{\partial^2 \phi}{\partial t^2} = -\frac{\rho}{\varepsilon}. \tag{11.1b}$$

where $c = 1/\sqrt{\varepsilon\mu}$, and we assume that the medium is homogeneous, i.e., $\varepsilon = \text{const}$ and $\mu = \text{const}$.

Mathematically, equations (11.1a) and (11.1b) are *inhomogeneous wave equations*. Their solutions can be written as a sum of any number of solutions of the homogeneous wave equation (representing electromagnetic waves, Ch. 10), and a particular solution of the inhomogeneous equation. In this section we show that the latter can be written in the form of the *retarded potentials*,

$$\boldsymbol{A}(\boldsymbol{r}, t) = \frac{\mu}{4\pi}\int_V \frac{\boldsymbol{j}(\boldsymbol{r}', t')}{|\boldsymbol{r} - \boldsymbol{r}'|} dV', \tag{11.2a}$$

$$\phi(\boldsymbol{r}, t) = \frac{1}{4\pi\varepsilon}\int_V \frac{\rho(\boldsymbol{r}', t')}{|\boldsymbol{r} - \boldsymbol{r}'|} dV', \tag{11.2b}$$

where

$$t' = t - \frac{|\boldsymbol{r} - \boldsymbol{r}'|}{c}. \tag{11.3}$$

The retarded potentials have a clear physical meaning. They describe how the current (or charge) density at point $\boldsymbol{r}'$ and time t' affects the potential at point $\boldsymbol{r}$ and time t. The difference between t and t' in (11.3) is the time it takes an electromagnetic disturbance to travel from $\boldsymbol{r}'$ to $\boldsymbol{r}$.

For steady currents and static charge distributions the potentials are time-independent. The solutions (11.2a) and (11.2b) are then identical to (6.29) used in magnetostatics, and the electrostatic potential (1.23) (without the surface charge term) that solves Poisson's equation (1.40).

<u>Proof.</u> To prove (11.2b), let us consider equation (11.1b) in which the density is due to a point charge $q(t)$ located at $\boldsymbol{r}'$,

$$\nabla^2 \phi - \frac{1}{c^2}\frac{\partial^2 \phi}{\partial t^2} = -\frac{q(t)}{\varepsilon}\delta(\boldsymbol{r} - \boldsymbol{r}'). \tag{11.4}$$

Introducing the relative position $\boldsymbol{R} = \boldsymbol{r} - \boldsymbol{r}'$, we see that the solution for a point charge at $\boldsymbol{R} = 0$ is spherically symmetric, i.e., $\phi = \phi(R, t)$, and only the radial part of the Laplacian gives a nonzero contribution. Hence we have

$$\frac{1}{R^2}\frac{\partial}{\partial R}\left(R^2 \frac{\partial \phi}{\partial R}\right) - \frac{1}{c^2}\frac{\partial^2 \phi}{\partial t^2} = -\frac{q(t)}{\varepsilon}\delta(\boldsymbol{R}), \tag{11.5}$$

where the right-hand side is zero everywhere except $\boldsymbol{R} = 0$.

The introduction of a new function χ,

$$\phi(R, t) = \frac{\chi(R, t)}{R},$$

transforms equation (11.5) into the one-dimensional wave equation,

$$\frac{\partial^2 \chi}{\partial R^2} - \frac{1}{c^2} \frac{\partial^2 \chi}{\partial t^2} = 0, \tag{11.6}$$

for $R \neq 0$. From Sec. 10.2 we know that its solution can be written as a sum of two arbitrary functions of $\xi = R - ct$ and $\eta = R + ct$ (10.20), or equivalently,

$$\chi(R, t) = f\left(t - \frac{R}{c}\right) + g\left(t + \frac{R}{c}\right). \tag{11.7}$$

Seeking a solution that describes a wave propagating *away* from the origin, we set $g = 0$ and have

$$\phi = \frac{f\left(t - \frac{R}{c}\right)}{R}. \tag{11.8}$$

To satisfy (11.5), this solution must behave as the potential of a point charge $q(t)$ for $R \to 0$ (where retardation can be neglected),

$$\phi \simeq \frac{q(t)}{4\pi\varepsilon R},$$

which gives $f(t) = q(t)/4\pi$. The solution of equation (11.4) thus is

$$\phi(\boldsymbol{r}, t) = \frac{q\left(t - \frac{|\boldsymbol{r} - \boldsymbol{r}'|}{c}\right)}{4\pi\varepsilon|\boldsymbol{r} - \boldsymbol{r}'|}. \tag{11.9}$$

The solution for the continuous charge density distribution $\rho(\boldsymbol{r}, t)$ on the right-hand side of (11.1b), is obtained by using the superposition principle, replacing

$$q(t') \longrightarrow \rho(\boldsymbol{r'}, t')dV',$$

in (11.9), and integrating over the volume, which gives (11.2b).

The equation for the vector potential (11.1a) is equivalent to three copies of equation (11.1b) (one for each of the x, y and z components of $\boldsymbol{A}$ and $\boldsymbol{j}$). Its solution (11.2a) is obtained by replacing $\rho(\boldsymbol{r'}, t')/\varepsilon$ in (11.2b) by $\mu\boldsymbol{j}(\boldsymbol{r'}, t')$.

Note that mathematically one can also set $f = 0$ in (11.7), which would lead to the so-called *advanced potentials* with $t' = t + |\boldsymbol{r} - \boldsymbol{r}'|/c$. The choice of either retarded or advanced potentials is determined by the boundary conditions. We usually consider problems in which charges are driven by some external forces (e.g., electromagnetic fields which satisfy the homogeneous wave equation). This motion produces additional fields that propagate away from the charges, which means that retarded potentials should be used.

11.2 Radiation by electric dipole

Consider a system of n point charges q_i with positions $\boldsymbol{r}_i(t)$, confined to a region of space near the origin. In this section we want to find the electromagnetic field produced by the moving charges at large distances $r \gg r_i$.

The electric dipole moment of this system is [cf. equation (1.44)],

$$\boldsymbol{p}(t) = \sum_{i=1}^{n} q_i \boldsymbol{r}_i(t), \tag{11.10}$$

and charge density and current density are

$$\rho(\boldsymbol{r}, t) = \sum_{i=1}^{n} q_i \delta(\boldsymbol{r} - \boldsymbol{r}_i(t)), \tag{11.11}$$

and

$$\boldsymbol{j}(\boldsymbol{r}, t) = \sum_{i=1}^{n} q_i \dot{\boldsymbol{r}}_i(t) \delta(\boldsymbol{r} - \boldsymbol{r}_i(t)), \tag{11.12}$$

where $\dot{\boldsymbol{r}}_i \equiv d\boldsymbol{r}_i/dt = \boldsymbol{v}_i$ is the velocity of ith charge [cf. equation (5.2)].

The vector potential produced by this charge distribution in vacuum ($\mu = \mu_0$) is obtained by substituting (11.12) into (11.2a),

$$\boldsymbol{A}(\boldsymbol{r}, t) = \frac{\mu_0}{4\pi} \int_V \frac{\sum_{i=1}^{n} q_i \dot{\boldsymbol{r}}_i(t') \delta(\boldsymbol{r}' - \boldsymbol{r}_i(t'))}{|\boldsymbol{r} - \boldsymbol{r}'|} dV', \tag{11.13}$$

where $t' = t - |\boldsymbol{r} - \boldsymbol{r}'|/c$. Assuming $r_i \ll r$, we can expand

$$t' = t - \frac{r}{c} + \frac{\boldsymbol{r} \cdot \boldsymbol{r}'}{rc} + \dots.$$

The third term on the right-hand hand side is $\Delta t \sim r_i/c$, which is the time it takes an electromagnetic wave to travel across the system of charges. During this time, the positions of the charges change by

$$|\Delta \boldsymbol{r}_i| = v_i \Delta t \sim \frac{v_i}{c} r_i.$$

If the charges move slowly compared to the speed of light c (i.e., $v_i \ll c$), we have $|\Delta \boldsymbol{r}_i| < r_i$. In this case the displacements of the charges over the time Δt can be neglected, and we can use the same time

$$t' = t - \frac{r}{c} \tag{11.14}$$

for all the charges in the system. Neglecting $\boldsymbol{r}'$ in comparison with $\boldsymbol{r}$ in the denominator of (11.13), we obtain

$$\begin{aligned} \boldsymbol{A}(\boldsymbol{r}, t) &= \frac{\mu_0}{4\pi r} \int_V \sum_{i=1}^{n} q_i \dot{\boldsymbol{r}}_i(t') \delta(\boldsymbol{r}' - \boldsymbol{r}_i(t')) dV', \\ &= \frac{\mu_0}{4\pi r} \sum_{i=1}^{n} q_i \dot{\boldsymbol{r}}_i(t'), \end{aligned}$$

with t' given by (11.14). Comparing with (11.10), we see that the vector potential of the system of charges is given by

$$\boldsymbol{A}(\boldsymbol{r}, t) = \frac{\mu_0}{4\pi r}\dot{\boldsymbol{p}}(t'). \tag{11.15}$$

At large distances r, the vector potential (11.15) behaves locally as that of a plane wave propagating in the radial direction $\boldsymbol{n} = \boldsymbol{r}/r$. We can then use (10.22) to find the magnetic field,

$$\boldsymbol{B} = \boldsymbol{n} \times \frac{d\boldsymbol{A}}{d\xi} = \boldsymbol{n} \times \frac{d\boldsymbol{A}}{dt}\frac{dt}{d\xi},$$

where $\xi = r - ct$ here, so that $d\xi/dt = -c$. This gives

$$\boldsymbol{B} = \frac{\mu_0}{4\pi cr}\,\ddot{\boldsymbol{p}}(t') \times \boldsymbol{n}. \tag{11.16}$$

According to Sec. 10.2, the electric field in the wave is perpendicular to $\boldsymbol{B}$ and $\boldsymbol{n}$ [see equation (10.23)]. Hence we find it as

$$\boldsymbol{E} = c\boldsymbol{B} \times \boldsymbol{n} = \frac{\mu_0}{4\pi r}[\,\ddot{\boldsymbol{p}}(t') \times \boldsymbol{n}] \times \boldsymbol{n}. \tag{11.17}$$

This shows that the radiation by the electric dipole is polarised in the plane of vectors $\ddot{\boldsymbol{p}}$ and $\boldsymbol{n}$, with vector $\boldsymbol{B}$ (11.16) being perpendicular to this plane.

The electromagnetic energy flux is given by the Poynting vector (9.13),

$$\boldsymbol{S} = \frac{1}{\mu_0}\boldsymbol{E} \times \boldsymbol{B} = \frac{\mu_0}{16\pi^2 cr^2}\{[\,\ddot{\boldsymbol{p}}(t') \times \boldsymbol{n}] \times \boldsymbol{n}\} \times [\,\ddot{\boldsymbol{p}}(t') \times \boldsymbol{n}],$$

which gives

$$\boldsymbol{S} = \frac{1}{16\pi^2\varepsilon_0 c^3 r^2}|\ddot{\boldsymbol{p}}(t') \times \boldsymbol{n}|^2\boldsymbol{n}. \tag{11.18}$$

If $\ddot{\boldsymbol{p}}$ is along the z axis, the magnitude of the Poynting vector is given by

$$S = \frac{|\ddot{\boldsymbol{p}}|^2 \sin^2\theta}{16\pi^2\varepsilon_0 c^3 r^2}, \tag{11.19}$$

where θ is the polar angle. Equation (11.19) shows that the radiation is a maximum in the plane perpendicular to $\ddot{\boldsymbol{p}}$ ($\theta = \pi/2$), and that there is no radiation in the direction parallel to $\ddot{\boldsymbol{p}}$ ($\theta = 0$ or π).

The total energy radiated in unit time (i.e., the radiation power) is found by integrating the Poynting vector over the surface of a sphere of radius r:

$$P = \frac{|\ddot{\boldsymbol{p}}|^2}{16\pi^2\varepsilon_0 c^3 r^2}\int_0^{2\pi}\int_0^{\pi}\sin^2\theta\, r^2 \sin\theta d\theta d\psi,$$

which gives[74]

$$P = \frac{|\ddot{\boldsymbol{p}}|^2}{6\pi\varepsilon_0 c^3}. \tag{11.20}$$

[74] $\int_0^\pi \sin^3\theta\, d\theta = \int_0^\pi (1-\cos^2\theta)\, d(-\cos\theta) = \int_{-1}^1 (1-x^2)dx = 4/3$.

If the dipole executes harmonic oscillations with amplitude $\boldsymbol{p}_0$ and frequency ω, $\boldsymbol{p} = \boldsymbol{p}_0 \cos \omega t$, the radiation power (11.20) averaged over one period is[75]

$$\langle P \rangle = \frac{p_0^2 \omega^4}{12\pi\varepsilon_0 c^3}. \tag{11.21}$$

If we consider the motion of a single electron with charge e (Sec. 1.1), then $\boldsymbol{p} = e\boldsymbol{r}$. Hence, the power radiated by the electron (11.20) is determined by its acceleration,[76]

$$P = \frac{e^2 |\ddot{\boldsymbol{r}}|^2}{6\pi\varepsilon_0 c^3}. \tag{11.22}$$

The fact that the radiation is determined by the electron's acceleration and not velocity, could be expected. Indeed, the velocity of a particle can always be made zero by considering it in a co-moving inertial frame of reference.

Using equation (11.22), one can estimate how long it will take a classical electron to spiral into the nucleus, as it keeps losing energy to radiation due to a nonzero centripetal acceleration it experiences in its orbital motion around the nucleus.

The fourth power of the frequency in equation (11.21) explains why the sky is blue. The electric charges in the atmosphere are driven by the electromagnetic waves (i.e., light) emitted by the Sun. They radiate much more strongly in the blue than in the red, because the frequency of the blue light is almost a factor of 2 greater than that of the red.

The equations derived in this section also explain why the light that comes from the sky is partially polarised. The light emitted by the Sun is unpolarised. Since the electromagnetic wave is transversal, the charges in the sky oscillate in all directions perpendicular to the direction from the Sun. However, the emission in the direction of the observer is strongest for the charges that oscillate perpendicular to the line of sight [see equation (11.19)]. As a result, the light scattered by the sky is partially polarised in the direction perpendicular to the plane through the Sun, the point in the sky and the eye of the observer. Wearing Polaroid® sunglasses allows anyone an easy check of this physics. Just put the glasses on and tilt your head sideways, while looking at the blue sky, and observe the change in its brightness.

---

[75] $\langle \cos^2 \omega t \rangle = \langle \sin^2 \omega t \rangle = 1/2$.

[76] This expression is sometimes called the Larmor formula, as it was first obtained by Joseph Larmor in his paper *On the theory of the magnetic influence on spectra; and on the radiation from moving ions*, Phil. Mag., Ser. 5, **44**, 503 (1897).