MAS207: ELECTROMAGNETISM

HENK VAN ELST
Astronomy Unit
School of Mathematical Sciences
Queen Mary, University of London
Mile End Road
London E1 4NS, United Kingdom
E–Mail: H.van.Elst@qmul.ac.uk

May 24, 2004

- These lecture notes are only intended for the use of students studying the course MAS207: Electromagnetism. Any distribution for any other purpose is not permitted.
- The selection of topics covered in these lecture notes is influenced by the following textbooks, lecture notes and websites:
  - *Classical Electrodynamics*, John D Jackson, Wiley & Sons, New York, ^{3}1999.
  - *Electrodynamics (Lectures on Theoretical Physics, Vol. III)*, Arnold Sommerfeld, Academic Press, New York, 1952.
  - *The Feynman Lectures on Physics Vol. II*, Richard P Feynman, Robert B Leighton, Matthew Sands, Addison-Wesley, Reading (MA), 1964.
  - *Foundations of Electromagnetic Theory*, John R Reitz, Frederick J Milford, Robert W Christy, Addison-Wesley, Reading (MA), 1979.
  - *A First Course in General Relativity*, Bernard F Schutz, Cambridge University Press, Cambridge, 1985.
  - *Electromagnetism and Special Relativity*, Reza Tavakol, Queen Mary, University of London, Lecture notes, 1984. Unavailable.
  - *Theorie B: Elektrodynamik*, Gottfried Falk, University of Karlsruhe, Notes taken in lectures, Summer Semester 1988 [in German]. Unavailable.
  - *Theorie B: Elektrodynamik*, Julius Wess, University of Karlsruhe, Lecture notes by Thomas Strohm, Summer Semester 1989 [in German]. Unavailable.
  - *Physik II: Elektrodynamik*, Friedrich Herrmann, University of Karlsruhe, Script, 1997 [in German]. Available online at URL: `www-tfp.physik.uni-karlsruhe.de/~didaktik/`.
  - Wolfram Research's *World of Science*, serviced by Eric W Weisstein, URL: `scienceworld.wolfram.com`.
  - *The MacTutor History of Mathematics archive*, URL: `www-history.mcs.st-andrews.ac.uk`.
  - *MAS209: Fluid Dynamics*, David Burgess and Henk van Elst, Queen Mary, University of London, Lecture notes, 1997–2001. Available from authors.
  - *Relativity and Cosmology*, Bernard Carr, Queen Mary, University of London, Lecture notes, 1989. Unavailable.

Contents

CONTENTS

Chapter 1
Introductory remarks

Many phenomena and applications we encounter in our everyday lives are grounded on **electromagnetic interactions**. Examples that spring to one's mind are: the switching on of an electric light bulb; the transmission of signals that let us communicate via radio, television or mobile phones; the inner workings of our PCs; or the generation of lightnings in a thunder storm. Electromagnetism (luckily) reveals itself to us as an inherently **linear** phenomenon of Nature, which means that the **linear superposition** of electromagnetic fields is possible. This is exploited to our advantage in many of our electromagnetic applications; e.g., many phone calls simultaneously going down the same wire, or different TV stations broadcasting at the same time without interference. Surprisingly, it seems that we have long stopped wondering about these issues which appear so familiar to us (or perhaps not?).

The theoretical understanding of electromagnetic interactions and the mathematical formulation of **electromagnetism** (or electrodynamics) as a **classical field theory** of physics have been completed during the nineteenth century, with major contributions by the French physicist Charles Augustin de Coulomb (1736–1806), the French mathematician and physicist André Marie Ampère (1775–1836), the English physicist Michael Faraday (1791–1867), the German physicist Heinrich Hertz (1857–1894), and, last but not least, the Scottish mathematician and physicist James Clerk Maxwell (1831–1879). The aim of this course is to provide an introduction to the mathematical formulation of **electromagnetism** and to discuss many examples — most of them attributed to **electrostatics**, **magnetostatics**, or **vacuum field configurations** — that arise as simple special subcases and can be solved analytically. Contrary to most text books, where **Maxwell's field equations**, a set of coupled *linear* first-order partial differential equations, are derived one after another following their historical development (inductive approach), we will here present them right at the beginning and then turn to discuss various subcases, as mentioned above (deductive approach). The hope is that this will help to demystify them (and indeed any other mathematical model used in theoretical physics) and become familiar with them at an earlier stage in the course than usual. In working with **Maxwell's field equations** we make use of many tools from **applied mathematics** and **mathematical physics**.

1.1 Field variables and Maxwell's field equations

Electromagnetic fields can be sourced by **electric charges** and **electric cur-**

rents, but, as we will see latter, they can also exist in "**vacuo**", where *no* such sources are present. Electric charge can take *both* positive and negative values. The SI unit for electric charge is the Coulomb ($1\,\mathrm{C}$). The charge carried by a single **electron** is $q_e = -1.60217733(49) \times 10^{-19}\,\mathrm{C}$. Its magnitude $|\,q_e\,|$ constitutes the *basic unit* of electric charge; in Nature electric charge is indeed *quantised*. Nevertheless, at a sufficiently macroscopic scale (compared to atomic scales) electric charge can be treated as continuously distributed due to the smallness of $|\,q_e\,|$.

The **macroscopic field variables** we use to describe the dynamics of electromagnetic fields and their interactions with continuous material media as functions of a **time coordinate** t and **position** $\boldsymbol{r}$ in Euclidian (flat) space, $\mathbb{R}^3$, are:

(i) $\boldsymbol{E} = \boldsymbol{E}(t, \boldsymbol{r})$ — the **electric field strength**, SI unit: $1\,\frac{\mathrm{kg\,m}}{\mathrm{s^2\,C}}$, also referred to as $1\,\mathrm{V(olt)/m}$,

(ii) $\boldsymbol{B} = \boldsymbol{B}(t, \boldsymbol{r})$ — the **magnetic field strength**, SI unit: $1\,\frac{\mathrm{kg}}{\mathrm{s\,C}}$, also referred to as $1\,\mathrm{T(esla)}$,

(iii) $\boldsymbol{D} = \boldsymbol{D}(t, \boldsymbol{r})$ — the **electric excitation**, SI unit: $1\,\frac{\mathrm{C}}{\mathrm{m^2}}$,

(iv) $\boldsymbol{H} = \boldsymbol{H}(t, \boldsymbol{r})$ — the **magnetic excitation**, SI unit: $1\,\frac{\mathrm{C}}{\mathrm{m\,s}}$,

(v) $\rho = \rho(t, \boldsymbol{r})$ — the **electric charge density**, SI unit: $1\,\frac{\mathrm{C}}{\mathrm{m^3}}$, and

(vi) $\boldsymbol{J} = \boldsymbol{J}(t, \boldsymbol{r})$ — the **electric current density**, SI unit: $1\,\frac{\mathrm{C}}{\mathrm{m^2 s}}$.

It should be emphasised that idealisations such as field strengths at a point must be viewed as a mathematical construct that permits a description of electromagnetic phenomena at the macroscopic level. They may fail to have a meaning at the microscopic level.

Maxwell's field equations for macroscopic field configurations are given as two sets of **partial differential equations** which are *linear* in the macroscopic field variables:

(a) the **time evolution equations** for $\boldsymbol{D}$ and $\boldsymbol{B}$,

$$\frac{\partial \boldsymbol{D}}{\partial t} - \boldsymbol{\nabla} \times \boldsymbol{H} = -\boldsymbol{J} \tag{1.1}$$

$$\frac{\partial \boldsymbol{B}}{\partial t} + \boldsymbol{\nabla} \times \boldsymbol{E} = \boldsymbol{0}\,, \tag{1.2}$$

which could be referred to as **Ampère–Maxwell law** and **Faraday's law**, respectively, and

(b) the **initial value constraint equations** for $\boldsymbol{D}$ and $\boldsymbol{B}$,

$$0 = \mathcal{C}_{\boldsymbol{D}} := \boldsymbol{\nabla} \cdot \boldsymbol{D} - \rho \tag{1.3}$$
$$0 = \mathcal{C}_{\boldsymbol{B}} := \boldsymbol{\nabla} \cdot \boldsymbol{B} \,, \tag{1.4}$$

which could be referred to as **Gauß' law** (named after the German mathematician and astronomer Carl Friedrich Gauß, 1777–1855) and **zero magnetic charges law**, respectively.

It can be shown that the constraint equations are preserved in time, i.e.,

$$\frac{\partial \mathcal{C}_{\boldsymbol{D}}}{\partial t} = 0 \tag{1.5}$$
$$\frac{\partial \mathcal{C}_{\boldsymbol{B}}}{\partial t} = 0 \,, \tag{1.6}$$

provided the condition (the **continuity equation**)

$$0 = \frac{\partial \rho}{\partial t} + \boldsymbol{\nabla} \cdot \boldsymbol{J} \tag{1.7}$$

holds. It is an empirical result that for all electric charge and current distributions occurring in Nature this is indeed the case. This result is better known as the concept of **conservation of electric charge**.

To understand why, conventionally, one maintains a *distinction* between the macroscopic field variables $\boldsymbol{E}$ and $\boldsymbol{D}$ on the one hand and $\boldsymbol{B}$ and $\boldsymbol{H}$ on the other, the following crude guideline may be helpful. One interprets $\boldsymbol{D}$ and $\boldsymbol{H}$ to be sourced by, respectively, unbounded macroscopic charges and currents in continuous material media only (represented through ρ and $\boldsymbol{J}$), while $\boldsymbol{E}$ and $\boldsymbol{B}$ are interpreted to be sourced by unbounded macroscopic charges and currents as well as bounded microscopic charges and currents (the total effect of the latter often being related to the field variables $\boldsymbol{P}$ and $\boldsymbol{M}$ which we will introduce below).

The fundamental electric and magnetic fields are $\boldsymbol{E}$ and $\boldsymbol{B}$, as they (a) determine the mechanical force $\boldsymbol{F}$ on a point charge q moving with velocity $\boldsymbol{v}$ in a given electromagnetic field according to **Lorentz's force law**

$$\boldsymbol{F} = q\,(\boldsymbol{E} + \boldsymbol{v} \times \boldsymbol{B}) \,, \tag{1.8}$$

(b) can be derived from **electromagnetic potentials**, as we will discuss in chapter 5, and (c) because of the way they behave under **Lorentz transformations** between different inertial reference frames, as we will discuss in chapter 6.

For the time evolution equations (1.1) and (1.2) amongst **Maxwell's field equations** to become a *closed* system of partial differential equations, they

need to be supplemented by **constitutive relations** of the form

$$\boldsymbol{D} = \boldsymbol{D}[\boldsymbol{E}, \boldsymbol{B}] \tag{1.9}$$
$$\boldsymbol{H} = \boldsymbol{H}[\boldsymbol{E}, \boldsymbol{B}] \,, \tag{1.10}$$

and a **generalised Ohm's law** (named after the German physicist Georg Simon Ohm, 1789–1854)

$$\boldsymbol{J} = \boldsymbol{J}[\boldsymbol{E}, \boldsymbol{B}] \,. \tag{1.11}$$

The determination of the precise form of these relations for a given macroscopic material medium is a highly complex matter and typically falls into the realm of **condensed matter physics** and **solid state physics**. To a first approximation, which is convenient for most practical purposes, one can assume continuous media to have *linear* electric and magnetic properties and negligible dispersion and losses so that linear constitutive relations of the form

$$\boldsymbol{D} = \epsilon_0 \, \boldsymbol{E} + \boldsymbol{P} \tag{1.12}$$
$$\boldsymbol{H} = \frac{1}{\mu_0} \, \boldsymbol{B} - \boldsymbol{M} \tag{1.13}$$

can be used. The new field variables thus introduced are $\boldsymbol{P} = \boldsymbol{P}(t, \boldsymbol{r})$ and $\boldsymbol{M} = \boldsymbol{M}(t, \boldsymbol{r})$ which are called, respectively, the **electric polarisation** and the **magnetisation** of a given macroscopic material medium. The constants $\epsilon_0 = (1/4\pi c^2) \times 10^7 \, \frac{\mathrm{C}^2}{\mathrm{kg\,m}}$, the **permittivity of vacuum**, and $\mu_0 = 4\pi \times 10^{-7} \, \frac{\mathrm{kg\,m}}{\mathrm{C}^2}$, the **permeability of vacuum**, are related to the **speed of light in vacuum**, $c = 299792458 \, \frac{\mathrm{m}}{\mathrm{s}}$, through $\epsilon_0 \mu_0 = c^{-2}$. **Ohm's law** (1826) takes the linear form

$$\boldsymbol{J} = \sigma \, \boldsymbol{E} \,, \tag{1.14}$$

with the positive constant σ known as the **electric conductivity**.

Assuming a known electric charge density ρ and a known electric current density $\boldsymbol{J}$, and using the particular constitutive relations (1.12) and (1.13), we can rewrite **Maxwell's field equations** (1.1)–(1.4) to yield (using the special relativistically more appropriate time coordinate ct)

$$\frac{\partial \boldsymbol{E}/c}{\partial ct} - \boldsymbol{\nabla} \times \boldsymbol{B} = -\mu_0 \left[\boldsymbol{J} + \frac{\partial (c\boldsymbol{P})}{\partial ct} + \boldsymbol{\nabla} \times \boldsymbol{M} \right] \tag{1.15}$$
$$\frac{\partial \boldsymbol{B}}{\partial ct} + \boldsymbol{\nabla} \times \boldsymbol{E}/c = \boldsymbol{0} \tag{1.16}$$
$$0 = \mathcal{C}_{\boldsymbol{E}} := \boldsymbol{\nabla} \cdot \boldsymbol{E}/c - \frac{1}{\epsilon_0 c^2} \left[c\rho - \boldsymbol{\nabla} \cdot (c\boldsymbol{P}) \right] \tag{1.17}$$
$$0 = \mathcal{C}_{\boldsymbol{B}} = \boldsymbol{\nabla} \cdot \boldsymbol{B} \,. \tag{1.18}$$

This is the form we will use throughout this course. *The* **theory of partial differential equations** *states that for given (known) sources $c\rho(ct, \boldsymbol{r})$, $\boldsymbol{J}(ct, \boldsymbol{r})$, $c\boldsymbol{P}(ct, \boldsymbol{r})$ and $\boldsymbol{M}(ct, \boldsymbol{r})$, there exist unique solutions $\boldsymbol{E}(ct, \boldsymbol{r})/c$ and $\boldsymbol{B}(ct, \boldsymbol{r})$*

to (1.15) *and* (1.16) *that depend continuously on given* **initial values**

$$
\begin{aligned}
{}_0\boldsymbol{E}/c &:= \boldsymbol{E}(ct = ct_0, \boldsymbol{r})/c && (1.19)\\
{}_0\boldsymbol{B} &:= \boldsymbol{B}(ct = ct_0, \boldsymbol{r})\ , && (1.20)
\end{aligned}
$$

which satisfy (1.17) *and* (1.18) *for given* **boundary conditions**. From a mathematical point of view one says that the **initial value problem** for **Maxwell's field equations** is **well-posed**.

1.2 Special cases

Special cases of (1.15)–(1.18) we are particularly interested in in this course are:

(a) **Electrostatics**: $\mathbf{0} = \boldsymbol{B} = \boldsymbol{M} = \boldsymbol{J} = \partial(c\boldsymbol{P})/\partial ct$, so $\mathbf{0} = \partial\boldsymbol{E}/c/\partial ct$, and the equations to be solved for given boundary conditions become

$$
\begin{aligned}
\mathbf{0} &= \boldsymbol{\nabla} \times \boldsymbol{E}/c && (1.21)\\
\boldsymbol{\nabla} \cdot \boldsymbol{E}/c &= \frac{1}{\epsilon_0 c^2}\left(c\rho - \boldsymbol{\nabla} \cdot (c\boldsymbol{P})\right)\ . && (1.22)
\end{aligned}
$$

(b) **Magnetostatics**: $0 = \rho$ and $\mathbf{0} = \boldsymbol{E} = \boldsymbol{P}$, so $\mathbf{0} = \partial\boldsymbol{B}/\partial ct$, and the equations to be solved for given boundary conditions become

$$
\begin{aligned}
\boldsymbol{\nabla} \times \boldsymbol{B} &= \mu_0\left(\boldsymbol{J} + \boldsymbol{\nabla} \times \boldsymbol{M}\right) && (1.23)\\
0 &= \boldsymbol{\nabla} \cdot \boldsymbol{B} && (1.24)\\
0 &= \boldsymbol{\nabla} \cdot \boldsymbol{J}\ . && (1.25)
\end{aligned}
$$

It is thus in the case of *static* field configurations that **Maxwell's field equations** determining $\boldsymbol{E}/c$ and $\boldsymbol{B}$ from known sources *decouple*, and electricity and magnetism are perceived as *distinct* physical phenomena.

(c) **Vacuum field configurations**: $0 = \rho$ and $\mathbf{0} = \boldsymbol{P} = \boldsymbol{M} = \boldsymbol{J}$, and so the equations to be solved for given initial values $\{\ {}_0\boldsymbol{E}/c, {}_0\boldsymbol{B}\ \}$ and boundary conditions become

$$
\begin{aligned}
\frac{\partial \boldsymbol{E}/c}{\partial ct} - \boldsymbol{\nabla} \times \boldsymbol{B} &= \mathbf{0} && (1.26)\\
\frac{\partial \boldsymbol{B}}{\partial ct} + \boldsymbol{\nabla} \times \boldsymbol{E}/c &= \mathbf{0} && (1.27)\\
0 &= \boldsymbol{\nabla} \cdot \boldsymbol{E}/c && (1.28)\\
0 &= \boldsymbol{\nabla} \cdot \boldsymbol{B}\ . && (1.29)
\end{aligned}
$$

In a reference frame with Cartesian coordinates where the vector analytical identity $\boldsymbol{\nabla} \times (\boldsymbol{\nabla} \times \boldsymbol{A}) \equiv \boldsymbol{\nabla}(\boldsymbol{\nabla} \cdot \boldsymbol{A}) - (\boldsymbol{\nabla} \cdot \boldsymbol{\nabla})\boldsymbol{A}$ holds [see (2.31) below], these can be combined to yield homogeneous **wave equations** for the Cartesian components of both $\boldsymbol{E}/c$ and $\boldsymbol{B}$ given by

$$
\begin{aligned}
-\frac{\partial^2 \boldsymbol{E}/c}{\partial (ct)^2} + (\boldsymbol{\nabla} \cdot \boldsymbol{\nabla})\boldsymbol{E}/c &= \mathbf{0} && (1.30)\\
-\frac{\partial^2 \boldsymbol{B}}{\partial (ct)^2} + (\boldsymbol{\nabla} \cdot \boldsymbol{\nabla})\boldsymbol{B} &= \mathbf{0}\ . && (1.31)
\end{aligned}
$$

1.3 Integral theorems and balance equations

In this section we briefly remind ourselves of some useful mathematical tools from MAS204 Calculus III, and we introduce the concept of a **balance equation**.

1.3.1 Gauß' and Stokes' integral theorems

(A) Consider in Euclidian space $\mathbb{R}^3$ a simply-connected **volume** G that is bounded by a **closed surface** ∂G, and a differentiable **vector field** $\boldsymbol{A} = \boldsymbol{A}(t, \boldsymbol{r})$ defined everywhere throughout the region that contains G. We assume that ∂G is fixed in time. We have[1]

Gauß' integral theorem:

$$\int\int\int_G \boldsymbol{\nabla} \cdot \boldsymbol{A} \, \mathrm{d}V = \int\int_{\partial G} \boldsymbol{A} \cdot \boldsymbol{n} \, \mathrm{d}A \,, \tag{1.32}$$

where $\boldsymbol{n}$ is the outward-pointing **unit normal** to ∂G. The theorem dates back to the early nineteenth century.

(B) Consider in Euclidian space $\mathbb{R}^3$ a simply-connected **surface** S that is bounded by an oriented **closed curve** ∂S, and a differentiable **vector field** $\boldsymbol{A} = \boldsymbol{A}(t, \boldsymbol{r})$ defined everywhere throughout the region that contains S. We assume that ∂S is fixed in time. We have the

Stokes' integral theorem:

$$\int\int_S (\boldsymbol{\nabla} \times \boldsymbol{A}) \cdot \boldsymbol{n} \, \mathrm{d}A = \oint_{\partial S} \boldsymbol{A} \cdot \mathrm{d}\boldsymbol{s} \,, \tag{1.33}$$

where $\boldsymbol{n}$ is the **unit normal** to S, directed in relation to the orientation on ∂S according to a right-hand convention. This theorem is named after the Irish mathematician and physicist George Gabriel Stokes (1819–1903), who established a proof during the mid-nineteenth century.

1.3.2 Balance equations

One of the unifying concepts in theoretical physics is the observation that dynamical interactions between two coupled physical systems are best described in terms of the exchange of physical quantities that flow from one system to the other, or vice versa. Let us discuss this issue on a more abstract level.

Let X represent a **physical quantity** whose magnitude scales with the **volume** (i.e., size) of a physical system. This is to say that the value of X doubles

[1] The surface integral sign on the right-hand side should really be a two-dimensional analog of the closed-loop line integral sign $\oint$, to emphasise that the integration is over a *closed* surface. However, it appears that LaTeX2ε does not hold such a symbol available (or at least we have not discovered it yet). Any help would be appreciated.

when the volume of the system doubles. A physical quantity with this property is referred to as **extensive**.[2] It distinguishes itself through the feature that it can form **densities** with respect to volumes and **current densities** with respect to surfaces. Note that extensive quantities can be **scalar-valued** or **vector-valued**. Examples of scalar-valued extensive quantities are **mass**, **electric charge**, **particle number**, **energy** or **entropy**, while examples of vector-valued physical quantities are **linear momentum** or **angular momentum**.

Let us consider in Euclidian space $\mathbb{R}^3$ a **volume** G, which is bounded by a **closed surface** ∂G. By assumption, this bounding surface shall be fixed in time. Then, if X is an extensive physical quantity, the

$$\left(\begin{array}{l}\textbf{Rate of change}\\ \textbf{in time} \text{ of the}\\ \text{amount of } X\\ \text{inside of } G\end{array}\right) = \left(\begin{array}{l}\textbf{Current } I_X\\ \text{of } X \text{ into } G\end{array}\right) + \left(\begin{array}{l}\textbf{Generation rate } \Sigma_X\\ \text{of } X \text{ inside of } G\end{array}\right) ,$$

stating that for X the **balance equation**

$$\frac{\mathrm{d}X}{\mathrm{d}t} = I_X + \Sigma_X \tag{1.34}$$

holds. Now, if for a *scalar-valued* X we introduce as differentiable functions of time and spatial position the scalar-valued X–**density** ρ_X, the vector-valued X–**current density** $\boldsymbol{j}_X$, and the scalar-valued X–**generation rate density** $\dot{\rho}_{X,\mathrm{gen}}$, then the balance equation (1.34) for X can be rewritten in the **integral form**

$$\frac{\mathrm{d}}{\mathrm{d}t}\int\int\int_G \rho_X \,\mathrm{d}V = -\int\int_{\partial G} \boldsymbol{j}_X \cdot \boldsymbol{n} \,\mathrm{d}A + \int\int\int_G \dot{\rho}_{X,\mathrm{gen}} \,\mathrm{d}V \ , \tag{1.35}$$

with $\boldsymbol{n}$ denoting the outward-pointing **unit normal** to the closed surface ∂G. The *minus sign* of the flux integral on the right-hand side arises because we record the current I_X that flows *into* the volume G. If we now employ Gauß' integral theorem (1.32) to this flux integral, and then bring all terms in the equation to one side, we get the balance equation for X in the form

$$0 = \int\int\int_G \left[\frac{\partial \rho_X}{\partial t} + \boldsymbol{\nabla}\cdot\boldsymbol{j}_X - \dot{\rho}_{X,\mathrm{gen}}\right] \mathrm{d}V \ .$$

Note that, as we assumed G to be fixed in time, we are allowed to pull the total time derivative through the volume integral sign; consequently under the integral sign this time derivative becomes a partial time derivative.

Now since our balance equation must hold for any arbitrary volume G of $\mathbb{R}^3$ that is fixed in time, it is the integrand in the expression we just derived that

[2] Physical quantities, on the other hand, that do *not* have this property are referred to as intensive. Examples of intensive quantities are mass density, pressure or temperature.

must vanish identically. We thus find the **differential form** of the balance equation for X to be given by

$$\frac{\partial \rho_X}{\partial t} + \boldsymbol{\nabla} \cdot \boldsymbol{j}_X = \dot{\rho}_{X,\mathrm{gen}} \; . \tag{1.36}$$

It should be pointed out that when X is a *vector-valued* extensive physical quantity, then an analogous balance equation can be formulated in terms of a vector-valued X–**density**, a tensor-valued X–**current density**, and a vector-valued X–**generation rate density**.

Remark: If for a specific physical quantity X empirical results show that the X–generation rate density $\dot{\rho}_{X,\mathrm{gen}}$ is *identically zero* (such as for mass, electric charge, energy, linear momentum and angular momentum), then (1.35) and (1.36) are referred to as **continuity equations** or **conservation equations**.

Physical dimensions:

$$[\,\rho_X\,] = \frac{[X]}{[\mathrm{length}]^3} \qquad [\,\boldsymbol{j}_X\,] = \frac{[X]}{[\mathrm{length}]^2[\mathrm{time}]} \qquad [\,\dot{\rho}_{X,\mathrm{gen}}\,] = \frac{[X]}{[\mathrm{length}]^3[\mathrm{time}]} \; .$$

1.3.3 Poynting's theorem

To see how this latest result of our's, (1.36), applies in the arena of **electromagnetism**, let us give an explicit example. Consider **Maxwell's field equations** in the form (1.15)–(1.18). If we scalar-multiply (1.15) by $\boldsymbol{E}/c$ and (1.16) by $\boldsymbol{B}$ and add both together, we obtain

$$\begin{aligned} &\frac{\partial}{\partial ct}\left(\tfrac{1}{2}\boldsymbol{E}/c \cdot \boldsymbol{E}/c\right) - \boldsymbol{E}/c \cdot (\boldsymbol{\nabla} \times \boldsymbol{B}) + \frac{\partial}{\partial ct}\left(\tfrac{1}{2}\boldsymbol{B} \cdot \boldsymbol{B}\right) + \boldsymbol{B} \cdot (\boldsymbol{\nabla} \times \boldsymbol{E}/c) \\ &= -\,\mu_0\, \boldsymbol{E}/c \cdot \left[\, \boldsymbol{J} + \frac{\partial (c\boldsymbol{P})}{\partial ct} + \boldsymbol{\nabla} \times \boldsymbol{M} \,\right] \; . \end{aligned} \tag{1.37}$$

Collecting together the time-derivative terms on the left-hand side and employing the vector analytical identity $\boldsymbol{\nabla} \cdot (\boldsymbol{A} \times \boldsymbol{B}) \equiv \boldsymbol{B} \cdot (\boldsymbol{\nabla} \times \boldsymbol{A}) - \boldsymbol{A} \cdot (\boldsymbol{\nabla} \times \boldsymbol{B})$ [cf. (2.28) below], this can be rewritten as

$$\begin{aligned} &\frac{\partial}{\partial ct}\left(\tfrac{1}{2}\boldsymbol{E}/c \cdot \boldsymbol{E}/c + \tfrac{1}{2}\boldsymbol{B} \cdot \boldsymbol{B}\right) + \boldsymbol{\nabla} \cdot (\boldsymbol{E}/c \times \boldsymbol{B}) \\ &= -\,\mu_0\, \boldsymbol{E}/c \cdot \left[\, \boldsymbol{J} + \frac{\partial (c\boldsymbol{P})}{\partial ct} + \boldsymbol{\nabla} \times \boldsymbol{M} \,\right] \; . \end{aligned} \tag{1.38}$$

Finally, multiplying through by c/μ_0, and defining the scalar-valued and vector-valued quantities (the deviating notation is used for historical reasons)

$$u \;:=\; \frac{1}{2\mu_0}\left(\boldsymbol{E}/c \cdot \boldsymbol{E}/c + \boldsymbol{B} \cdot \boldsymbol{B}\right) \;=\; \rho_{\mathrm{energy}}(\mathrm{EMF}) \tag{1.39}$$

$$\boldsymbol{S} \;:=\; \frac{c}{\mu_0}\left(\boldsymbol{E}/c \times \boldsymbol{B}\right) \;=\; \boldsymbol{j}_{\mathrm{energy}}(\mathrm{EMF}) \; , \tag{1.40}$$

we arrive at

$$\frac{\partial u}{\partial t} + \boldsymbol{\nabla} \cdot \boldsymbol{S} = -\, c\, \boldsymbol{E}/c \cdot \left[\boldsymbol{J} + \frac{\partial (c\boldsymbol{P})}{\partial ct} + \boldsymbol{\nabla} \times \boldsymbol{M} \right] . \tag{1.41}$$

This relation was established by the English physicist John Poynting (1852–1914) in 1884 and is therefore known as **Poynting's theorem**. It is the local balance equation which describes how **energy** is transferred in the interaction between an electromagnetic field and a macroscopic material medium for which the constitutive relations (1.12) and (1.13) hold. If the overall sign of the term on the right-hand side is negative, energy is flowing out of the electromagnetic field; if it is positive, energy is flowing into the electromagnetic field. Note that in a **vacuum** [cf. (1.26)–(1.29)] the term on the right-hand side vanishes altogether. For historical reasons, the energy current density of the electromagnetic field is referred to as the **Poynting vector field**, $\boldsymbol{S}$; the product $\boldsymbol{E} \cdot \boldsymbol{J}$ is called **Joule heat** (named after the English physicist James Joule, 1818–1889). The latter is a measure for the work done by an electromagnetic field on moving charges per unit volume and per unit time.

A local balance equation similar to (1.41) can be obtained that describes the exchange of linear or angular momentum between an electromagnetic field and a macroscopic material medium. The derivation of the balance equation for linear momentum will be discussed in chapter 5.

1.4 Maxwell's field equations in integral form

Historically the laws governing electromagnetic phenomena were discovered through many ingenious experiments. The results so obtained are all of an *integral* nature, i.e., they make statements about electromagnetic processes in extended regions of Euclidian space $\mathbb{R}^3$ of finite size. These regions are assumed to be fixed in time. The *differential*, local, formulation of the **laws of electromagnetism** was presented in complete form only later by Maxwell. Here we take the reverse route to obtain the original integral form of **Maxwell's field equations** from the differential ones given by (1.1)–(1.4).

(a) **Ampère–Maxwell law**: We integrate both sides of (1.1) over a surface S, with unit normal $\boldsymbol{n}$, that is bounded by the closed oriented curve ∂S (right-hand convention).

$$\int\!\!\int_S (\boldsymbol{\nabla} \times \boldsymbol{H}) \cdot \boldsymbol{n}\, \mathrm{d}A = \int\!\!\int_S \boldsymbol{J} \cdot \boldsymbol{n}\, \mathrm{d}A + \int\!\!\int_S \frac{\partial \boldsymbol{D}}{\partial t} \cdot \boldsymbol{n}\, \mathrm{d}A\,. \tag{1.42}$$

By use of Stokes' integral theorem (1.33) and by pulling the partial derivative with respect to t through the surface integral (this is possible as we assume the surface S to be fixed in time), this can be converted into

$$\oint_{\partial S} \boldsymbol{H} \cdot \mathrm{d}\boldsymbol{s} = \int\!\!\int_S \boldsymbol{J} \cdot \boldsymbol{n}\, \mathrm{d}A + \frac{\mathrm{d}}{\mathrm{d}t} \int\!\!\int_S \boldsymbol{D} \cdot \boldsymbol{n}\, \mathrm{d}A\,. \tag{1.43}$$

The Ampère–Maxwell law thus states:

> The **circulation** of the magnetic excitation $\boldsymbol{H}$ along the closed oriented curve ∂S bounding the surface S is equal to the sum of the total **electric current** $I(S) := \int\int_S \boldsymbol{J} \cdot \boldsymbol{n}\, \mathrm{d}A$ through S and the rate of change in time of the **flux** of the electric excitation $\boldsymbol{D}$ through S.

(b) **Faraday's law**: We integrate both sides of (1.2) over a surface S, with unit normal $\boldsymbol{n}$, that is bounded by the closed oriented curve ∂S (right-hand convention).

$$\int\int_S (\boldsymbol{\nabla} \times \boldsymbol{E}) \cdot \boldsymbol{n}\, \mathrm{d}A = -\int\int_S \frac{\partial \boldsymbol{B}}{\partial t} \cdot \boldsymbol{n}\, \mathrm{d}A \tag{1.44}$$

Again, by use of Stokes' integral theorem (1.33) and by pulling the partial derivative with respect to t through the surface integral, this can be converted into

$$\oint_{\partial S} \boldsymbol{E} \cdot \mathrm{d}\boldsymbol{s} = -\frac{\mathrm{d}}{\mathrm{d}t} \int\int_S \boldsymbol{B} \cdot \boldsymbol{n}\, \mathrm{d}A\,. \tag{1.45}$$

Faraday's law thus states:

> The **circulation** of the electric field strength $\boldsymbol{E}$ along the closed oriented curve ∂S bounding the surface S is equal to the negative of the rate of change in time of the **flux** of the magnetic field strength $\boldsymbol{B}$ through S.

The minus sign is a reflection of **Lenz's rule** , which is to be discussed at a later stage.

(c) **Gauß' law**: We integrate (1.3) over a volume G that is bounded by the closed surface ∂G, with outward-pointing unit normal $\boldsymbol{n}$.

$$\int\int\int_G \boldsymbol{\nabla} \cdot \boldsymbol{D}\, \mathrm{d}V = \int\int\int_G \rho\, \mathrm{d}V\,. \tag{1.46}$$

By use of Gauß' integral theorem (1.32), this can be converted into

$$\int\int_{\partial G} \boldsymbol{D} \cdot \boldsymbol{n}\, \mathrm{d}A = \int\int\int_G \rho\, \mathrm{d}V\,. \tag{1.47}$$

Gauß' law thus states:

> The **flux** of the electric excitation $\boldsymbol{D}$ through the closed surface ∂G bounding the volume G is equal to the total **electric charge** $Q(G) := \int\int\int_G \rho\, \mathrm{d}V$ contained in G.

(d) **Zero magnetic charges law**: We integrate (1.4) over a volume G that is bounded by the closed surface ∂G, with outward-pointing unit normal $\boldsymbol{n}$.

$$\int\int\int_G \boldsymbol{\nabla} \cdot \boldsymbol{B} \, \mathrm{d}V = 0 \, . \tag{1.48}$$

Again, by use of Gauß' integral theorem (1.32), this can be converted into

$$\int\int_{\partial G} \boldsymbol{B} \cdot \boldsymbol{n} \, \mathrm{d}A = 0 \, . \tag{1.49}$$

The zero magnetic charges law thus states:

> The **flux** of the magnetic field strength $\boldsymbol{B}$ through the closed surface ∂G bounding the volume G is **zero**, irrespective of what volume G in $\mathbb{R}^3$ is chosen.

There are *no* magnetic charges found in Nature.

(e) **Continuity equation**: It follows from the considerations in subsection 1.3.2 that the integral form of (1.7) is given by

$$\frac{\mathrm{d}}{\mathrm{d}t} \int\int\int_G \rho \, \mathrm{d}V = - \int\int_{\partial G} \boldsymbol{J} \cdot \boldsymbol{n} \, \mathrm{d}A \, . \tag{1.50}$$

The rate of change in time of the total **electric charge** contained in a volume G is equal to the total **electric current** into G through the bounding closed surface ∂G,

$$\frac{\mathrm{d}Q(G)}{\mathrm{d}t} = I(\partial G) \, ,$$

[cf. (1.34) with $\Sigma_X \equiv 0$].

Chapter 2
Mathematical techniques

On the mathematical side, our study of **electromagnetism** is in terms of scalar fields (e.g., the electric charge density) and vector fields (e.g., the electric and magnetic field strengths) defined on Euclidian space $\mathbb{R}^3$, and the relationships between them. Since these relationships are mostly *differential* in nature, we make extensive use of **vector calculus**. This chapter provides a brief review of the mathematical techniques we will employ in this study.

First we will review vector calculus for different coordinate systems on $\mathbb{R}^3$. Then we will introduce **Dirac's delta function**. And, finally, we will discuss the general expansion of a square integrable real-valued function $f(x)$ over a given interval in terms of **complete sets of orthonormal functions**, the most familiar ones being **Fourier series expansions** and **Fourier integral representations**.

2.1 Vector calculus in Cartesian coordinates

Let us remind ourselves of the vector analytical differential operators in a right-handed oriented **Cartesian coordinate basis** of $\mathbb{R}^3$, $\{\, \boldsymbol{e}_x,\, \boldsymbol{e}_y,\, \boldsymbol{e}_z \,\}$, with coordinates $\{x, y, z\}$, where

$$1 = |\, \boldsymbol{e}_x \,| = |\, \boldsymbol{e}_y \,| = |\, \boldsymbol{e}_z \,| \,.$$

If on a domain $D \subset \mathbb{R}^3$ we have ϕ as a differentiable scalar-valued function of position $\boldsymbol{r} = (x, y, z)^T$, and $\boldsymbol{A} = (A_x, A_y, A_z)^T$ as a differentiable vector-valued function of $\boldsymbol{r}$, then the vector analytical differential operators assume the explicit forms

Gradient operator:

$$\boldsymbol{\nabla}\phi \;=\; \frac{\partial\phi}{\partial x}\,\boldsymbol{e}_x + \frac{\partial\phi}{\partial y}\,\boldsymbol{e}_y + \frac{\partial\phi}{\partial z}\,\boldsymbol{e}_z \,. \tag{2.1}$$

Divergence operator:

$$\boldsymbol{\nabla}\cdot\boldsymbol{A} \;=\; \frac{\partial A_x}{\partial x} + \frac{\partial A_y}{\partial y} + \frac{\partial A_z}{\partial z} \,. \tag{2.2}$$

Curl operator:

$$\boldsymbol{\nabla}\times\boldsymbol{A} \;=\; \left[\frac{\partial A_z}{\partial y} - \frac{\partial A_y}{\partial z}\right]\boldsymbol{e}_x + \left[\frac{\partial A_x}{\partial z} - \frac{\partial A_z}{\partial x}\right]\boldsymbol{e}_y$$

$$+ \left[\frac{\partial A_y}{\partial x} - \frac{\partial A_x}{\partial y}\right] \boldsymbol{e}_z \, . \tag{2.3}$$

Laplace operator:

$$(\boldsymbol{\nabla} \cdot \boldsymbol{\nabla})\phi \;\; = \;\; \frac{\partial^2 \phi}{\partial x^2} + \frac{\partial^2 \phi}{\partial y^2} + \frac{\partial^2 \phi}{\partial z^2} \, . \tag{2.4}$$

2.2 Vector calculus in orthogonal curvilinear coordinates

Assume on $\mathbb{R}^3$ given a right-handed oriented **orthogonal coordinate basis** $\{\, \boldsymbol{e}_1, \, \boldsymbol{e}_2, \, \boldsymbol{e}_3 \,\}$, with coordinates $\{x_1, x_2, x_3\}$, where orthogonal means that $\boldsymbol{e}_i \cdot \boldsymbol{e}_j = 0$ for $i \neq j$, and $i, j = 1, 2, 3$. From this a **normalised orthogonal basis**[1] $\{\, \hat{\boldsymbol{e}}_1, \, \hat{\boldsymbol{e}}_2, \, \hat{\boldsymbol{e}}_3 \,\}$ can be obtained, where normalised means that

$$1 = \hat{\boldsymbol{e}}_1 \cdot \hat{\boldsymbol{e}}_1 = \hat{\boldsymbol{e}}_2 \cdot \hat{\boldsymbol{e}}_2 = \hat{\boldsymbol{e}}_3 \cdot \hat{\boldsymbol{e}}_3 \, .$$

Introducing the quantities

$$g_{11} := \boldsymbol{e}_1 \cdot \boldsymbol{e}_1 \, , \quad g_{22} := \boldsymbol{e}_2 \cdot \boldsymbol{e}_2 \, , \quad g_{33} := \boldsymbol{e}_3 \cdot \boldsymbol{e}_3 \, ,$$

the two sets of basis vectors are related by the definitions

$$\hat{\boldsymbol{e}}_1 := \frac{1}{\sqrt{g_{11}}} \, \boldsymbol{e}_1 \, , \quad \hat{\boldsymbol{e}}_2 := \frac{1}{\sqrt{g_{22}}} \, \boldsymbol{e}_2 \, , \quad \hat{\boldsymbol{e}}_3 := \frac{1}{\sqrt{g_{33}}} \, \boldsymbol{e}_3 \, .$$

Let, on a domain $D \subset \mathbb{R}^3$, $\phi = \phi(x_1, x_2, x_3)$ be a differentiable scalar-valued function, and $\boldsymbol{A} = \boldsymbol{A}(x_1, x_2, x_3)$ a differentiable vector-valued function. We have

$$\boldsymbol{A} = {}^c\!A_1 \, \boldsymbol{e}_1 + {}^c\!A_2 \, \boldsymbol{e}_2 + {}^c\!A_3 \, \boldsymbol{e}_3 = A_1 \, \hat{\boldsymbol{e}}_1 + A_2 \, \hat{\boldsymbol{e}}_2 + A_3 \, \hat{\boldsymbol{e}}_3 \, ,$$

where

$$A_1 = \sqrt{g_{11}} \, {}^c\!A_1 \, , \quad A_2 = \sqrt{g_{22}} \, {}^c\!A_2 \, , \quad A_3 = \sqrt{g_{33}} \, {}^c\!A_3 \, .$$

One can show that with respect to the normalised orthogonal basis $\{\, \hat{\boldsymbol{e}}_1, \, \hat{\boldsymbol{e}}_2, \, \hat{\boldsymbol{e}}_3 \,\}$, the vector analytical differential operators are given by

Gradient operator:

$$\boldsymbol{\nabla} \phi \;\; = \;\; \frac{1}{\sqrt{g_{11}}} \, \frac{\partial \phi}{\partial x_1} \, \hat{\boldsymbol{e}}_1 + \frac{1}{\sqrt{g_{22}}} \, \frac{\partial \phi}{\partial x_2} \, \hat{\boldsymbol{e}}_2 + \frac{1}{\sqrt{g_{33}}} \, \frac{\partial \phi}{\partial x_3} \, \hat{\boldsymbol{e}}_3 \tag{2.5}$$

[1] This is, in general, a *non-coordinate basis*, i.e., it *cannot* be generated from a transformation of an (orthogonal) coordinate basis $\{\, \boldsymbol{e}_1, \, \boldsymbol{e}_2, \, \boldsymbol{e}_3 \,\}$ with a Jacobian matrix of a specific coordinate transformation.

Divergence operator:

$$\begin{aligned}\boldsymbol{\nabla}\cdot\boldsymbol{A} &= \frac{1}{\sqrt{g_{11}g_{22}g_{33}}} \\ &\quad\times\left[\frac{\partial}{\partial x_1}(\sqrt{g_{22}g_{33}}\,A_1)+\frac{\partial}{\partial x_2}(\sqrt{g_{33}g_{11}}\,A_2)+\frac{\partial}{\partial x_3}(\sqrt{g_{11}g_{22}}\,A_3)\right]\end{aligned} \tag{2.6}$$

Curl operator:

$$\begin{aligned}\boldsymbol{\nabla}\times\boldsymbol{A} &= \frac{1}{\sqrt{g_{11}g_{22}g_{33}}}\left[\left(\frac{\partial(\sqrt{g_{33}}\,A_3)}{\partial x_2}-\frac{\partial(\sqrt{g_{22}}\,A_2)}{\partial x_3}\right)\sqrt{g_{11}}\,\hat{\boldsymbol{e}}_1\right. \\ &\quad+\left(\frac{\partial(\sqrt{g_{11}}\,A_1)}{\partial x_3}-\frac{\partial(\sqrt{g_{33}}\,A_3)}{\partial x_1}\right)\sqrt{g_{22}}\,\hat{\boldsymbol{e}}_2 \\ &\quad+\left.\left(\frac{\partial(\sqrt{g_{22}}\,A_2)}{\partial x_1}-\frac{\partial(\sqrt{g_{11}}\,A_1)}{\partial x_2}\right)\sqrt{g_{33}}\,\hat{\boldsymbol{e}}_3\right]\end{aligned} \tag{2.7}$$

Laplace operator:

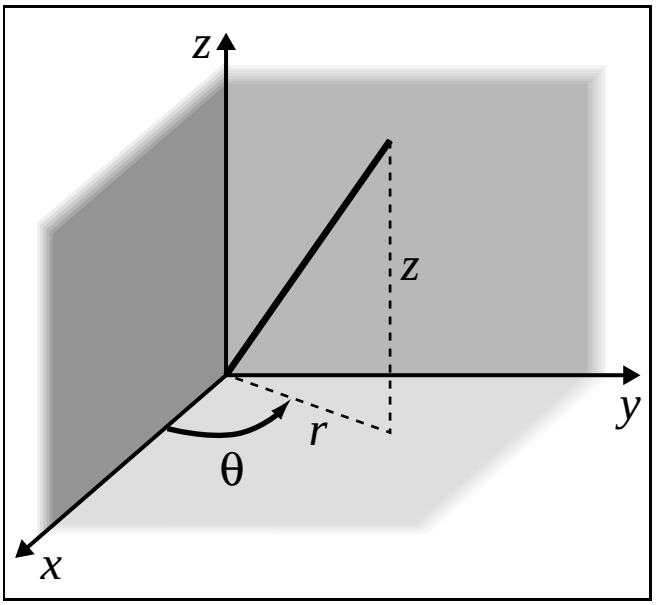
$$\begin{aligned}(\boldsymbol{\nabla}\cdot\boldsymbol{\nabla})\phi &= \frac{1}{\sqrt{g_{11}g_{22}g_{33}}} \\ &\quad\times\left[\frac{\partial}{\partial x_1}\left(\sqrt{\frac{g_{22}g_{33}}{g_{11}}}\,\frac{\partial\phi}{\partial x_1}\right)+\frac{\partial}{\partial x_2}\left(\sqrt{\frac{g_{33}g_{11}}{g_{22}}}\,\frac{\partial\phi}{\partial x_2}\right)\right. \\ &\quad+\left.\frac{\partial}{\partial x_3}\left(\sqrt{\frac{g_{11}g_{22}}{g_{33}}}\,\frac{\partial\phi}{\partial x_3}\right)\right]\,,\end{aligned} \tag{2.8}$$

respectively. The last result follows from (2.6) when $A_1 = \frac{1}{\sqrt{g_{11}}}\,\partial\phi/\partial x_1$, $A_2 = \frac{1}{\sqrt{g_{22}}}\,\partial\phi/\partial x_2$ and $A_3 = \frac{1}{\sqrt{g_{33}}}\,\partial\phi/\partial x_3$ are used. We are interested in the explicit form these vector analytical differential operators assume for two frequently employed kinds of **orthogonal curvilinear coordinates**.

2.2.1 Cylindrical polar coordinates

Cartesian coordinates $\{x, y, z\}$ on $\mathbb{R}^3$ are related to **cylindrical polar coordinates** $\{r, \varphi, z\}$ according to

$$x = r\cos\varphi\,, \quad y = r\sin\varphi\,, \quad z = z\,, \tag{2.9}$$



where the coordinates $\{r, \varphi, z\}$ vary in the intervals $r \geq 0$, $0 \leq \varphi \leq 2\pi$ and $-\infty < z < +\infty$. The coordinate r gives the magnitude of the position vector $\boldsymbol{r}$ projected onto a plane $z =$ const, while z is the magnitude of $\boldsymbol{r}$ projected onto the z–axis. The coordinate φ is the azimuthal angle subtended by the projection of $\boldsymbol{r}$ onto a plane $z =$ const and the positive (x, z)–half plane, measured anti-clockwise. *Note* that sometimes θ is used to denote the azimuthal angle, instead of φ.

The right-handed oriented coordinate basis $\{\, \boldsymbol{e}_r,\, \boldsymbol{e}_\varphi,\, \boldsymbol{e}_z \,\}$ is given in terms of the Cartesian coordinate basis $\{\, \boldsymbol{e}_x,\, \boldsymbol{e}_y,\, \boldsymbol{e}_z \,\}$ by

$$\begin{aligned} \boldsymbol{e}_r &= \cos\varphi\, \boldsymbol{e}_x + \sin\varphi\, \boldsymbol{e}_y \\ \boldsymbol{e}_\varphi &= -\, r \sin\varphi\, \boldsymbol{e}_x + r \cos\varphi\, \boldsymbol{e}_y \quad . \\ \boldsymbol{e}_z &= \boldsymbol{e}_z \end{aligned} \tag{2.10}$$

For cylindrical polar coordinates we thus have

$$\sqrt{g_{11}} = \sqrt{g_{rr}} = 1\,, \qquad \sqrt{g_{22}} = \sqrt{g_{\varphi\varphi}} = r\,, \qquad \sqrt{g_{33}} = \sqrt{g_{zz}} = 1\,,$$

so that we define a normalised orthogonal basis $\{\, \hat{\boldsymbol{e}}_r,\, \hat{\boldsymbol{e}}_\varphi,\, \hat{\boldsymbol{e}}_z \,\}$ by

$$\hat{\boldsymbol{e}}_r := \boldsymbol{e}_r\,, \qquad \hat{\boldsymbol{e}}_\varphi := \frac{1}{r}\, \boldsymbol{e}_\varphi\,, \qquad \hat{\boldsymbol{e}}_z := \boldsymbol{e}_z\,. \tag{2.11}$$

Then we obtain from (2.5)–(2.8), respectively,

Gradient operator:

$$\boldsymbol{\nabla}\phi = \frac{\partial\phi}{\partial r}\, \hat{\boldsymbol{e}}_r + \frac{1}{r}\, \frac{\partial\phi}{\partial\varphi}\, \hat{\boldsymbol{e}}_\varphi + \frac{\partial\phi}{\partial z}\, \hat{\boldsymbol{e}}_z \tag{2.12}$$

Divergence operator:

$$\boldsymbol{\nabla} \cdot \boldsymbol{A} = \frac{1}{r}\, \frac{\partial}{\partial r}(r\, A_r) + \frac{1}{r}\, \frac{\partial A_\varphi}{\partial\varphi} + \frac{\partial A_z}{\partial z} \tag{2.13}$$

Curl operator:

$$\begin{aligned}\boldsymbol{\nabla}\times\boldsymbol{A} &= \left(\frac{1}{r}\,\frac{\partial A_z}{\partial\varphi}-\frac{\partial A_\varphi}{\partial z}\right)\hat{\boldsymbol{e}}_r+\left(\frac{\partial A_r}{\partial z}-\frac{\partial A_z}{\partial r}\right)\hat{\boldsymbol{e}}_\varphi \\ &\quad+\frac{1}{r}\left(\frac{\partial}{\partial r}(r\,A_\varphi)-\frac{\partial A_r}{\partial\varphi}\right)\hat{\boldsymbol{e}}_z\end{aligned} \tag{2.14}$$

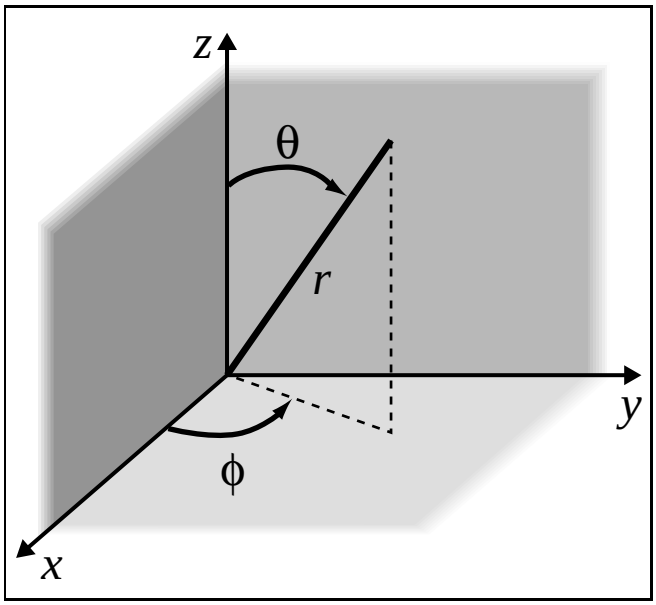
Laplace operator:

$$(\boldsymbol{\nabla}\cdot\boldsymbol{\nabla})\phi = \frac{1}{r}\,\frac{\partial}{\partial r}\left(r\,\frac{\partial\phi}{\partial r}\right)+\frac{1}{r^2}\,\frac{\partial^2\phi}{\partial\varphi^2}+\frac{\partial^2\phi}{\partial z^2}\,. \tag{2.15}$$

2.2.2 Spherical polar coordinates

Cartesian coordinates $\{x,y,z\}$ on $\mathbb{R}^3$ are related to **spherical polar coordinates** $\{r,\vartheta,\varphi\}$ according to

$$x=r\sin\vartheta\cos\varphi\,,\quad y=r\sin\vartheta\sin\varphi\,,\quad z=r\cos\vartheta\,, \tag{2.16}$$



where the coordinates $\{r,\vartheta,\varphi\}$ vary in the intervals $r\geq 0$, $0\leq\vartheta\leq\pi$, and $0\leq\varphi\leq 2\pi$. The coordinate r gives the magnitude of the position vector $\boldsymbol{r}$, while the coordinate ϑ denotes the angle between $\boldsymbol{r}$ and the z–axis. The coordinate φ is the azimuthal angle subtended by the projection of $\boldsymbol{r}$ onto the (x,y)–plane and the positive (x,z)–half plane, measured anti-clockwise.

The right-handed oriented coordinate basis $\{\,\boldsymbol{e}_r,\,\boldsymbol{e}_\vartheta,\,\boldsymbol{e}_\varphi\,\}$ is given in terms of the Cartesian coordinate basis $\{\,\boldsymbol{e}_x,\,\boldsymbol{e}_y,\,\boldsymbol{e}_z\,\}$ by

$$\begin{aligned}\boldsymbol{e}_r &= \sin\vartheta\cos\varphi\,\boldsymbol{e}_x+\sin\vartheta\sin\varphi\,\boldsymbol{e}_y+\cos\vartheta\,\boldsymbol{e}_z\\ \boldsymbol{e}_\vartheta &= r\cos\vartheta\cos\varphi\,\boldsymbol{e}_x+r\cos\vartheta\sin\varphi\,\boldsymbol{e}_y-r\sin\vartheta\,\boldsymbol{e}_z\\ \boldsymbol{e}_\varphi &= -\,r\sin\vartheta\sin\varphi\,\boldsymbol{e}_x+r\sin\vartheta\cos\varphi\,\boldsymbol{e}_y\end{aligned}\quad. \tag{2.17}$$

For spherical polar coordinates we thus have

$$\sqrt{g_{11}} = \sqrt{g_{rr}} = 1 \ , \quad \sqrt{g_{22}} = \sqrt{g_{\vartheta\vartheta}} = r \ , \quad \sqrt{g_{33}} = \sqrt{g_{\varphi\varphi}} = r \, \sin\vartheta \ ,$$

so that we define a normalised orthogonal basis $\{ \, \hat{\boldsymbol{e}}_r, \, \hat{\boldsymbol{e}}_\vartheta, \, \hat{\boldsymbol{e}}_\varphi \, \}$ by

$$\hat{\boldsymbol{e}}_r := \boldsymbol{e}_r \ , \quad \hat{\boldsymbol{e}}_\vartheta := \frac{1}{r} \, \boldsymbol{e}_\vartheta \ , \quad \hat{\boldsymbol{e}}_\varphi := \frac{1}{r \sin\varphi} \, \boldsymbol{e}_\varphi \ . \tag{2.18}$$

Then we obtain from (2.5)–(2.8), respectively,

Gradient operator:

$$\boldsymbol{\nabla}\phi \ = \ \frac{\partial\phi}{\partial r} \, \hat{\boldsymbol{e}}_r + \frac{1}{r} \, \frac{\partial\phi}{\partial\vartheta} \, \hat{\boldsymbol{e}}_\vartheta + \frac{1}{r \sin\vartheta} \, \frac{\partial\phi}{\partial\varphi} \, \hat{\boldsymbol{e}}_\varphi \tag{2.19}$$

Divergence operator:

$$\boldsymbol{\nabla} \cdot \boldsymbol{A} \ = \ \frac{1}{r^2} \, \frac{\partial}{\partial r}(r^2 \, A_r) + \frac{1}{r \sin\vartheta} \, \frac{\partial}{\partial\vartheta}(\sin\vartheta \, A_\vartheta) + \frac{1}{r \sin\vartheta} \, \frac{\partial A_\varphi}{\partial\varphi} \tag{2.20}$$

Curl operator:

$$\begin{aligned} \boldsymbol{\nabla} \times \boldsymbol{A} \ = \ & \frac{1}{r \sin\vartheta} \left(\frac{\partial}{\partial\vartheta}(\sin\vartheta \, A_\varphi) - \frac{\partial A_\vartheta}{\partial\varphi} \right) \hat{\boldsymbol{e}}_r \\ & + \frac{1}{r} \left(\frac{1}{\sin\vartheta} \, \frac{\partial A_r}{\partial\varphi} - \frac{\partial}{\partial r}(r \, A_\varphi) \right) \hat{\boldsymbol{e}}_\vartheta + \frac{1}{r} \left(\frac{\partial}{\partial r}(r \, A_\vartheta) - \frac{\partial A_r}{\partial\vartheta} \right) \hat{\boldsymbol{e}}_\varphi \end{aligned} \tag{2.21}$$

Laplace operator:

$$\begin{aligned} (\boldsymbol{\nabla} \cdot \boldsymbol{\nabla})\phi \ = \ & \frac{1}{r^2} \, \frac{\partial}{\partial r} \left(r^2 \, \frac{\partial\phi}{\partial r} \right) + \frac{1}{r^2 \sin\vartheta} \, \frac{\partial}{\partial\vartheta} \left(\sin\vartheta \, \frac{\partial\phi}{\partial\vartheta} \right) \\ & + \frac{1}{r^2 \sin^2\vartheta} \, \frac{\partial^2\phi}{\partial\varphi^2} \ . \end{aligned} \tag{2.22}$$

2.3 Vector analytical identities

Certain vector analytical identities will be of use in this course, which do hold *independent* of the choice of coordinate system we pick on $\mathbb{R}^3$. Let, on a domain $D \subset \mathbb{R}^3$, $\phi = \phi(\boldsymbol{r})$ and $\psi = \psi(\boldsymbol{r})$ be differentiable scalar-valued functions, and $\boldsymbol{A} = \boldsymbol{A}(\boldsymbol{r})$ and $\boldsymbol{B} = \boldsymbol{B}(\boldsymbol{r})$ be differentiable vector-valued functions. Then

$$\boldsymbol{\nabla} \cdot (\phi \, \boldsymbol{A}) \ \equiv \ \phi \, (\boldsymbol{\nabla} \cdot \boldsymbol{A}) + \boldsymbol{A} \cdot (\boldsymbol{\nabla}\phi) \tag{2.23}$$

$$\boldsymbol{\nabla} \times (\phi \, \boldsymbol{A}) \ \equiv \ \phi \, (\boldsymbol{\nabla} \times \boldsymbol{A}) - \boldsymbol{A} \times (\boldsymbol{\nabla}\phi) \tag{2.24}$$

$$\boldsymbol{\nabla} \times (\boldsymbol{\nabla}\phi) \ \equiv \ \boldsymbol{0} \tag{2.25}$$

$$\boldsymbol{\nabla} \cdot (\boldsymbol{\nabla} \times \boldsymbol{A}) \ \equiv \ 0 \tag{2.26}$$

$$\boldsymbol{\nabla}(\phi\,\psi) \ \equiv \ \psi \, \boldsymbol{\nabla}\phi + \phi \, \boldsymbol{\nabla}\psi \tag{2.27}$$

$$\boldsymbol{\nabla}\cdot(\boldsymbol{A}\times\boldsymbol{B}) \equiv \boldsymbol{B}\cdot(\boldsymbol{\nabla}\times\boldsymbol{A}) - \boldsymbol{A}\cdot(\boldsymbol{\nabla}\times\boldsymbol{B}) \tag{2.28}$$

$$\begin{aligned}\boldsymbol{\nabla}\times(\boldsymbol{A}\times\boldsymbol{B}) &\equiv (\boldsymbol{B}\cdot\boldsymbol{\nabla})\boldsymbol{A} - (\boldsymbol{A}\cdot\boldsymbol{\nabla})\boldsymbol{B} \\ &\quad + \boldsymbol{A}\,(\boldsymbol{\nabla}\cdot\boldsymbol{B}) - \boldsymbol{B}\,(\boldsymbol{\nabla}\cdot\boldsymbol{A})\end{aligned} \tag{2.29}$$

$$\begin{aligned}\boldsymbol{\nabla}(\boldsymbol{A}\cdot\boldsymbol{B}) &\equiv \boldsymbol{A}\times(\boldsymbol{\nabla}\times\boldsymbol{B}) + \boldsymbol{B}\times(\boldsymbol{\nabla}\times\boldsymbol{A}) \\ &\quad + (\boldsymbol{A}\cdot\boldsymbol{\nabla})\boldsymbol{B} + (\boldsymbol{B}\cdot\boldsymbol{\nabla})\boldsymbol{A}\,.\end{aligned} \tag{2.30}$$

Note that with respect to a *Cartesian* coordinate basis (only) the additional identity

$$\boldsymbol{\nabla}\times(\boldsymbol{\nabla}\times\boldsymbol{A}) \equiv \boldsymbol{\nabla}(\boldsymbol{\nabla}\cdot\boldsymbol{A}) - (\boldsymbol{\nabla}\cdot\boldsymbol{\nabla})\boldsymbol{A} \tag{2.31}$$

holds.

2.4 Consequences of Gauß' integral theorem

Suppose again that $\phi = \phi(\boldsymbol{r})$ and $\psi = \psi(\boldsymbol{r})$ are differentiable scalar-valued functions on a domain $D \subset \mathbb{R}^3$. Now suppose that $\boldsymbol{A} = \psi\,\boldsymbol{\nabla}\phi$. Then, using this particular $\boldsymbol{A}$ in Gauß' integral theorem (1.32), we find on a region $G \subset D$ the identity

$$\int\!\!\int\!\!\int_G \psi\,\boldsymbol{\nabla}\cdot(\boldsymbol{\nabla}\phi)\,\mathrm{d}V \equiv \int\!\!\int_{\partial G} \psi\,(\boldsymbol{n}\cdot\boldsymbol{\nabla})\phi\,\mathrm{d}A - \int\!\!\int\!\!\int_G (\boldsymbol{\nabla}\psi)\cdot(\boldsymbol{\nabla}\phi)\,\mathrm{d}V\,. \tag{2.32}$$

This is known as **Green's first integral formula** (named after the English mathematician, physicist and miller George Green, 1793–1841). Interchanging the roles of ϕ and ψ in (2.32), and subtracting the two expressions so obtained from one another, we arrive at

$$\int\!\!\int\!\!\int_G [\,\psi\,\boldsymbol{\nabla}\cdot(\boldsymbol{\nabla}\phi) - \phi\,\boldsymbol{\nabla}\cdot(\boldsymbol{\nabla}\psi)\,]\,\mathrm{d}V \equiv \int\!\!\int_{\partial G} [\,\psi\,(\boldsymbol{n}\cdot\boldsymbol{\nabla})\phi - \phi\,(\boldsymbol{n}\cdot\boldsymbol{\nabla})\psi\,]\,\mathrm{d}A\,, \tag{2.33}$$

i.e., **Green's second integral formula**.

In a similarly straightforward fashion Gauß' integral theorem (1.32) yields for $\boldsymbol{A} = \boldsymbol{\nabla}\phi$ the identity

$$\int\!\!\int\!\!\int_G (\boldsymbol{\nabla}\cdot\boldsymbol{\nabla})\phi\,\mathrm{d}V \equiv \int\!\!\int_{\partial G} (\boldsymbol{n}\cdot\boldsymbol{\nabla}\phi)\,\mathrm{d}A\,. \tag{2.34}$$

Next, take $\boldsymbol{A} = \phi\,\boldsymbol{e}$, where $\boldsymbol{e}$ is an arbitrary *constant* vector on $D \subset \mathbb{R}^3$. Inserting in (1.32) and using (2.23), we obtain

$$0 = \boldsymbol{e}\cdot\left[\int\!\!\int\!\!\int_G \boldsymbol{\nabla}\phi\,\mathrm{d}V - \int\!\!\int_{\partial G} \phi\,\boldsymbol{n}\,\mathrm{d}A\right]\,, \tag{2.35}$$

as $\boldsymbol{e}$ is assumed constant. Now, because $\boldsymbol{e}$ is arbitrary (and non-vanishing), we must have

$$\int\!\!\int_{\partial G} \phi\,\boldsymbol{n}\,\mathrm{d}A \equiv \int\!\!\int\!\!\int_G (\boldsymbol{\nabla}\phi)\,\mathrm{d}V\,. \tag{2.36}$$

Note the intimate relation of this result to the *definition* of the **gradient** of a differentiable scalar-valued function $\phi = \phi(\boldsymbol{r})$ given by

$$\boldsymbol{\nabla}\phi := \lim_{V\to 0} \frac{\int\int_{\partial G} \phi\, \boldsymbol{n}\, \mathrm{d}A}{V} \,.$$

Finally, take $\boldsymbol{A} = \boldsymbol{e} \times \boldsymbol{B}$, with $\boldsymbol{e}$ is an arbitrary *constant* vector and $\boldsymbol{B}$ a differentiable vector on $D \subset \mathbb{R}^3$. Inserting in (1.32) and using (2.28), we obtain

$$0 = \boldsymbol{e} \cdot \left[\int\int\int_G (\boldsymbol{\nabla} \times \boldsymbol{B})\, \mathrm{d}V + \int\int_{\partial G} (\boldsymbol{B} \times \boldsymbol{n})\, \mathrm{d}A \right] \,, \tag{2.37}$$

as $\boldsymbol{e}$ is assumed constant. Now, because $\boldsymbol{e}$ is arbitrary (and non-vanishing), we must have

$$\int\int_{\partial G} \boldsymbol{B} \times \boldsymbol{n}\, \mathrm{d}A \equiv - \int\int\int_G (\boldsymbol{\nabla} \times \boldsymbol{B})\, \mathrm{d}V \,. \tag{2.38}$$

Note the intimate relation of this result to the *definition* of the **curl** of a differentiable vector-valued function $\boldsymbol{B} = \boldsymbol{B}(\boldsymbol{r})$ given by

$$\boldsymbol{\nabla} \times \boldsymbol{B} := - \lim_{V\to 0} \frac{\int\int_{\partial G} \boldsymbol{B} \times \boldsymbol{n}\, \mathrm{d}A}{V} \,.$$

2.5 Dirac's delta function

Let us assume that $f(\boldsymbol{r})$ is a continuously differentiable real-valued function on a region $G \subset \mathbb{R}^3$. Then, by definition, **Dirac's delta function** (named after the English physicist Paul Adrien Maurice Dirac, 1902–1984) has the property

$$\int\int\int_G f(\boldsymbol{r}')\, \delta(\boldsymbol{r} - \boldsymbol{r}')\, \mathrm{d}V' = \begin{cases} f(\boldsymbol{r}) & \text{for } \boldsymbol{r} = \boldsymbol{r}' \in G \\ 0 & \text{for } \boldsymbol{r} = \boldsymbol{r}' \notin G \end{cases} \,. \tag{2.39}$$

With the special choice $f(\boldsymbol{r}) = 1$, this yields

$$\int\int\int_G \delta(\boldsymbol{r} - \boldsymbol{r}')\, \mathrm{d}V' = \begin{cases} 1 & \text{for } \boldsymbol{r} = \boldsymbol{r}' \in G \\ 0 & \text{for } \boldsymbol{r} = \boldsymbol{r}' \notin G \end{cases} \,. \tag{2.40}$$

Assuming compact support for $f(\boldsymbol{r})$ on G (i.e., sufficiently rapid fall-off behaviour as $r \to \infty$), a gradient of **Dirac's delta function** can be interpreted according to

$$\int\int\int_G f(\boldsymbol{r}')\, \boldsymbol{\nabla}_{\boldsymbol{r}'} \delta(\boldsymbol{r} - \boldsymbol{r}')\, \mathrm{d}V' = - \boldsymbol{\nabla}_{\boldsymbol{r}'} f(\boldsymbol{r}') \,\big|_{\boldsymbol{r}'=\boldsymbol{r}} \,; \tag{2.41}$$

$\boldsymbol{\nabla}_{\boldsymbol{r}'}$ here denotes the gradient with respect to $\boldsymbol{r}'$. Note that $\delta(\boldsymbol{r} - \boldsymbol{r}') = \delta(\boldsymbol{r}' - \boldsymbol{r})$ applies, i.e., **Dirac's delta function** is symmetric.

2.6 Orthonormal function expansions

Many **linear differential equations** that occur in **mathematical physics** (as, e.g., **Poisson's** and **Laplace's equations** of electrostatics to be discussed below) give rise to solutions that can be represented as expansions over **complete sets of orthonormal functions** on a given interval. The particular complete set that one uses depends on the geometry of the problem at hand.

Consider, in *one* spatial dimension, an interval (a, b) in a variable x, with a complete set of real or complex functions $U_n(x)$, $n = 1, 2, \ldots, \infty$, that are square integrable and orthonormal on (a, b). That is, they satisfy the

Orthonormality condition:

$$\int_a^b U_n^*(x)\, U_m(x)\, \mathrm{d}x = \delta_{nm} \;, \tag{2.42}$$

with δ_{nm} denoting the **Kronecker symbol** (named after the German mathematician Leopold Kronecker, 1823–1891), and the

Completeness condition:

$$\sum_{n=1}^{\infty} U_n^*(x')\, U_n(x) = \delta(x' - x) \;, \tag{2.43}$$

with $\delta(x' - x)$ denoting **Dirac's delta function**. Then an arbitrary real-valued function $f(x)$ that is square integrable on (a, b) can be expanded in an infinite series of the $U_n(x)$ according to

$$f(x) = \sum_{n=1}^{\infty} a_n\, U_n(x) \tag{2.44}$$

$$a_n = \int_a^b U_n^*(x)\, f(x)\, \mathrm{d}x \;. \tag{2.45}$$

The constant expansion coefficients a_n are said to represent the **spectrum** of $f(x)$ with respect to the $U_n(x)$.

That the right-hand side of (2.44) does indeed provide a rigorous representation of $f(x)$ can be seen as follows. Using (2.45) in (2.44), we find

$$f(x) = \sum_{n=1}^{\infty} \left[\int_a^b U_n^*(x')\, f(x')\, \mathrm{d}x' \right] U_n(x) \;,$$

the right-hand side of which can be rewritten as

$$\sum_{n=1}^{\infty} \left[\int_a^b U_n^*(x')\, f(x')\, \mathrm{d}x' \right] U_n(x) = \int_a^b f(x') \left[\sum_{n=1}^{\infty} U_n^*(x')\, U_n(x) \right] \mathrm{d}x' \;.$$

But then, by (2.43), we have

$$\int_a^b f(x') \left[\sum_{n=1}^{\infty} U_n^*(x')\, U_n(x) \right] \mathrm{d}x' = \int_a^b f(x')\, \delta(x' - x)\, \mathrm{d}x' = f(x) \;,$$

which ends this demonstration.

Note that the concept of orthonormal function expansions can be conveniently extended to square integrable functions in *three* spatial dimensions.

2.6.1 Fourier series expansions

The *most famous* set of orthonormal functions are the normalised sines and cosines,

$$\sqrt{\frac{2}{a}}\,\sin\left(\frac{n\,2\pi}{a}\,x\right)\;,\quad \sqrt{\frac{2}{a}}\,\cos\left(\frac{n\,2\pi}{a}\,x\right)\;,\qquad n=1,2,\ldots,\infty$$

used in a **Fourier series expansion** (named after the French mathematician Jean Baptiste Joseph Fourier, 1768–1830) of a real-valued function $f(x)$ that is periodic over an x–interval $(-a/2,a/2)$. Namely,

$$\begin{aligned} f(x) &= \tfrac{1}{2}\,A_0+\sqrt{\frac{2}{a}}\sum_{n=1}^{\infty}\left[A_n\,\cos\left(\frac{n\,2\pi}{a}\,x\right)+B_n\,\sin\left(\frac{n\,2\pi}{a}\,x\right)\right]\\ A_n &= \sqrt{\frac{2}{a}}\int_{-a/2}^{a/2} f(x)\,\cos\left(\frac{n\,2\pi}{a}\,x\right)\mathrm{d}x\\ B_n &= \sqrt{\frac{2}{a}}\int_{-a/2}^{a/2} f(x)\,\sin\left(\frac{n\,2\pi}{a}\,x\right)\mathrm{d}x\;. \end{aligned}$$

Alternatively, using the normalised complex exponentials

$$U_n(x)=\frac{1}{\sqrt{a}}\,\mathrm{e}^{i\frac{n\,2\pi}{a}\,x}\;,\quad n=0,\pm1,\pm2,\ldots,\pm\infty\;,$$

$f(x)$ can be expressed by the expansion

$$f(x) = \frac{1}{\sqrt{a}}\sum_{n=-\infty}^{+\infty} A_n\,\mathrm{e}^{i\frac{n\,2\pi}{a}\,x} \tag{2.46}$$

$$A_n = \frac{1}{\sqrt{a}}\int_{-a/2}^{a/2}\mathrm{e}^{-i\frac{n\,2\pi}{a}\,x}\,f(x)\,\mathrm{d}x\;. \tag{2.47}$$

2.6.2 Fourier integral representations

In the *limit* that we let the x–interval $(-a/2,a/2)$ become infinite, i.e., $a\to\infty$, while simultaneously making the transitions

$$\begin{aligned} \frac{n\,2\pi}{a}\,x &\longrightarrow k\\ \sum_{n=-\infty}^{+\infty} &\longrightarrow \int_{-\infty}^{\infty}\mathrm{d}n \;=\; \frac{a}{2\pi}\int_{-\infty}^{\infty}\mathrm{d}k\\ A_n &\longrightarrow \sqrt{\frac{2\pi}{a}}\,A(k) \end{aligned}$$

from discretely to continuously varying quantities, we obtain the **Fourier integral representation** of a square integrable real-valued function $f(x)$ ("of period infinity") over the interval $(-\infty, +\infty)$ given by

$$f(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} A(k)\, \mathrm{e}^{ikx}\, \mathrm{d}k \tag{2.48}$$

$$A(k) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \mathrm{e}^{-ikx}\, f(x)\, \mathrm{d}x\ . \tag{2.49}$$

The orthonormality condition for the continuous set of square integrable functions $U(k,x) = \frac{1}{\sqrt{2\pi}}\, \mathrm{e}^{ikx}$ on $(-\infty, +\infty)$ reads

$$\frac{1}{2\pi} \int_{-\infty}^{\infty} \mathrm{e}^{i(k-k')x}\, \mathrm{d}x = \delta(k-k')\ , \tag{2.50}$$

while the completeness condition reads

$$\frac{1}{2\pi} \int_{-\infty}^{\infty} \mathrm{e}^{ik(x-x')x}\, \mathrm{d}k = \delta(x-x')\ . \tag{2.51}$$

Hence, both of these integral expressions provide convenient representations of **Dirac's delta function** discussed in section 2.5.

We conclude by stating that straightforward extensions of the Fourier integral representation to square integrable functions in *three* spatial dimensions do exist.

Chapter 3
Electrostatics

As already indicated in section 1.2, **electrostatics** investigates the special case of **Maxwell's field equations** when all field variables are assumed to be independent of the time coordinate t, and, moreover, we impose the restrictions $\mathbf{0} = \boldsymbol{B} = \boldsymbol{M} = \boldsymbol{J}$. In addition, we will presently assume also that $\boldsymbol{P} = \mathbf{0}$. Then it follows from (1.15)–(1.18) that

$$\boldsymbol{\nabla} \times \boldsymbol{E} = \mathbf{0} \tag{3.1}$$

$$\boldsymbol{\nabla} \cdot \boldsymbol{E} = \frac{\rho}{\epsilon_0} \,. \tag{3.2}$$

The investigations of this chapter will revolve entirely around simultaneously solving (3.1) and (3.2), which prescribe to us the values of the curl and the divergence of the vector field $\boldsymbol{E}$ in terms of a given static continuous distribution of electric charges represented by ρ. The imposition of boundary conditions on $\boldsymbol{E}$ will ensure that solutions we obtain to (3.1) and (3.2) will be unique. We can say that **electrostatics** is the study of the fields and interactions of static electric charges and conducting boundaries.

As the speed of light in vacuum, c, has no relevance in **electrostatics**, we will drop it from our equations throughout this chapter.

3.1 Coulomb's law and superposition principle

Discussions on **electrostatics** traditionally start by introducing the following experimental result.

Coulomb's law:

> Between two point charges q_1 and q_2 at rest relative to each other there acts a mechanical force directly proportional to the product of the charges and inversely proportional to the square of the distance between them. The force is along the straight line from one charge to the other. It is attractive for unlike charges but repulsive for like charges.

In quantitative terms the **Coulomb force** felt by q_1 at position $\boldsymbol{r}_1$ due to the presence of q_2 at position $\boldsymbol{r}_2$ is expressed by

$$\boldsymbol{F}(\boldsymbol{r}_1) = \frac{1}{4\pi\epsilon_0}\,\frac{q_1 q_2}{|\boldsymbol{r}_1 - \boldsymbol{r}_2|^2}\,\hat{\boldsymbol{e}} = -\,\boldsymbol{F}(\boldsymbol{r}_2)\;; \tag{3.3}$$

the **unit vector** pointing from q_2 to q_1 is defined by $\hat{\boldsymbol{e}} := \boldsymbol{r}_1 - \boldsymbol{r}_2/|\boldsymbol{r}_1 - \boldsymbol{r}_2|$. Note that $\boldsymbol{F}(\boldsymbol{r}_2)$ on q_2 is equal in magnitude but opposite in direction to $\boldsymbol{F}(\boldsymbol{r}_1)$.

Comparing this result to the electrostatic subcase of **Lorentz's force law** (1.8), we presently have $\boldsymbol{F}(\boldsymbol{r}_1) = q_1\,\boldsymbol{E}(\boldsymbol{r}_1)$, so that the electric field strength at position $\boldsymbol{r}_1$ is given by

$$\boldsymbol{E}(\boldsymbol{r}_1) = \frac{1}{4\pi\epsilon_0}\,\frac{q_2}{|\boldsymbol{r}_1 - \boldsymbol{r}_2|^2}\,\hat{\boldsymbol{e}}\;. \tag{3.4}$$

It is the underlying **linearity** of electromagnetic interactions that made Coulomb also realise experimentally that a **superposition principle** holds for his electrostatic force law. Therefore, the electric field strength at position $\boldsymbol{r}$ due to a system of n point charges q_i ($i = 1, \ldots, n$) at rest at positions $\boldsymbol{r}_i$ is just given by the vectorial sum

$$\boldsymbol{E}(\boldsymbol{r}) = \frac{1}{4\pi\epsilon_0}\,\sum_{i=1}^{n}\,\frac{q_i}{|\boldsymbol{r} - \boldsymbol{r}_i|^2}\,\hat{\boldsymbol{e}}_i\;, \tag{3.5}$$

where $\hat{\boldsymbol{e}}_i := \boldsymbol{r} - \boldsymbol{r}_i/|\boldsymbol{r} - \boldsymbol{r}_i|$. It is important to note that in writing expression (3.4) [or (3.5)] for $\boldsymbol{E}$, we implicitly introduced the idealised concept of a **test point charge** q_1 [or q] that does itself feel $\boldsymbol{E}$ but that does *not* contribute to the generation of $\boldsymbol{E}$. That is, we have assumed that q_1 itself is not part of a given distribution of electric charges.

To illustrate these results, let us give two numerical examples. The magnitude of the electric field strength generated by a single **electron** at rest at $1\,\mathrm{m}$ distance is $|\boldsymbol{E}_e|_{r=1\mathrm{m}} = 1.44 \times 10^{-9}\,\frac{\mathrm{kg\,m}}{\mathrm{s^2\,C}}$. Likewise, the magnitude of the electric field strength generated by a charge of $1\,\mathrm{C}$ at rest at $1\,\mathrm{m}$ distance is $|\boldsymbol{E}_{1\mathrm{C}}|_{r=1\mathrm{m}} = 8.99 \times 10^{9}\,\frac{\mathrm{kg\,m}}{\mathrm{s^2\,C}}$.

As mentioned before, at a macroscopic level it is adequate to consider *continuous* distributions of electric charges over given volumes and surfaces. Thus, if ΔQ is the amount of electric charge inside a small volume ΔV, or on a small surface ΔA, then a **volume charge density** ρ and a **surface charge density** σ are defined by the limits

$$\rho := \lim_{\Delta V \to 0}\,\frac{\Delta Q}{\Delta V} \qquad\qquad \sigma := \lim_{\Delta A \to 0}\,\frac{\Delta Q}{\Delta A}\;, \tag{3.6}$$

respectively.[1] So when, in a finite spatial region $G \subset \mathbb{R}^3$ bounded by a closed surface ∂G, there is given a continuous charge distribution of density ρ in G

[1] In reality, ΔV and ΔA cannot really go to zero, since eventually ΔQ would have to vary discontinuously due to the fact that electric charge in Nature is quantised. However, depending on the characteristic scale of a given problem, for practical purposes it is generally sufficient if ΔV and ΔA approach orders of magnitudes between $10^{-24}\,\mathrm{m}^3$ to $10^{-30}\,\mathrm{m}^3$ and $10^{-16}\,\mathrm{m}^2$ to $10^{-20}\,\mathrm{m}^2$, respectively.

and density σ on ∂G, the electric field strength at position $\boldsymbol{r}$ amounts to

$$\boldsymbol{E}(\boldsymbol{r}) = \frac{1}{4\pi\epsilon_0}\int\int\int_G \frac{\rho(\boldsymbol{r}')}{|\boldsymbol{r}-\boldsymbol{r}'|^2}\,\hat{\boldsymbol{e}}'\,\mathrm{d}V' + \frac{1}{4\pi\epsilon_0}\int\int_{\partial G}\frac{\sigma(\boldsymbol{r}')}{|\boldsymbol{r}-\boldsymbol{r}'|^2}\,\hat{\boldsymbol{e}}'\,\mathrm{d}A'\,, \tag{3.7}$$

where $\hat{\boldsymbol{e}}' := \boldsymbol{r}-\boldsymbol{r}'/|\boldsymbol{r}-\boldsymbol{r}'|$. Note that the integrand of each of the two terms on the right-hand side becomes singular when the observation point position $\boldsymbol{r}$ coincides with the source integration position $\boldsymbol{r}'$. This cannot happen when $\boldsymbol{r}$ lies outside of G (and ∂G). The case when $\boldsymbol{r}$ lies inside of G will be addressed below.

Multiplying (3.7) with the value of a test point charge q_1 results in a generalised form of **Coulomb's force law**.

3.1.1 Gauß' law

We conclude this section by recalling **Gauß' law** (cf. chapter 1), which is of fundamental importance to applications in **electrostatics** (and beyond). It states that the total flux of the electric field strength through a *closed* surface ∂G is proportional to the total electric charge contained in the volume G inside of ∂G, i.e.,

$$\int\int_{\partial G}\boldsymbol{E}\cdot\boldsymbol{n}\,\mathrm{d}A = \frac{1}{\epsilon_0}\int\int\int_G \rho\,\mathrm{d}V = \frac{Q(G)}{\epsilon_0}\,. \tag{3.8}$$

Let us use this relation to obtain the magnitude of the electric field strength $\boldsymbol{E}$ of a point charge q at rest. By the isotropy of Euclidian space, $\boldsymbol{E}$ cannot have a preferred direction for a point charge. Thus, we have $\boldsymbol{E} = E(r)\,\hat{\boldsymbol{e}}_r$. Now introducing a "*Gaußian sphere*", of radius r and with outward-pointing unit normal $\boldsymbol{n} = \hat{\boldsymbol{e}}_r$, so that q lies at its centre, we find

$$\int\int_{\text{sphere}}\boldsymbol{E}\cdot\boldsymbol{n}\,\mathrm{d}A = \int_0^{2\pi}\left[\int_0^{\pi}E(r)\,r^2\,\sin\vartheta\,\mathrm{d}\vartheta\right]\,\mathrm{d}\varphi = \frac{q}{\epsilon_0}\,,$$

from which we obtain

$$E(r) = \frac{1}{4\pi\epsilon_0}\,\frac{q}{r^2}\,,$$

which is the result we were looking for.

3.2 Electrostatic scalar potential

We now make the following observation. The rules of vector calculus tell us that the vector factors in the integrands of (3.7) can be expressed as gradients given by

$$\frac{\boldsymbol{r}-\boldsymbol{r}'}{|\boldsymbol{r}-\boldsymbol{r}'|^3} = -\boldsymbol{\nabla}\left(\frac{1}{|\boldsymbol{r}-\boldsymbol{r}'|}\right)\,. \tag{3.9}$$

But since the gradient operation here is with respect to the observation point variable $\boldsymbol{r}$ (and *not* the source integration variable $\boldsymbol{r}'$), the $\boldsymbol{\nabla}$–operator in (3.7)

can be brought outside of the integral signs to obtain

$$\boldsymbol{E}(\boldsymbol{r}) = -\frac{1}{4\pi\epsilon_0}\boldsymbol{\nabla}\int\int\int_G \frac{\rho(\boldsymbol{r}')}{|\boldsymbol{r}-\boldsymbol{r}'|}\,\mathrm{d}V' - \frac{1}{4\pi\epsilon_0}\boldsymbol{\nabla}\int\int_{\partial G}\frac{\sigma(\boldsymbol{r}')}{|\boldsymbol{r}-\boldsymbol{r}'|}\,\mathrm{d}A' \,. \tag{3.10}$$

It is clear at this stage that, by identity (2.25), this expression *does* satisfy the electrostatic condition (3.1), namely $\boldsymbol{\nabla}\times\boldsymbol{E}=\boldsymbol{0}$. This is because we have found that $\boldsymbol{E}$ can be obtained from a, by assumption, continuously differentiable scalar field ϕ as

$$\boldsymbol{E}(\boldsymbol{r}) = -\boldsymbol{\nabla}\phi(\boldsymbol{r}) \,. \tag{3.11}$$

The scalar field $\phi(\boldsymbol{r})$ is called the **electrostatic scalar potential**; it has SI unit $1\ \frac{\mathrm{kg\,m^2}}{\mathrm{s^2 C}}$, also referred to as $1\ \mathrm{V(olt)}$.

Relating (3.11) and (3.10), we thus have

$$\phi(\boldsymbol{r}) = \frac{1}{4\pi\epsilon_0}\int\int\int_G \frac{\rho(\boldsymbol{r}')}{|\boldsymbol{r}-\boldsymbol{r}'|}\,\mathrm{d}V' + \frac{1}{4\pi\epsilon_0}\int\int_{\partial G}\frac{\sigma(\boldsymbol{r}')}{|\boldsymbol{r}-\boldsymbol{r}'|}\,\mathrm{d}A' + \text{const} \,; \tag{3.12}$$

the electrostatic scalar potential is determined only up to an arbitrary additive constant. Physically, however, it is *differences* in the values of ϕ that are of relevance, as we will see below. Note that it is generally simpler, and certainly more economical, to solve just *one* integral expression for ϕ rather than three for the components of $\boldsymbol{E}$. Knowing ϕ for a given charge distribution, $\boldsymbol{E}$ follows from (3.11).

For the system of n point charges q_i ($i=1,\ldots,n$) of (3.5), the electrostatic scalar potential is given by

$$\phi(\boldsymbol{r}) = \frac{1}{4\pi\epsilon_0}\sum_{i=1}^{n}\frac{q_i}{|\boldsymbol{r}-\boldsymbol{r}_i|} + \text{const} \,. \tag{3.13}$$

A physical interpretation of the electrostatic scalar potential arises through the following consideration. Let us calculate the **work** done *against* the Coulomb forces in transporting a test point charge q in a given electrostatic field $\boldsymbol{E}(\boldsymbol{r})$ from a field point A to another field point B along an arbitrary path. The position-dependent Coulomb force felt by q is $\boldsymbol{F}(\boldsymbol{r}) = q\,\boldsymbol{E}(\boldsymbol{r})$. Using this relation, we find that the work done between A and B amounts to

$$W = -\int_A^B \boldsymbol{F}\cdot\mathrm{d}\boldsymbol{s} = -q\int_A^B \boldsymbol{E}\cdot\mathrm{d}\boldsymbol{s} \,. \tag{3.14}$$

But, by (3.11), this can be written as

$$W = q\int_A^B \boldsymbol{\nabla}\phi\cdot\mathrm{d}\boldsymbol{s} = q\int_A^B \mathrm{d}\phi = q\,[\,\phi(\boldsymbol{r}_B)-\phi(\boldsymbol{r}_A)\,] \,. \tag{3.15}$$

From this we see that $q\,\phi$ can be interpreted as the **potential energy** of the test point charge q in the electrostatic field described by ϕ.

It follows directly from (3.14) and (3.15) that when we consider a *closed* path C, i.e., a path such that A and B coincide, the magnitude of the work done per unit charge is

$$\left|\frac{W}{q}\right| = \left|\oint_C \boldsymbol{E}\cdot\mathrm{d}\boldsymbol{s}\right| = 0 = \left|\int\int_S (\boldsymbol{\nabla}\times\boldsymbol{E})\cdot\boldsymbol{n}\,\mathrm{d}A\right| \,, \tag{3.16}$$

i.e., zero; on the right-hand side Stokes' integral theorem (1.33) was used to obtain the flux of $\boldsymbol{\nabla} \times \boldsymbol{E}$ through the surface S spanned by C. We see that electrostatic fields provide examples of so-called **conservative fields**, for which the work done in transports between two field points A and B does *not* depend on the path chosen to get from A to B.

3.3 Poisson's and Laplace's equations

As stated before, (3.1) implies the existence for $\boldsymbol{E}$ of a continuously differentiable electrostatic scalar potential ϕ, so that

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

When we substitute this into (3.2), we obtain

$$(\boldsymbol{\nabla}\cdot\boldsymbol{\nabla})\phi(\boldsymbol{r}) = -\,\frac{\rho(\boldsymbol{r})}{\epsilon_0}\ . \tag{3.17}$$

This constitutes **Poisson's equation** (named after the French mathematician Siméon Denis Poisson, 1781–1840), which is the central field equation to be solved in **electrostatics**. It serves to determine the electrostatic scalar potential that is generated by a continuous charge distribution which is localised in some finite region G of Euclidian space $\mathbb{R}^3$. **Poisson's equation** also plays the central role in Newton's theory of gravitation (cf. the lectures on MAS107 Newtonian Dynamics and Gravitation).

In those regions of $\mathbb{R}^3$ where the electric charge density is zero, (3.17) reduces to **Laplace's equation** (named after the French physicist and mathematician Pierre-Simon Laplace, 1749–1827)

$$(\boldsymbol{\nabla}\cdot\boldsymbol{\nabla})\phi(\boldsymbol{r}) = 0\ . \tag{3.18}$$

Note that **Laplace's equation** also arises in fluid dynamics, namely when one considers irrotational and incompressible fluid flows (cf. the lectures on MAS209 Fluid Dynamics).

We want to emphasise that in order to obtain *unique* solutions to **Poisson's equation** for given $\rho(\boldsymbol{r})$ (or for **Laplace's equation** when $\rho(\boldsymbol{r}) = 0$), we need to supplement it by prescribed **boundary conditions** which we demand $\phi(\boldsymbol{r})$ to satisfy. These typically specify the behaviour of $\phi(\boldsymbol{r})$ either at spatial infinity, or on the bounding surface ∂G of a finite spatial region $G \subset \mathbb{R}^3$ that does or does not contain a charge distribution.

In order to discuss solutions to **Poisson's equation**, let us make the following two observations. Acting with the Laplace operator $(\boldsymbol{\nabla}\cdot\boldsymbol{\nabla})$ on $1/|\boldsymbol{r}-\boldsymbol{r}'|$, we have

$$(\boldsymbol{\nabla}\cdot\boldsymbol{\nabla})\left(\frac{1}{|\boldsymbol{r}-\boldsymbol{r}'|}\right) = \begin{cases} 0 & \text{for } \boldsymbol{r} \neq \boldsymbol{r}' \\ \text{singular} & \text{for } \boldsymbol{r} = \boldsymbol{r}' \end{cases}\ . \tag{3.19}$$

Moreover, by use of Gauß' integral theorem (1.32), we can show that, given a finite spatial region $G \subset \mathbb{R}^3$ with bounding surface ∂G, the volume integral of the expression $(\boldsymbol{\nabla} \cdot \boldsymbol{\nabla})(1/|\boldsymbol{r} - \boldsymbol{r}'|)$ over G yields

$$\int\int\int_G (\boldsymbol{\nabla} \cdot \boldsymbol{\nabla}) \left(\frac{1}{|\boldsymbol{r} - \boldsymbol{r}'|} \right) \mathrm{d}V = \int\int_{\partial G} (\boldsymbol{n} \cdot \boldsymbol{\nabla}) \left(\frac{1}{|\boldsymbol{r} - \boldsymbol{r}'|} \right) \mathrm{d}A = -4\pi \, . \tag{3.20}$$

This result follows, e.g., from evaluating the surface integral in **spherical polar coordinates** for a very small sphere of radius a centred on the origin of the reference frame.

For our further discussion it is now convenient to employ **Dirac's delta function** that we introduced in section 2.5.

3.3.1 Formal solutions

With the help of the properties of **Dirac's delta function**, we can formally summarise our results (3.19) and (3.20) by

$$(\boldsymbol{\nabla} \cdot \boldsymbol{\nabla}) \left(\frac{1}{|\boldsymbol{r} - \boldsymbol{r}'|} \right) = -4\pi \, \delta(\boldsymbol{r} - \boldsymbol{r}') \, . \tag{3.21}$$

Thus, $1/|\boldsymbol{r} - \boldsymbol{r}'|$ can be interpreted as being proportional to the scalar potential of a **unit point charge**. Let us multiply both sides of (3.21) by a charge density $\rho(\boldsymbol{r})$, and then formally integrate over all space, i.e., $\mathbb{R}^3$. We thus obtain

$$\begin{aligned} &\int\int\int_{\mathbb{R}^3} \rho(\boldsymbol{r}')(\boldsymbol{\nabla} \cdot \boldsymbol{\nabla}) \left(\frac{1}{|\boldsymbol{r} - \boldsymbol{r}'|} \right) \mathrm{d}V' \\ &= -4\pi \int\int\int_{\mathbb{R}^3} \rho(\boldsymbol{r}') \, \delta(\boldsymbol{r} - \boldsymbol{r}') \, \mathrm{d}V' \\ &= -4\pi \, \rho(\boldsymbol{r}) \, . \end{aligned} \tag{3.22}$$

Now pulling the $(\boldsymbol{\nabla} \cdot \boldsymbol{\nabla})$–operator out of the integral on the left-hand side (which we are allowed to do, as its action is with respect to the observation point variable $\boldsymbol{r}$, and *not* the source integration variable $\boldsymbol{r}'$), and comparing the result with **Poisson's equation** (3.17), we find for the electrostatic scalar potential the "*infinite space solution*"

$$\phi(\boldsymbol{r}) = \frac{1}{4\pi\epsilon_0} \int\int\int_{\mathbb{R}^3} \frac{\rho(\boldsymbol{r}')}{|\boldsymbol{r} - \boldsymbol{r}'|} \, \mathrm{d}V' + \text{const} \, . \tag{3.23}$$

However, since the volume integral is over the whole of $\mathbb{R}^3$ (i.e., there are *no* boundary surfaces involved), this solution of **Poisson's equation** will rarely be useful in actual applications.

For electrostatic problems involving *finite* regions of $\mathbb{R}^3$, with or without a given charge distribution contained inside, and prescribed boundary conditions on the bounding surfaces, **Poisson's equation** (3.17) can be converted into an *integral equation* by means of the Green's second integral formula (2.33). Consider a volume $G \subset \mathbb{R}^3$ with bounding surface ∂G. Choosing in (2.33) $\psi = 1/|\boldsymbol{r} - \boldsymbol{r}'|$, and taking ϕ as the electrostatic scalar potential that satisfies

(3.17), we obtain with (3.21), for the case that the observation point $\boldsymbol{r}$ is *within* G,

$$\begin{aligned}\phi(\boldsymbol{r}) &= \frac{1}{4\pi\epsilon_0}\int\int\int_G \frac{\rho(\boldsymbol{r}')}{|\boldsymbol{r}-\boldsymbol{r}'|}\,\mathrm{d}V' \\ &\quad + \frac{1}{4\pi}\int\int_{\partial G}\frac{1}{|\boldsymbol{r}-\boldsymbol{r}'|}\,(\boldsymbol{n}'\cdot\boldsymbol{\nabla}_{\boldsymbol{r}'})\phi(\boldsymbol{r}')\,\mathrm{d}A' \\ &\quad - \frac{1}{4\pi}\int\int_{\partial G}\phi(\boldsymbol{r}')\,(\boldsymbol{n}'\cdot\boldsymbol{\nabla}_{\boldsymbol{r}'})\left(\frac{1}{|\boldsymbol{r}-\boldsymbol{r}'|}\right)\mathrm{d}A'\,. \quad (3.24)\end{aligned}$$

When $\boldsymbol{r}$ is *not* within G, the left-hand side of (3.24) is zero.

The result (3.24) does *not* yet provide a *solution* to **Poisson's equation** (3.17) (for a given charge distribution ρ), as the surface integrals on the right-hand side still contain the (so far) *unknown* values on ∂G of *both* ϕ itself and its normal derivative $(\boldsymbol{n}\cdot\boldsymbol{\nabla})\phi = -E_{\boldsymbol{n}}$. So we have to find ways of obtaining additional information on what values these quantities can possibly take. In any case, when ∂G is pushed out to spatial infinity and we assume that the normal component of the electric field strength [$E_{\boldsymbol{n}} = -(\boldsymbol{n}\cdot\boldsymbol{\nabla})\phi$] on ∂G falls off faster than $1/|\boldsymbol{r}-\boldsymbol{r}'|$, the expression (3.24) reduces to the "infinite space solution" (3.23).

3.3.2 Boundary conditions and uniqueness of solutions

One could think that if values on ∂G were prescribed for *both* ϕ and its normal derivative $(\boldsymbol{n}\cdot\boldsymbol{\nabla})\phi = -E_{\boldsymbol{n}}$ (in addition to giving $\rho(\boldsymbol{r})$ on G), then (3.24) could lead to unique solutions of **Poisson's equation**. However, one can show that the prescription of such so-called **Cauchy boundary conditions** (named after the French mathematician Augustin Louis Cauchy, 1789–1857) leads to inconsistencies and, hence, *not* to any solutions. Instead one finds that one has the freedom to specify only one *or* the other. This then leads to the definition of the following two alternative sets of boundary conditions:

(a) **Dirichlet boundary conditions** (named after the German mathematician Johann Peter Gustav Dirichlet, 1805–1859): we prescribe on ∂G the value of the electrostatic scalar potential $\phi(\boldsymbol{r})$ itself,

(b) **von Neumann boundary conditions** (named after the Hungarian-American mathematician John von Neumann, 1903–1957): we prescribe on ∂G the value of the normal derivative of $\phi(\boldsymbol{r})$, $(\boldsymbol{n}\cdot\boldsymbol{\nabla})\phi(\boldsymbol{r})$, which is (up to a sign) the normal component of the electric field strength, $E_{\boldsymbol{n}} := \boldsymbol{n}\cdot\boldsymbol{E}$.

That either choice of boundary conditions on ∂G leads to *unique* solutions of **Poisson's equation** (3.17) in G can be shown as follows.

Suppose that in G there exist two solutions, ϕ_1 and ϕ_2, to (3.17) satisfying the *same* boundary conditions on ∂G. Defining $U := \phi_2 - \phi_1$, we have

$$(\boldsymbol{\nabla}\cdot\boldsymbol{\nabla})U = 0 \quad \text{in} \quad G\,, \quad \text{and} \qquad (3.25)$$

$$U = 0 \quad \textit{or} \quad (\boldsymbol{n}\cdot\boldsymbol{\nabla})U = 0 \quad \text{on} \quad \partial G\,. \qquad (3.26)$$

Then, by the Green's first integral formula (2.32) with $\phi = \psi = U$, we have

$$\int\int\int_G U\,(\boldsymbol{\nabla}\cdot\boldsymbol{\nabla})U\,\mathrm{d}V = \int\int_{\partial G} U\,(\boldsymbol{n}\cdot\boldsymbol{\nabla})U\,\mathrm{d}A - \int\int\int_G |\,\boldsymbol{\nabla}U\,|^2\,\mathrm{d}V\ ,$$

and so

$$\int\int\int_G |\,\boldsymbol{\nabla}U\,|^2\,\mathrm{d}V = 0\ , \tag{3.27}$$

which implies $\boldsymbol{\nabla}U = 0$ in G, i.e., $U = \text{const}$ in G. Thus,

(a) for Dirichlet boundary conditions, where $U = 0$ on ∂G, $U = \text{const} = 0 \Rightarrow \phi_2 = \phi_1$ in G,

(b) for von Neumann boundary conditions, where $(\boldsymbol{n}\cdot\boldsymbol{\nabla})U = 0$ on ∂G, $U = \text{const} \Rightarrow \phi_2 = \phi_1 + \text{const}$ in G.

Hence, *both* kinds of boundary conditions lead to unique solutions of (3.17) in G (in the second case up to an additive constant).

3.3.3 Boundary value problems and Green's functions

The function $1/|\boldsymbol{r}-\boldsymbol{r}'|$ is a special member of a whole class of functions $G(\boldsymbol{r},\boldsymbol{r}')$ that satisfy in a finite spatial region $G \subset \mathbb{R}^3$ with bounding surface ∂G the equation

$$(\boldsymbol{\nabla}\cdot\boldsymbol{\nabla})G(\boldsymbol{r},\boldsymbol{r}') = -\,4\pi\,\delta(\boldsymbol{r}-\boldsymbol{r}')\ . \tag{3.28}$$

Functions in this class are called **Green's functions** and have the form

$$G(\boldsymbol{r},\boldsymbol{r}') = \frac{1}{|\boldsymbol{r}-\boldsymbol{r}'|} + F(\boldsymbol{r},\boldsymbol{r}')\ , \tag{3.29}$$

where $F(\boldsymbol{r},\boldsymbol{r}')$ satisfies in G **Laplace's equation**,

$$(\boldsymbol{\nabla}\cdot\boldsymbol{\nabla})F(\boldsymbol{r},\boldsymbol{r}') = 0\ . \tag{3.30}$$

To obtain *unique* solutions to **Poisson's equation** (3.17) in G, with either Dirichlet or von Neumann boundary conditions on ∂G, we can now use Green's second integral formula (2.33) with $\psi = G(\boldsymbol{r},\boldsymbol{r}')$ and use the *new freedom* given through $F(\boldsymbol{r},\boldsymbol{r}')$ to set one of the two surface integrals that so arise to zero, or at least we can try to simplify these surface integrals according to our needs.

Equation (2.33) with $\psi = G(\boldsymbol{r},\boldsymbol{r}')$ and (3.28) yields for the case that the observation point $\boldsymbol{r}$ is *within* G

$$\begin{aligned}\phi(\boldsymbol{r}) \;=\; & \frac{1}{4\pi\epsilon_0}\int\int\int_G \rho(\boldsymbol{r}')\,G(\boldsymbol{r},\boldsymbol{r}')\,\mathrm{d}V' \\ & + \frac{1}{4\pi}\int\int_{\partial G} G(\boldsymbol{r},\boldsymbol{r}')\,(\boldsymbol{n}'\cdot\boldsymbol{\nabla}_{\boldsymbol{r}'})\phi(\boldsymbol{r}')\,\mathrm{d}A' \\ & - \frac{1}{4\pi}\int\int_{\partial G} \phi(\boldsymbol{r}')\,(\boldsymbol{n}'\cdot\boldsymbol{\nabla}_{\boldsymbol{r}'})G(\boldsymbol{r},\boldsymbol{r}')\ \mathrm{d}A'\ ,\end{aligned} \tag{3.31}$$

which generalises the integral equation (3.24). The freedom available in $G(\boldsymbol{r},\boldsymbol{r}')$ through $F(\boldsymbol{r},\boldsymbol{r}')$ can now be used to our advantage as follows:

(a) for **Dirichlet boundary conditions** the simplest choice of **Green's function** is such that

$$G_{\mathrm{D}}(\boldsymbol{r},\boldsymbol{r}') = 0 \quad \text{for } \boldsymbol{r}' \in \partial G \; . \tag{3.32}$$

Then, having prescribed the value of $\phi(\boldsymbol{r})$ on ∂G, we get from (3.31) in G the **solution**

$$\begin{aligned} \phi(\boldsymbol{r}) &= \frac{1}{4\pi\epsilon_0}\int\int\int_G \rho(\boldsymbol{r}')\, G_{\mathrm{D}}(\boldsymbol{r},\boldsymbol{r}')\,\mathrm{d}V' \\ &\quad - \frac{1}{4\pi}\int\int_{\partial G} \phi(\boldsymbol{r}')\,(\boldsymbol{n}'\cdot\boldsymbol{\nabla}_{\boldsymbol{r}'})G_{\mathrm{D}}(\boldsymbol{r},\boldsymbol{r}')\,\mathrm{d}A' \; . \end{aligned} \tag{3.33}$$

(b) for **von Neumann boundary conditions** the simplest choice of **Green's function** is such that

$$(\boldsymbol{n}'\cdot\boldsymbol{\nabla}_{\boldsymbol{r}'})G_{\mathrm{N}}(\boldsymbol{r},\boldsymbol{r}') = -\frac{4\pi}{S} \quad \text{for } \boldsymbol{r}' \in \partial G \; , \tag{3.34}$$

with S the total area of the bounding surface ∂G, since by (3.28) $\int\int_{\partial G}(\boldsymbol{n}\cdot\boldsymbol{\nabla})G\,\mathrm{d}A = -\,4\pi$. Then, having prescribed the value of $(\boldsymbol{n}\cdot\boldsymbol{\nabla})\phi(\boldsymbol{r})$ on ∂G, we get from (3.31) in G the **solution**

$$\begin{aligned} \phi(\boldsymbol{r}) &= \langle\phi\rangle_{\partial G} + \frac{1}{4\pi\epsilon_0}\int\int\int_G \rho(\boldsymbol{r}')\, G_{\mathrm{N}}(\boldsymbol{r},\boldsymbol{r}')\,\mathrm{d}V' \\ &\quad + \frac{1}{4\pi}\int\int_{\partial G} G_{\mathrm{N}}(\boldsymbol{r},\boldsymbol{r}')\,(\boldsymbol{n}'\cdot\boldsymbol{\nabla}_{\boldsymbol{r}'})\phi(\boldsymbol{r}')\,\mathrm{d}A' \; , \end{aligned} \tag{3.35}$$

where $\langle\phi\rangle_{\partial G}$ denotes the average value of ϕ over the whole of ∂G.

The symmetry property $G(\boldsymbol{r},\boldsymbol{r}') = G(\boldsymbol{r}',\boldsymbol{r})$ is automatic for $G_{\mathrm{D}}(\boldsymbol{r},\boldsymbol{r}')$; for $G_{\mathrm{N}}(\boldsymbol{r},\boldsymbol{r}')$ it can be imposed as an additional requirement.

As regards applications of the present **Green's function formalism** to boundary value problems for **Poisson equation** (3.17), we should note that both $G_{\mathrm{D}}(\boldsymbol{r},\boldsymbol{r}')$ and $G_{\mathrm{N}}(\boldsymbol{r},\boldsymbol{r}')$ are explicitly known only in a few special cases. Typically they are *not* expressible in terms of elementary functions (if they can be found at all) as they very much depend on the shape of the bounding surface ∂G that is involved in a specific boundary value problem. Hence, it is numerical approximations to $G_{\mathrm{D}}(\boldsymbol{r},\boldsymbol{r}')$ and $G_{\mathrm{N}}(\boldsymbol{r},\boldsymbol{r}')$ [and so to $\phi(\boldsymbol{r})$] that are at the heart of actual applications.

3.4 Further remark on boundary conditions

As eluded to in subsection 3.3.2, to obtain unique solutions, the electrostatic field equations (3.2) and (3.1) need to be supplemented by appropriate boundary conditions. Starting from their respective integral forms (3.8) and (the left-hand side of) (3.16), by introducing a small "*Gaußian pillbox*" and a small "*Stokesian loop*" that both penetrate an interface which separates a region 1 of

space from a region 2 and which carries an idealised **electric surface charge density** σ, one can show that these boundary conditions must take the form

$$\boldsymbol{n} \cdot (\boldsymbol{E}_2 - \boldsymbol{E}_1) = \frac{\sigma}{\epsilon_0} \qquad \boldsymbol{n} \times (\boldsymbol{E}_2 - \boldsymbol{E}_1) = \boldsymbol{0} \,. \tag{3.36}$$

Here $\boldsymbol{n}$ denotes the unit normal to the interface pointing from region 1 into region 2. The boundary conditions thus state that, at an interface that separates two regions of space, the **normal component** of the electrostatic field strength may be *discontinuous* while the **tangential components** must be *continuous*.

One finds that the *same* boundary conditions also apply to *time-dependent fields* which interact with continuous material media of *linear* electric properties [and so are described by (1.15)–(1.18)] for which $\boldsymbol{P} = \boldsymbol{0}$.

3.5 Energy density of electrostatic fields

In this section we want to discuss the notion of **energy** in the context of electrostatic fields. We already know from (3.15) that it takes the **work** $W = q\,[\,\phi(\boldsymbol{r}_B) - \phi(\boldsymbol{r}_A)\,]$ against the Coulomb forces to transport a test point charge q in a given electrostatic field with scalar potential $\phi(\boldsymbol{r})$ from position $\boldsymbol{r}_A$ to $\boldsymbol{r}_B$. We interpreted this work as the difference in the potential energy of q at these two positions. So when we assume that $\phi(\boldsymbol{r})$ satisfies the boundary condition

$$\lim_{|\,\boldsymbol{r}\,|\to\infty} \phi(\boldsymbol{r}) = 0 \,, \tag{3.37}$$

and we imagine to bring in q from "infinity" to position $\boldsymbol{r}$, this work becomes simply $W = q\,\phi(\boldsymbol{r})$.

Now let us assume that we assemble n **point charges** q_i by successively bringing them in from "infinity" to positions $\boldsymbol{r}_i$. We want to know the value of the **total work** required by this process. To move q_1 to position $\boldsymbol{r}_1$ does not require any work at all, because, by assumption, there was no field present initially, i.e., initially $\phi(\boldsymbol{r}) = 0$. Then, in the presence of q_1, the scalar potential at position $\boldsymbol{r}$ is

$$\phi(\boldsymbol{r}) = \frac{1}{4\pi\epsilon_0}\,\frac{q_1}{|\boldsymbol{r} - \boldsymbol{r}_1|} \,.$$

Hence, the work to be done to bring in q_2 to position $\boldsymbol{r}_2$ is

$$W = q_2\,\frac{1}{4\pi\epsilon_0}\,\frac{q_1}{|\boldsymbol{r}_2 - \boldsymbol{r}_1|} \,,$$

while, since the superposition principle holds, the scalar potential in the presence of both q_1 and q_2 at position $\boldsymbol{r}$ is

$$\phi(\boldsymbol{r}) = \frac{1}{4\pi\epsilon_0}\,\frac{q_1}{|\boldsymbol{r} - \boldsymbol{r}_1|} + \frac{1}{4\pi\epsilon_0}\,\frac{q_2}{|\boldsymbol{r} - \boldsymbol{r}_2|} \,.$$

When now q_3 is added to the configuration by bringing it in from "infinity" to position $\boldsymbol{r}_3$, the total work done to assemble these three point charges amounts to

$$W = q_2\,\frac{1}{4\pi\epsilon_0}\,\frac{q_1}{|\boldsymbol{r}_2 - \boldsymbol{r}_1|} + q_3\,\frac{1}{4\pi\epsilon_0}\,\frac{q_1}{|\boldsymbol{r}_3 - \boldsymbol{r}_1|} + q_3\,\frac{1}{4\pi\epsilon_0}\,\frac{q_2}{|\boldsymbol{r}_3 - \boldsymbol{r}_2|} \,.$$

It thus follows that in order to assemble all n point charges q_i (by moving them from "infinity" to positions $\boldsymbol{r}_i$, respectively), the total work to be done is expressed by

$$W = \frac{1}{8\pi\epsilon_0} \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{q_i\, q_j}{|\boldsymbol{r}_i - \boldsymbol{r}_j|} \qquad i \neq j \; ; \tag{3.38}$$

the extra factor of $1/2$ is needed since by symmetry of the double sum each point charge is counted twice. This expression accounts only for the interaction potential energy between the point charges (and so, depending on their signs, can be negative), and not for any infinite "self-energy" contributions that would arise when $i = j$. We may also write (3.38) as

$$W = \frac{1}{2} \sum_{i=1}^{n} q_i\, \phi_i \; , \tag{3.39}$$

with ϕ_i the electrostatic potential felt by q_i at $\boldsymbol{r}_i$ due to the remaining $n-1$ point charges.

Now let us turn to consider a **continuous charge distribution** with compact support in $\mathbb{R}^3$ (i.e., $\rho(\boldsymbol{r})$ goes to zero as $|\boldsymbol{r}| \to \infty$). Again, we want to assume that the resultant electrostatic scalar potential satisfies the boundary condition (3.37). Formally the total work to be done to assemble the continuous charge distribution is obtained from (3.38) by making the replacements

$$q_i \longrightarrow \rho(\boldsymbol{r})\, \mathrm{d}V \qquad q_j \longrightarrow \rho(\boldsymbol{r}')\, \mathrm{d}V' \; ,$$

and converting the sums into integrals over all of $\mathbb{R}^3$ according to

$$\sum_{i=1}^{n} \longrightarrow \int\int\int_{\mathbb{R}^3} \; .$$

We thus get

$$W = \frac{1}{8\pi\epsilon_0} \int\int\int_{\mathbb{R}^3} \int\int\int_{\mathbb{R}^3} \frac{\rho(\boldsymbol{r})\, \rho(\boldsymbol{r}')}{|\boldsymbol{r} - \boldsymbol{r}'|}\, \mathrm{d}V\, \mathrm{d}V' \; ,$$

an expression which *does* contain "self-energy" contributions and therefore will be found below to be positive definite. (The apparent failure of this expression due to the above-mentioned "infinities" is resolved at the **microscopic level** by the existence of a fundamental **quantum of electric charge**.) Using the "infinite space solution" (3.23) for $\phi(\boldsymbol{r})$ [subject to the boundary condition (3.37)], W can be rewritten as

$$W = \tfrac{1}{2} \int\int\int_{\mathbb{R}^3} \rho(\boldsymbol{r})\, \phi(\boldsymbol{r})\, \mathrm{d}V \; . \tag{3.40}$$

This result expresses the **electrostatic potential energy** in terms of the Coulomb interactions within the continuous charge distribution (as does (3.38) for the case of discrete charges).

An alternative viewpoint sees the potential energy as being stored in the electrostatic field surrounding the charge distribution. Here we substitute for $\rho(\boldsymbol{r})$ in (3.40) from **Poisson's equation** (3.17) to obtain

$$W = -\frac{\epsilon_0}{2} \int\int\int_{\mathbb{R}^3} \phi(\boldsymbol{r})\, (\boldsymbol{\nabla}\cdot\boldsymbol{\nabla})\phi(\boldsymbol{r})\, \mathrm{d}V \ .$$

Green's first integral formula (2.32) then states that this is equal to

$$W = \frac{\epsilon_0}{2} \int\int\int_{\mathbb{R}^3} |\, \boldsymbol{\nabla}\phi(\boldsymbol{r})\,|^2\, \mathrm{d}V + \frac{\epsilon_0}{2} \int\int_{\partial(\mathbb{R}^3)} \phi(\boldsymbol{r})\, (\boldsymbol{n}\cdot\boldsymbol{\nabla})\phi(\boldsymbol{r})\, \mathrm{d}A \ ,$$

where, in view of the assumed boundary condition (3.37) for $\phi(\boldsymbol{r})$, the value of the surface integral is zero. Using (3.11) to express $\boldsymbol{\nabla}\phi(\boldsymbol{r})$, we finally obtain for the electrostatic potential energy the result

$$W = \frac{\epsilon_0}{2} \int\int\int_{\mathbb{R}^3} |\, \boldsymbol{E}(\boldsymbol{r})\,|^2\, \mathrm{d}V \ . \tag{3.41}$$

This says that the electrostatic potential energy is entirely carried by the electrostatic field $\boldsymbol{E}(\boldsymbol{r})$, with a (positive definite) energy density in space given by

$$u(\boldsymbol{r}) = \frac{\epsilon_0}{2}\, |\, \boldsymbol{E}(\boldsymbol{r})\,|^2 \ . \tag{3.42}$$

[Compare this expression to the electrostatic subcase of (1.39), i.e., $u = (\boldsymbol{E}/c\cdot\boldsymbol{E}/c)/(2\mu_0)$; it is clearly the same.] Thus, as a carrier of energy, clearly *physical reality* must be ascribed to an electrostatic field (and all electromagnetic fields in general).

3.6 Method of images

Boundary value problems in **electrostatics** that involve **point charges** in the presence of **boundary surfaces**, as, e.g., provided by conductors at either zero or non-zero *fixed* potential, can sometimes be solved without explicit recourse to Green's function method of subsection 3.3.3 for obtaining unique solutions to **Poisson's equation**. Indeed, when the geometry of a boundary surface is sufficiently simple, the related boundary value problem can be treated as though the boundary was not present at all by *simulating* the boundary conditions it imposes by a small number of (point) charges of appropriate magnitude, suitably placed at positions *external* to the region of interest. The distribution of these so-called **image charges** must be of such a nature that their individual electrostatic scalar potentials add up of to form an equipotential surface of the same shape and of the same value as the original boundary surface. The (imagined) replacement of the original boundary value problem by a larger field region with image charges but *no* boundaries is called the **method of images**.

3.6.1 Point charge near a grounded conducting plane

To illustrates this technique, let us discuss as a first example the boundary value problem for a point charge at rest in front of an infinitely extended, *grounded*, conducting plane. Assuming the conducting plane to be grounded provides the boundary condition

$$\phi(\boldsymbol{r})|_{\text{on plane}} = 0 \,, \tag{3.43}$$

i.e., the electrostatic scalar potential has to vanish when the observation point $\boldsymbol{r}$ is located *on* the surface of the conducting plane.

Let us suppose that the point charge is q_+, and that its location in a reference frame with **Cartesian coordinates** is given by the fixed position $\boldsymbol{a} = a\,\boldsymbol{e}_z$ normal to the conducting plane. (For convenience we place the origin of our reference frame at the base of $\boldsymbol{a}$ so that the plane is given by $z = 0$.) Now because the plane in question is conducting, q_+ will induce a total charge $q_- = -\,q_+$ on its surface, of a position-dependent surface charge density σ. It is clear that by the planar symmetry of the present configuration the image charge q_- must be located at the position $-\,a\,\boldsymbol{e}_z$ *outside* the field region $z \geq 0$ for which we want to determine $\phi(\boldsymbol{r})$.

By the superposition principle, the electrostatic scalar potential generated by q_+ and q_- must be of the form

$$\phi(\boldsymbol{r}) = \phi_{q_+}(\boldsymbol{r}) + \phi_{q_-}(\boldsymbol{r}) \,.$$

With $q_- = -\,q_+$, and using (3.13) with $n = 2$ and const $= 0$, it is thus given by

$$\phi(\boldsymbol{r}) = \frac{q_+}{4\pi\epsilon_0}\left[\frac{1}{|\,\boldsymbol{r}-\boldsymbol{a}\,|} - \frac{1}{|\,\boldsymbol{r}+\boldsymbol{a}\,|}\right] , \tag{3.44}$$

which is valid for $z \geq 0$. Evaluating the moduli terms in the denominators, and recalling that $(\boldsymbol{r}\cdot\boldsymbol{a}) = ra\cos\gamma$, we rewrite (3.44) as

$$\phi(\boldsymbol{r}) = \frac{q_+}{4\pi\epsilon_0}\left[\frac{1}{[\,r^2+a^2-2(\boldsymbol{r}\cdot\boldsymbol{a})\,]^{1/2}} - \frac{1}{[\,r^2+a^2+2(\boldsymbol{r}\cdot\boldsymbol{a})\,]^{1/2}}\right] . \tag{3.45}$$

This already gives what we were looking for: the solution to our boundary value problem in the field region $z \geq 0$. We clearly see that we have $\phi(\boldsymbol{r}) = 0$ when $z = 0 \Rightarrow (\boldsymbol{r}\cdot\boldsymbol{a}) = 0$ (because $\gamma = \pm\,\pi/2$), i.e., when the observation point $\boldsymbol{r}$ is located *on* the surface of the conducting plane.

Proceeding further, we obtain with (3.11) from (3.44) the expression

$$\boldsymbol{E}(\boldsymbol{r}) = \frac{q_+}{4\pi\epsilon_0}\left[\frac{1}{|\,\boldsymbol{r}-\boldsymbol{a}\,|^2}\,\frac{\boldsymbol{r}-\boldsymbol{a}}{|\,\boldsymbol{r}-\boldsymbol{a}\,|} - \frac{1}{|\,\boldsymbol{r}+\boldsymbol{a}\,|^2}\,\frac{\boldsymbol{r}+\boldsymbol{a}}{|\,\boldsymbol{r}+\boldsymbol{a}\,|}\right] . \tag{3.46}$$

for the electric field strength in the observation point $\boldsymbol{r}$.

The last result can be used to address the particularly interesting question: how does the surface charge density σ induced by q_+ vary on the conducting plane? For this purpose let us imagine a rectangular "*Gaußian box*" (with a surface ∂G enclosing a small volume G) which cuts through the conducting plane. By Gauß' law (3.8), the total flux of the electric field strength $\boldsymbol{E}$ (as

generated by q_+ and q_-) through the surface of the box is proportional to the total electric charge $Q(G)$ contained inside. But in the present case $Q(G)$ is given by $Q(G) = \int\int_{\text{area}} \sigma \, \mathrm{d}A$, where "area" stands for the crossectional area A cut out by the box from the plane. By the geometry of the box, A is also the area of the two faces that are *parallel* to the plane. Hence, from (3.8), we thus find

$$(\boldsymbol{n} \cdot \boldsymbol{E})\, A = \frac{\sigma}{\epsilon_0}\, A \, ;$$

because of the planar symmetry of the configuration, the electric field strength given by (3.46) is *constant* on the front face of the box; on the back face (in the external region $z < 0$ behind the plane) $\boldsymbol{E} = \boldsymbol{0}$; and there are *no* contributions from the side faces since there $(\boldsymbol{n} \cdot \boldsymbol{E}) = 0$. Now evaluating this result *on* the surface of the plane ($z = 0 \Rightarrow (\boldsymbol{r} \cdot \boldsymbol{a}) = 0$), i.e., in the limit of a "*Gaußian box*" of zero height, we find with (3.46) and $\boldsymbol{n} = \boldsymbol{e}_z$

$$(\boldsymbol{n} \cdot \boldsymbol{E})|_{z=0} = -\,\frac{1}{4\pi\epsilon_0}\,\frac{2q_+\, a}{(r^2 + a^2)^{3/2}} = \frac{\sigma}{\epsilon_0} \, ,$$

so that the surface charge density on the conducting plane as a function of $r = |\,\boldsymbol{r}\,|$ is given by

$$\sigma(r) = \epsilon_0\, (\boldsymbol{n} \cdot \boldsymbol{E})|_{z=0} = -\,\frac{1}{2\pi}\,\frac{q_+\, a}{(r^2 + a^2)^{3/2}} \, .$$

With this result, using plane polar coordinates $\{r, \varphi\}$, we can check (as must be the case) that the total charge induced on the infinitely extended conducting plane amounts to

$$\int\int_{\text{whole plane}} \sigma(r) \, \mathrm{d}A = -\,\frac{q_+\, a}{2\pi} \int_0^{2\pi} \int_0^{\infty} \frac{r \, \mathrm{d}r \, \mathrm{d}\varphi}{(r^2 + a^2)^{3/2}} = -\, q_+ = q_- \, .$$

Finally, the Coulomb force felt by the point charge q_+ at position $a\, \boldsymbol{e}_z$ due to the induced image charge q_- at position $-\, a\, \boldsymbol{e}_z$ is simply given, from (3.3), by

$$\boldsymbol{F} = -\,\frac{1}{4\pi\epsilon_0}\,\frac{q_+^2}{(2a)^2}\, \boldsymbol{e}_z \, .$$

3.6.2 Point charge near a grounded conducting sphere

A second example where the method of images can be successfully applied is the case of a point charge q at rest *outside* a grounded conducting sphere. Let us assume that the sphere has radius a and that it is centred on the origin of a reference frame in $\mathbb{R}^3$ with **spherical polar coordinates**, $\{r, \vartheta, \varphi\}$. The position of q shall be $\boldsymbol{r}_1$, with $r_1 \geq a$. As before, assuming that the conducting surface in our problem is grounded imposes a specific boundary condition on the electrostatic scalar potential, namely

$$\phi(\boldsymbol{r})|_{r=a} = 0 \, . \tag{3.47}$$

To solve the boundary value problem for $\phi(\boldsymbol{r})$ in the field region $r \geq a$, we place an image charge q' at position $\boldsymbol{r}_1'$ *inside* the sphere, which, by the symmetry of the configuration, must be located on the straight line connecting the

origin of the reference frame to q. We determine the exact values for q' and r_1' by applying (3.47) to $\phi(\boldsymbol{r})$.

By the superposition principle, $\phi(\boldsymbol{r})$ has the form

$$\phi(\boldsymbol{r}) = \phi_q(\boldsymbol{r}) + \phi_{q'}(\boldsymbol{r}) \,,$$

which, by (3.13) with $n = 2$ and const $= 0$, is

$$\phi(\boldsymbol{r}) = \frac{1}{4\pi\epsilon_0} \left[\frac{q}{|\,\boldsymbol{r} - \boldsymbol{r}_1\,|} + \frac{q'}{|\,\boldsymbol{r} - \boldsymbol{r}_1'\,|} \right] .$$

Now introducing unit vectors $\hat{\boldsymbol{r}}$ and $\hat{\boldsymbol{r}}_1$ by writing $\boldsymbol{r} = r\,\hat{\boldsymbol{r}}$, $\boldsymbol{r}_1 = r_1\,\hat{\boldsymbol{r}}_1$, and $\boldsymbol{r}_1' = r_1'\,\hat{\boldsymbol{r}}_1$, we get

$$\phi(\boldsymbol{r}) = \frac{1}{4\pi\epsilon_0} \left[\frac{q}{r\,|\,\hat{\boldsymbol{r}} - (r_1/r)\,\hat{\boldsymbol{r}}_1\,|} + \frac{q'}{r_1'\,|\,(r/r_1')\,\hat{\boldsymbol{r}} - \hat{\boldsymbol{r}}_1\,|} \right]$$

Evaluating this solution Ansatz for $\phi(\boldsymbol{r})$ at $r = a$, we find

$$\phi(\boldsymbol{r})\,|_{r=a} = \frac{1}{4\pi\epsilon_0} \left[\frac{q}{a\,|\,\hat{\boldsymbol{r}} - (r_1/a)\,\hat{\boldsymbol{r}}_1\,|} + \frac{q'}{r_1'\,|\,(a/r_1')\,\hat{\boldsymbol{r}} - \hat{\boldsymbol{r}}_1\,|} \right] .$$

It follows that in order to satisfy the boundary condition (3.47), it is required to choose q' and r_1' as

$$q' = -\,\frac{r_1'}{a}\,q \qquad r_1' = \frac{a^2}{r_1} \qquad \Rightarrow \qquad q' = -\,\frac{a}{r_1}\,q \,. \tag{3.48}$$

Note that these relations state that as the point charge q is brought *closer* to the surface of the sphere ($r_1 \to a$ from above), the image charge q' *grows* in magnitude and, at the same time, *moves out* from the centre of the sphere.

With (3.48), our solution for the electrostatic scalar potential which is valid in the field region $r \geq 0$ thus reads

$$\phi(\boldsymbol{r}) = \frac{q}{4\pi\epsilon_0} \left[\frac{1}{|\,r\,\hat{\boldsymbol{r}} - r_1\,\hat{\boldsymbol{r}}_1\,|} - \frac{1}{a\,|\,(rr_1/a^2)\,\hat{\boldsymbol{r}} - \hat{\boldsymbol{r}}_1\,|} \right] ,$$

or, alternatively,

$$\phi(\boldsymbol{r}) = \frac{(q/4\pi\epsilon_0)}{|\,\boldsymbol{r} - \boldsymbol{r}_1\,|} - \frac{a\,(q/4\pi\epsilon_0)}{r_1\,|\,\boldsymbol{r} - (a/r_1)^2\,\boldsymbol{r}_1\,|} \,. \tag{3.49}$$

Expanding the moduli terms in the denominators by using $(\boldsymbol{r}\cdot\boldsymbol{r}_1) = rr_1\cos\gamma$, our result can be rewritten as

$$\phi(\boldsymbol{r}) = \frac{(q/4\pi\epsilon_0)}{\left(r^2 + r_1^2 - 2rr_1\cos\gamma\right)^{1/2}} - \frac{(q/4\pi\epsilon_0)}{\left(r^2r_1^2/a^2 + a^2 - 2rr_1\cos\gamma\right)^{1/2}} \,. \tag{3.50}$$

This form shows most clearly that our solution does indeed satisfy the boundary condition (3.47). Note that $\phi = 0$ also holds when $r_1 = a$, i.e., when the point charge q is placed on the surface of the grounded sphere.

As in the previous example, the presence of the point charge q induces a position-dependent surface charge density σ on the conducting sphere, the net effect of which is perceived by q as originating from the single point charge q'. Again, by Gauß' law (in differential form), this surface charge density can be obtained by evaluating the normal derivative of the electrostatic scalar potential on the surface of the sphere according to

$$\sigma = \epsilon_0 \, (\hat{\boldsymbol{r}} \cdot \boldsymbol{E})|_{r=a} = -\,\epsilon_0 \, (\hat{\boldsymbol{r}} \cdot \boldsymbol{\nabla}\phi)|_{r=a} = -\,\epsilon_0 \left.\frac{\partial \phi}{\partial r}\right|_{r=a} .$$

Applying this relation to (3.50) yields

$$\begin{aligned} \sigma(r_1,\gamma) &= -\,\frac{q}{4\pi\, a^2}\left(\frac{a}{r_1}\right)\frac{1-(a/r_1)^2}{\left(1+(a/r_1)^2-2(a/r_1)\cos\gamma\right)^{3/2}} \\ &= -\,\frac{q}{4\pi}\,\frac{(r_1^2-a^2)}{a\left(r_1^2+a^2-2r_1 a\cos\gamma\right)^{3/2}} . \end{aligned} \tag{3.51}$$

The induced charge q' is distributed over the surface of the sphere in a way [given by (3.51)] that ensures mechanical equilibrium amongst all electrostatic Coulomb forces that are active. Without going into details, we remark that integration of σ over the whole surface of the sphere gives

$$\int\!\!\int_{\text{sphere surface}} \sigma \, \mathrm{d}A = \int_0^{2\pi}\!\!\int_0^{\pi} \sigma(\vartheta,\varphi)\, a^2 \, \mathrm{d}\vartheta \, \mathrm{d}\varphi = -\,\frac{a}{r_1}\, q = q' ,$$

as must be the case.

We conclude this example by calculating the Coulomb force that the point charge q experiences due to the image charge q' which is induced on the surface of the sphere. This can be obtained directly from (3.3). Using

$$q_1 = q \qquad q_2 = -\,\frac{a}{r_1}\, q \qquad \boldsymbol{r}_1 - \boldsymbol{r}_2 = r_1\left(1-\frac{a^2}{r_1^2}\right)\hat{\boldsymbol{r}}_1 ,$$

we thus find

$$\boldsymbol{F} = -\,\frac{1}{4\pi\epsilon_0}\,\frac{q^2}{a^2}\left(\frac{a}{r_1}\right)^3\left(1-\frac{a^2}{r_1^2}\right)^{-2}\hat{\boldsymbol{r}}_1 .$$

We observe that for large separations between the point charge q and the sphere ($r_1 \gg a$) the magnitude of the Coloumb force effectively obeys an inverse cube law, $|\boldsymbol{F}| \propto r_1^{-3}$, while for small separations ($r_1 \approx a$) this magnitude assumes a (more familiar) inverse square law behaviour, $|\boldsymbol{F}| \propto r_1^{-2}(1-a^2/r_1^2)^{-2}$.

3.6.3 Point charge near a charged, insulated, conducting sphere

Suppose we start from the configuration of the previous example. However, we now want to assume that we first insulate the conducting sphere (of radius a, and with induced charge q'), and subsequently add to it a charge $(Q-q')$ to make the total charge residing on its surface equal to Q. As the image charge

q' at position $\boldsymbol{r}'_1$ already balances the electrostatic Coulomb forces due to the point charge q at position $\boldsymbol{r}_1$, the charge $(Q-q')$ will be evenly spread over the surface of the sphere, and perceived from any observation point $\boldsymbol{r}$ with $r > a$ as located at the *origin* of the sphere. The issue, again, is to find the solution for the resultant electrostatic scalar potential in the field region $r \geq a$.

By the superposition principle, $\phi(\boldsymbol{r})$ must be of the form

$$\phi(\boldsymbol{r}) = \phi_q(\boldsymbol{r}) + \phi_{q'}(\boldsymbol{r}) + \phi_{(Q-q')}(\boldsymbol{r}) \; .$$

Adding the scalar potential for a point charge $(Q-q')$ at $\boldsymbol{r} = \boldsymbol{0}$ to (3.49) yields with (3.48)

$$\phi(\boldsymbol{r}) = \frac{1}{4\pi\epsilon_0} \left[\frac{q}{|\,\boldsymbol{r} - \boldsymbol{r}_1\,|} - \frac{a\,q}{r_1\,|\,\boldsymbol{r} - (a/r_1)^2\,\boldsymbol{r}_1\,|} + \frac{Q + (a/r_1)\,q}{|\,\boldsymbol{r}\,|} \right] ; \tag{3.52}$$

this result provides the solution valid for $r \geq a$ we were looking for.

The mechanical force felt by the point charge q in the field of the the charge $(Q-q')$ and the induced image charge q' is derived from Coulomb's law. With the distance between q and $(Q-q')$ equal to r_1, and the distance between q and q' equal to $r_1[1-(a/r_1)^2]$, we thus find (after a few algebraic manipulations)

$$\boldsymbol{F} = \frac{1}{4\pi\epsilon_0}\,\frac{q}{r_1^2} \left[Q - q \left(\frac{a}{r_1}\right)^3 \left(2 - \frac{a^2}{r_1^2}\right) \left(1 - \frac{a^2}{r_1^2}\right)^{-2} \right] \hat{\boldsymbol{r}}_1 \; .$$

Note that for $r_1 \gg a$, i.e., when we place the point charge q at a distance from the sphere that is large compared to its radius, this result reduces to the familiar inverse square law for the Coulomb force between point charges q and Q.

3.6.4 Dirichlet Green's function for a conducting sphere

The discussion of the previous two examples naturally leads us to consider how electrostatic boundary value problems are solved in terms of the **Green's function formalism** of subsection 3.3.3 when the boundary surface is a *conducting sphere*. Prescribing the value of the electrostatic scalar potential itself on the spherical surface means one is opting for **Dirichlet boundary conditions**. Now supposing again that the sphere has radius a, the **Dirichlet Green's function** for a unit point charge at position $\boldsymbol{r}'$ and its image at position $(a/r')^2\boldsymbol{r}'$, which satisfies (3.28) and is valid at observation points $\boldsymbol{r}$ with $r \geq a$, can be inferred from (3.49) to be

$$G_{\mathrm{D}}(\boldsymbol{r}, \boldsymbol{r}') = \frac{1}{|\boldsymbol{r} - \boldsymbol{r}'|} - \frac{a}{r'\,|\boldsymbol{r} - (a/r')^2\,\boldsymbol{r}'|} \; .$$

We can expand the moduli terms in the denominators of this expression so that it reads

$$G_{\mathrm{D}}(\boldsymbol{r}, \boldsymbol{r}') = \frac{1}{(r^2 + r'^2 - 2rr'\cos\gamma)^{1/2}} - \frac{1}{(r^2r'^2/a^2 + a^2 - 2rr'\cos\gamma)^{1/2}} \; ; \tag{3.53}$$

here γ denotes the angle subtended between the source position $\boldsymbol{r}'$ and the observation point $\boldsymbol{r}$. In this form of we easily recognise (i) the automatic symmetry of $G_{\rm D}(\boldsymbol{r},\boldsymbol{r}')$ under exchange of $\boldsymbol{r}$ and $\boldsymbol{r}'$, and (ii) that we have $G_{\rm D}(\boldsymbol{r},\boldsymbol{r}')=0$ when either $r=a$ or $r'=a$, in line with the defining condition (3.32) of a Dirichlet Green's function.

In order to get in a position where we can use formula (3.33) to obtain a **solution** to **Poisson's equation** in the field region $r \geq a$ (for given Dirichlet boundary conditions, and here also $\rho=0$), we also need to know the value of the **normal derivative** of $G_{\rm D}(\boldsymbol{r},\boldsymbol{r}')$ on the surface of the sphere. Note that in the present situation the normal vector $\boldsymbol{n}'$ points *into* the sphere, i.e, *out of* the field region of interest, $r \geq a$; thus $\boldsymbol{n}' = -\,\hat{\boldsymbol{r}}'$. Taking this into account, we obtain from (3.53)

$$\begin{aligned}(\boldsymbol{n}'\cdot\boldsymbol{\nabla}_{\boldsymbol{r}'})G_{\rm D}(\boldsymbol{r},\boldsymbol{r}')\,\big|_{r'=a} &= -\left.\frac{\partial G_{\rm D}(\boldsymbol{r},\boldsymbol{r}')}{\partial r'}\right|_{r'=a} \\ &= -\frac{(r^2-a^2)}{a\,(r^2+a^2-2ra\cos\gamma)^{3/2}}\,, \qquad (3.54)\end{aligned}$$

which, up to a constant of proportionality, just corresponds to the surface charge density σ that a unit point charge induces on the surface of the conducting sphere [cf. (3.51) with $(q/4\pi)\to 1$ and $r_1\to r$]. Using this result (and setting $\rho=0$), we finally find from (3.33)

$$\phi(\boldsymbol{r}) = \frac{1}{4\pi}\int_0^{2\pi}\int_0^{\pi}\phi(a,\vartheta',\varphi')\,\frac{a\,(r^2-a^2)}{(r^2+a^2-2ra\cos\gamma)^{3/2}}\,\sin\vartheta'\,{\rm d}\vartheta'\,{\rm d}\varphi'\,, \qquad (3.55)$$

which provides the solution to **Laplace's equation** (3.18) in the field region $r \geq a$. In terms of the angular coordinates $\{\vartheta,\varphi;\vartheta',\varphi'\}$, we have $\cos\gamma = \cos\vartheta\cos\vartheta' + \sin\vartheta\sin\vartheta'\cos(\varphi-\varphi')$. To emphasise once more, the information on specific boundary behaviour to be given in the present procedure (so that the surface integral can actually be performed) is the variation of the electrostatic scalar potential *on* the surface of the sphere, $\phi(a,\vartheta',\varphi')$.

3.6.5 Conducting sphere in a uniform electrostatic field

The last example demonstrating an application of the method of images that we want to discuss is the electrostatic potential due to a conducting sphere at rest in a uniform electrostatic field. Let us assume that the uniform field is oriented along the z–axis of a reference frame with **Cartesian coordinates** so that $\boldsymbol{E}_0 = E_0\,\boldsymbol{e}_z$. Let us also assume that the conducting sphere has radius a. If the sphere is placed at the origin, then the configuration has **azimuthal symmetry**, i.e., with respect to **spherical polar coordinates** $\{r,\vartheta,\varphi\}$ on $\mathbb{R}^3$ the resultant electrostatic potential satisfies

$$\frac{\partial\phi}{\partial\varphi} = 0\,.$$

It is easy to write down the electrostatic potential for the uniform field alone. In **Cartesian coordinates** this is just $\phi_{\rm uniform}(\boldsymbol{r}) = -\,E_0\,z$. However, to

address the boundary value problem for the setting that we outlined, a different, approximative, route proves more fruitful. We therefore assume that we have a field generated by two point charges $+Q$ and $-Q$ which are at rest at positions $z = -R$ and $z = +R$. It follows that in the field region near the origin (where $r \ll R$) we have $\boldsymbol{E}_0 \approx (2Q/4\pi\epsilon_0 R^2)\,\boldsymbol{e}_z$. In the limit

$$R \to \infty \qquad Q \to \infty \qquad \text{with} \quad \frac{Q}{R^2} = \text{const}\,,$$

this becomes an exact result. If now the conducting sphere of radius a is positioned at the origin, the field induces a charge density σ on the sphere's surface. As we know from the example in subsection 3.6.2, this induction effect can be modelled by placing two image charges of magnitudes $-Qa/R$ and $+Qa/R$ at positions $z = -a^2/R$ and $z = +a^2/R$ *inside* the sphere. By the superposition principle, $\phi(\boldsymbol{r})$ thus has the form

$$\phi(\boldsymbol{r}) = \phi_{+Q}(\boldsymbol{r}) + \phi_{-Q}(\boldsymbol{r}) + \phi_{-Q'}(\boldsymbol{r}) + \phi_{+Q'}(\boldsymbol{r})\,.$$

Using **spherical polar coordinates** on $\mathbb{R}^3$, this becomes

$$\begin{aligned}\phi \;=\; \frac{Q}{4\pi\epsilon_0}\,\Bigg[\,&\frac{1}{(r^2+R^2+2rR\cos\vartheta)^{1/2}} - \frac{1}{(r^2+R^2-2rR\cos\vartheta)^{1/2}} \\ &- \frac{a}{R\,(r^2+(a^2/R)^2+2r(a^2/R)\cos\vartheta)^{1/2}} \\ &+ \frac{a}{R\,(r^2+(a^2/R)^2-2r(a^2/R)\cos\vartheta)^{1/2}}\,\Bigg]\;;\end{aligned}$$

note that $\partial\phi/\partial\varphi = 0$ applies indeed. Next, employing a Taylor expansion of the denominators, to first order in powers of (r/R) and (a^2/rR), we get

$$\begin{aligned}\phi(r,\vartheta) \;&=\; \frac{1}{4\pi\epsilon_0}\,\left[-2\,\frac{Q}{R^2}\,r\cos\vartheta + 2\,\frac{Q}{R^2}\,\frac{a^3}{r^2}\,\cos\vartheta\right] + \ldots \\ &=\; -\,\frac{Q}{2\pi\epsilon_0\,R^2}\,r\cos\vartheta\,\left[1-\frac{a^3}{r^3}\right] + \ldots \\ &=\; -\,E_0\,r\,\left[1-\frac{a^3}{r^3}\right]\,\cos\vartheta + \ldots\end{aligned}$$

As in the limit $R \to \infty$ all the higher order terms (the "dots") go to zero, we finally obtain the solution

$$\phi(r,\varphi) = -\,E_0\,r\cos\vartheta + E_0\,\frac{a^3}{r^2}\,\cos\vartheta\;; \tag{3.56}$$

it is valid in the field region $r \geq a$. This result just corresponds to the linear superposition of the scalar potentials of each of the uniform electrostatic field and of the dipolar field due to the surface charge which is induced on the conducting sphere.

For the surface charge density we have

$$\sigma = -\,\epsilon_0\,(\hat{\boldsymbol{r}}\cdot\boldsymbol{\nabla}\phi)|_{r=a} = -\,\epsilon_0\,\left.\frac{\partial\phi}{\partial r}\right|_{r=a} = 3\epsilon_0\,E_0\cos\vartheta\,.$$

Integrating σ over the entire surface of the sphere leads to

$$\int\int_{\text{sphere surface}} \sigma \, \mathrm{d}A = 3\epsilon_0 \, E_0 \int_0^{2\pi} \int_0^{\pi} \cos\vartheta \, a^2 \sin\vartheta \, \mathrm{d}\vartheta \, \mathrm{d}\varphi = 0 \, ;$$

the total charge induced on the conducting sphere by the uniform electrostatic field is zero (a result which, by the azimuthal symmetry of the setting, was to be expected). We can conclude from this that it makes no difference whether the conducting sphere in the problem is grounded or not.

3.7 Separation of variables method

For most (sets of) partial differential equations that one encounters in mathematical physics there does *not* exist a general procedure to obtain their **general solutions**. Hence, the next best thing that comes to mind to address this issue is to look instead for **special solutions** to these equations. One systematic approach that has emerged from this view is to employ the **separation of variables method**. This looks for families of solutions to the partial differential equations at hand which are constructed from products of functions each of which depends only on *one* of the independent variables (coordinates) involved of the given problem. One thus generates from a set of **partial differential equations** an associated set of **ordinary differential equations**, the latter of which is, hopefully, easier to solve. In the present section we want to employ this technique to obtain solutions to **Laplace's equation**

$$(\boldsymbol{\nabla} \cdot \boldsymbol{\nabla})\phi = 0$$

for sourcefree electrostatic field regions in $\mathbb{R}^3$ in two different circumstances, namely, (i) when the coordinate system is **Cartesian**, and (ii) when **spherical polar coordinates** are used.

3.7.1 Laplace's equation in Cartesian coordinates

As we know from chapter 2, in **Cartesian coordinates** $\{x, y, z\}$ on $\mathbb{R}^3$ **Laplace's equation** takes the explicit form

$$\frac{\partial^2\phi}{\partial x^2} + \frac{\partial^2\phi}{\partial y^2} + \frac{\partial^2\phi}{\partial z^2} = 0 \, . \tag{3.57}$$

We now want to look for solutions to (3.57) that are of the very special **separable form**

$$\phi(x, y, z) = X(x) \, Y(y) \, Z(z) \, ,$$

where each of $X(x)$, $Y(y)$ and $Z(z)$ are at least twice continuously differentiable real-valued functions of their arguments. When we take this solution Ansatz and plug it into (3.57), the result we get upon division by the product $X(x) \, Y(y) \, Z(z)$ is

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

It is clear from the present form that the first term on the left-hand side depends only on x, the second only on y, and the third only on z. Hence, for this equation to hold for *arbitrary* values of the coordinates x, y, and z, each of the three terms must be *separately constant*. Thus, introducing constants $\alpha, \beta, \gamma \in \mathbb{R}$ according to

$$\frac{1}{X}\,\frac{\mathrm{d}^2 X}{\mathrm{d}x^2} = -\,\alpha^2 \qquad \frac{1}{Y}\,\frac{\mathrm{d}^2 Y}{\mathrm{d}y^2} = -\,\beta^2 \qquad \frac{1}{Z}\,\frac{\mathrm{d}^2 Z}{\mathrm{d}z^2} = \gamma^2\,,$$

we have obtained a set of three decoupled ordinary differential equations (of second order) subject to the condition

$$\alpha^2 + \beta^2 = \gamma^2\,.$$

The solutions to these three ordinary differential equations are obviously of the form

$$X \propto \mathrm{e}^{\pm i\alpha x} \qquad Y \propto \mathrm{e}^{\pm i\beta y} \qquad Z \propto \mathrm{e}^{\pm\sqrt{\alpha^2+\beta^2}\,z}\,, \tag{3.58}$$

and so we have for the electrostatic scalar potential

$$\phi(x,y,z) \propto \mathrm{e}^{\pm i\alpha x}\,\mathrm{e}^{\pm i\beta y}\,\mathrm{e}^{\pm\sqrt{\alpha^2+\beta^2}\,z}\,. \tag{3.59}$$

The real-valued constants α and β in this expression need to be determined from given boundary conditions [cf. subsection 3.3.2] that $\phi(x,y,z)$ is required to satisfy.

To illustrate our result, let us consider an actual example of a boundary value problem in **Cartesian coordinates** for **Laplace's equation** (3.57). Suppose there is a rectangular box, made from mutually isolated planar metal plates, which has linear dimensions a, b, and c in the x–, y–, and z–directions, respectively. Let us assume that all sides of the box are held at zero potential, except for to top one at $z = c$, which we want to assume to be at potential $V(x,y)$. We would like to determine the electrostatic scalar potential $\phi(x,y,z)$ resulting from these boundary conditions in the field region *inside* of the box.

Demanding that $\phi = 0$ holds at each of $x = 0$, $y = 0$, and $z = 0$, leads with (3.58) to

$$X \propto \sin(\alpha x) \qquad Y \propto \sin(\beta y) \qquad Z \propto \sinh(\sqrt{\alpha^2+\beta^2}\,z)\,.$$

When, in addition, we demand that $\phi = 0$ shall also hold at $x = a$ and $y = b$, this requires that the periodicity conditions $\alpha a = n\pi$ and $\beta b = m\pi$ be satisfied. So by defining

$$\alpha_n := \frac{n\pi}{a} \qquad \beta_m := \frac{m\pi}{b} \qquad \gamma_{nm} := \pi\sqrt{\left(\frac{n}{a}\right)^2 + \left(\frac{m}{b}\right)^2}\,,$$

we now have from (3.59)

$$\phi(x,y,z) \propto \sin(\alpha_n x)\,\sin(\beta_m y)\,\sinh(\gamma_{mn} z) := \phi_{nm}(x,y,z)\,. \tag{3.60}$$

Now we have reached a stage where we are left with fixing only one final boundary condition, namely that $\phi = V(x,y)$ at $z = c$. Let us proceed by

making the following Ansatz for $\phi(x,y,z)$: we use the functions $\phi_{nm}(x,y,z)$ defined in (3.60) to write down a linear expansion of $\phi(x,y,z)$ with constant coefficients A_{nm} according to

$$\phi(x,y,z) = \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} A_{nm}\, \phi_{nm}(x,y,z) \; . \tag{3.61}$$

Imposing on this expansion of $\phi(x,y,z)$ the condition $\phi = V(x,y)$ at $z = c$, we find

$$V(x,y) = \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} A_{nm}\, \sin(\alpha_n x)\, \sin(\beta_m y)\, \sinh(\gamma_{mn} c) \; .$$

We note that this expression is nothing but a **double Fourier series expansion** (cf. section 2.6) for the (given) function $V(x,y)$, with constant expansion coefficients A_{nm} determined according to

$$A_{nm} = \frac{4}{ab \sinh(\gamma_{mn} c)} \int_0^a \mathrm{d}x \int_0^b \mathrm{d}y\, V(x,y)\, \sin(\alpha_n x)\, \sin(\beta_m y) \; . \tag{3.62}$$

It is these A_{nm} we have just found which provide exactly those constant expansion coefficients we need for (3.61) to form a separable solution to **Laplace's equation** (3.57) that is valid in the interior of the rectangular box under the presently given boundary conditions.

Note that separable solutions to the present boundary value problem for (3.57) with *more* than one side of the rectangular box at non-zero potential can be obtained by linear superpostion of the appropriate number of different solutions of the kind (3.61) and (3.62).

3.7.2 Laplace's equation in spherical polar coordinates

It follows from the considerations in chapter 2 that, when written down in **spherical polar coordinates** $\{r, \vartheta, \varphi\}$ on $\mathbb{R}^3$, **Laplace's equation** (3.18) reads

$$\frac{1}{r} \frac{\partial^2 (r\,\phi)}{\partial r^2} + \frac{1}{r^2 \sin\vartheta} \frac{\partial}{\partial \vartheta} \left(\sin\vartheta \frac{\partial \phi}{\partial \vartheta} \right) + \frac{1}{r^2 \sin^2\vartheta} \frac{\partial^2 \phi}{\partial \varphi^2} = 0 \; ; \tag{3.63}$$

this form of the $(\boldsymbol{\nabla} \cdot \boldsymbol{\nabla})$–operator makes use of the identity

$$\frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial \phi}{\partial r} \right) \equiv \frac{1}{r} \frac{\partial^2 (r\,\phi)}{\partial r^2} \; .$$

Now let us look for solutions to (3.63) of the **separable form**

$$\phi(r,\vartheta,\varphi) = \frac{U(r)}{r}\, P(\vartheta)\, Q(\varphi) \; .$$

Upon substitution of this solution Ansatz in (3.63), and conveniently rearranging terms, we obtain

$$r^2 \sin^2\vartheta \left[\frac{1}{U} \frac{\mathrm{d}^2 U}{\mathrm{d}r^2} + \frac{1}{P r^2 \sin\vartheta} \frac{\mathrm{d}}{\mathrm{d}\vartheta} \left(\sin\vartheta \frac{\mathrm{d}P}{\mathrm{d}\vartheta} \right) \right] + \frac{1}{Q} \frac{\mathrm{d}^2 Q}{\mathrm{d}\varphi^2} = 0 \; .$$

We observe that the term on the left-hand side containing the square brackets depends only on r and ϑ, while the last term depends only on φ. Hence, for this equation to hold for *arbitrary* values of the coordinates r, ϑ, and φ, each of the two terms must be *separately constant*. It is standard to set the two terms equal to m^2 and $-m^2$, $m \in \mathbb{R}$, respectively.

For the Q–term we thus find the decoupled ordinary differential equation (of second order)

$$\frac{1}{Q}\frac{\mathrm{d}^2 Q}{\mathrm{d}\varphi^2} = -m^2 \quad \Rightarrow \quad Q(\varphi) \propto \mathrm{e}^{\pm im\varphi} \,. \tag{3.64}$$

When the range of the coordinate φ is $0 \leq \varphi \leq 2\pi$, the periodicity condition $Q(\varphi + 2\pi) = Q(\varphi)$ must be satisfied in order for $Q(\varphi)$ to be a **single-valued** function. It follows that in this case m can only take *integer* values. We note that the solutions $Q_m(\varphi) = \mathrm{e}^{im\varphi}$ (with $m \in \mathbb{R}$) to (3.64) form a complete set of orthogonal functions in the index m on the φ–interval $0 \leq \varphi \leq 2\pi$ (cf. subsection 2.6.1).

Proceeding with the square brackets term, we have, after some rearranging,

$$r^2 \frac{1}{U}\frac{\mathrm{d}^2 U}{\mathrm{d}r^2} + \frac{1}{P\sin\vartheta}\frac{\mathrm{d}}{\mathrm{d}\vartheta}\left(\sin\vartheta\,\frac{\mathrm{d}P}{\mathrm{d}\vartheta}\right) - \frac{m^2}{\sin^2\vartheta} = 0 \,.$$

In this form, the first term on the left-hand side depends only on r, while the remaining two depend only on ϑ. Again, for *arbitrary* values of the coordinates r and ϑ, each of the two parts must be *separately constant*. It is standard to set them equal to $l(l+1)$ and $-l(l+1)$, $l \in \mathbb{R}$, respectively. We thus find the two decoupled ordinary differential equations (of second order)

$$\frac{\mathrm{d}^2 U}{\mathrm{d}r^2} - \frac{l(l+1)}{r^2}\,U = 0 \tag{3.65}$$

$$\frac{1}{\sin\vartheta}\frac{\mathrm{d}}{\mathrm{d}\vartheta}\left(\sin\vartheta\,\frac{\mathrm{d}P}{\mathrm{d}\vartheta}\right) + \left[\,l(l+1) - \frac{m^2}{\sin^2\vartheta}\,\right] P = 0 \,. \tag{3.66}$$

The general solution to (3.65), which determines the radial dependence of $\phi(r,\vartheta,\varphi)$, is given by

$$U(r) = A\,r^{l+1} + B\,r^{-l} \qquad A, B \in \mathbb{R} \,. \tag{3.67}$$

The term proportional to r^{l+1} is **regular** at the origin ($r = 0$), while it becomes **singular** at spatial infinity ($r \to \infty$). The term proportional to r^{-l}, on the other hand, shows the the opposite behaviour. The real-valued constants A and B need to be determined from given boundary conditions [cf. subsection 3.3.2] that $\phi(r,\vartheta,\varphi)$ is required to satisfy. The real-valued constants l and integers m, however, so far remain undetermined.

To discuss the general solution to the more complicated equation (3.66), a separate consideration is required.

Legendre's equation and Legendre polynomials

Equation (3.66), which determines the ϑ–dependence of $\phi(r,\vartheta,\varphi)$, assumes a more transparent form when we rewrite it in terms of the *new* independent

variable

$$x = \cos\vartheta \,, \tag{3.68}$$

implying a transition $P(\vartheta) \to P(x)$. We thus obtain

$$\frac{\mathrm{d}}{\mathrm{d}x}\left[(1-x^2)\,\frac{\mathrm{d}P}{\mathrm{d}x}\right] + \left[l(l+1) - \frac{m^2}{1-x^2}\right]\,P = 0 \,. \tag{3.69}$$

This constitutes the **generalised Legendre equation** (named after the French mathematician Adrien-Marie Legendre, 1752–1833); its solutions are provided by the **associated Legendre functions** which are defined on the finite x–interval $-1 \leq x \leq +1$. We will give the explicit form of the latter a little later, after we have discussed the following subcase of (3.69).

When $m^2 = 0$, (3.69) reduces to **Legendre's equation**

$$\frac{\mathrm{d}}{\mathrm{d}x}\left[(1-x^2)\,\frac{\mathrm{d}P}{\mathrm{d}x}\right] + l(l+1)\,P = 0 \,, \tag{3.70}$$

a linear ordinary differential equation of second order for an at least twice continuously differentiable real-valued function $P(x)$. It can be solved by employing Frobenius' method (named after the German mathematician Ferdinand Georg Frobenius, 1849–1917), i.e., by assuming that its solutions can be represented by a power series of the form

$$P(x) = x^r \sum_{n=0}^{\infty} a_n\, x^n \,,$$

where we need to determine the real-valued parameter r and the real-valued constant expansion coefficients a_n. One finds that in order to get solutions on the x–interval $-1 \leq x \leq +1$ that are single-valued, finite, and continuous, the parameter l can only take positive integer values or the value zero. This has the consequence that the power series expansion for $P(x)$ terminates. The solutions, $P_l(x)$, to (3.70) one so obtains are called the **Legendre polynomials of order** l, meaning that the highest-order term in x they contain is proportional to x^l.

The Legendre polynomials can be compactly represented through **Rodrigues' formula** as

$$P_l(x) = \frac{1}{2^l l!}\,\frac{\mathrm{d}^l}{\mathrm{d}x^l}(x^2-1)^l \,. \tag{3.71}$$

The first few $P_l(x)$ read

$$P_0(x) = 1 \qquad P_1(x) = x \qquad P_2(x) = \frac{1}{2}(3x^2-1)$$
$$P_3(x) = \frac{1}{2}(5x^3-3x) \qquad P_4(x) = \frac{1}{8}(35x^4-30x^2+3) \,.$$

We make the following two remarks: (i) the $P_l(x)$ alternately constitute even and odd polynomials in x, i.e., they satisfy

$$P_{l=2n}(-x) = P_{l=2n}(x) \quad P_{l=2n+1}(-x) = -\,P_{l=2n+1}(x) \quad n = 0, 1, 2, \ldots \,,$$

and (ii) we have

$$P_l(x=1)=1 \qquad P_l(x=-1)=(-1)^l\,.$$

The Legendre polynomials $P_l(x)$ are particularly interesting because they form a complete set of *orthogonal* functions on the x–interval $-1\le x\le +1$. The **orthogonality condition** in this case is given by

$$\int_{-1}^{1} P_{l'}(x)\,P_l(x)\,\mathrm{d}x=\frac{2}{2l+1}\,\delta_{l'l}\,. \tag{3.72}$$

Hence, by introducing

$$U_l(x):=\sqrt{\frac{2l+1}{2}}\,P_l(x)\,,$$

we do obtain *normalised* Legendre polynomials of order l. The **completeness condition** for the $U_l(x)$ then reads

$$\sum_{l=0}^{\infty} U_l(x')\,U_l(x)=\delta(x'-x)\,. \tag{3.73}$$

We thus can use either the $U_l(x)$ or the $P_l(x)$ to expand any real-valued function $f(x)$ that is square integrable on the x–interval $-1\le x\le +1$ in an infinite series with constant coefficients by

$$f(x)=\sum_{l=0}^{\infty} a_l\,U_l(x)=\sum_{l=0}^{\infty}\sqrt{\frac{2l+1}{2}}\,a_l\,P_l(x) \tag{3.74}$$

$$a_l=\int_{-1}^{1} U_l(x)\,f(x)\,\mathrm{d}x=\sqrt{\frac{2l+1}{2}}\int_{-1}^{1} P_l(x)\,f(x)\,\mathrm{d}x\,. \tag{3.75}$$

This ends our discussion of **Legendre's equation**.

Solutions with azimuthal symmetry

In this paragraph we want to discuss separable solutions to **Laplace's equation** (3.63) that have **azimuthal symmetry**. This is to say that they shall not depend on the coordinate φ. It is clear that in this case we must have

$$\frac{\partial\phi}{\partial\varphi}=0\,, \tag{3.76}$$

so that by (3.64) it follows that $m=0$. Solutions with this symmetry property can thus be expanded in terms of products between the fundamental solutions r^l and $r^{-(l+1)}$ to (3.65) and the Legendre polynomials P_l according to

$$\phi(r,\vartheta)=\sum_{l=0}^{\infty}\left[\,A_l\,r^l+B_l\,r^{-(l+1)}\,\right]\,P_l(\cos\vartheta)\,. \tag{3.77}$$

The constant expansion coefficients A_l and B_l in this expression are to be determined from boundary conditions which we expect to be given in a real application.

Suppose, e.g., that we have a conducting sphere of radius a that is centred on the origin of our reference frame. We want to assume that the electrostatic scalar potential varies on the surface of the conducting sphere as $V(\vartheta)$. Imposing the boundary condition

$$\phi(r,\vartheta)|_{r=a} = V(\vartheta) \;, \tag{3.78}$$

let us now look for solutions $\phi(r,\vartheta)$ that are *regular* in the *interior* of the sphere, i.e., in the field region $r \leq a$. It follows that regularity, in particular at $r = 0$, is given when

$$B_l = 0 \quad \text{for all } l \;.$$

This fixes *one* boundary condition. Next, evaluating $\phi(r,\vartheta)$ at $r = a$ and equating it to $V(\vartheta)$, we find

$$V(\vartheta) = \sum_{l=0}^{\infty} A_l \, a^l \, P_l(\cos\vartheta) \;.$$

But this is nothing but a Legendre series expansion for the (known) function $V(\vartheta)$ on the ϑ–interval $0 \leq \vartheta \leq \pi$ [cf. (3.74)], so that the constant expansion coefficients A_l are just [cf. (3.75)]

$$A_l = \frac{2l+1}{2a^l} \int_0^{\pi} P_l(\cos\vartheta)\, V(\vartheta)\, \sin\vartheta \, \mathrm{d}\vartheta \;.$$

To make this discussion more graphical, let us take $V(\vartheta)$ to be a function that is *odd* on the ϑ–interval $0 \leq \vartheta \leq \pi$, explicitly given by

$$V(\vartheta) = \begin{cases} +V & 0 \leq \vartheta \leq \pi/2 \\ -V & \pi/2 \leq \vartheta \leq \pi \end{cases} \quad , \quad V = \text{const} > 0 \;. \tag{3.79}$$

This represents a configuration with two insulated conducting hemispheres at equal but opposite constant potential, In this case, we find that the **interior solution** for $\phi(r,\vartheta)$ becomes

$$\phi(r,\vartheta) = V \left[\frac{3}{2}\left(\frac{r}{a}\right) P_1(\cos\vartheta) - \frac{7}{8}\left(\frac{r}{a}\right)^3 P_3(\cos\vartheta) + \frac{11}{16}\left(\frac{r}{a}\right)^5 P_5(\cos\vartheta) + \ldots \right] \;, \tag{3.80}$$

which, because $V(\vartheta)$ was assumed to be an odd function, only contains Legendre polynomials of *odd* order.

We note that if, in the present example, we were interested instead in the **exterior solution**, i.e., a solution in the field region $r \geq a$ that is regular at spatial infinity, then a similar analysis would show that all we had to do was to replace in (3.80) the terms $(r/a)^l$ by $(a/r)^{l+1}$. With (3.79), the (regular) exterior solution would still be an expression that only contains Legendre polynomials of odd order.

Spherical harmonics

Of course, the property (3.76) does not apply to a *general* electrostatic scalar potential. Hence, to obtain separable solutions to the full **Laplace's equation** in **spherical polar coordinates**, (3.63), we have to solve the full **generalised Legendre equation** (3.66), which has $m^2 > 0$. Having rewritten (3.66) in terms of the independent variable x defined in (3.68), one can show, in a fashion that is practically analogous to solving **Legendre's equation**, that solutions on the x–interval $-1 \leq x \leq +1$ that are single-valued, finite, and continuous must have the parameter l equal to a positive integer or zero and the parameter m equal to one of the integer values $-l, -(l-1), \ldots, 0, \ldots, +(l+1), +l$. The solutions to (3.66) which have these features are referred to as the **associated Legendre functions**, $P_l^m(x)$. In a generalisation of Rodrigues' formula, (3.71), they can be represented by

$$P_l^m(x) = \frac{(-1)^m}{2^l l!}\,(1-x^2)^{m/2}\,\frac{\mathrm{d}^{l+m}}{\mathrm{d}x^{l+m}}(x^2-1)^l \qquad (3.81)$$
$$l = 0, 1, 2, \ldots \qquad\qquad m = -l, \ldots, +l\ .$$

One can use this expression to show that $P_l^{-m}(x)$ and $P_l^m(x)$ are related by

$$P_l^{-m}(x) = (-1)^m\,\frac{(l-m)!}{(l+m)!}\,P_l^m(x)\ . \qquad (3.82)$$

It is important to note that for *fixed* value of the index m the $P_l^m(x)$ form a complete set of orthogonal functions in the index l on the x–interval $-1 \leq x \leq +1$.

It proves very convenient in **mathematical physics** to combine the solutions $Q_m(\varphi)$ to (3.64) with the solutions $P_l^m(\cos\vartheta)$ to (3.66) and construct a complete set of orthonormal functions over the **unit sphere** (which is given through the coordinate ranges $r = 1$, $0 \leq \vartheta \leq \pi$, and $0 \leq \varphi \leq 2\pi \Rightarrow -1 \leq \cos\vartheta \leq +1$). These so-called **spherical harmonics**, commonly denoted by $Y_{lm}(\vartheta,\varphi)$, are defined by

$$Y_{lm}(\vartheta,\varphi) := \sqrt{\frac{2l+1}{4\pi}\,\frac{(l-m)!}{(l+m)!}}\,P_l^m(\cos\vartheta)\,\mathrm{e}^{im\varphi}\ . \qquad (3.83)$$

It follows from (3.82) that $Y_{l,-m}(\vartheta,\varphi)$ and $Y_{l,m}(\vartheta,\varphi)$ are related by

$$Y_{l,-m}(\vartheta,\varphi) = (-1)^m\,Y_{lm}^*(\vartheta,\varphi)\ ; \qquad (3.84)$$

here $Y_{lm}^*(\vartheta,\varphi)$ denotes the complex conjugate of $Y_{lm}(\vartheta,\varphi)$. The **orthonormality condition** for the spherical harmonics is given by

$$\int_0^{2\pi}\mathrm{d}\varphi\int_0^{\pi}\sin\vartheta\,\mathrm{d}\vartheta\,Y_{l'm'}^*(\vartheta,\varphi)\,Y_{lm}(\vartheta,\varphi) = \delta_{l'l}\,\delta_{m'm}\ , \qquad (3.85)$$

while the **completeness condition** reads

$$\sum_{l=0}^{\infty}\sum_{m=-l}^{+l}Y_{lm}^*(\vartheta',\varphi')\,Y_{lm}(\vartheta,\varphi) = \delta(\cos\vartheta'-\cos\vartheta)\,\delta(\varphi'-\varphi)\ . \qquad (3.86)$$

The first few spherical harmonics, in explicit form, are

(a) $l = 0$

$$Y_{00} = \frac{1}{\sqrt{4\pi}}$$

(b) $l = 1$

$$Y_{11} = -\sqrt{\frac{3}{8\pi}}\,\sin\vartheta\,\mathrm{e}^{i\varphi}$$
$$Y_{10} = \sqrt{\frac{3}{4\pi}}\,\cos\vartheta$$

(c) $l = 2$

$$Y_{22} = \frac{1}{4}\sqrt{\frac{15}{2\pi}}\,\sin^2\vartheta\,\mathrm{e}^{i2\varphi}$$
$$Y_{21} = -\sqrt{\frac{15}{8\pi}}\,\sin\vartheta\cos\vartheta\,\mathrm{e}^{i\varphi}$$
$$Y_{20} = \frac{1}{2}\sqrt{\frac{5}{4\pi}}\left(3\cos^2\vartheta - 1\right)$$

(d) $l = 3$

$$Y_{33} = -\frac{1}{4}\sqrt{\frac{35}{4\pi}}\,\sin^3\vartheta\,\mathrm{e}^{i3\varphi}$$
$$Y_{32} = \frac{1}{4}\sqrt{\frac{105}{2\pi}}\,\sin^2\vartheta\cos\vartheta\,\mathrm{e}^{i2\varphi}$$
$$Y_{31} = -\frac{1}{4}\sqrt{\frac{21}{4\pi}}\,\sin\vartheta\left(5\cos^2\vartheta - 1\right)\mathrm{e}^{i\varphi}$$
$$Y_{30} = \frac{1}{2}\sqrt{\frac{7}{4\pi}}\left(5\cos^3\vartheta - 3\cos\vartheta\right)\ .$$

The spherical harmonics that are *not* listed here, namely the $Y_{l,-m}(\vartheta,\varphi)$ for each of $l = 1, 2, 3$, can be easily derived using formula (3.84).

With the spherical harmonics in place, we can now expand any real-valued function $g(\vartheta,\varphi)$ that is square integrable over the unit sphere in an infinite series with constant coefficients according to

$$g(\vartheta,\varphi) = \sum_{l=0}^{\infty}\sum_{m=-l}^{+l} a_{lm}\,Y_{lm}(\vartheta,\varphi) \tag{3.87}$$

$$a_{lm} = \int_0^{2\pi}\int_0^{\pi} Y^*_{lm}(\vartheta,\varphi)\,g(\vartheta,\varphi)\,\sin\vartheta\,\mathrm{d}\vartheta\,\mathrm{d}\varphi\ . \tag{3.88}$$

These formulae find widespread applications in **mathematical physics**. To give three prominent examples: (i) they feature in **quantum physics** in the solution process of the stationary Schrödinger equation with a Coulomb electrostatic scalar potential to determine the electron orbits for the hydrogen atom (cf. the lectures on MAS217 Quantum Theory), (ii) they arise in the models

of **helio-** and **asteroseismology** to describe the different modes of internal oscillations observed in the Sun and nearby stars (i.e., "Sun- and starquakes"; cf. the lectures on MAS402 Astrophysical Fluid Dynamics), or (iii) they are used in **cosmology** to give a mathematical representation of the (minute) temperature variations that are measured to an astonishing precision in the cosmic background radiation, the thermal radiation believed to be the afterglow of a universal big bang in the finite past for space, time and all the matter and fields in the Universe (cf. the lectures on MAS401 Advanced Cosmology).

In our case, namely **electrostatics**, we use the spherical harmonics [and so (3.87) and (3.88)] to express separable solutions to boundary value problems for **Laplace's equation** in **spherical polar coordinates**, (3.63), as series expansions. That is, we combine the $Y_{lm}(\vartheta,\varphi)$ with the fundamental solutions r^l and $r^{-(l+1)}$ to (3.65) so that a general expression for an **electrostatic scalar potential** that is separable is given by

$$\phi(r,\vartheta,\varphi) = \sum_{l=0}^{\infty} \sum_{m=-l}^{+l} \left[a_{lm}\, r^l + b_{lm}\, r^{-(l+1)} \right] Y_{lm}(\vartheta,\varphi) \; . \tag{3.89}$$

Unique solutions $\phi(r,\vartheta,\varphi)$ that are separable are obtained upon specification of explicit boundary conditions, appropriate for a given problem, according to subsection 3.3.2. If, e.g., $\phi(r,\vartheta,\varphi)$ is specified on some given spherical surface of radius a, then the constant expansion coefficients a_{lm} and b_{lm} can be determined by evaluating (3.89) for $r = a$ and using (3.88).

3.8 Multipole expansions and multipole moments

The final topic we want to discuss in this chapter is how we can represent the electrostatic scalar potential in a sourcefree field region that results from a static distribution of electric charges which is *localised* in a region $G \subset \mathbb{R}^3$ of finite spatial extent.

Let us assume that we have a distribution of electric charges of density ρ, confined to the *interior*, G, of a sphere of radius a which is centred on the origin of the reference frame we set up. We can express the electrostatic scalar potential due to this charge distribution at an observation point $\boldsymbol{r}$ *exterior* to the sphere ($r = |\boldsymbol{r}| > a$) in terms of an expansion over the spherical harmonics $Y_{lm}(\vartheta,\varphi)$ and the fundamental solution $r^{-(l+1)}$ (which is regular for $r \to \infty$) as

$$\phi(\boldsymbol{r}) = \frac{1}{4\pi\epsilon_0} \sum_{l=0}^{\infty} \sum_{m=-l}^{+l} \frac{4\pi}{2l+1}\, q_{lm}\, \frac{Y_{lm}(\vartheta,\varphi)}{r^{l+1}} \; . \tag{3.90}$$

This is called a **multipole expansion** of $\phi(\boldsymbol{r})$; the *constant* expansion coefficients q_{lm} are referred to as the **electric multipole moments** of the localised charge distribution ρ.

Let us now turn to discuss how to *determine* the q_{lm} in terms of the properties of ρ. Let us start from the integral expression

$$\phi(\boldsymbol{r}) = \frac{1}{4\pi\epsilon_0} \int\int\int_G \frac{\rho(\boldsymbol{r}')}{|\boldsymbol{r}-\boldsymbol{r}'|}\, \mathrm{d}V' \; ; \tag{3.91}$$

this is (3.12) with $\sigma = 0$ and const $= 0$. Without going into further details, we remark that it can be shown that the electrostatic scalar potential at an observation point $\boldsymbol{r}$ due to a **unit point charge** located at position $\boldsymbol{r}'$, i.e., $1/|\,\boldsymbol{r} - \boldsymbol{r}'\,|$ [cf. (3.21)], is given in terms of the $Y_{lm}(\vartheta, \varphi)$ and the $r^{-(l+1)}$ as

$$\frac{1}{|\,\boldsymbol{r} - \boldsymbol{r}'\,|} = 4\pi \sum_{l=0}^{\infty} \sum_{m=-l}^{+l} \frac{1}{2l+1} \frac{r_<^l}{r_>^{l+1}} Y_{lm}^*(\vartheta', \varphi')\, Y_{lm}(\vartheta, \varphi) \; .$$

In this expansion $r_<$ denotes the smaller and $r_>$ the larger between $|\boldsymbol{r}|$ and $|\boldsymbol{r}'|$. With, in the present case, $r_< = r'$ and $r_> = r$, we thus find that we can rewrite (3.91) as

$$\phi(\boldsymbol{r}) = \frac{1}{\epsilon_0} \sum_{l=0}^{\infty} \sum_{m=-l}^{+l} \frac{1}{2l+1} \left[\int\int\int_G Y_{lm}^*(\vartheta', \varphi')\, r'^l\, \rho(\boldsymbol{r}')\, \mathrm{d}V' \right] \frac{Y_{lm}(\vartheta, \varphi)}{r^{l+1}} \; .$$

At this stage, a comparison with (3.90) shows us that the constant multipole moments q_{lm} of ρ have to be given by

$$q_{lm} = \int\int\int_G Y_{lm}^*(\vartheta', \varphi')\, r'^l\, \rho(\boldsymbol{r}')\, \mathrm{d}V' \; . \tag{3.92}$$

Note that it follows with (3.84) that $q_{l,-m}$ and q_{lm} are related by

$$q_{l,-m} = (-1)^m\, q_{lm}^* \; . \tag{3.93}$$

Using **Cartesian coordinates**, left us briefly list in explicit form the first few multipole moments. They are

(a) $l = 0$: **Electric monopole moment**

$$\begin{aligned} q_{00} &= \frac{1}{\sqrt{4\pi}} \int\int\int_G \rho(\boldsymbol{r}')\, \mathrm{d}V' \\ &= \frac{1}{\sqrt{4\pi}}\, q \end{aligned}$$

(b) $l = 1$: **Electric dipole moments**

$$\begin{aligned} q_{11} &= -\sqrt{\frac{3}{8\pi}} \int\int\int_G (x' - iy')\, \rho(\boldsymbol{r}')\, \mathrm{d}V' \\ &= -\sqrt{\frac{3}{8\pi}}\, (p_x - ip_y) \\ q_{10} &= \sqrt{\frac{3}{4\pi}} \int\int\int_G z'\, \rho(\boldsymbol{r}')\, \mathrm{d}V' \\ &= \sqrt{\frac{3}{4\pi}}\, p_z \end{aligned}$$

(c) $l = 2$: **Electric quadrupole moments**

$$q_{22} = \frac{1}{4} \sqrt{\frac{15}{2\pi}} \int\int\int_G (x' - iy')^2\, \rho(\boldsymbol{r}')\, \mathrm{d}V'$$

$$
\begin{aligned}
&= \frac{1}{12}\sqrt{\frac{15}{2\pi}}\,(Q_{xx} - Q_{yy} - i\,2Q_{xy}) \\
q_{21} &= -\sqrt{\frac{15}{8\pi}}\int\int\int_G z'\,(x' - iy')\,\rho(\boldsymbol{r}')\,\mathrm{d}V' \\
&= -\frac{1}{3}\sqrt{\frac{15}{8\pi}}\,(Q_{zx} - iQ_{yz}) \\
q_{20} &= \frac{1}{2}\sqrt{\frac{5}{4\pi}}\int\int\int_G (3z'^2 - r'^2)\,\rho(\boldsymbol{r}')\,\mathrm{d}V' \\
&= \frac{1}{2}\sqrt{\frac{5}{4\pi}}\,Q_{zz}\,.
\end{aligned}
$$

The multipole moments that are *not* listed here, namely the $q_{l,-m}$ for each of $l = 1, 2$, can be easily derived using formula (3.93). Note that at multipole order l there exist, altogether, $2l + 1$ different multipole moments.

We need to explain some of the notation just used in the expressions for the multipole moments of the lowest three orders. As before, the **total charge** contained in G, or, in the present context, the **electric monopole moment scalar** of ρ, is defined as

$$q := \int\int\int_G \rho(\boldsymbol{r}')\,\mathrm{d}V'\,.$$

Next, the **electric dipole moment vector** of ρ is defined by

$$\boldsymbol{p} := \int\int\int_G \boldsymbol{r}'\,\rho(\boldsymbol{r}')\,\mathrm{d}V'\,.$$

And, last, the **electric quadrupole moment tensor** of ρ is defined by

$$\boldsymbol{Q} = 3\int\int\int_G \left(\boldsymbol{r}'\otimes\boldsymbol{r}' - \frac{1}{3}\,(\boldsymbol{r}'\cdot\boldsymbol{r}')\,\mathbf{1}\right)\rho(\boldsymbol{r}')\,\mathrm{d}V'\,;$$

this tensor is symmetric and tracefree. The mathematical symbol "$\otimes$" here denotes the tensor product between any two vector fields on $\mathbb{R}^3$, while $\mathbf{1}$ represents the unit tensor on $\mathbb{R}^3$, which in **Cartesian coordinates** can be written as a diagonal matrix with components $\mathrm{diag}\,(1, 1, 1)$ (i.e., as the unit (3×3)-matrix).

In terms of these quantities, we can alternatively write (3.90) as

$$\phi(\boldsymbol{r}) = \frac{1}{4\pi\epsilon_0}\left[\frac{q}{r} + \frac{\boldsymbol{p}\cdot\boldsymbol{r}}{r^3} + \frac{1}{2}\,\frac{(\boldsymbol{Q}\cdot\boldsymbol{r})\cdot\boldsymbol{r}}{r^5} + \ldots\right]\,. \tag{3.94}$$

With this result, the electric field strength at position $\boldsymbol{r}$ in the exterior of G, due to the charge distribution ρ in G, follows, as usual, from (3.11) as the negative gradient $\boldsymbol{E}(\boldsymbol{r}) = -\,\boldsymbol{\nabla}\phi(\boldsymbol{r})$.

We finally remark that there exists a theorem which states that for an arbitrary localised electric charge distribution ρ the values of the $2l + 1$ moments of the first *non-vanishing* multipole are *independent* of the choice of origin for the

reference frame in use, while the values of all higher multipole moments do in general depend on this choice.

Let us illustrate this property by the following example. Suppose we have within a finite spatial region $G \subset \mathbb{R}^3$ a localised electric charge distribution ρ such that

$$q = 0 \qquad \boldsymbol{p} = \mathbf{0} \,. \tag{3.95}$$

Then upon shifting the origin of our reference frame from $\mathbf{0}$ to a new position $\boldsymbol{a}$, we find that the quadrupole tensor $\tilde{\boldsymbol{Q}}$ defined with respect to position $\boldsymbol{a}$ is given by

$$\begin{aligned}
\tilde{\boldsymbol{Q}} &= 3 \int\int\int_G \left[(\boldsymbol{r}' - \boldsymbol{a}) \otimes (\boldsymbol{r}' - \boldsymbol{a}) - \frac{1}{3} (\boldsymbol{r}' - \boldsymbol{a}) \cdot (\boldsymbol{r}' - \boldsymbol{a}) \, \mathbf{1} \right] \\
&\qquad \times \rho(\boldsymbol{r}' - \boldsymbol{a}) \, \mathrm{d}V' \\
&= \boldsymbol{Q} - 3 \left[\boldsymbol{p} \otimes \boldsymbol{a} - \frac{1}{3} (\boldsymbol{p} \cdot \boldsymbol{a}) \, \mathbf{1} \right] - 3 \left[\boldsymbol{a} \otimes \boldsymbol{p} - \frac{1}{3} (\boldsymbol{a} \cdot \boldsymbol{p}) \, \mathbf{1} \right] \\
&\qquad + 3 \left[\boldsymbol{a} \otimes \boldsymbol{a} - \frac{1}{3} (\boldsymbol{a} \cdot \boldsymbol{a}) \, \mathbf{1} \right] q \\
&= \boldsymbol{Q} \,;
\end{aligned}$$

the last result follows on account of (3.95).

Chapter 4
Magnetostatics

It was indicated before in section 1.2, that **magnetostatics** investigates the special case of **Maxwell's field equations** when all field variables are assumed to be independent of the time coordinate t, and, moreover, we impose the restrictions $\mathbf{0} = \boldsymbol{E} = \boldsymbol{P}$. In addition, we will presently assume also that $\boldsymbol{M} = \mathbf{0}$. Then it follows from (1.15)–(1.18) that

$$\boldsymbol{\nabla} \times \boldsymbol{B} = \mu_0 \boldsymbol{J} \tag{4.1}$$

$$\boldsymbol{\nabla} \cdot \boldsymbol{B} = 0 \,. \tag{4.2}$$

By identity (2.26), we infer from (4.1) the result

$$\boldsymbol{\nabla} \cdot \boldsymbol{J} = 0 \,. \tag{4.3}$$

This also follows from the continuity equation, (1.7), in the case when $\partial\rho/\partial t = 0$, i.e., when electric charges are in steady motion; one commonly refers to the latter phenomenon as **steady currents**. The zero divergence condition on $\boldsymbol{J}$ means that for steady currents the electric charges must flow along paths that close back on themselves so that $\boldsymbol{J}$ does not have sources nor sinks. This is in complete analogy with (1.4) [and (4.2)], which states that, likewise, the magnetic field strength $\boldsymbol{B}$ can only have *closed* field lines. (However, for $\boldsymbol{B}$ this property is generally applicable, i.e., in particular, when time-dependent field configurations are considered.)

As will soon become apparent, there are numerous formal parallels between **electrostatics** and **magnetostatics** (though there also are many differences, that are mainly of a physical nature). This will help us develop a better intuition for the mathematical relations we are about to encounter. For example, (4.1) and (4.2) here do prescribe to us the values of the curl and the divergence of $\boldsymbol{B}$ in terms of a given $\boldsymbol{J}$, as did (3.1) and (3.2) for $\boldsymbol{E}$ in terms of ρ. The imposition of boundary conditions on $\boldsymbol{B}$ will then ensure that solutions we obtain to (4.1) and (4.2) will be unique. We can say that **magnetostatics** is the study of the fields and interactions of steady electric currents and conducting boundaries.

4.1 Ampère's force law and Biot–Savart law

In 1819, the Danish physicist and philosopher Hans Oersted (1777–1851) made the observation that current-carrying wires produce mechanical forces

on permanent magnetic dipoles (such as compass needles) when the latter are placed in the immediate vicinity of the former. The explanation given for this effect was that a current through a wire generates a magnetic field around it which in turn exerts a mechanical torque on nearby permanent magnetic dipoles if they are not aligned with the direction of the magnetic field. A year later, the French physicists Jean-Baptiste Biot (1774–1862) and Félix Savart (1791–1841) were the first to establish a quantitative relationship between the strength of a given steady current and the strength of the magnetic field it generates. These results were systematically refined and extended in a series of experiments by Ampère (1820-1825), who, in particular, showed that there were also mechanical forces acting when two current-carrying thin wires are placed next to each other. Ampère derived a force law that quantified this interaction. It states, e.g., that when the currents through two parallel thin wires flow in the same direction, the mechanical force between them is attractive, but when they flow in opposite directions, the mechanical force is repulsive. We now want to look at these quantitative relations in a little more detail.

Let us consider a thin wire that was placed in a static magnetic field of strength $\boldsymbol{B}$. We assume that through this wire runs a steady electric current of strength I. At a microscopic level, a steady current consists of electric charges q that move with constant velocity $\boldsymbol{v}$. It now follows from **Lorentz's force law** (1.8) (with $\boldsymbol{E} = \boldsymbol{0}$) that each charge q in the wire experiences a mechanical force

$$\boldsymbol{F} = q(\boldsymbol{v} \times \boldsymbol{B}) \,,$$

which acts in a direction that is *perpendicular* to both $\boldsymbol{v}$ and $\boldsymbol{B}$. Now consider a small segment of the wire, the length and direction of which may be represented by an infinitesimal line element $\mathrm{d}\boldsymbol{s}$. In particular, we have $\mathrm{d}\boldsymbol{s} \parallel \boldsymbol{v}$. Suppose the wire contains n charges per unit volume. Then it the mechanical force on the wire line element $\mathrm{d}\boldsymbol{s}$ due to the external magnetostatic field $\boldsymbol{B}$ is given by $\mathrm{d}\boldsymbol{F} = nA|\mathrm{d}\boldsymbol{s}|q(\boldsymbol{v} \times \boldsymbol{B})$, where A denotes the crossectional area of the wire and $nA|\mathrm{d}\boldsymbol{s}|q$ is the total charge contained within the wire line element. But as $\mathrm{d}\boldsymbol{s} \parallel \boldsymbol{v}$, we have $|\mathrm{d}\boldsymbol{s}|\boldsymbol{v} = |\boldsymbol{v}|\mathrm{d}\boldsymbol{s}$, and so $\mathrm{d}\boldsymbol{F} = nAq|\boldsymbol{v}|(\mathrm{d}\boldsymbol{s} \times \boldsymbol{B})$, which is equal to

$$\mathrm{d}\boldsymbol{F} = I\,\mathrm{d}\boldsymbol{s} \times \boldsymbol{B} \,, \tag{4.4}$$

because the strength of the steady current through $\mathrm{d}\boldsymbol{s}$ is $I = nAq|\boldsymbol{v}|$. Using this result, we find that the mechanical force on a closed loop circuit, C, carrying a steady current I, which was placed in an external magnetostatic field $\boldsymbol{B}$, is thus given by

$$\boldsymbol{F} = I \oint_C \mathrm{d}\boldsymbol{s} \times \boldsymbol{B} \,. \tag{4.5}$$

Ampère's experimental results on the mechanical forces between two thin wire loops, C_1 and C_2, carrying, respectively, steady currents I_1 and I_2, are summarised in **Ampère's force law**

$$\boldsymbol{F}_1 = \frac{\mu_0}{4\pi} I_1 I_2 \oint_{C_1} \oint_{C_2} \frac{\mathrm{d}\boldsymbol{s}_1 \times \mathrm{d}\boldsymbol{s}_2 \times \hat{\boldsymbol{e}}}{|\boldsymbol{r}_1 - \boldsymbol{r}_2|^2} = -\,\boldsymbol{F}_2 \,; \tag{4.6}$$

a unit vector has been defined by $\hat{\boldsymbol{e}} := \boldsymbol{r}_1 - \boldsymbol{r}_2/|\boldsymbol{r}_1 - \boldsymbol{r}_2|$. In this expression $\boldsymbol{F}_1$ is the force exerted on the wire loop C_1 due to the current I_2 through the wire loop C_2, which, by Newton's third law, is equal but opposite to the force $\boldsymbol{F}_2$ exerted on the wire loop C_2. Ampère's force law is the analogue to Coulomb's force law (3.3) of electrostatics. It says that the mechanical force acting between two current-carrying wire loops is directly proportional to the product of the current strengths, I_1 and I_2, and inversely proportional to the square of the distance $|\boldsymbol{r}_1 - \boldsymbol{r}_2|$ between the wire loops.

When we compare Ampère's force law (4.6) to the force formula (4.5), it becomes apparent that these two relations imply the result

$$\boldsymbol{B}(\boldsymbol{r}_1) = \frac{\mu_0}{4\pi}\, I_2 \oint_{C_2} \frac{\mathrm{d}\boldsymbol{s}_2 \times \hat{\boldsymbol{e}}}{|\boldsymbol{r}_1 - \boldsymbol{r}_2|^2} \tag{4.7}$$

for the external magnetostatic field that the wire loop C_1 sees as a consequence of the steady current I_2 in C_2. Equation (4.7) is referred to as the **Biot–Savart law**. It is the magnetostatic analogue of (3.4). Unfortunately, it is only of *limited practical* use as, in general, the loop integral is rather difficult to perform; unless the geometry of the loop configuration given is very simple.

In differential form the Biot–Savart law is given by

$$\mathrm{d}\boldsymbol{B}(\boldsymbol{r}) = \frac{\mu_0}{4\pi}\, I\, \frac{\mathrm{d}\boldsymbol{s} \times \hat{\boldsymbol{e}}'}{|\boldsymbol{r} - \boldsymbol{r}'|^2}\,, \tag{4.8}$$

with a unit vector defined by $\hat{\boldsymbol{e}}' := \boldsymbol{r} - \boldsymbol{r}'/|\boldsymbol{r} - \boldsymbol{r}'|$. It expresses the infinitesimal contribution to the magnetostatic field strength by the steady current through an infinitesimal line element of a given current-carrying wire loop.

Before we turn to discuss an application of the Biot–Savart law, let us make the following comments. At a macroscopic level it is adequate to consider *continuous* distributions of electric currents over given volumes and surfaces. Thus, if $(\Delta I)\,\hat{\boldsymbol{e}}$ is the electric current inside a small volume ΔV, or on a small surface ΔA (with $\hat{\boldsymbol{e}}$ giving the direction of the current), then a **volume current density $\boldsymbol{J}$** and a (idealised) **surface current density $\boldsymbol{K}$** are defined by the limits

$$\boldsymbol{J} := \lim_{\Delta V \to 0} \frac{(\Delta I)\,\hat{\boldsymbol{e}}}{\Delta V} \qquad\qquad \boldsymbol{K} := \lim_{\Delta A \to 0} \frac{(\Delta I)\,\hat{\boldsymbol{e}}}{\Delta A}\,, \tag{4.9}$$

respectively.[1] So when, in a finite spatial region $G \subset \mathbb{R}^3$ bounded by a closed surface ∂G, there is given a continuous current distribution of density $\boldsymbol{J}$ in G and density $\boldsymbol{K}$ on ∂G, the magnetic field strength at position $\boldsymbol{r}$ amounts to

$$\boldsymbol{B}(\boldsymbol{r}) = \frac{\mu_0}{4\pi} \int\int\int_G \frac{\boldsymbol{J}(\boldsymbol{r}')}{|\boldsymbol{r} - \boldsymbol{r}'|^2} \times \hat{\boldsymbol{e}}'\, \mathrm{d}V' + \frac{\mu_0}{4\pi} \int\int_{\partial G} \frac{\boldsymbol{K}(\boldsymbol{r}')}{|\boldsymbol{r} - \boldsymbol{r}'|^2} \times \hat{\boldsymbol{e}}'\, \mathrm{d}A'\,, \tag{4.10}$$

where $\hat{\boldsymbol{e}}' := \boldsymbol{r} - \boldsymbol{r}'/|\boldsymbol{r} - \boldsymbol{r}'|$. This is the magnetostatic analogue of (3.7).

[1] We repeat the remark made before in chapter 3, that in reality ΔV and ΔA cannot really go to zero, since eventually ΔI would have to vary discontinuously due to the fact that electric charge in Nature is quantised.

4.1.1 Magnetic field due to current-carrying straight wire

Now let us discuss an application of the Biot–Savart law. Let us determine the magnetic field strength due to a steady current through an infinitely long straight wire. To facilitate matters, we assume that the wire be oriented along the z–axis of a reference frame with **cylindrical polar coordinates**. Then $\boldsymbol{r}$ is an observation point that has a perpendicular distance r from the wire. Let us consider along the wire an infinitesimal line element $\mathrm{d}\boldsymbol{r}'$, the centre of which be at position $\boldsymbol{r}'$. Within the current setup, the position difference vector $\boldsymbol{r} - \boldsymbol{r}'$ makes an angle ϑ with the z–axis.

From (4.8) we have for the infinitesimal contribution to the magnetostatic field strength by the infinitesimal line element

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

We need to find transparent expression for each of the vector-valued quantities on the right-hand side so that we can integrate over all contributions $\mathrm{d}\boldsymbol{B}(\boldsymbol{r})$ along the wire, keeping the observation point $\boldsymbol{r}$ fixed, but varying the position $\boldsymbol{r}'$ of $\mathrm{d}\boldsymbol{r}'$. By the geometry of the configuration, we have $\boldsymbol{r} = r\,\hat{\boldsymbol{e}}_r$, $\boldsymbol{r}' = r'\,\hat{\boldsymbol{e}}_z$ and so $\mathrm{d}\boldsymbol{r}' = \mathrm{d}r'\,\hat{\boldsymbol{e}}_z$. Moreover, a little trigonometry shows that the relations

$$r' = r\,\frac{1}{\tan(\pi - \vartheta)} = -\,r\,\frac{1}{\tan\vartheta} \qquad \Rightarrow \qquad \mathrm{d}r' = \frac{r\,\mathrm{d}\vartheta}{\sin^2\vartheta}\ ,$$

and

$$|\boldsymbol{r} - \boldsymbol{r}'| = r\,\frac{1}{\sin(\pi - \vartheta)} = r\,\frac{1}{\sin\vartheta}\ ,$$

and

$$\hat{\boldsymbol{e}}_z \times (\boldsymbol{r} - \boldsymbol{r}') = |\boldsymbol{r} - \boldsymbol{r}'|\,\sin\vartheta\,\hat{\boldsymbol{e}}_\varphi = r\,\hat{\boldsymbol{e}}_\varphi$$

hold. Thus, with the help of these expressions, performing now the integration along the wire, we find by Biot–Savart for the resultant magnetic field strength

$$\begin{aligned}
\boldsymbol{B}(\boldsymbol{r}) &= \frac{\mu_0}{4\pi}\, I \int_{-\infty}^{\infty} \frac{\mathrm{d}r'\,\hat{\boldsymbol{e}}_z \times (\boldsymbol{r} - \boldsymbol{r}')}{|\boldsymbol{r} - \boldsymbol{r}'|^3} \\
&= \frac{\mu_0}{4\pi}\,\frac{I}{r}\left[\int_0^{\pi} \sin\vartheta\,\mathrm{d}\vartheta\right]\,\hat{\boldsymbol{e}}_\varphi \\
&= \frac{\mu_0}{4\pi}\,\frac{2I}{r}\,\hat{\boldsymbol{e}}_\varphi\ .
\end{aligned}$$

We see that the lines of constant $\boldsymbol{B}$ are concentric circles that are centred on the z–axis (the wire). Their orientation is clockwise with respect to the direction in which the current is flowing. We observe that the magnitude of $\boldsymbol{B}$ falls off inversely proportional to r.

4.1.2 Ampère's law

Ampère's law of **magnetostatics** states that the circulation of a given magnetic field strength along a given closed oriented curve ∂S is proportional to

the total steady current flowing through the surface S spanned by ∂S, i.e.,

$$\oint_{\partial S} \boldsymbol{B} \cdot \mathrm{d}\boldsymbol{s} = \mu_0 \int\int_S \boldsymbol{J} \cdot \boldsymbol{n}\, \mathrm{d}A = \mu_0\, I(S)\ . \tag{4.11}$$

This provides an immediate analogue to Gauß' law, (3.8), of electrostatic. The difference, of course, is that volumes and their bounding closed surfaces in that law are here replaced by surfaces and their bounding closed curves.

Let us now use (4.11) to derive once more the magnitude of the magnetic field strength of a long straight wire carrying a steady electric current. As before, the wire be oriented along the z–axis of a reference frame with **cylindrical polar coordinates**. By the azimuthal symmetry of this configuration and the translational symmetry along the wire, $\boldsymbol{B}$ cannot depend on the coordinates φ and z. Likewise, it cannot depend on the directions of $\hat{\boldsymbol{e}}_r$ and $\hat{\boldsymbol{e}}_z$. We thus have $\boldsymbol{B} = B(r)\, \hat{\boldsymbol{e}}_\varphi$. Upon introduction of a "*Stokesian circle*" that is centred on the wire, of radius r and with tangent $\mathrm{d}\boldsymbol{s} = r\, \mathrm{d}\varphi\, \hat{\boldsymbol{e}}_\varphi$, we find

$$\oint_{\text{circle}} \boldsymbol{B} \cdot \mathrm{d}\boldsymbol{s} = \int_0^{2\pi} B(r)\, r\, \mathrm{d}\varphi = 2\pi\, B(r)\, r\ = \mu_0\, I\ ,$$

i.e.,

$$B(r) = \frac{\mu_0}{4\pi}\, \frac{2I}{r}\ .$$

As it turns out, for this example the approach via Ampère's law to obtain $B(r)$ proves much more straightforward than the route via the Biot–Savart law.

Finally, we want to take a brief look at a simple numerical example. Suppose we have a steady current of strength $I = 1\,\frac{\mathrm{C}}{\mathrm{s}}$ flowing through the straight wire of the above-mentioned configuration. Then, as can be easily verified, the magnitude of the resultant magnetic field strength at $1\,\mathrm{m}$ distance from the wire is $|\boldsymbol{B}_{1\mathrm{C/s}}|_{r=1\mathrm{m}} = 2 \times 10^{-7}\, \frac{\mathrm{kg}}{\mathrm{s\,C}}$.

4.2 Magnetostatic vector potential

Let us rewrite (4.10), using (3.9), to obtain

$$\boldsymbol{B}(\boldsymbol{r}) = \frac{\mu_0}{4\pi}\, \boldsymbol{\nabla} \times \int\int\int_G \frac{\boldsymbol{J}(\boldsymbol{r}')}{|\boldsymbol{r} - \boldsymbol{r}'|}\, \mathrm{d}V' + \frac{\mu_0}{4\pi}\, \boldsymbol{\nabla} \times \int\int_{\partial G} \frac{\boldsymbol{K}(\boldsymbol{r}')}{|\boldsymbol{r} - \boldsymbol{r}'|}\, \mathrm{d}A'\ . \tag{4.12}$$

It is clear at this stage that, by identity (2.26), this expression does immediately satisfy the magnetostatic condition (4.2), namely $\boldsymbol{\nabla} \cdot \boldsymbol{B} = \boldsymbol{0}$. This is because we have found that the magnetostatic field strength $\boldsymbol{B}(\boldsymbol{r})$ can be obtained from a, by assumption, continuously differentiable vector field $\boldsymbol{A}(\boldsymbol{r})$ as

$$\boldsymbol{B}(\boldsymbol{r}) = \boldsymbol{\nabla} \times \boldsymbol{A}(\boldsymbol{r})\ . \tag{4.13}$$

This vector field is called the **magnetostatic vector potential**; it has SI unit $1\,\frac{\mathrm{kg\,m}}{\mathrm{s\,C}}$. Because of the vector analytical curl-derivative, the directions of $\boldsymbol{B}$ and $\boldsymbol{A}$ will always be *perpendicular* to each other.

Combining (4.12) and (4.13), we thus have that

$$\boldsymbol{A}(\boldsymbol{r}) = \frac{\mu_0}{4\pi}\int\int\int_G \frac{\boldsymbol{J}(\boldsymbol{r}')}{|\boldsymbol{r}-\boldsymbol{r}'|}\,\mathrm{d}V' + \frac{\mu_0}{4\pi}\int\int_{\partial G}\frac{\boldsymbol{K}(\boldsymbol{r}')}{|\boldsymbol{r}-\boldsymbol{r}'|}\,\mathrm{d}A' + \boldsymbol{\nabla}\chi(\boldsymbol{r})\;, \tag{4.14}$$

i.e., by the vector analytical identity (2.25), the magnetostatic vector potential is determined only up to the gradient of an arbitrary twice continuously differentiable real-valued function $\chi(\boldsymbol{r})$.

4.2.1 Gauge freedom

In real applications, the *non-uniqueness* in the definition of $\boldsymbol{A}$ according to (4.13) can be exploited to one's advantage. One generally speaks of the **gauge freedom** associated with the choice of $\boldsymbol{A}$.

A **gauge transformation** from a magnetostatic vector potential $\boldsymbol{A}$ to a magnetostatic vector potential $\tilde{\boldsymbol{A}}$, which both result in the *same* magnetostatic field strength $\boldsymbol{B}$, is defined by

$$\tilde{\boldsymbol{A}} := \boldsymbol{A} + \boldsymbol{\nabla}\chi\;; \tag{4.15}$$

the real-valued function $\chi(\boldsymbol{r})$ is arbitrary and twice continuously differentiable. It follows from (4.15) that

$$\boldsymbol{\nabla}\cdot\tilde{\boldsymbol{A}} = \boldsymbol{\nabla}\cdot\boldsymbol{A} + (\boldsymbol{\nabla}\cdot\boldsymbol{\nabla})\chi\;;$$

so, e.g., by choosing χ such that it satisfies

$$(\boldsymbol{\nabla}\cdot\boldsymbol{\nabla})\chi = -\,\boldsymbol{\nabla}\cdot\boldsymbol{A}\;,$$

we can always introduce a *special* magnetostatic vector potential $\tilde{\boldsymbol{A}}$ that has the property

$$\boldsymbol{\nabla}\cdot\tilde{\boldsymbol{A}} = 0\;. \tag{4.16}$$

This is referred to as the **Coulomb gauge condition**.

With (4.13), the magnetostatic field equation (4.1) converts to

$$\boldsymbol{\nabla}\times\boldsymbol{\nabla}\times\boldsymbol{A}(\boldsymbol{r}) = \mu_0\,\boldsymbol{J}(\boldsymbol{r})\;, \tag{4.17}$$

providing an analogue of **Poisson's equation** (3.17) in electrostatics. In a reference frame with **Cartesian coordinates** (and only for such coordinates), by using the vector analytical identity (2.31) and by imposing the Coulomb gauge condition (4.16) on $\boldsymbol{A}$ (dropping the "tilde"), this reduces to

$$(\boldsymbol{\nabla}\cdot\boldsymbol{\nabla})\boldsymbol{A}(\boldsymbol{r}) = -\,\mu_0\,\boldsymbol{J}(\boldsymbol{r})\;. \tag{4.18}$$

This is the central field equation we will have to solve. In fact, it provides *three* **Poisson's equations** for the Cartesian components A_x, A_y and A_z of $\boldsymbol{A}$, with the Cartesian components J_x, J_y and J_z of $\boldsymbol{J}$ as respective sources. Hence, we can apply all the knowledge on solving **Poisson's equation** gained in chapter 3.

Let us discuss a little more the notion of the magnetostatic vector potential. As a simple illustration, let us work out $\boldsymbol{A}$ for a *uniform* magnetostatic field

as generated, e.g., by a very long coil (solenoid) that is carrying a steady current. We assume that the magnetostatic field is oriented along the z–axis of a reference frame with **cylindrical polar coordinates** so that $\boldsymbol{B} = B_0\,\hat{\boldsymbol{e}}_z$ with $B_0 = \text{const}$. Now consider a circular closed oriented curve ∂S of radius r in the plane perpendicular to and centred on the z–axis; the sense of orientation on ∂S be given by the right-hand convention. Then, by Stokes' integral theorem (1.33), the circulation of $\boldsymbol{A}$ along ∂S is related to the flux of $\boldsymbol{\nabla} \times \boldsymbol{A}$ through the surface S spanned by ∂S according to

$$\oint_{\partial S} \boldsymbol{A} \cdot \mathrm{d}\boldsymbol{s} = \int\!\!\int_S (\boldsymbol{\nabla} \times \boldsymbol{A}) \cdot \boldsymbol{n}\,\mathrm{d}A\,.$$

But this is just

$$\oint_{\partial S} \boldsymbol{A} \cdot \mathrm{d}\boldsymbol{s} = \int\!\!\int_S \boldsymbol{B} \cdot \boldsymbol{n}\,\mathrm{d}A\,,$$

using (4.13). By the axial symmetry of the present configuration and its independence of position along the z–axis, as well as the fact that $\boldsymbol{A}$ is always perpendicular to $\boldsymbol{B}$, it follows that both $\boldsymbol{A}$ and the line element $\mathrm{d}\boldsymbol{s}$ must be parallel to $\hat{\boldsymbol{e}}_\varphi$. Hence, we obtain

$$\int_0^{2\pi} A(r)\,r\,\mathrm{d}\varphi = 2\pi r\,A(r) = B_0 \int_0^{2\pi}\!\!\int_0^r r'\,\mathrm{d}r'\,\mathrm{d}\varphi = B_0\,\pi r^2\,.$$

This yields for the magnitudes of $\boldsymbol{A}$ and $\boldsymbol{B}$

$$A(r) = \tfrac{1}{2}\,B_0\,r\,,$$

and, adding in the directional information that is quite simple in the current axially symmetric case,

$$\boldsymbol{A}(\boldsymbol{r}) = A(r)\,\hat{\boldsymbol{e}}_\varphi = \tfrac{1}{2}\,(B_0 r)\,(\hat{\boldsymbol{e}}_z \times \hat{\boldsymbol{e}}_r)\,. \tag{4.19}$$

This ends the present example.

4.3 Remark on boundary conditions

To obtain unique solutions, the magnetostatic field equations (4.1) and (4.2) need to be supplemented by appropriate boundary conditions. Starting from their respective integral forms (4.11) and (1.49), by introducing a small "*Gaußian pillbox*" and a small "*Stokesian loop*" that both penetrate an interface which separates a region 1 of space from a region 2 and which carries an idealised **electric surface current density** $\boldsymbol{K}$, one can show that these boundary conditions must take the form

$$\boldsymbol{n} \cdot (\boldsymbol{B}_2 - \boldsymbol{B}_1) = 0\,, \qquad \boldsymbol{n} \times (\boldsymbol{B}_2 - \boldsymbol{B}_1) = \mu_0\,\boldsymbol{K}\,. \tag{4.20}$$

Here $\boldsymbol{n}$ denotes the unit normal to the interface pointing from region 1 into region 2. The boundary conditions thus state that at an interface that separates two regions of space the **normal component** of the magnetostatic field

strength must be *continuous* while the **tangential components** may be *discontinuous*.

One finds that the *same* boundary conditions also apply to *time-dependent fields* which interact with continuous material media of *linear* magnetic properties for which $\boldsymbol{M} = \boldsymbol{0}$.

In regions of space where the steady current density *vanishes* such that

$$\boldsymbol{J}(\boldsymbol{r}) = \boldsymbol{0} \quad \Rightarrow \quad \boldsymbol{\nabla} \times \boldsymbol{B}(\boldsymbol{r}) = \boldsymbol{0}$$

applies, a **magnetostatic scalar potential** ψ may be introduced according to

$$\boldsymbol{B}(\boldsymbol{r}) = -\,\boldsymbol{\nabla}\psi(\boldsymbol{r})\ . \tag{4.21}$$

Then, combined with (4.2), ψ is determined by solving **Laplace's equation**

$$(\boldsymbol{\nabla}\cdot\boldsymbol{\nabla})\psi(\boldsymbol{r}) = 0\ , \tag{4.22}$$

subject to boundary conditions (4.20).

4.4 Energy density of magnetostatic fields

The **total work** to be done against the Ampère forces to assemble a configuration of n **rigid single loop circuits**, carrying steady currents I_i, at positions $\boldsymbol{r}_i$ is given by [cf. (3.39)]

$$W = \frac{1}{2}\sum_{i=1}^{n} I_i\,\Phi_i\ , \tag{4.23}$$

where Φ_i denotes the magnetic flux through the loop at $\boldsymbol{r}_i$ due to the field generated by the remaining $n-1$ steady currents. For the latter we can write

$$\Phi_i = \int\!\!\int_{S_i} \boldsymbol{B}\cdot\boldsymbol{n}\,\mathrm{d}A = \oint_{\partial S_i} \boldsymbol{A}\cdot\mathrm{d}\boldsymbol{s}_i\ ;$$

hence, we obtain

$$W = \frac{1}{2}\sum_{i=1}^{n} I_i \oint_{\partial S_i} \boldsymbol{A}\cdot\mathrm{d}\boldsymbol{s}_i\ . \tag{4.24}$$

The transition to a **continuous current distribution** with compact support in $\mathbb{R}^3$ (i.e., $\boldsymbol{J}$ falls off to zero sufficiently fast towards spatial infinity) is obtained as follows. First, we assume that in the present problem $\boldsymbol{A}$ satisfies the boundary condition

$$\lim_{|\boldsymbol{r}|\to\infty} \boldsymbol{A}(\boldsymbol{r}) = 0\ . \tag{4.25}$$

Then, formally, the total work to be done to assemble the continuous current distribution is obtained from (4.24) by replacing

$$I_i\,\mathrm{d}\boldsymbol{s}_i \longrightarrow \boldsymbol{J}(\boldsymbol{r})\,\mathrm{d}V\ ,$$

and converting

$$\sum_{i=1}^{n}\oint_{\partial S_i} \longrightarrow \int\!\!\int\!\!\int_{\mathbb{R}^3}\ .$$

We thus obtain

$$W = \tfrac{1}{2} \int\int\int_{\mathbb{R}^3} \boldsymbol{J}(\boldsymbol{r}) \cdot \boldsymbol{A}(\boldsymbol{r}) \, \mathrm{d}V \; . \tag{4.26}$$

This result expresses the **magnetostatic potential energy** in terms of the Ampère interactions within the continuous current distribution (as does (4.23) in the case of discrete currents). It provides the magnetostatic analogue to (3.40).

An alternative viewpoint sees the potential energy as being stored in the magnetostatic field surrounding the current distribution. Here we substitute for $\boldsymbol{J}(\boldsymbol{r})$ in (4.26) from the magnetostatic field equation (4.1) to obtain

$$W = \frac{1}{2\mu_0} \int\int\int_{\mathbb{R}^3} \boldsymbol{A}(\boldsymbol{r}) \cdot [\, \boldsymbol{\nabla} \times \boldsymbol{B}(\boldsymbol{r}) \,] \, \mathrm{d}V \; .$$

By the vector analytical identity (2.28), this is equivalent to

$$W = \frac{1}{2\mu_0} \int\int\int_{\mathbb{R}^3} \boldsymbol{B}(\boldsymbol{r}) \cdot [\, \boldsymbol{\nabla} \times \boldsymbol{A}(\boldsymbol{r}) \,] \, \mathrm{d}V - \frac{1}{2\mu_0} \int\int\int_{\mathbb{R}^3} \boldsymbol{\nabla} \cdot [\, \boldsymbol{A}(\boldsymbol{r}) \times \boldsymbol{B}(\boldsymbol{r}) \,] \, \mathrm{d}V \; .$$

Then, by (4.13), and employing Gauß' integral theorem (1.32), we find

$$W = \frac{1}{2\mu_0} \int\int\int_{\mathbb{R}^3} |\, \boldsymbol{B}(\boldsymbol{r}) \,|^2 \, \mathrm{d}V - \frac{1}{2\mu_0} \int\int_{\partial\mathbb{R}^3} [\, \boldsymbol{A}(\boldsymbol{r}) \times \boldsymbol{B}(\boldsymbol{r}) \,] \cdot \boldsymbol{n}(\boldsymbol{r}) \, \mathrm{d}A \; .$$

Finally, with the boundary condition (4.25), we have

$$W = \frac{1}{2\mu_0} \int\int\int_{\mathbb{R}^3} |\, \boldsymbol{B}(\boldsymbol{r}) \,|^2 \, \mathrm{d}V \; , \tag{4.27}$$

which says the potential energy is entirely carried by the magnetostatic field $\boldsymbol{B}(\boldsymbol{r})$, with a (positive definite) energy density in space given by

$$u(\boldsymbol{r}) = \frac{1}{2\mu_0} \, |\, \boldsymbol{B}(\boldsymbol{r}) \,|^2 \; . \tag{4.28}$$

Equations (4.27) and (4.28) are the magnetostatic analogues to (3.41) and (3.42) [also compare this expression to the magnetostatic subcase of (1.39), i.e., $u = (\boldsymbol{B} \cdot \boldsymbol{B})/(2\mu_0)$.] We conclude that, as a carrier of energy, clearly *physical reality* must be ascribed to a magnetostatic field (and all electromagnetic fields in general).

Chapter 5
Time-dependent electric and magnetic fields

5.1 Energy and linear momentum balance equations for electromagnetic fields

We already derived the **energy balance equation** for electromagnetic fields interacting with charged and current-carrying macroscopic material media in subsection 1.3.3. For continuous media that are *linear* in their electric and magnetic properties and that have negligible dispersion and losses so that **Maxwell's field equations** in the form (1.15)–(1.18) apply, it is given by **Poynting's theorem** as

$$\frac{\partial u}{\partial t} + \boldsymbol{\nabla} \cdot \boldsymbol{S} = - c \left[\boldsymbol{J} + \frac{\partial (c\boldsymbol{P})}{\partial ct} + \boldsymbol{\nabla} \times \boldsymbol{M} \right] \cdot \boldsymbol{E}/c \,, \tag{5.1}$$

with

$$u := \frac{1}{2\mu_0} \left(\boldsymbol{E}/c \cdot \boldsymbol{E}/c + \boldsymbol{B} \cdot \boldsymbol{B} \right) \qquad \boldsymbol{S} := \frac{c}{\mu_0} \left(\boldsymbol{E}/c \times \boldsymbol{B} \right) \tag{5.2}$$

denoting the scalar-valued **energy density** and vector-valued **energy current density** of electromagnetic fields, respectively. The terms on the right-hand side of (5.1) make up the **energy generation rate density** for electromagnetic fields, which describes the work done per unit volume per unit time by electromagnetic fields on the charges and currents present within any given macroscopic material media. Due to the minus sign, this quantity represents the energy *outflow* per unit volume per unit time from a given time-dependent electromagnetic field into a given macroscopic material medium [cf. (1.36)]. The energy balance equation describes the conversion of **electromagnetic energy** into **mechanical** and/or **thermal energy**.

Now let us find out what the respective balance equation for the **linear momentum** carried by electromagnetic fields looks like. This requires a few subtle mathematical manipulations of **Maxwell's field equations**. Again, let us consider them in the form (1.15)–(1.18) which applies for continuous media that are linear in their electric and magnetic properties and that have negligible dispersion and losses. Let us first vector-multiply (1.15) from the left with $(-\boldsymbol{B})$ and (1.16) from the left with $\boldsymbol{E}/c$. Then we multiply (1.17) by $(-\boldsymbol{E}/c)$ and (1.18) by $(-\boldsymbol{B})$. Now let us add all four resulting equations, collect to-

gether similar terms, rearrange them, and divide by μ_0. We thus obtain

$$\begin{aligned}
&\frac{1}{\mu_0}\,\frac{\partial}{\partial ct}(\boldsymbol{E}/c\times\boldsymbol{B})\\
&-\frac{1}{\mu_0}\,[\,\boldsymbol{E}/c\,(\boldsymbol{\nabla}\cdot\boldsymbol{E}/c)-\boldsymbol{E}/c\times(\boldsymbol{\nabla}\times\boldsymbol{E}/c)+\boldsymbol{B}\,(\boldsymbol{\nabla}\cdot\boldsymbol{B})-\boldsymbol{B}\times(\boldsymbol{\nabla}\times\boldsymbol{B})\,]\\
&\qquad\qquad=-\,[\,c\rho-\boldsymbol{\nabla}\cdot(c\boldsymbol{P})\,]\,\boldsymbol{E}/c-\left[\,\boldsymbol{J}+\frac{\partial(c\boldsymbol{P})}{\partial ct}+\boldsymbol{\nabla}\times\boldsymbol{M}\,\right]\times\boldsymbol{B}\,,
\end{aligned}$$

which already has effectively the form (1.36) of a balance equation as introduced in subsection 1.3.2. This is so because by defining the vector-valued **linear momentum density** of electromagnetic fields as

$$\boldsymbol{g}:=\frac{1}{\mu_0 c}\,(\boldsymbol{E}/c\times\boldsymbol{B})=\frac{\boldsymbol{S}}{c^2}\,, \tag{5.3}$$

and the tensor-valued **linear momentum current density** of electromagnetic fields as

$$\boldsymbol{T}:=-\frac{1}{\mu_0}\,(\boldsymbol{E}/c\otimes\boldsymbol{E}/c+\boldsymbol{B}\otimes\boldsymbol{B})+\frac{1}{2\mu_0}\,(\boldsymbol{E}/c\cdot\boldsymbol{E}/c+\boldsymbol{B}\cdot\boldsymbol{B})\,\mathbf{1}\,, \tag{5.4}$$

it can be shown that we can rewrite the **linear momentum balance equation** in the explicit form

$$\boxed{\frac{\partial\boldsymbol{g}}{\partial t}+\boldsymbol{\nabla}\cdot\boldsymbol{T}=-\,[\,c\rho-\boldsymbol{\nabla}\cdot(c\boldsymbol{P})\,]\,\boldsymbol{E}/c-\left[\,\boldsymbol{J}+\frac{\partial(c\boldsymbol{P})}{\partial ct}+\boldsymbol{\nabla}\times\boldsymbol{M}\,\right]\times\boldsymbol{B}\,.} \tag{5.5}$$

The terms on the right-hand side of (5.5) make up the **linear momentum generation rate density** for electromagnetic fields, which, in more familiar terms, describes the **mechanical forces** per unit volume exerted by electromagnetic fields on the charges and currents present within any given macroscopic material media. Due to the minus sign, this quantity represents the linear momentum *outflow* per unit volume per unit time from a given time-dependent electromagnetic field into a given macroscopic material medium [cf. (1.36)]. Note that the balance equations (5.5) and (5.1) state that an outflow of linear momentum from an electromagnetic field into a macroscopic material medium is intimately coupled to an analogous outflow of energy.

As an example, let us consider the (highly idealised) situation of a continuous medium with $\mathbf{0}=\boldsymbol{P}=\boldsymbol{M}$ and an electric current density that is purely *convective*, i.e., the charges in the medium move with a velocity $\boldsymbol{v}$ so that $\boldsymbol{J}=\rho\,\boldsymbol{v}$. Then the terms on the right-hand side of (5.5) reduce to

$$-\,c\rho\,\left[\,\boldsymbol{E}/c+\left(\frac{\boldsymbol{v}}{c}\right)\times\boldsymbol{B}\,\right]\,.$$

By (1.8), we recognise this as the negative of a Lorentz force per unit volume. It is the action of such a Lorentz force what is effectively happening when an electromagnetic signal is incident on the reception antenna of a radio or TV

set or a mobile phone. Energy and linear momentum are transferred from the electromagnetic signal to the electrons in the antenna at a rate of, respectively, $c\rho\,(\boldsymbol{v}/c)\cdot\boldsymbol{E}$ and $c\rho\,[\,\boldsymbol{E}/c+(\boldsymbol{v}/c)\times\boldsymbol{B}\,]$. On the basis of such everyday experiences, almost everyone would agree on that as electromagnetic fields manifestly carry energy and linear momentum (as well as angular momentum), this provides compelling evidence for the physical reality of these fields.

Historically the tensor-valued quantity defined in (5.4) is known as the **Maxwell stress tensor** of electromagnetic fields (the mathematical symbol "$\otimes$" denotes the tensor product between vector fields in $\mathbb{R}^3$, $\mathbf{1}$ represents the unit tensor in $\mathbb{R}^3$). Its physical dimension is [lin. momentum]/[area]/[time] which is equal to [force]/[area] which in turn is equal to [pressure]; its SI unit is $1\ \frac{\mathrm{kg}}{\mathrm{m\,s^2}}$. The Maxwell stress tensor expresses the fact that there exist **mechanical tensions** in electromagnetic fields along their field lines (the *longitudinal* direction) and **mechanical pressures** perpendicular to their field lines (the *transverse* direction).

Let us illustrate this property by the following simple example taken from magnetostatics. Let us look at a homogeneous magnetostatic field oriented in a reference frame with **Cartesian coordinates** such that

$$\boldsymbol{B} = B_0\,\boldsymbol{e}_z \qquad B_0 = \text{const}\,.$$

Then, from (5.4), we find that the Maxwell stress tensor for this magnetostatic field is given by

$$\boldsymbol{T} = \frac{1}{2\mu_0}\begin{bmatrix} B_0^2 & 0 & \\ 0 & B_0^2 & 0 \\ 0 & 0 & -B_0^2 \end{bmatrix}\,.$$

Clearly, the component of $\boldsymbol{T}$ that is purely longitudinal to the field direction is negative, $T_{zz} = -B_0^2/(2\mu_0)$, which means there resides a "negative pressure" along the field lines, conventionally referred to as a mechanical tension. On the other hand, the components of $\boldsymbol{T}$ that are purely transverse to the field direction are positive, $T_{xx} = T_{yy} = B_0^2/(2\mu_0)$, indicating the existence of mechanical pressures (repulsion) between neighbouring field lines.

Let us conclude this section with a consideration of some wider focus. We realise by comparing definitions (5.2) and (5.3) that Maxwell's theory of **electromagnetism** teaches us the relation

$$\left(\begin{array}{c}\textbf{Energy current density}\\ \text{of electromagnetic fields}\end{array}\right) = \left(\begin{array}{c}\textbf{Linear momentum density}\\ \text{of electromagnetic fields}\end{array}\right)\times c^2\,,$$

where $c^2 := 1/(\epsilon_0\mu_0)$. On the other hand, we find in the Newtonian theory of the **mechanics of continuous material media** that (cf., e.g., the lectures on MAS209 Fluid Dynamics)

$$\begin{aligned}\left(\begin{array}{c}\textbf{Linear momentum density}\\ \text{of matter}\end{array}\right) &= \left(\begin{array}{c}\textbf{Mass density}\\ \text{of matter}\end{array}\right)\times(\text{velocity})\\ &= \left(\begin{array}{c}\textbf{Mass current density}\\ \text{of matter}\end{array}\right)\,.\end{aligned}$$

Thus, if, by brute force, one wanted to formally adapt the Newtonian theory of **continuous material media** to Maxwell's theory of **electromagnetism**, by necessity the relation

$$(\textbf{Energy current density}) = (\textbf{Mass current density}) \times c^2 \qquad (5.6)$$

had to be satisfied. In cases with vanishing pressure this could be reinterpreted in terms of the Einsteinian relation

$$(\textbf{Energy}) = (\textbf{Mass}) \times c^2 \, . \qquad (5.7)$$

5.2 Inhomogeneous wave equations

The most striking feature of Maxwell's theory of **electromagnetism**, besides its impressive unification of all electric and magnetic phenomena known at the time, was Maxwell's prediction in 1864 of the existence of **electromagnetic waves** which, in vacuum, should propagate at the **speed of light**. The existence of electromagnetic waves was confirmed about two decades later in 1888 in an ingenious set of experiments by Hertz. It is apparent to what extent these discoveries have had an influence on the development of the "civilised world" ever since.

Let us briefly look at how **Maxwell's field equations** lead to wave propagation equations. For this purpose we will again resort to the form (1.15)–(1.18) of this set of coupled first-order partial differential equations that holds for continuous media that are linear in their electric and magnetic properties and that have negligible dispersion and losses. By taking derivatives with respect to $\partial/\partial ct$ of each of (1.15) and (1.16), and then substituting back in from the other one to evaluate the mixed time–space derivatives that arise, we get

$$\begin{aligned} \frac{\partial^2 \boldsymbol{E}/c}{\partial (ct)^2} + \boldsymbol{\nabla} \times \boldsymbol{\nabla} \times \boldsymbol{E}/c &= -\mu_0 \left[\frac{\partial \boldsymbol{J}}{\partial ct} + \frac{\partial^2 (c\boldsymbol{P})}{\partial (ct)^2} + \frac{\partial (\boldsymbol{\nabla} \times \boldsymbol{M})}{\partial ct} \right] \\ \frac{\partial^2 \boldsymbol{B}}{\partial (ct)^2} + \boldsymbol{\nabla} \times \boldsymbol{\nabla} \times \boldsymbol{B} &= \mu_0 \left[\boldsymbol{\nabla} \times \boldsymbol{J} + \frac{\partial [\boldsymbol{\nabla} \times (c\boldsymbol{P})]}{\partial ct} + \boldsymbol{\nabla} \times \boldsymbol{\nabla} \times \boldsymbol{M} \right] . \end{aligned}$$

If we further assume the linear constitutive relation

$$\boldsymbol{J} = \sigma \, \boldsymbol{E} \, , \qquad \sigma = \text{const} \, ,$$

to hold for the electric current density [cf. (1.14)], and we introduce a reference frame with **Cartesian coordinates** so that the vector analytical identity (2.31) can be applied, we find for $\boldsymbol{E}/c$ the equation

$$\begin{aligned} &- \frac{\partial^2 \boldsymbol{E}/c}{\partial (ct)^2} + (\boldsymbol{\nabla} \cdot \boldsymbol{\nabla}) \boldsymbol{E}/c - \mu_0 c \sigma \, \frac{\partial \boldsymbol{E}/c}{\partial ct} \\ &= \mu_0 \, \boldsymbol{\nabla} (c\rho) - \mu_0 \left[- \frac{\partial^2 (c\boldsymbol{P})}{\partial (ct)^2} + (\boldsymbol{\nabla} \cdot \boldsymbol{\nabla}) (c\boldsymbol{P}) \right] + \mu_0 \, \boldsymbol{\nabla} \times \left(\frac{\partial \boldsymbol{M}}{\partial ct} \right) \, , \end{aligned} \qquad (5.8)$$

and for $\boldsymbol{B}$ the equation

$$-\frac{\partial^2 \boldsymbol{B}}{\partial (ct)^2} + (\boldsymbol{\nabla}\cdot\boldsymbol{\nabla})\boldsymbol{B} - \mu_0 c\sigma \frac{\partial \boldsymbol{B}}{\partial ct} = -\mu_0 \left[\boldsymbol{\nabla} \times \left(\frac{\partial (c\boldsymbol{P})}{\partial ct} \right) + \boldsymbol{\nabla}(\boldsymbol{\nabla}\cdot\boldsymbol{M}) - (\boldsymbol{\nabla}\cdot\boldsymbol{\nabla})\boldsymbol{M} \right] . \quad (5.9)$$

Both relations constitute linear **inhomogeneous wave equations** for the Cartesian components of $\boldsymbol{E}/c$ and $\boldsymbol{B}$, with "driving terms" given by the changes in time and space of each of the electromagnetic sources $c\rho$, $\boldsymbol{J}$, $c\boldsymbol{P}$ and $\boldsymbol{M}$.

In the rest of this course we will assume (highly idealised) continuous material media for which

$$\mathbf{0} = c\boldsymbol{P} = \boldsymbol{M} . \quad (5.10)$$

5.3 Plane wave solutions in vacuum

In a region of space exterior to (and, let as also assume, far away from) some time-dependent electromagnetic sources we have $c\rho = 0$ and $\boldsymbol{J} = \mathbf{0}$, thus, with also (5.10) holding, providing conditions that we say constitute a **vacuum**. **Maxwell's field equations** are then given by (1.26)–(1.29). The vacuum constraint equations (1.28) and (1.29) yield for both the electric and magnetic field strengths the conditions that they have to have zero divergence, i.e., they have to have *closed* field lines. Choosing a reference frame with **Cartesian coordinates** so that the vector analytical identity (2.31) holds, $\boldsymbol{E}/c$ and $\boldsymbol{B}$ are now to be determined as solutions of the **homogeneous wave equations** [cf. (1.30) and (1.31)]

$$-\frac{\partial^2 \boldsymbol{E}/c}{\partial (ct)^2} + (\boldsymbol{\nabla}\cdot\boldsymbol{\nabla})\boldsymbol{E}/c = \mathbf{0} \quad (5.11)$$

$$-\frac{\partial^2 \boldsymbol{B}}{\partial (ct)^2} + (\boldsymbol{\nabla}\cdot\boldsymbol{\nabla})\boldsymbol{B} = \mathbf{0} . \quad (5.12)$$

Let us now *assume* the existence of **plane wave solutions** to these equations. This is to say we want to assume that we have no dependence of the fields $\boldsymbol{E}/c$ and $\boldsymbol{B}$ on the spatial coordinates x and y so that

$$\boldsymbol{E}/c = \boldsymbol{E}(ct, z)/c \qquad \boldsymbol{B} = \boldsymbol{B}(ct, z) .$$

Then the vacuum constraint equations (1.28) and (1.29) reduce presently to

$$\frac{\partial E_z/c}{\partial z} = 0 \qquad \frac{\partial B_z}{\partial z} = 0 .$$

On the other hand, the vacuum time evolution equations for E_z/c and B_z become [cf. (1.26) and (1.27)], on account of the assumed independence of the fields from the spatial coordinates x and y,

$$\frac{\partial E_z/c}{\partial ct} = 0 \qquad \frac{\partial B_z}{\partial ct} = 0 .$$

Hence, without loss of generality, we can set

$$E_z/c = 0 \qquad B_z = 0 \, .$$

This result leaves us with each of the Cartesian components E_x/c, E_y/c, B_x and B_y to be determined individually as solutions of a homogeneous wave equation in the coordinates ct and z. Note, however, that by (1.26) and (1.27) these components are not independent from each other. Let us start by looking at the solutions for E_x/c and E_y/c. First, however, we want to take a slight detour.

The **general solution** for a real-valued function $f = f(ct, z)$ that is twice continuously differentiable in ct and z and that satisfies the linear **homogeneous wave equation**

$$-\frac{\partial^2 f}{\partial (ct)^2} + \frac{\partial^2 f}{\partial z^2} = 0 \tag{5.13}$$

is given by

$$f(ct, z) = f_1(z - ct) + f_2(z + ct) \, . \tag{5.14}$$

This solution is interpreted as a **linear superposition** of one plane wave front propagating at speed c in the *positive* z–direction (represented by f_1) and a second plane wave front propagating at speed c in the *negative* z–direction (represented by f_2). As both f_1 and f_2 do *not* change their initial shape in the process of propagation, it follows that these plane wave solutions are in fact **non-dispersive**!

That (5.14) does indeed provide the general solution to (5.13) is most clearly seen as follows. Let us first rewrite (5.13) by factorising the partial derivatives acting on f so that the equation assumes the alternative form

$$-\left(\frac{\partial}{\partial ct} + \frac{\partial}{\partial z}\right)\left(\frac{\partial}{\partial ct} - \frac{\partial}{\partial z}\right) f = -\left(\frac{\partial}{\partial ct} - \frac{\partial}{\partial z}\right)\left(\frac{\partial}{\partial ct} + \frac{\partial}{\partial z}\right) f = 0 \, . \tag{5.15}$$

This then, upon introduction of the auxiliary (so-called) **null coordinates**[1]

$$u := z - ct \qquad v := z + ct \tag{5.16}$$

so that

$$\frac{\partial}{\partial u} = -\frac{1}{2}\left(\frac{\partial}{\partial ct} - \frac{\partial}{\partial z}\right) \qquad \frac{\partial}{\partial v} = \frac{1}{2}\left(\frac{\partial}{\partial ct} + \frac{\partial}{\partial z}\right) \, , \tag{5.17}$$

becomes

$$\frac{\partial^2 f}{\partial u \partial v} = \frac{\partial^2 f}{\partial v \partial u} = 0 \, .$$

We recognise immediately that the last second-order linear partial differential equation is solved by

$$f(u, v) = f_1(u) + f_2(v) \, .$$

[1] This constitutes a simple linear transformation from the coordinates ct and z to the coordinates u and v.

Substituting back in for u and v from (5.16) gives the result (5.14). It follows from this consideration that the two real-valued functions f_1 and f_2 need only be *once* continuously differentiable in their respective arguments for the second-order linear partial differential equation (5.13) to be well defined (see the remarks below).

To obtain a unique solution to (5.13), we need to specify **initial values** for $f(ct, z)$ and its first partial derivative with respect to ct given, e.g., by

$$f(ct,z)|_{ct=ct_0} = a_1(z) \qquad \left.\frac{\partial f(ct,z)}{\partial ct}\right|_{ct=ct_0} = a_2(z)\ , \tag{5.18}$$

with $a_1(z)$ and $a_2(z)$ two arbitrary real-valued functions that are once continuously differentiable in z.

Note that the homogeneous wave equation (5.13) leaves unspecified the second partial derivatives

$$\left(\frac{\partial}{\partial ct} - \frac{\partial}{\partial z}\right)\left(\frac{\partial}{\partial ct} - \frac{\partial}{\partial z}\right) f$$

and

$$\left(\frac{\partial}{\partial ct} + \frac{\partial}{\partial z}\right)\left(\frac{\partial}{\partial ct} + \frac{\partial}{\partial z}\right) f\ ,$$

which are the second partial derivatives of $f(ct, z)$ in the directions of the unit normals to the plane surfaces $\{z - ct = \text{const}; x, y\ arbitrary\}$ and $\{z + ct = \text{const}; x, y\ arbitrary\}$, respectively. That these derivatives of $f(ct, z)$ really are left unspecified is particularly apparent from the factorised form of the homogeneous wave equation (5.15). In terms of the null coordinates (5.16), these derivatives are given by $\partial^2 f/\partial u^2$ and $\partial^2 f/\partial v^2$, which, with (5.14), correspond to

$$\frac{\partial^2 f_1(u)}{\partial u^2} \qquad \text{and} \qquad \frac{\partial^2 f_2(v)}{\partial v^2}\ .$$

Now as no further conditions on these two particular derivatives are given, this means they may in fact be **discontinuous** across, respectively, the plane surfaces $\{z - ct = \text{const}; x, y\ arbitrary\}$ and $\{z + ct = \text{const}; x, y\ arbitrary\}$. This reflects the freedom in the choice of the two initial data functions $a_1(z)$ and $a_2(z)$.

The purpose of this detailed account of the solutions to the linear homogeneous wave equation (5.13) was to show that the solutions for E_x/c and E_y/c are thus given by

$$E_x(ct,z)/c = f_1(z-ct) + f_2(z+ct) \tag{5.19}$$

$$E_y(ct,z)/c = g_1(z-ct) + g_2(z+ct)\ , \tag{5.20}$$

with f_1, f_2, g_1 and g_2 each arbitrary once continuously differentiable real-valued functions of their respective arguments.

The solutions for the components B_x and B_y now follow from simultaneously solving the vacuum time evolution equations (1.26) and (1.27). For B_x we find the conditions

$$\frac{\partial B_x}{\partial ct} = \frac{\partial E_y/c}{\partial z} = \frac{\partial g_1}{\partial u} + \frac{\partial g_2}{\partial v} \qquad \frac{\partial B_x}{\partial z} = \frac{\partial E_y/c}{\partial ct} = -\frac{\partial g_1}{\partial u} + \frac{\partial g_2}{\partial v}\ ,$$

which, without loss of generality, is solved by

$$B_x(ct,z) \;=\; -\,g_1(z-ct)+g_2(z+ct)\,, \tag{5.21}$$

while for B_y we find the conditions

$$\frac{\partial B_y}{\partial ct}=-\,\frac{\partial E_x/c}{\partial z}=-\,\frac{\partial f_1}{\partial u}-\frac{\partial f_2}{\partial v} \qquad\qquad \frac{\partial B_y}{\partial z}=-\,\frac{\partial E_x/c}{\partial ct}=\frac{\partial f_1}{\partial u}-\frac{\partial f_2}{\partial v}\,,$$

which, without loss of generality, is solved by

$$B_y(ct,z) \;=\; f_1(z-ct)-f_2(z+ct)\,. \tag{5.22}$$

We can thus summarise our results on **plane wave solutions** to **Maxwell's field equations** in **vacuum** (1.26)–(1.29) as

$$\boldsymbol{E}/c=\begin{bmatrix} f_1(z-ct)+f_2(z+ct)\\ g_1(z-ct)+g_2(z+ct)\\ 0\end{bmatrix} \qquad \boldsymbol{B}=\begin{bmatrix} -\,g_1(z-ct)+g_2(z+ct)\\ f_1(z-ct)-f_2(z+ct)\\ 0\end{bmatrix}\,.$$

Note that the fields $\boldsymbol{E}/c$ and $\boldsymbol{B}$ are *transverse* to the directions of propagation of the plane waves, which, in the present example, were assumed to be along the $\boldsymbol{e}_z$–directions.

We see that, altogether, we have *four* arbitrary functions (of position z) to be specified as the initial data for the vacuum electromagnetic fields presently discussed. These correspond, e.g., to the initial values of each of f_1, f_2, g_1 and g_2. One can show that also for **fully general electromagnetic fields** there are four arbitrary functions (of position $\boldsymbol{r}$) to be specified as the initial data. These can be interpreted as setting up electromagnetic waves of *two* possible **polarisation states** propagating in *two* possible **spatial directions** relative to a reference spatial direction. One says that general electromagnetic fields have *four* **dynamical degrees of freedom**.[2] It is this property of electromagnetic fields which reflects the possibility of communicating arbitrary information between different parts of space by means of propagating electromagnetic waves. Of course, this is the fundamental aspect behind many everyday applications such as TV and radio broadcasting or speaking on mobile phones.

To conclude our discussion on **vacuum plane wave solutions** to **Maxwell's field equations**, let us briefly calculate for our solutions (5.19)–(5.22) (i) the energy density u, which, by (5.2), is

$$u=\frac{1}{2\mu_0}\,\left[\,(f_1+f_2)^2+(f_1-f_2)^2+(g_1+g_2)^2+(g_1-g_2)^2\,\right]\,,$$

and (ii) the energy current density $\boldsymbol{S}$, which, by (5.2), is

$$\boldsymbol{S}=\frac{c}{\mu_0}\,\left[\,(f_1+f_2)\,(f_1-f_2)+(g_1+g_2)\,(g_1-g_2)\,\right]\,\boldsymbol{e}_z\,.$$

As to be expected, the latter is aligned with the spatial direction of propagation.

[2] In the Hamiltonian understanding of dynamics, where each pair of conjugate coordinate and momentum constitutes one dynamical degree of freedom of a system, electromagnetic fields are said to have two dynamical degrees of freedom.

Next, if we confine ourselves to considering plane waves propagating in the *positive* $\boldsymbol{e}_z$–direction only (so that $0 = f_2 = g_2$), we find from these relations

$$u = \frac{1}{2\mu_0}\,(f_1^2 + g_1^2) \qquad\qquad \boldsymbol{S} = (uc)\,\boldsymbol{e}_z\ ,$$

and $\boldsymbol{E}/c \cdot \boldsymbol{B} = 0$. In this case the fields are mutually perpendicular, as well as being perpendicular to the spatial direction of propagation.

Finally, for a plane wave propagating in positive $\boldsymbol{e}_z$–direction that is *linearly polarised*, these results specialise further by setting either $f_1 = 0$ or $g_1 = 0$.

5.4 Electromagnetic scalar and vector potentials

In solving **Maxwell's field equations** for static electric charge configurations in section 3 or for configurations of steady electric currents in section 4, the introduction of, respectively, an electrostatic scalar potential and a magnetostatic vector potential proved mathematically helpful. In this section we want to discuss the generalisation of this **method of potentials** to general time-dependent field configurations.

Let us use **Maxwell's field equations** in the form (1.15)–(1.18), with $\mathbf{0} = c\boldsymbol{P} = \boldsymbol{M}$. From the zero magnetic charges law (1.18) and (2.26) we still find the relation

$$\boldsymbol{B} = \boldsymbol{\nabla} \times \boldsymbol{A}\ . \tag{5.23}$$

We now call $\boldsymbol{A}$ the **electromagnetic vector potential**. Then, upon substitution of (5.23) into Faraday's law (1.16), we find the condition

$$\boldsymbol{\nabla} \times \left(\boldsymbol{E}/c + \frac{\partial \boldsymbol{A}}{\partial ct} \right) = \mathbf{0}\ ,$$

which, by (2.25), says that the terms in the bracket must add up to form the gradient of some differentiable scalar function. In line with the electrostatic case of chapter 3, let us hence introduce the **electromagnetic scalar potential** ϕ/c so that we have

$$\boldsymbol{E}/c = -\,\boldsymbol{\nabla}\phi/c - \frac{\partial \boldsymbol{A}}{\partial ct}\ . \tag{5.24}$$

Through introducing the electromagnetic scalar and vector potentials, **Maxwell's field equations** (1.16) and (1.18) are now identically satisfied ("$0 = 0$"). Both ϕ/c and $\boldsymbol{A}$ have SI unit $1\ \frac{\mathrm{kg\,m}}{\mathrm{s\,C}}$.

Next, from Gauß' law (1.17) we find

$$-\,(\boldsymbol{\nabla}\cdot\boldsymbol{\nabla})\phi/c - \frac{\partial(\boldsymbol{\nabla}\cdot\boldsymbol{A})}{\partial ct} = \frac{c\rho}{\epsilon_0 c^2}\ , \tag{5.25}$$

while from the Ampère–Maxwell law (1.15) we obtain

$$-\,\frac{\partial^2 \boldsymbol{A}}{\partial (ct)^2} - \boldsymbol{\nabla} \times \boldsymbol{\nabla} \times \boldsymbol{A} - \frac{\partial(\boldsymbol{\nabla}\phi/c)}{\partial ct} = -\,\mu_0\,\boldsymbol{J}\ . \tag{5.26}$$

By now we have reduced four coupled first-order linear partial differential equations for $\boldsymbol{E}/c$ and $\boldsymbol{B}$ to two coupled second-order linear partial differential equations for ϕ/c and $\boldsymbol{A}$. We can proceed to *decouple* them by making the following observation.

5.4.1 Gauge transformations

The electromagnetic potentials ϕ/c and $\boldsymbol{A}$ are *not* uniquely defined. A simultaneous **gauge transformation** of the form

$$\tilde{\phi}/c = \phi/c - \frac{\partial \chi}{\partial ct} \tag{5.27}$$

$$\tilde{\boldsymbol{A}} = \boldsymbol{A} + \boldsymbol{\nabla}\chi \tag{5.28}$$

with $\chi(ct, \boldsymbol{r})$ an arbitrary twice continuously differentiable real-valued function, leaves both of the physically relevant fields $\boldsymbol{E}/c$ and $\boldsymbol{B}$ unchanged, as seen from (5.24) and (5.23), respectively. Due to this **gauge freedom** in the definition of ϕ/c and $\boldsymbol{A}$, we can choose these potentials according to what may be convenient in a given situation.

Let us take a derivative with respect to ct of (5.27) and add to it the divergence of (5.28). It follows that

$$\frac{\partial \tilde{\phi}/c}{\partial ct} + \boldsymbol{\nabla}\cdot\tilde{\boldsymbol{A}} = \frac{\partial \phi/c}{\partial ct} + \boldsymbol{\nabla}\cdot\boldsymbol{A} - \frac{\partial^2 \chi}{\partial (ct)^2} + (\boldsymbol{\nabla}\cdot\boldsymbol{\nabla})\chi \,.$$

Hence, e.g., we can choose $\chi(ct, \boldsymbol{r})$ so that it satisfies the particular linear inhomogeneous wave equation

$$-\frac{\partial^2 \chi}{\partial (ct)^2} + (\boldsymbol{\nabla}\cdot\boldsymbol{\nabla})\chi = -\left[\frac{\partial \phi/c}{\partial ct} + \boldsymbol{\nabla}\cdot\boldsymbol{A}\right] ,$$

leading to

$$\frac{\partial \tilde{\phi}/c}{\partial ct} + \boldsymbol{\nabla}\cdot\tilde{\boldsymbol{A}} = 0 \,. \tag{5.29}$$

But this means that we can always introduce *special* electromagnetic potentials $\tilde{\phi}/c$ and $\tilde{\boldsymbol{A}}$ that satisfy the condition (5.29), which is referred to as the **Lorentz gauge condition**.

We note that imposing the Lorentz gauge condition on $\tilde{\phi}/c$ and $\tilde{\boldsymbol{A}}$ still does not determine them uniquely. There exists a special **remaining gauge freedom** in the choice of $\tilde{\phi}/c$ and $\tilde{\boldsymbol{A}}$ of the form

$$\bar{\phi}/c = \tilde{\phi}/c - \frac{\partial \bar{\chi}}{\partial ct} \qquad\qquad \bar{\boldsymbol{A}} = \tilde{\boldsymbol{A}} + \boldsymbol{\nabla}\bar{\chi} \,,$$

with a twice differentiable real-valued function $\bar{\chi}(ct, \boldsymbol{r})$ that satisfies the linear homogeneous wave equation

$$-\frac{\partial^2 \bar{\chi}}{\partial (ct)^2} + (\boldsymbol{\nabla}\cdot\boldsymbol{\nabla})\bar{\chi} = 0 \,.$$

This special remaining gauge freedom is particularly useful for time-dependent **vacuum configurations**, where we can always achieve $\bar{\phi}/c = 0$ so that $\bar{\boldsymbol{A}}$ satisfies the **Coulomb gauge condition** $\boldsymbol{\nabla}\cdot\bar{\boldsymbol{A}} = 0$, a special subcase of the **Lorentz gauge condition** (5.29).

5.4.2 Retarded solutions

Returning to our coupled equations (5.25) and (5.26), if we now choose electromagnetic scalar and vector potentials ϕ/c and $\boldsymbol{A}$ that satisfy the **Lorentz gauge condition** (5.29), and we remember that $(\epsilon_0 c^2)^{-1} = \mu_0$, then, with respect to a reference frame with **Cartesian coordinates** so that the vector analytical identity (2.31) can be employed, these two equations formally combine to give the linear **inhomogeneous wave equation**

$$\left[-\frac{\partial^2}{\partial(ct)^2} + (\boldsymbol{\nabla}\cdot\boldsymbol{\nabla}) \right] \left[\begin{array}{c} \phi/c \\ \boldsymbol{A} \end{array} \right] = -\mu_0 \left[\begin{array}{c} c\rho \\ \boldsymbol{J} \end{array} \right] \tag{5.30}$$

for ϕ/c and the Cartesian components of $\boldsymbol{A}$, for given ("*known*") electromagnetic sources $c\rho(ct,\boldsymbol{r})$ and $\boldsymbol{J}(ct,\boldsymbol{r})$. If we *assume* continuous electric charge and current distributions $c\rho$ and $\boldsymbol{J}$ to be localised in a volume G, then the solutions to (5.30) evaluated at an observation point $\boldsymbol{r}$ exterior to G at a time t are formally given by

$$\phi(ct,\boldsymbol{r})/c = \frac{\mu_0}{4\pi}\int\int\int_G \frac{c\rho(ct-|\boldsymbol{r}-\boldsymbol{r}'|,\boldsymbol{r}')}{|\boldsymbol{r}-\boldsymbol{r}'|}\,\mathrm{d}V' \tag{5.31}$$

$$\boldsymbol{A}(ct,\boldsymbol{r}) = \frac{\mu_0}{4\pi}\int\int\int_G \frac{\boldsymbol{J}(ct-|\boldsymbol{r}-\boldsymbol{r}'|,\boldsymbol{r}')}{|\boldsymbol{r}-\boldsymbol{r}'|}\,\mathrm{d}V'\,. \tag{5.32}$$

These are referred to as the **retarded solutions** of (5.30). The electromagnetic potentials at position $\boldsymbol{r}$ at time t depend on the state of the electromagnetic sources at $\boldsymbol{r}'$ in G at the *earlier* time

$$t - \frac{|\boldsymbol{r}-\boldsymbol{r}'|}{c}\,, \tag{5.33}$$

where $|\boldsymbol{r}-\boldsymbol{r}'|$ is the spatial distance between the observation point $\boldsymbol{r}$ and the source point $\boldsymbol{r}'$ in G, and $|\boldsymbol{r}-\boldsymbol{r}'|/c$ is the time it takes to travel this distance at the speed of light. The time coordinate (5.33) is commonly known as **retarded time**.

Chapter 6
Lorentz invariance of Maxwell's field equations

6.1 Historical remarks

It is Maxwell's unifying theory of **electromagnetism** that has to be regarded as the cradle of the **special theory of relativity**. After Maxwell's prediction of the existence of **electromagnetic waves** which propagate at the **speed of light** in 1864, and the subsequent experimental confirmation of this prediction by Hertz in 1888, it was rather natural that the leading physicists of this time wondered what the physical "medium" was which carried such electromagnetic waves — just like, e.g., pressure waves are carried by continuous material media. This hypothetical all-pervasive "medium" was referred to as the "**ether**". The strongest motivation for resolving the question on the "**ether**" was given by the long-held tradition that the **laws of Nature** ought to be **invariant** (i.e., unchanged) under the **Galilean transformations** of Newtonian physics. These are the special class of *linear* coordinate transformations that link observations relative to one **inertial** (i.e., non-accelerating and non-rotating) **reference frame** $\mathcal{R}$ to those relative to a second **inertial reference frame** $\mathcal{R}'$. In different terms, this invariance principle demands that from a physical point of view the choice of one particular **inertial reference frame** must be as good as the choice of any other one; all **inertial reference frames** must be physically *equivalent*.

Employing **Cartesian coordinate systems** in both reference frames and supposing that $\mathcal{R}'$ moves with *constant* velocity $\boldsymbol{v} := v\,\boldsymbol{e}_z$ in the positive $\boldsymbol{e}_z$–direction of $\mathcal{R}$ (with the coordinate axes of $\mathcal{R}$ and $\mathcal{R}'$ aligned), the linear **Galilean transformations** are given by

$$\begin{aligned} t' &= t & x' &= x \\ z' &= z - v\,t & y' &= y \end{aligned} \qquad v = \text{const}\,, \tag{6.1}$$

where t and t' denote time coordinates. **Galilean transformations** leave **time intervals** and **spatial distances** invariant, i.e.,

$$\left| t' - t'_0 \right| = \left| t - t_0 \right|\,, \qquad \left| \boldsymbol{r}' - \boldsymbol{r}'_0 \right| = \left| \boldsymbol{r} - \boldsymbol{r}_0 \right|\,,$$

and so respect Newton's concepts of "**absolute time**" and "**absolute space**". Moreover, they have the consequence that a **velocity** $\boldsymbol{u}$ relative to $\mathcal{R}$ changes to a **velocity**

$$\boldsymbol{u}' = \boldsymbol{u} - \boldsymbol{v} \tag{6.2}$$

relative to $\mathcal{R}'$. Hence, the expectation at the time was that the **speed of light** should have a *different* value relative to *different* **inertial reference frames**.

The *finiteness* of the **speed of light** was long known by then. Its magnitude in **vacuum**, c, was first determined by the Danish astronomer Olaf Römer (1644–1710) who studied the time delay in the appearance of Jupiter's brightest moon, Io, in 1676. The first accurate measurement of c in a laboratory on Earth was achieved only much later by the French physicist Armand Fizeau (1819–1896) in 1849. In addition, Fizeau also found that light which is travelling in a moving substance (such as a liquid) is **dragged** along in the direction of motion of the substance.

Another well known consequence of the finiteness of the **speed of light** was the **aberration** of starlight due to the velocity of the Earth on its orbit around the Sun, which the English astronomer James Bradley (1693–1762) first reported on in around 1728. One detects an overall error, ϑ, for the angular position of a star, with $\tan\vartheta \approx (v/c)$, where v is the magnitude of the *tangential* component of the velocity of the Earth relative to the star in question.

In the light of these facts the German-American physicist Albert Abraham Michelson (1852–1931) and the American chemist Edward W Morley (1838–1923) spent about a decade trying to measure a *difference* in the value of c with respect to **inertial reference frames** that are in **relative motion**, publishing results in 1887. Supposing the "**ether**" had zero velocity in the rest reference frame of the stars (i.e., Newton's "**absolute space**") and referring to Bradley's observation of **aberration**, they tried in particular to measure the velocity of the Earth relative to the "**ether**". However, all of their dedicated and ingenious attempts to detect this velocity failed, i.e., the **Michelson–Morley experiments** led to a **null result**. It thus seemed that the **speed of light** was constant relative to *all* **inertial reference frames**.

In geometrical terms the results of the **Michelson–Morley experiments** can be summarised as follows. An observer located in an **inertial reference frame** $\mathcal{R}$ sees a light ray which is emitted (in a vacuum) from a position $\boldsymbol{r}_0$ at a time ct_0 to reach a position $\boldsymbol{r}$ at a time ct. This light ray satisfies the distance-travelled-in-a-given-time relation

$$0 = -\,|\,ct - ct_0\,|^2 + |\,\boldsymbol{r} - \boldsymbol{r}_0\,|^2 \ ,$$

for a given value of c. Likewise, a second observer located in an **inertial reference frame** $\mathcal{R}'$ which is moving with *constant* velocity $\boldsymbol{v}$ relative to $\mathcal{R}$ sees the *same* light ray, now emitted from a position $\boldsymbol{r}'_0$ at a time ct'_0, to reach a position $\boldsymbol{r}'$ at a time ct'. For the observer in $\mathcal{R}'$ the light ray satisfies an analogous distance-travelled-in-a-given-time relation given by

$$0 = -\,\left|\,ct' - ct'_0\,\right|^2 + \left|\,\boldsymbol{r}' - \boldsymbol{r}'_0\,\right|^2 \ ,$$

for the *same* value c as perceived by $\mathcal{R}$. This thus yields the universal result

$$0 = -\,\left|\,ct' - ct'_0\,\right|^2 + \left|\,\boldsymbol{r}' - \boldsymbol{r}'_0\,\right|^2 = -\,|\,ct - ct_0\,|^2 + |\,\boldsymbol{r} - \boldsymbol{r}_0\,|^2 \ .$$

The implication is that the **speed of light** *does not depend on the choice of* **inertial reference frame**.

6.2 Special theory of relativity

It was the then unknown young Albert Einstein (1879–1955) who, motivated by the disturbing fact of the non-invariance of **Maxwell's field equations** under **Galilean transformations**, suggested in 1905 that consequently Newton's concepts of "**absolute time**" and "**absolute space**" could no longer be maintained, but had to be abandoned in favour of a "*new*" law of Nature, namely that the **speed of light in vacuum** c was a fundamental constant. The idea of a "medium" that was the carrier of **electromagnetic waves** (and so light) did no longer make sense; the all-familiar set of linear **Galilean transformations** (6.1) relating different **inertial reference frames** had to be replaced by a new set of linear coordinate transformations which respected the constancy of c.

Einstein strongly emphasised two **postulates** on which he built the **special theory of relativity**. These are the

> **Principle of relativity**: The physical laws of Nature, and so the results of all experiments, are independent of the uniform translational motion of any inertial reference frame $\mathcal{R}$.

— this principle had been introduced a long time ago by the Italian scientist and philosopher Galileo Galilei (1564–1642) in around 1592 —, and the

> **Universality of the speed of light**: The speed of light in vacuum relative to any inertial reference frame $\mathcal{R}$ is $c = 299792458\ \frac{\mathrm{m}}{\mathrm{s}}$, regardless of the motion of the light's source relative to $\mathcal{R}$.

(Note that the **special theory of relativity** *neglects* any **gravitational interactions** between physical systems.) In place of Newton's "**absolute time**" and "**absolute space**", the **special theory of relativity** considers instead a 4-dimensional "distance" between two **events**, $(ct, \boldsymbol{r})$ and $(ct_0, \boldsymbol{r}_0)$, defined by

$$(\Delta s)^2 := -\,|\,ct - ct_0\,|^2 + |\,\boldsymbol{r} - \boldsymbol{r}_0\,|^2\ , \tag{6.3}$$

as a quantity which is *independent* of the **inertial reference frame** with respect to which it is evaluated. This quantity is referred to as the special relativistic **line element**; as a squared distance its SI unit is $1\,\mathrm{m}^2$. One says that when for two events $(ct, \boldsymbol{r})$ and $(ct_0, \boldsymbol{r}_0)$

$$(\Delta s)^2 \begin{cases} < 0\ , & \text{these events are } \textit{timelike} \text{ separated.} \\ = 0\ , & \text{these events are } \textit{lightlike} \text{ separated.} \\ > 0\ , & \text{these events are } \textit{spacelike} \text{ separated.} \end{cases}$$

For a given event $(ct_0, \boldsymbol{r}_0)$, the set of events $(ct, \boldsymbol{r})$ for which $(\Delta s)^2 = 0$ forms a 3-dimensional hyperboloid in a 4-dimensional so-called **Minkowski space-time** (named after the German mathematician Hermann Minkowski, 1864–1909). This 3-dimensional hyperboloid is referred to as the **light cone** at the event $(ct_0, \boldsymbol{r}_0)$.

It makes for an illuminating exercise to show that coordinate transformations

$$ct' = ct'(ct, \boldsymbol{r}) \qquad \boldsymbol{r}' = \boldsymbol{r}'(ct, \boldsymbol{r})$$

between two **inertial reference frames** $\mathcal{R}$ and $\mathcal{R}'$ which are (i) *linear*, and which (ii) leave the special relativistic **line element** *invariant* so that

$$(\Delta s')^2 = (\Delta s)^2 \tag{6.4}$$

holds, must be of the particular form of the coordinate transformations devised by the Dutch physicist Hendrik Antoon Lorentz (1853–1928) in 1903. Using **Cartesian coordinates** in both $\mathcal{R}$ and $\mathcal{R}'$, the special relativistic **line element** is expressed by, respectively,

$$(\Delta s)^2 = -(ct - ct_0)^2 + (x - x_0)^2 + (y - y_0)^2 + (z - z_0)^2$$

and

$$(\Delta s')^2 = -(ct' - ct'_0)^2 + (x' - x'_0)^2 + (y' - y'_0)^2 + (z' - z'_0)^2 \,.$$

Then, supposing that $\mathcal{R}'$ moves with *constant* velocity $\boldsymbol{v} := v\,\boldsymbol{e}_z$ in the positive $\boldsymbol{e}_z$–direction of $\mathcal{R}$ (with all coordinate axes aligned), the linear **Lorentz transformations** are given by

$$\begin{array}{ll} ct' = \Gamma\,(ct - (v/c)\,z) & x' = x \\ z' = \Gamma\,(z - (v/c)\,ct) & y' = y \end{array} \,, \tag{6.5}$$

with the **Lorentz factor** Γ defined by

$$\Gamma(v) := \frac{1}{\sqrt{1 - (v/c)^2}} \qquad v = \text{const}\,. \tag{6.6}$$

To illustrate the **Lorentz invariance** of the special relativistic **line element** (6.3), let us consider the propagation of a spherical light signal emitted in $\mathcal{R}$ from the origin $\boldsymbol{r}_0 = \boldsymbol{0}$ at a time $t_0 = 0$. After a time t has elapsed, the spherical light front will have reached positions $\boldsymbol{r}$ so that, in **Cartesian coordinates**, for every point on the light front

$$(\Delta s)^2 = 0 = -(ct)^2 + x^2 + y^2 + z^2$$

applies. Assuming that the origin $\boldsymbol{r}'_0 = \boldsymbol{0}$ of $\mathcal{R}'$ coincides with that of $\mathcal{R}$ at the instant of emission, and that likewise at this instant $\mathcal{R}'$'s clock reads $t'_0 = 0$, we have

$$\begin{aligned} (\Delta s')^2 \;=\; 0 \quad &= \quad -(ct)'^2 + x'^2 + y'^2 + z'^2 \\ &= \quad -\frac{1}{1-(v/c)^2}\,\left[\,(ct)^2 - 2(v/c)ct\,z + (v/c)^2 z^2\,\right] + x^2 + y^2 \\ &\qquad\qquad + \frac{1}{1-(v/c)^2}\,\left[\,z^2 - 2(v/c)z\,ct + (v/c)^2 (ct)^2\,\right] \\ &= \quad (\Delta s)^2\,; \end{aligned}$$

this is the result we expected.

Under a **Lorentz transformation** of the form (6.5), the **first partial derivatives** transform according to

$$\begin{array}{ll} \partial/\partial ct' = \Gamma\,(\partial/\partial ct + (v/c)\,\partial/\partial z) & \partial/\partial x' = \partial/\partial x \\ \partial/\partial z' = \Gamma\,(\partial/\partial z + (v/c)\,\partial/\partial ct) & \partial/\partial y' = \partial/\partial y \end{array} \quad . \tag{6.7}$$

It was realised already by people before Einstein that **Maxwell's field equations** were **form-invariant** (in mathematical terms: **covariant**) under **Lorentz transformations** (given that the electromagnetic field variables transformed appropriately), whereas **Newton's equations of motion** for point particles and continuous material media were *not*. However, before Einstein's work it was rather unclear what the implications of this situation were.

We conclude by stating that fundamental **Lorentz invariants** of all continuous bodies are constituted by their **rest mass** m_0 and their **total electric charge** q.

6.3 Lorentz transformations of electromagnetic field variables

By the **principle of relativity**, **Maxwell's field equations** are required to remain **form-invariant** under **Lorentz transformations** between two arbitrary **inertial reference frames** $\mathcal{R}$ and $\mathcal{R}'$. That is, we need to have

$$\left.\begin{array}{ll} \partial \boldsymbol{E}/c/\partial ct - \boldsymbol{\nabla}\times\boldsymbol{B} = -\mu_0\,\boldsymbol{J} & \\ \partial \boldsymbol{B}/\partial ct + \boldsymbol{\nabla}\times\boldsymbol{E}/c = \boldsymbol{0} & \\ 0 = \boldsymbol{\nabla}\cdot\boldsymbol{E}/c - \mu_0\,c\rho & \\ 0 = \boldsymbol{\nabla}\cdot\boldsymbol{B} & \text{in } \mathcal{R} \end{array}\right\} \Leftrightarrow \left\{\begin{array}{ll} \partial \boldsymbol{E}'/c/\partial ct' - \boldsymbol{\nabla}'\times\boldsymbol{B}' = -\mu_0\,\boldsymbol{J}' & \\ \partial \boldsymbol{B}'/\partial ct' + \boldsymbol{\nabla}'\times\boldsymbol{E}'/c = \boldsymbol{0} & \\ 0 = \boldsymbol{\nabla}'\cdot\boldsymbol{E}'/c - \mu_0\,c\rho' & \\ 0 = \boldsymbol{\nabla}'\cdot\boldsymbol{B}' & \text{in } \mathcal{R}' \end{array}\right. .$$

It thus follows with (6.7) that the **Cartesian components** of the electric and the magnetic fields and the electromagnetic sources must transform according to the rules

$$\begin{array}{ll} E'_x/c = \Gamma\,(E_x/c - (v/c)\,B_y) & B'_x = \Gamma\,(B_x + (v/c)\,E_y/c) \\ E'_y/c = \Gamma\,(E_y/c + (v/c)\,B_x) & B'_y = \Gamma\,(B_y - (v/c)\,E_x/c) \\ E'_z/c = E_z/c & B'_z = B_z \end{array} \quad , \tag{6.8}$$

and

$$\begin{array}{ll} c\rho' = \Gamma\,(c\rho - (v/c)\,J_z) & J'_x = J_x \\ J'_z = \Gamma\,(J_z - (v/c)\,c\rho) & J'_y = J_y \end{array} \quad , \tag{6.9}$$

respectively. Note that for $\boldsymbol{E}/c$ and $\boldsymbol{B}$ only components *transverse* to the relative motion between $\mathcal{R}$ and $\mathcal{R}'$ are affected by **Lorentz transformations**.

It can be easily shown from the result just obtained, that the quantities

$$-\,|\,\boldsymbol{E}/c\,|^2+|\,\boldsymbol{B}\,|^2 \qquad\qquad \boldsymbol{E}/c\cdot\boldsymbol{B}$$

are **Lorentz invariants**, as is

$$-\,(c\rho)^2+|\,\boldsymbol{J}\,|^2\,.$$

Let us look at the consequences of the transformation rules (6.8) through the following illustrative example. Let us assume to be given a homogeneous electrostatic field $\boldsymbol{E}/c$ in $\mathcal{R}$. We consider two cases:

(i) $\boldsymbol{E}/c=(E_0/c)\,\boldsymbol{e}_z$, $(E_0/c)=$ const.

In this case, with (6.8), we find that an observer at rest in $\mathcal{R}'$ perceives electric and magnetic fields given by

$$\boldsymbol{E}'/c=(E_0/c)\,\boldsymbol{e}_z \qquad\qquad \boldsymbol{B}'=\boldsymbol{0}\,,$$

i.e., the observer in $\mathcal{R}'$ does not make an experience different compared to the electrostatic field perceived by the observer in $\mathcal{R}$. This is because $\mathcal{R}'$'s relative motion is directed *along* the field lines of $\boldsymbol{E}/c$.

(ii) $\boldsymbol{E}/c=(E_0/c)\,\boldsymbol{e}_x$, $(E_0/c)=$ const.

In this case, with (6.8), we find that an observer at rest in $\mathcal{R}'$ perceives electric and magnetic fields given by

$$\boldsymbol{E}'/c=\Gamma\,(E_0/c)\,\boldsymbol{e}_x \qquad\qquad \boldsymbol{B}'=-\,(v/c)\,\Gamma\,(E_0/c)\,\boldsymbol{e}_y\,.$$

Because $\mathcal{R}'$'s relative motion is now directed *transverse* to the field lines of $\boldsymbol{E}/c$, the observer in $\mathcal{R}'$ feels a *stronger* electrostatic field than the observer in $\mathcal{R}$, on which, moreover, an additional magnetostatic field is superimposed. Note that the directions of the fields and of the motion are all mutually perpendicular.

Nevertheless, in both cases (i) and (ii), we have

$$-\,|\,\boldsymbol{E}'/c\,|^2+|\,\boldsymbol{B}'\,|^2=-\,|\,\boldsymbol{E}/c\,|^2+|\,\boldsymbol{B}\,|^2=-\,(E_0/c)^2$$

and

$$\boldsymbol{E}'/c\cdot\boldsymbol{B}'=\boldsymbol{E}/c\cdot\boldsymbol{B}=0\,.$$

We conclude this section by stating that at the level of the electromagnetic scalar and vector potentials the **Lorentz transformation** rules are given by

$$\begin{aligned}\phi'/c&=\Gamma\,(\phi/c-(v/c)\,A_z) & A'_x&=A_x\\ A'_z&=\Gamma\,(A_z-(v/c)\,\phi/c) & A'_y&=A_y\end{aligned}\quad, \tag{6.10}$$

with

$$-\,(\phi/c)^2+|\,\boldsymbol{A}\,|^2$$

being a **Lorentz invariant**.

6.4 Length contraction, time dilation and velocity-composition law

A few curious, nevertheless experimentally well established results follow from the **special theory of relativity**. One is that the length scales of bodies in **relative motion** with respect to an observer appear to shrink along the direction of motion. This effect, which is a direct manifestation of the relativity of space and time, is known as the **length contraction**, and was first discussed by the Irish physicist George Francis FitzGerald (1851–1901) in 1889, and subsequently by Lorentz. In more detail, let us consider a rod to be at rest in an **inertial reference frame** $\mathcal{R}'$ where, appropriately oriented, it has length $z'_2 - z'_1 := \ell_0$. By the **Lorentz transformation** rules (6.5), we have

$$z'_2 = \Gamma\,(z_2 - (v/c)\,ct_2) \qquad\qquad z'_1 = \Gamma\,(z_1 - (v/c)\,ct_1)\ .$$

If we now take the difference $z'_2 - z'_1$ and evaluate it at a simultaneous time $ct_2 - ct_1 = 0$ in $\mathcal{R}$, we obtain for $z_2 - z_1 := \ell$ the result

$$\ell(v) = \sqrt{1-(v/c)^2}\,\ell_0\ , \tag{6.11}$$

i.e., $\ell(v)$ is *shorter* than ℓ_0 for $0 < v < c$.

In an analogous fashion we can illustrate the effect of **time dilation**. If we consider a clock to be at rest in $\mathcal{R}'$ at position z'_0, then the time interval between two events (t'_1, z'_0) and (t'_2, z'_0) at z'_0 is measured by the clock as $t'_2 - t'_1 := \tau_0$. By the *inverse* of the **Lorentz transformation** rules (6.5), we have

$$t_2 = \Gamma\left(t'_2 + (v/c)\,z'_0/c\right) \qquad\qquad t_1 = \Gamma\left(t'_1 + (v/c)\,z'_0/c\right)\ .$$

It follows that for $t_2 - t_1 := \tau$ we obtain

$$\tau(v) = \frac{\tau_0}{\sqrt{1-(v/c)^2}}\ , \tag{6.12}$$

so that $\tau(v)$ is *longer* than τ_0 for $0 < v < c$.

Lastly, **Lorentz transformations** of the form (6.5) have the consequence that a velocity $\boldsymbol{u} = (u_x, u_y, u_z)^T$ relative to $\mathcal{R}$ changes to

$$\begin{pmatrix} u'_x \\ u'_y \\ u'_z \end{pmatrix} = \frac{1}{1 - u_x v/c^2} \begin{pmatrix} u_x - v \\ \Gamma^{-1}\,u_y \\ \Gamma^{-1}\,u_z \end{pmatrix} \tag{6.13}$$

relative to $\mathcal{R}'$. This is the special relativistic **velocity-composition law**.

6.5 Equations of motion for compact bodies

When only *non-gravitational* **forces** are active, **Newton's equations of motion** of mechanics, for a compact body of **rest mass** $m_0 = \text{const} > 0$, can

be rendered **form-invariant** under **Lorentz transformation** by replacing the expression $m_0\boldsymbol{v}$ for the **linear momentum** by the relativistic expression

$$\boldsymbol{p} = \frac{m_0\boldsymbol{v}}{\sqrt{1-(v/c)^2}}\,, \tag{6.14}$$

with $\boldsymbol{v}$ denoting the **velocity** of the compact body. We thus obtain

$$\frac{\mathrm{d}\boldsymbol{p}}{\mathrm{d}t} = \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{m_0\boldsymbol{v}}{\sqrt{1-(v/c)^2}}\right) = \boldsymbol{F}\,. \tag{6.15}$$

Note that when $\boldsymbol{v} \neq$ **const** $\Rightarrow |\boldsymbol{v}| = v \neq$ const, the *chain rule* has to be employed when performing the total derivative with respect to t.

Let us now derive, as an example, the **equations of motion** for a **point particle** of **rest mass** m_0 and with **electric charge** q moving through an **electromagnetic field** with $\boldsymbol{E}/c$ and $\boldsymbol{B}$, which is thus subjected to the **Lorentz force** (1.8). In this case, the **equations of motion** read

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{m_0\boldsymbol{v}}{\sqrt{1-(v/c)^2}}\right) = cq\,\left[\,\boldsymbol{E}/c + \left(\frac{\boldsymbol{v}}{c}\right)\times\boldsymbol{B}\,\right]\,.$$

To specialise further, let us assume that we have

$$\boldsymbol{E}/c = \boldsymbol{0} \qquad \boldsymbol{B} = B_0\,\boldsymbol{e}_z \qquad B_0 = \text{const}\,,$$

and so we may also assume $v_z = 0$, as in the present situation the point particle feels no force acting in the $\boldsymbol{e}_z$–direction. These assumptions lead to the coupled **equations of motion**

$$\begin{aligned}
\left[1-(v/c)^2\right]^{-3/2}\,\left[\left[1-(v_y/c)^2\right]\frac{\mathrm{d}v_x}{\mathrm{d}t} + v_xv_y/c^2\,\frac{\mathrm{d}v_y}{\mathrm{d}t}\right] &= \frac{qB_0}{m_0}\,v_y \\
\left[1-(v/c)^2\right]^{-3/2}\,\left[\left[1-(v_x/c)^2\right]\frac{\mathrm{d}v_y}{\mathrm{d}t} + v_xv_y/c^2\,\frac{\mathrm{d}v_x}{\mathrm{d}t}\right] &= -\,\frac{qB_0}{m_0}\,v_x\,,
\end{aligned}$$

where now, because $v_z = 0$, we have $(v/c)^2 = (v_x/c)^2 + (v_y/c)^2$. Clearly, even in this special case, the equations of motion do not look very simple at all, and their solution requires the application of some advanced analytical methods. Here, we do not want to discuss this problem any further. This ends this example.

We conclude by noting that the quantity

$$-\,m_0^2c^2 = -\,(E/c)^2 + |\,\boldsymbol{p}\,|^2\,,$$

with $m_0 =$ const the **rest mass** of a compact body, E its **total energy**, and $\boldsymbol{p}$ its **linear momentum**, is a **Lorentz invariant**. For bodies with non-zero rest mass ($m_0 > 0$) such as **electrons**, this relation can be inverted to give

$$E(|\boldsymbol{p}|) = m_0c^2\,\sqrt{1+\frac{|\boldsymbol{p}|^2}{m_0^2c^2}}\,. \tag{6.16}$$

[Note that Dirac also allowed for the possible minus sign of E in (6.16).] Alternatively, substituting from (6.14), (6.16) can be expressed by Einstein's famous relativistic formula

$$E(v) = m(v)c^2 = \frac{m_0 c^2}{\sqrt{1-(v/c)^2}} \,. \tag{6.17}$$

In cases where $|\boldsymbol{p}|^2 \ll m_0^2 c^2$, we can approximate (6.16) by

$$E(|\boldsymbol{p}|) \approx m_0 c^2 \left[1 + \tfrac{1}{2} \frac{|\boldsymbol{p}|^2}{m_0^2 c^2} + \ldots \right] = m_0 c^2 + \frac{|\boldsymbol{p}|^2}{2m_0} + \ldots \,.$$

On the other hand, for particles of zero rest mass ($m_0 = 0$) such as **photons**, we find

$$E(|\boldsymbol{p}|) = c\,|\boldsymbol{p}| \,.$$