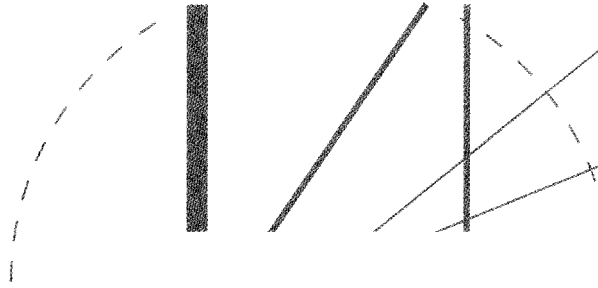# Electrodynamics: An Introduction Including Quantum Effects

Harald J W Müller-Kirsten

Electrodynamics: An Introduction Including Quantum Effects

This page intentionally left blank

Electrodynamics: An Introduction Including Quantum Effects

Harald J W Müller-Kirsten

University of Kaiserslautern, Germany

World Scientific

ELECTRODYNAMICS: AN INTRODUCTION INCLUDING QUANTUM EFFECTS

Contents

This page intentionally left blank

Preface

Electrodynamics is one of the pillars of physics with arms reaching through classical and atomic physics far into high energy particle physics and present day string theory. With its unification of electric and magnetic phenomena and subsequently with classical mechanics in Special Relativity it pointed towards a unification of all physical theories in a future theory of "everything".

Classical electrodynamics is the electrodynamics of macroscopic phenomena and is summarised essentially in the four Maxwell equations which, however, permeate in one way or another into microscopic phenomena. Although beset with tantalising divergence problems, the quantised form of Maxwell's electrodynamics, that is quantum electrodynamics, has so far defied all mathematical purists by yielding highly precise values of such vital physical quantities as the Lamb shift and magnetic moments. In its nonabelian extension this quantised electrodynamics led to quantum chromodynamics and beyond, thereby permitting its unification with strong and weak interactions. It is expected that the ultimate theory, possibly a form of higher dimensional string theory, will one day permit its unification also with quantum gravity. This theoretical endeavour, starting with Maxwell's unification of electric and magnetic phenomena and now reaching as far as cosmology, is one of the most amazing intellectual achievements of mankind.

A countless number of treatises has been written on electrodynamics, including applications in all directions. These cover the entire range from elementary introductions to authoritative reference tomes. The present text is but one more, and like any other is imprinted with the author's likes and preferences. In repeatedly teaching electrodynamics the author set himself the task of making the presentation as easily comprehensible as possible, and to direct the view beyond the traditional domain of pure macroscopic Maxwell theory.

Rightly or not, and irrespective of what others think, the author claims that the companion text of a first course in a basic subject must be readable, must explain the issues at hand in detail and, not the least, must present

any nontrivial calculational steps. The author remembers from his time as a student the relief derived from a book or script which explains what others passed over (knowingly or not), if only as a reassurance that one's own understanding is correct. A student's struggle with a book may flatter the vanity of the author, but will more likely result in the book being put aside. The present text therefore attempts to avoid as far as possible the "*it-can–easily–be–shown–jumps*" familiar from everyday literature. With more than a hundred mostly worked and typical examples interspersed, the text might also serve as a "*Teach Yourself Course*" in electrodynamics. To what extent this succeeds only the reader can decide.

The further attempt pursued in the text (frequently in examples) is to open some views to other domains which are usually excluded in traditional presentations of electrodynamics. This applies in particular to intimately related quantum effects (mostly of Schrödinger equation or harmonic oscillator type), electric-magnetic duality, reference to higher dimensions and to gravity, all of which the student of today encounters elsewhere much earlier than in the past, and hence might expect some mention in the text.*

I thank Dr. K. K. Yim of World Scientific for his untiring help and a countless number of suggestions for improving the manuscript.

<div align="right">H.J.W. Müller–Kirsten</div>

---

*Only textbook literature consulted repeatedly at various times in the course of preparation of this text is cited in the bibliography. Many other sources have, of course, also been used at various points but their specific relevance has been lost track of as the years passed by. Hence no attempt is made to even refer to all important and standard works on the subject. An extensive list of this literature can be found in the book of Jackson [3]. Literature related to very specific topics or sources is cited in footnotes at appropriate points in the text.

Chapter 1

Introduction

1.1 What is Electrodynamics?

Electrodynamics as a theory deals with electric charges which move in space and with electromagnetic fields that are produced by the charges and again interact with charges. Electrodynamics is described by Maxwell's equations whose most important consequences are: (a) the electromagnetic nature of light (Maxwell), (b) the emission of electromagnetic waves by an oscillating dipole (H. Hertz), and (c) the unification of electric and magnetic forces. Unlike in Maxwell's days one does not refer to an aether anymore — it was Einstein who concluded that this does not exist — and instead uses the concept of fields in space. Einstein's Special Theory of Relativity unifies electrodynamics with classical mechanics.

1.2 Presentation of Macroscopic Electrodynamics

In general a first course in electrodynamics, or electricity and magnetism as it is also called, is preceded by a course in classical mechanics which even today is not always combined with the Special Theory of Relativity (as would be desirable). In addition the concept of a field in space and the distinction between directly observable field quantities and at best indirectly observable field potentials is at that stage still too vague to permit an immediate relativistic field theoretic approach to appear plausible. Moreover, in general the term electrodynamics is usually restricted to macroscopically observable phenomena, so that a quantised treatment with operator-valued fields is beyond its scope, the latter being dealt with in quantum electrodynamics. Thus the classical fields are c-numbers. As a consequence of this restriction, and also in order to establish classical electrodynamics as a theory in its own right,

1

the fundamental equations of this theory, Maxwell's equations, are usually
not derived from Hamilton's principle as a prior course in methods of clas-
sical mechanics might suggest. In addition the law of force between moving
charges is considerably more complicated than that between pointlike masses
so that an analogous procedure is not immediately advisable. Nonetheless
Hamilton's variational principle is of such basic significance that it permits,
of course, the derivation of Maxwell's equations as the Euler–Lagrange equa-
tions (and their consequences) of an appropriate variational principle. Thus
here we do not adopt this procedure immediately (i.e. till Chapter 18).
What are the other methods which suggest themselves? Perusing the liter-
ature, one observes two main procedures. The approach with emphasis on
logic starts from an axiomatic presentation of Maxwell's equations, whereas
the other more historical and phenomenological approach abstracts these
from observations. A textbook which adopts the former procedure is that of
Sommerfeld [1] who follows H. Hertz in this respect, but whereas H. Hertz
starts from the differential form of Maxwell's equations, Sommerfeld chooses
the vector integral form. Another text that follows this procedure is, for in-
stance, that of Lim [2]. Books which choose the second procedure are those
of Jackson [3] and Greiner [4].

In the axiomatic formulation of classical or macroscopic electrodynamics
the *two principal axioms* are Faraday's law of induction and Ampère's flux
theorem. In integral form and in units of the internationally agreed system
of units (ISU), these are

$$\frac{d}{dt} \int_F \mathbf{B} \cdot d\mathbf{F} = - \oint_{C(F)} \mathbf{E} \cdot d\mathbf{l} \tag{1.1}$$

and

$$\frac{d}{dt} \int_F \mathbf{D} \cdot d\mathbf{F} + \int_F \mathbf{j} \cdot d\mathbf{F} = \oint_{C(F)} \mathbf{H} \cdot d\mathbf{l}, \tag{1.2}$$

where for a given fixed surface of integration F,

$$\frac{d}{dt} \int_F \mathbf{D} \cdot d\mathbf{F} = \int_F \frac{\partial \mathbf{D}}{\partial t} \cdot d\mathbf{F},$$

since (by assumption) F does not change with t.*

---

*Recall, for instance the analogy, that as shown in Whittaker and Watson [7], p. 67,

$$\frac{d}{d\alpha} \int_a^b f(x, \alpha) dx = f(b, \alpha)\frac{db}{d\alpha} - f(a, \alpha)\frac{da}{d\alpha} + \int_a^b \frac{\partial f(x, \alpha)}{\partial \alpha} dx,$$

which shows the α-dependence of boundary values.

In the first or Faraday's law, $\mathbf{B}$ is the magnetic induction and $\mathbf{E}$ the electric field strength. The law says that every change of the magnetic flux[†] with time through an open, double-sided area F (a plane is single-sided) with boundary given by the closed curve $C(F)$ — this flux being the number of lines of force in Faraday's considerations — generates an equal but oppositely directed circuit potential, the *electromotive force* $\oint \mathbf{E} \cdot d\mathbf{l}$, along the boundary $C(F)$. An open surface can, for instance, be visualised as a container without a lid, and a closed surface as one closed with the lid. In the former case the rim of the opening corresponds to the curve $C(F)$. The word *double-sided* implies that the surface possesses direction normals directed towards inside or outside regions. A closed surface has no boundary, but we can imagine on it a closed curve $C(F)$ which divides the surface into two regions.

In the second or Ampère's law, $\mathbf{D}$ is the so-called dielectric displacement ($\mathbf{D} = \epsilon_0 \mathbf{E} + \mathbf{P}, \epsilon_0$ the dielectric constant of vacuum, $\mathbf{P}$ the polarisation vector or dipole moment per unit volume of the medium), $\mathbf{j}$ is the density of the electron current and $\mathbf{H}$ the magnetic field strength. The law says that analogous to Faraday's law the time change of an electric flux through an area F is equal to a magnetic circuit potential in the boundary curve $C(F)$ in the same direction. The expression $\partial \mathbf{D}/\partial t$ obviously has the dimension of a current density. The expression

$$\frac{d}{dt} \int_F \mathbf{D} \cdot d\mathbf{F}$$

is called *Maxwell's displacement current.*

The two principal axioms are supplemented by two *subsidiary axioms* which are consequences of the principal axioms with implementation of some empirical findings. We arrive at these by considering a boundary curve $C(F)$ enclosing the area F, and by permitting $C(F)$ to shrink to zero, so that the line integrals vanish, i.e.

$$\lim_{C(F) \to 0} \frac{d}{dt} \int_F \mathbf{B} \cdot d\mathbf{F} = 0, \tag{1.3}$$

$$\lim_{C(F) \to 0} \left\{ \frac{d}{dt} \int_F \mathbf{D} \cdot d\mathbf{F} + \int_F \mathbf{j} \cdot d\mathbf{F} \right\} = 0. \tag{1.4}$$

The area F thus loses its boundary and becomes closed. In the case of magnetic dipoles (irrespectively of whether one considers magnets or current loops) every line of force leaving the area F is accompanied by a corresponding line of force toward it. The integral extended over the entire closed

---

[†]Flux is in general the scalar product of a vector field with an area.

surface thus yields the value zero, i.e.

$$\int_{F_{\text{closed}}} \mathbf{B} \cdot d\mathbf{F} = 0. \tag{1.5}$$

This is a general result in view of the empirical fact that no isolated magnetic poles exist and hence no region with only ingoing or only outgoing lines of force. With the definition of the divergence

$$div\mathbf{A}(\mathbf{r}) = \lim_{V \to 0} \frac{1}{V} \int_{F(V)} \mathbf{A} \cdot d\mathbf{F} \tag{1.6}$$

we obtain the differential form of Eq. (1.5), i.e.

$$div\mathbf{B}(\mathbf{r}) \equiv \boldsymbol{\nabla} \cdot \mathbf{B}(\mathbf{r}) = 0. \tag{1.7}$$

One says: $\mathbf{B}(\mathbf{r})$ is *divergenceless*.

Equation (1.4) permits an analogous consideration but with different consequences. Equation (1.2), or rather (1.4), applies to cases in which the electron current

$$\int_{F(V)} \mathbf{j} \cdot d\mathbf{F}$$

leaves the surface F enclosing the volume $V(F)$. This implies that the volume with total charge q loses charge at the rate

$$\int_{F(V)} \mathbf{j} \cdot d\mathbf{F} = -\frac{dq}{dt}. \tag{1.8}$$

The minus sign indicates that $V(F)$ is losing charge. Put differently, as

$$\frac{dq}{dt} + \int_{F(V)} \mathbf{j} \cdot d\mathbf{F} = 0, \tag{1.9}$$

the equation says that the amount of charge in $V(F)$ remains constant (this is called charge conservation). Here one has to be careful, because if we write

$$\frac{dq}{dt} = + \int_{F} \mathbf{j} \cdot d\mathbf{F},$$

then q represents the amount of charge which passes through the area F per second *with no reference to* V. We can now write Eq. (1.4) as

$$\lim_{C(F) \to 0} \left\{ \frac{d}{dt} \int_{F} \mathbf{D} \cdot d\mathbf{F} - \frac{dq}{dt} \right\} = 0, \tag{1.10}$$

which implies

$$\int_{F_{\text{closed}}} \mathbf{D} \cdot d\mathbf{F} = q. \tag{1.11}$$

We could have added a constant of integration on the right side. But for charge $q = 0$ the field $\mathbf{D} = 0$, and thus this constant must be zero. The differential form of Eq. (1.11) again follows with (1.6); i.e.

$$\boldsymbol{\nabla} \cdot \mathbf{D} = \rho, \quad \rho = \frac{q}{V}. \tag{1.12}$$

Here ρ is the charge density. Equations (1.11) and (1.12) are known as the Gauss law of electrodynamics. We can now write Eq. (1.8):

$$\frac{d}{dt} \int_V \rho dV = - \int_{F(V)} \mathbf{j} \cdot d\mathbf{F}. \tag{1.13}$$

Going to an infinitesimal volume and using the Gauss divergence theorem, this implies

$$\int_V \boldsymbol{\nabla} \cdot \mathbf{j} dV = \int_{F(V)} \mathbf{j} \cdot d\mathbf{F} \tag{1.14}$$

and hence the *equation of continuity*

$$\frac{\partial \rho(\mathbf{r}, t)}{\partial t} + \boldsymbol{\nabla} \cdot j = 0. \tag{1.15}$$

A method of deriving electrodynamics essentially from this equation and the Lorentz force (the force exerted on a charge in an electromagnetic field) — as an alternative approach to electrodynamics — can be found in a paper of Bopp.[‡]

Returning to comments at the beginning we mention that attempts have repeatedly been made to derive Maxwell's equations from the Coulomb law and the Special Theory of Relativity. Since it is known that the General Theory of Relativity contains the generalisation of Newton's law of gravitation, i.e. a law which has the same form as the Coulomb law for charges, such an attempt cannot succeed without additional assumptions. A discussion of this topic can be found in the book of Jackson [3]. An entirely different approach which sees the only logical foundation of Maxwell's theory in a relativistic treatment of radiation theory is the book of Page and Adams [5]. In the following we adopt the historical and phenomenological approach starting with electrostatics, since this seems to be more suitable for an understanding of the physics of electrodynamics, particularly in a first course on the subject.

[‡]F. Bopp, *Z. Physik* **169** (1962) 45.

1.3   On the Choice of Units

In textbooks on electrodynamics nowadays two systems of units are in use: The more common internationally used MKSA-system of units with meter (m), kilogramme (kg), second (s) and ampere (A), and the still frequently used system of Gaussian units which is based on the c.g.s. units, i.e. centimeter (c), gramme (g) and second (s) (the current is then given in statampere, $1\,\mathrm{A} = 3\times 10^9$ statampere). Either system has its advantages — the former has been agreed upon internationally and is therefore used particularly in applications, and the latter, the Gaussian system, is somewhat theory-oriented and therefore frequently used in considerations of singly charged particles. Here we employ the first system but will refer occasionally also to the other system. We add the following comments.

In discussions on units the word "dimension" is frequently referred to. The dimension of a physical quantity is nothing absolute. It is possible to choose units (like the so-called *natural units*) in which Planck's action quantum and the velocity of light in vacuum have the value 1 and are dimensionless. Since velocity $=$ distance/time, one can then express lengths in units of time, i.e. seconds, or more commonly time intervals in meters. This arbitrariness was already pointed out by Planck. Some authors even amuse themselves today about dimensional considerations.[§] The example shows that also the number of fundamental units is a matter of choice (in "natural units" every quantity can be expressed in a power of length). In earlier days the units used in the literature on electrodynamics were the so-called electrostatic units (e.s.u.) and electromagnetic units (e.m.u.) which today are indirectly contained in the Gaussian system. In essence, in the Gaussian system of units electric quantities are expressed in e.s.u. and magnetic quantities in e.m.u. while the change to Gaussian units requires a factor c or $1/c$ (c the velocity of light in vacuum), which can be determined experimentally (cf. Appendix B). For instance

$$\mathbf{B}_{\mathrm{e.s.u.}} = \frac{1}{c}\mathbf{B}_{\mathrm{e.m.u.}}.$$

In particular one has in these units for the dielectric constant of the vacuum ϵ_0 and for the magnetic permeability of the vacuum μ_0

$$\epsilon_0(\mathrm{e.s.u.}) = 1, \quad \epsilon_0(\mathrm{e.m.u.}) = \frac{1}{c^2}, \quad \mu_0(\mathrm{e.s.u.}) = \frac{1}{c^2}, \quad \mu_0(\mathrm{e.m.u.}) = 1.$$

[§]See e.g. A. O'Rahilly, Vol. I [6], p. 65, where the author says: *"Maxwell invented a second system... But this system is never employed, it merely occurs in those pages of textbooks which profess to deal with something called 'dimensions'."* See also p. 68.

We do not enter into a deeper discussion of these units here (sometimes it is not clear whether a quantity is electric or magnetic).

The choice of a system of units in electrodynamics depends on the choice of the magnitude as well as the dimension of two arbitrary constants, as an exhaustive investigation of Maxwell's equations shows.[¶] The appearance of one constant k can immediately be seen by looking at the Coulomb law in electrostatics, which determines the force $\mathbf{F}_{12}$ between two point charges q_1, q_2, separated by a distance $\mathbf{r}$, i.e.

$$\mathbf{F}_{12} = k q_1 q_2 \frac{\mathbf{r}}{r^3}.$$

For length, mass and time, we choose here as agreed upon internationally the ISU (international system of units, internationally abbreviated SI, see Appendix B) and thus the units meter (m), kilogramme (kg), second (s) (correspondingly in the Gaussian system centimeter (cm), gramme (g), second (s)). Depending on the choice of dimension and magnitude of k we obtain different units for the charge q. The electric field strength $\mathbf{E}(\mathbf{r})$ at the distance $\mathbf{r}$ away from the charge q is

$$\mathbf{E}(\mathbf{r}) = k \frac{q}{r^3} \mathbf{r}.$$

Thus the field $\mathbf{E}(\mathbf{r})$ can be defined as force per unit charge.

In the consideration of magnetic phenomena we have to deal with currents. Currents I are defined with respect to charges q, $I = dq/dt$. The connection between electric and magnetic phenomena or between $\mathbf{E}$ and $\mathbf{B}$ introduces another constant for the choice of units of $\mathbf{B}$. This constant can be introduced in a number of ways. With the help of the theorem of Stokes (see later), Eq. (1.1) can be written as

$$\nabla \times \mathbf{E} + k^* \frac{\partial \mathbf{B}}{\partial t} = 0,$$

where k^* is chosen to be 1 and dimensionless in the MKSA-system (the constants k and k^* can be chosen arbitrarily in magnitude and dimension).

In accordance with the ISU, i.e. the MKSA-system of units, the unit of current is taken as the *ampere* (the spelling in English being this, ampere, which is not part of SI). One ampere is defined as that amount of current which runs through two parallel, straight, infinitely long, thin conductors which are one meter apart when the force between these is 2×10^{-7} newton/m (per meter in length) (the expression for this attractive force will be derived later). Since current × time is charge, we obtain the definition of the unit of charge:

[¶]See Jackson [3], p. 811.

$$1 \text{ coulomb (C)} = 1 \text{ ampere-second (A s)}.$$

It remains to determine the dimensions of $\mathbf{D}$ and $\mathbf{H}$. For a large number of different materials the following *"matter"* or *"material equations"* are valid:

$$\mathbf{D} = \epsilon\mathbf{E}, \quad \mathbf{B} = \mu\mathbf{H}.$$

ϵ is the dielectric constant of the material and μ its magnetic permeability. Both have magnitudes and dimensions depending on the chosen system of units. The vacuum values ϵ_0, μ_0 obey, as we shall see, the important relation

$$\epsilon_0\mu_0 = \text{const.}$$

In the Gaussian system of units ϵ_0, μ_0 are taken as 1 and dimensionless. The constant is then 1. In the MKSA-system the constant is found to be $1/c^2$, where c is the velocity of light in vacuum, i.e.

$$c = 2.998 \times 10^8 \text{ m/s}.$$

In this case

$$\epsilon_0 = \frac{10^7}{4\pi c^2} \text{ farad/meter}$$

with the dimension of current^2 time^4 mass^{-1} length^{-3}, or $\epsilon_0 = 8.854 \times 10^{-12}$ A s/(V m) or C^2/(joule-meter), where 1 volt (V) $= 1$ A^{-1} m^2 kg s^{-2} and 1 joule $= 1$ newton-meter (N m). In addition

$$\mu_0 = 4\pi \times 10^{-7} \quad \text{newton-ampere}^{-2} \quad (\text{N A}^{-2}),$$

or $\mu_0 = 1.257 \times 10^{-6}$ with dimension of mass length current^{-2} time^{-2}.

In the *Gaussian system* the constants k, k^* are chosen as follows:

$$k = 1 \quad \text{(dimensionless)}, \quad k^* = \frac{1}{c} \quad \text{(dimension time/length)}.$$

In the MKSA-system the vacuum constants k, k^* are chosen as

$$k = \frac{1}{4\pi\epsilon_0} = 10^{-7}c^2$$

with dimension kg m^3 s^{-4} A^{-2} and

$$k^* = 1$$

(dimensionless). We see that in comparison with the Gaussian system the MKSA-system requires the use of ϵ_0, μ_0.

Chapter 2

Electrostatics – Basic Aspects

2.1   Introductory Remarks

In this chapter the Coulomb law is introduced, and various static consequences are investigated, such as electrical screening and diverse applications. The latter require also the introduction of essential mathematics, in particular that of the delta distribution.

2.2   The Coulomb Law

Electrostatics is the theory of static charges, which means that moving charges, i.e. currents, are not considered. The fundamental phenomenological law which is the basis of electrostatics is *Coulomb's law*, which determines the force acting between two point charges q_1, q_2 and which in vacuum is given by

$$\mathbf{F}_{12} = k\frac{q_1 q_2}{r^3}\mathbf{r}.$$

Here $\mathbf{r} = \mathbf{r}_1 - \mathbf{r}_2$ is the vectorial separation of the charges q_i at positions $\mathbf{r}_i$. k is the constant fixed by the choice of units. In the ISU, as explained in Chapter 1,

$$k = \frac{1}{4\pi\epsilon_0}$$

in vacuum. The force is given in newtons (N), the charge in coulombs (C) and the separation in meters (m). The electric field strength $\mathbf{E}(\mathbf{r})$ at a point $\mathbf{r}$ as in Fig. 2.1 is defined as the Coulomb force acting on a unit charge $(+1\mathrm{C})$

9

at this point, i.e.

$$\mathbf{E}(\mathbf{r}) = k\frac{q_1(\mathbf{r} - \mathbf{r}_1)}{|\mathbf{r} - \mathbf{r}_1|^3}.$$

In the ISU or MKSA-system of units the electric field strength is given in units of N/C or V/m. In the *Gaussian system* the charge is given in *statcoulombs*, the distance in cm, the force in dynes, and the electric field strength in statvolt/cm (1 C $= 3 \times 10^9$ statcoulombs, 1 statcoulomb $= 10$ e.s.u.; 1 newton $= 10^5$ dynes).
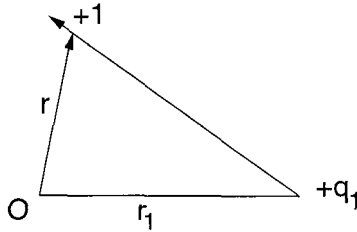
Fig. 2.1 Point charges q_1 and $+1$.

It should be remembered that a charge q can be positive or negative. Conventionally electric lines of force are drawn as leaving positive charges and directed towards negative charges. In this sense positive charges are regarded as sources of the electric fields and negative charges as sinks.

One distinguishes between discrete charge distributions (charges at distinct points) and continuous charge distributions (charges spread over spatial domains). Correspondingly one defines a charge distribution or charge density $\rho(\mathbf{r})$ as charge per unit volume. Considering an infinitesimal volume element, we can use the above expression for the field strength at a point $\mathbf{r}$ due to a point charge q_1 at a point $\mathbf{r}_1$ to obtain the field strength at a point $\mathbf{r}$ resulting from superposition, i.e. summation or integration of the contributions of point or pointlike charges in a domain V of space, i.e.

$$\mathbf{E}(\mathbf{r}) = k\int_V \rho(\mathbf{r}')\frac{(\mathbf{r} - \mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|^3}d\mathbf{r}', \quad dV' \equiv d\mathbf{r}'.$$

Here V is a volume which does not enclose the observation point at $\mathbf{r}$. The lines of force of $\mathbf{E}$, as shown in Fig. 2.2, are continuous for one and the same medium (having the same ϵ). Since $E \propto k$, this does not apply in the case of different media. But if we define $D := E/k$, called *dielectric induction* or *displacement*, the lines of D are continuous at interfaces between different media. The law of force

$$F_{12} \propto \frac{q_1 q_2}{r^2}$$

for point charges q_i was discovered experimentally by Coulomb, who published his observations in 1785.
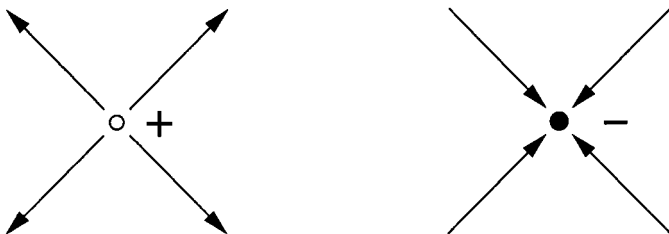
Fig. 2.2 Charges and lines of force.

Later it became known that the law had been observed earlier by Cavendish who, however, never published his findings. Coulomb used for his observations a so-called torsion balance consisting of a vertical rod or fibre with another fibre attached horizontally at its lower end. At one end of this horizontal fibre a small chargeable sphere was mounted and at the other end a piece of paper to stabilise the system in a rotation. If another charge of the same magnitude is brought close to the mounted charge, at a distance r say, the latter experiences a repulsion evidenced by a rotation of the horizontal fibre through an angle θ. The corresponding torque is proportional to θ. By selecting various values of r ($nr, n = 1, 2, 3, \dots$) Coulomb arrived at the law now named after him for the case of repulsive charges. For charges which attract each other Coulomb modified his experiment. His experiments did not establish the proportionality to the charges themselves. This was more or less assumed in analogy to masses in Newton's law of gravitation. It was Gauss who later recognised that Coulomb's law permits the definition of charge. In recent times the Coulomb law has, of course, been subjected to much more critical investigations. Thus the power of -2 of r has been verified up to deviations of less than 10^{-9}.

Cavendish, on the other hand, assumed a law of force of the form

$$F_{12} \propto \frac{1}{r^n}, \quad E = k\frac{q}{r^n},$$

and then designed his experiment to determine n. His experimental setup consisted of two concentric conducting spherical shells connected by a conducting wire from one to the other. He charged the exterior shell and then removed the wire together with the exterior shell. The interior shell then provided no evidence of a field in its neighbourhood. From this Cavendish concluded that $n = 2$, because in this case the interior of a charged spherical

shell has $E = 0$ (thus in the case $n = 2$ the charge covered the exterior shell, as we explain later; see also Example 2.2).

2.3   The Electrostatic Potential

We now demonstrate that $\mathbf{E}$ can be derived from a potential, the *electrostatic potential*. Since

$$\boldsymbol{\nabla}_r \frac{1}{|\mathbf{r} - \mathbf{r}'|} = -\frac{(\mathbf{r} - \mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|^3}, \quad \boldsymbol{\nabla} = \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z} \right),$$

we have

$$\mathbf{E}(\mathbf{r}) = -k \int \rho(\mathbf{r}') \boldsymbol{\nabla}_r \frac{1}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}' \equiv -\boldsymbol{\nabla}\phi,$$

i.e.

$$\mathbf{E} = -\boldsymbol{\nabla}\phi,$$

where ϕ is the potential

$$\phi(\mathbf{r}) = k \int \frac{\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}'.$$

Since *curl grad* $= 0$, it follows that

$$\boldsymbol{\nabla} \times \mathbf{E} = 0,$$

i.e. the electric field is *curl-free* (or whirl- or vortex-free; *curl* is also often written *rot* for "rotation") .

Next we consider the *flux* $\mathbf{E} \cdot d\mathbf{F}$ of the vector field $\mathbf{E}$ through the surface area element $d\mathbf{F}(\mathbf{r})$ at $\mathbf{r}$ of a surface F surrounding the charges. Then, cf. Fig. 2.3,

$$\mathbf{E} \cdot d\mathbf{F}(\mathbf{r}) = k \int_V \rho(\mathbf{r}') \frac{(\mathbf{r} - \mathbf{r}') \cdot d\mathbf{F}(\mathbf{r})}{|\mathbf{r} - \mathbf{r}'|^3} d\mathbf{r}'.$$

Let V be the volume $V(F)$ enclosed by F; the case of $V \neq V(F)$ has to be dealt with separately. It follows that

$$\int_F \mathbf{E} \cdot d\mathbf{F} = k \int_{V(F)} \rho(\mathbf{r}') d\mathbf{r}' \int_F \frac{(\mathbf{r} - \mathbf{r}') \cdot d\mathbf{F}(\mathbf{r})}{|\mathbf{r} - \mathbf{r}'|^3}.$$

But

$$|\mathbf{r} - \mathbf{r}'|^2 d\Omega = \frac{(\mathbf{r} - \mathbf{r}') \cdot d\mathbf{F}(\mathbf{r})}{|\mathbf{r} - \mathbf{r}'|}.$$
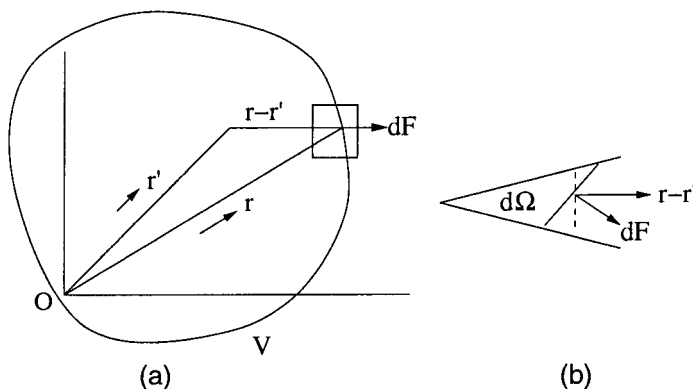
Fig. 2.3 (a) The surface and (b) the solid angle.

Hence

$$\int_F \mathbf{E} \cdot d\mathbf{F} = k \int_{V(F)} \rho(\mathbf{r}')d\mathbf{r}' \int_F d\Omega = 4\pi k \int_V \rho(\mathbf{r}')d\mathbf{r}', \qquad (2.1)$$

provided the point $\mathbf{r}'$ (charge element $\rho(\mathbf{r}')$) is within V. This relation is the integral form of the Gauss law $\nabla \cdot \mathbf{E} = 4\pi k\rho$. If $\mathbf{r}'$ is outside of V, that is if we wish to know the flux through a closed surface F in the case when the field is due to charges outside of F, we have

$$\int_F d\Omega = \Omega_1 - \Omega_2 = 0,$$

as we can see from Fig. 2.4, since the consideration of the solid angles of a closed volume such as that of a sphere subtended at a charge q implies:

$$\Omega_1 + (-\Omega_1) = 0.$$

In this consideration the volume V is a purely mathematical volume, so to speak. If the material of the volume $V(F)$ is such that the lines of force of $\mathbf{E}$ penetrate into V, this result implies that every ingoing line must also leave the volume again. If the closed surface F, which does not enclose a charge, is made of conducting material (see later), then the field $\mathbf{E}$ in the interior is zero.
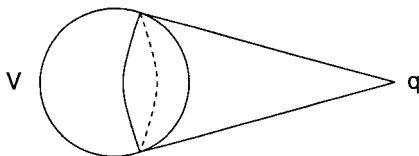
Fig. 2.4 Charge outside of V.

With the help of the *Gauss divergence theorem* (here assumed known) we obtain from Eq. (2.1):

$$\int_F \mathbf{E} \cdot d\mathbf{F} = \int_V \boldsymbol{\nabla} \cdot \mathbf{E} d\mathbf{r} = 4\pi k \int_V \rho(\mathbf{r}) d\mathbf{r} = 4\pi k \int_V \Sigma_i q_i \delta(\mathbf{r} - \mathbf{r}_i) d\mathbf{r},$$

depending on whether the charge distribution is continuous or discrete. It follows that

$$\int_V (\boldsymbol{\nabla} \cdot \mathbf{E} - 4\pi k \rho(\mathbf{r})) d\mathbf{r} = 0.$$

Since this is valid for any volume V, i.e. $\int_V = \int_{V'} = \cdots = 0$, it follows that

$$\boldsymbol{\nabla} \cdot \mathbf{E} = 4\pi k \rho.$$

Thus, *charges in space are sources* of the electric field.

2.4   The Equations of Electrostatics

We have now obtained the *three important equations of electrostatics*:

$$\mathbf{E} = -\boldsymbol{\nabla}\phi,$$

$$\boldsymbol{\nabla} \cdot \mathbf{E} = 4\pi k \rho, \quad \text{or} \quad \int \mathbf{E} \cdot d\mathbf{F} = 4\pi k \Sigma_i q_i, \tag{2.2a}$$

$$\boldsymbol{\nabla} \times \mathbf{E} = 0.$$

Later we shall assume that these equations are also valid when $\rho, \mathbf{E}, \phi$ are time-dependent (but not leading to currents). From the first two equations we obtain the *Poisson equation*

$$\triangle\phi = -4\pi k \rho. \tag{2.2b}$$

The equation for $\rho = 0$, i.e. $\triangle\phi = 0$, is known as the *Laplace equation*.

We recapitulate the following important formulae which are needed in various applications, also in this text:

(a) In *Cartesian coordinates*

$$\boldsymbol{\nabla} \cdot \mathbf{E} = \frac{\partial E_x}{\partial x} + \frac{\partial E_y}{\partial y} + \frac{\partial E_z}{\partial z},$$

$$\triangle\phi = \boldsymbol{\nabla} \cdot \boldsymbol{\nabla}\phi = \frac{\partial^2 \phi}{\partial x^2} + \frac{\partial^2 \phi}{\partial y^2} + \frac{\partial^2 \phi}{\partial z^2},$$

(b) in *cylindrical coordinates*

$$\boldsymbol{\nabla} \cdot \mathbf{E} = \frac{1}{r}\frac{\partial}{\partial r}(rE_r) + \frac{1}{r}\frac{\partial E_\theta}{\partial \theta} + \frac{\partial E_z}{\partial z},$$

$$\triangle\phi = \frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial\phi}{\partial r}\right) + \frac{1}{r^2}\frac{\partial^2\phi}{\partial\theta^2} + \frac{\partial^2\phi}{\partial z^2},$$

(c) in *spherical polar coordinates* (spherical coordinates)

$$\mathbf{\nabla}\cdot\mathbf{E} = \frac{1}{r^2}\frac{\partial}{\partial r}(r^2 E_r) + \frac{1}{r\sin\theta}\frac{\partial}{\partial\theta}(\sin\theta E_\theta) + \frac{1}{r\sin\theta}\frac{\partial E_\phi}{\partial\phi},$$

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

Example 2.1: Discrete and continuous charge distributions

Let the following potential be given:

$$\phi(r) = kq\frac{e^{-\alpha r}}{r}\left(1 + \frac{\alpha r}{2}\right).$$

Show that there are both discrete and continuous charge distributions with a vanishing total charge.

Solution: In the case of the given spherically symmetric potential or electric field, the Gauss law is

$$\frac{1}{r^2}\frac{\partial}{\partial r}(r^2 E_r) = 4\pi k\rho(r),$$

so that

$$r^2 E_r = \int_0^r 4\pi k\rho(r)r^2 dr \equiv kQ(r),$$

where $Q(r)$ is the charge enclosed in a sphere of radius r. With

$$E_r = -\frac{d\phi(r)}{dr},$$

it follows that

$$Q(r) = -\frac{r^2}{k}\frac{d\phi(r)}{dr}.$$

For the given potential $\phi(r)$ we have

$$Q(r) = -4\pi\epsilon_0 r^2\frac{d\phi(r)}{dr} = -qe^{-\alpha r}\left[\frac{\alpha r}{2} - (1 + \alpha r)\left(1 + \frac{\alpha r}{2}\right)\right].$$

For $r \to 0$ we obtain the discrete charge

$$Q(r) \to q, \tag{2.3a}$$

whereas for $r \to \infty$

$$Q(r) \to 0.$$

The continuous charge distribution is given by $\rho(r)$, for which we obtain from the first integral

$$4\pi\rho(r)r^2 = \frac{dQ(r)}{dr},$$

i.e.

$$\rho(r) = \frac{1}{4\pi r^2}\frac{dQ(r)}{dr} = -\frac{1}{8\pi}q\alpha^3 e^{-\alpha r}.$$

This can be verified by computing the sum of the distribution, i.e.

$$\int_0^\infty \rho(r)4\pi r^2 dr = -q. \tag{2.3b}$$

The vanishing of the sum of results (2.3a) and (2.3b) implies that the total charge in space is zero. Applied to the neutral hydrogen atom, we can interpret the discrete charge as that of the proton and the continuous charge distribution as that of the electron.

Example 2.2: Theoretical explanation of Cavendish's experiment

Assuming that the electric field $\mathbf{E}$ of a point charge q is given by the radially symmetric expression

$$E = \frac{1}{4\pi\epsilon_0}\frac{q}{r^n},$$

show that the field $\mathbf{E}$ at a point r without or correspondingly within a spherical shell of radius r_0 and with surface charge density σ are given by the following expressions:

$$r > r_0: \quad E = \frac{\sigma r_0}{2\epsilon_0 r^2(3-n)(1-n)}[(r+r_0)^{2-n}\{r(2-n)-r_0\}-(r-r_0)^{2-n}\{r(2-n)+r_0\}],$$

$$r < r_0: \quad E = \frac{\sigma r_0}{2\epsilon_0 r^2(3-n)(1-n)}[(r+r_0)^{2-n}\{r(2-n)-r_0\}+(r-r_0)^{2-n}\{r(2-n)+r_0\}].$$

What can be deduced from this for $n < 2, n = 2, n > 2$?

Solution: Let $Q/4\pi r_0^2$ be the charge per unit area of the spherical shell with radius r_0. The quantity to be calculated is the electric field at the point P in Fig. 2.5. The charge on an infinitesimal element of area dF is $\sigma dF = \sigma r_0^2 \sin\theta d\theta d\psi$. From the formula for the law of force: Field at P equals

$$E_P = \text{charge}/4\pi\epsilon_0(\text{distance})^n,$$

we deduce that here

$$dE_P = \frac{\sigma r_0^2 \sin\theta d\theta d\psi}{4\pi\epsilon_0(r^2+r_0^2-2rr_0\cos\theta)^{n/2}}.$$

Integrating over the entire sphere we obtain the total field strength at P. From the symmetry of the geometry we infer that contributions perpendicular to OP cancel each other so that only the component along OP remains.
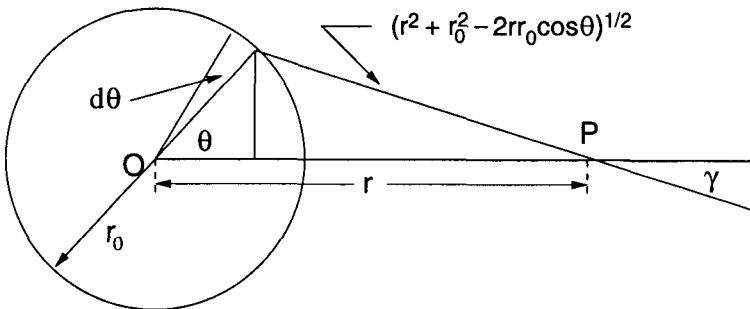
Fig. 2.5 Field point P outside the spherical shell.

We obtain this by multiplying dE_P by $\cos\gamma$ where (cf. Fig. 2.5):

$$\cos\gamma = \frac{r-r_0\cos\theta}{\sqrt{r^2+r_0^2-2rr_0\cos\theta}}.$$

We thus obtain for the field at P:

$$E_P = \int_0^{2\pi} \int_0^{\pi} \frac{\sigma r_0^2 (r - r_0 \cos\theta) \sin\theta d\theta d\psi}{4\pi\epsilon_0 (r^2 + r_0^2 - 2rr_0 \cos\theta)^{(n+1)/2}}.$$

With $z = \cos\theta$ the integral is of the form

$$\int_{-1}^{1} \frac{(a - bz)dz}{(a^2 + b^2 - 2abz)^{(n+1)/2}}.$$

But

$$\int \frac{dz}{(A + Bz)^{(n+1)/2}} = -\frac{2}{B(n-1)} \frac{1}{(A + Bz)^{(n-1)/2}}$$

and

$$\int \frac{zdz}{(A + Bz)^{(n+1)/2}} = \frac{1}{B} \int \frac{(A + Bz - A)dz}{(A + Bz)^{(n+1)/2}}$$

$$= -\frac{A}{B} \int \frac{dz}{(A + Bz)^{(n+1)/2}} + \frac{1}{B} \int \frac{dz}{(A + Bz)^{(n-1)/2}}.$$

If E_P is evaluated with the help of these integrals, one obtains the expressions for the field given at the beginning. We then see that for $n = 2$, the field E_P vanishes if $r < r_0$. This is the theoretical explanation of Cavendish's experiment.

2.5 Dirac's Delta Distribution

A solution of the Poisson equation $\triangle\phi = -4\pi k\rho$ is the Poisson integral (already encountered above):

$$\phi(\mathbf{r}) = k \int \frac{\rho(\mathbf{r}')d\mathbf{r}'}{|\mathbf{r} - \mathbf{r}'|}. \tag{2.4}$$

Before we show that this expression satisfies the Poisson equation, we introduce the so-called *δ-distribution* (an introduction is given in Appendix A). This quantity was introduced by Dirac and was originally called *delta function*. Actually it is an improper function which is today called *distribution*, and is defined as a functional by the property

$$f(a) = \int_{-\infty}^{\infty} f(x)\delta(x - a)dx.$$

$\delta(x)$ can be looked at as the limit of a function which vanishes everywhere except at $x = 0$, where it is singular. Thus, e.g.

$$\delta(x) = \lim_{\epsilon \to 0} \frac{1}{2\pi} \int_{-\infty}^{\infty} dk e^{-\epsilon|k|} e^{ikx}$$

$$= \lim_{\epsilon \to 0} \frac{1}{\pi} \frac{\epsilon}{\epsilon^2 + x^2}$$

$$= \begin{cases} 0 & : \quad x \neq 0, \\ \infty & : \quad x = 0, \end{cases} \tag{2.5}$$

or

$$\delta(x) = \lim_{a \to 0} \frac{1}{2\pi} \int_{-\infty}^{\infty} dk e^{-ak^2} e^{ikx}$$

$$= \lim_{a \to 0} \frac{1}{2\sqrt{\pi a}} e^{-\frac{x^2}{4a}}. \tag{2.6}$$

The delta function has the integral representation

$$\delta(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} dk e^{ikx}. \tag{2.7}$$

Important properties of $\delta(x)$ are:

$$\delta(x) = \delta(-x),$$
$$\delta'(x) = -\delta'(-x),$$
$$x\delta(x) = 0,$$
$$x\delta'(x) = -\delta(x),$$
$$\delta(ax) = \frac{1}{a}\delta(x), \quad (a > 0). \tag{2.8}$$

These relations can be verified by multiplication by a continuous differentiable function and followed by an integration (as in Appendix A). Furthermore the following relations can be shown to hold:

$$\delta(x^2 - a^2) = \frac{1}{2a}[\delta(x - a) + \delta(x + a)], \quad (a > 0),$$
$$f(x)\delta(x) = f(0)\delta(x),$$
$$\theta'(x) = \delta(x), \tag{2.9}$$

where the θ–function or *Heaviside step function* is defined by

$$\theta(x) = \begin{cases} 1 & : \quad x > 0, \\ 0 & : \quad x < 0. \end{cases} \tag{2.10}$$

Equations (2.9) are explicitly verified in Appendix A. We observe that

$$\delta(\mathbf{r}) \equiv \delta(x)\delta(y)\delta(z).$$

A solution of the inhomogeneous equation

$$\triangle G(\mathbf{r}) = \delta(\mathbf{r})$$

is called *Green's function* (we shall consider these in detail later). Once $G(\mathbf{r})$ is known, the inhomogeneous solution $\phi(\mathbf{r})$ of $\triangle\phi = -4k\pi\rho$ follows from the relation

$$\phi(\mathbf{r}) = -k \int G(\mathbf{r} - \mathbf{r}')4\pi\rho(\mathbf{r}')d\mathbf{r}' \tag{2.11}$$

apart from a solution of the homogeneous equation $\triangle\phi = 0$ which could be added subject to the same boundary conditions as for G. The relation (2.11) is readily verified, because if we apply the Laplace operator $\triangle$ to this, we obtain

$$\triangle\phi(\mathbf{r}) = -k \int \delta(\mathbf{r} - \mathbf{r}')4\pi\rho(\mathbf{r}')d\mathbf{r}' = -4k\pi\rho(\mathbf{r}).$$

For $\phi(\mathbf{r})$ to satisfy the integral relation (2.4) we must have

$$G(\mathbf{r} - \mathbf{r}') = -\frac{1}{4\pi|\mathbf{r} - \mathbf{r}'|}, \tag{2.12}$$

i.e.

$$\triangle\frac{1}{|\mathbf{r}|} = -4\pi\delta(\mathbf{r}). \tag{2.13}$$

This expression makes sense only for $\mathbf{r} \neq 0$. Now,

$$r = \sqrt{x^2 + y^2 + z^2}, \quad \triangle = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}$$

and

$$\frac{\partial}{\partial x}\left(\frac{1}{r}\right) = -\frac{x}{r^3}, \quad \frac{\partial^2}{\partial x^2}\left(\frac{1}{r}\right) = -\frac{1}{r^3} + \frac{3x^2}{r^5},$$

so that

$$\triangle\frac{1}{r} = -\frac{3}{r^3} + \frac{3r^2}{r^5} = 0.$$

We can see by integration that the factor -4π in (2.13) is correct:

$$\int_V \triangle\frac{1}{r}d\mathbf{r} = \int_V \boldsymbol{\nabla} \cdot \boldsymbol{\nabla}\frac{1}{r}d\mathbf{r} = \int_F \boldsymbol{\nabla}\frac{1}{r} \cdot d\mathbf{F}$$

$$= -\int_F \frac{\mathbf{r}}{r^3} \cdot d\mathbf{F} = -\int \frac{dF}{r^2} = -\int d\Omega = -4\pi.$$

This verifies Eq. (2.13). Equation (2.12) shows that the Green's function is the potential of a negative unit charge at the source point multiplied by ϵ_0 since with

$$\rho(\mathbf{r}') = -\epsilon_0\delta(\mathbf{r}')$$

we have

$$\phi(\mathbf{r}') = -k \int G(\mathbf{r} - \mathbf{r}') 4\pi \rho(\mathbf{r}') d\mathbf{r}' = G(\mathbf{r}).$$

We see also that the Green's function G is the response of a system (here represented by the Laplacian, but this could also be some other second order differential operator, such as that of an equation of motion in classical mechanics or of a Schrödinger equation in quantum mechanics) to a delta function-like source. One can also refer to output and input. In the expression for the potential $\phi(\mathbf{r})$ as an integral over the Green's function G the values of $-4k\pi\rho(\mathbf{r}')$ at points over which one integrates play the role of weight factors. The overall expression thus represents the sum or superposition of the responses of the system thus weighted to delta function-like sources or inputs.

Example 2.3: The Planck mass

Show that in a world with n spatial dimensions Newton's gravitational constant G raised to the power $1/(2-n)$ has in natural units the dimension of mass, which defines the so-called *Planck mass M_P*, i.e.

$$M_P := G^{-1/(2-n)},$$

for $\hbar = 1, c = 1$ and is dimensionless.

Solution: In n spatial dimensions

$$\Delta = \sum_{i=1}^{n} \frac{\partial^2}{\partial x_i^2},$$

and Newton's gravitational potential ψ is just like the Coulomb potential given by the solution of $\Delta\psi = 0$ with $r = \sqrt{\sum_{i=1}^{n} x_i^2} \neq 0$. For $\psi = 1/r^m$ we have

$$\frac{\partial\psi}{\partial x_i} = -\frac{mx_i}{r^{m+2}}, \qquad \sum_{i=1}^{n} \frac{\partial^2\psi}{\partial x_i^2} = -\frac{mn}{r^{m+2}} + \frac{m(m+2)}{r^{m+2}}.$$

For $\Delta\psi$ to be zero, we must have $m = n - 2$. The gravitational constant G is now defined by the potential $V = -G \cdot \text{mass}/r^{n-2}$. In natural units this potential has the dimension of mass M (recall that in Einstein's formula energy $=$ mass times c^2), if G has the dimension of $M^{-(n-2)}$.

Example 2.4: The delta function in various coordinates

What is $\delta(\mathbf{r})$ in three dimensional cylindrical and spherical polar coordinates?

Solution: Since $\int d\mathbf{r}' \delta(\mathbf{r} - \mathbf{r}') = 1$ and in cylindrical coordinates ρ, ψ, z and spherical polar coordinates r, θ, φ the volume element $d\mathbf{r} = \rho d\rho dz d\psi$ and $d\mathbf{r} = r^2 dr \sin\theta d\theta d\varphi$ respectively, we have

$$\delta(\mathbf{r} - \mathbf{r}')|_{\text{cyl}} = \frac{1}{\rho'} \delta(\rho - \rho') \delta(\psi - \psi') \delta(z - z')$$

and

$$\delta(\mathbf{r} - \mathbf{r}')|_{\text{spher}} = \frac{1}{r'^2} \delta(r - r') \delta(\cos\theta - \cos\theta') \delta(\varphi - \varphi')$$

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

Example 2.5: Representations of delta functions

By appropriate evaluation of the integral

$$\delta(\mathbf{x}) = \lim_{\epsilon \to 0} \frac{1}{(2\pi)^n} \int d^n k e^{-\epsilon |\mathbf{k}|} e^{i\mathbf{k} \cdot \mathbf{x}}$$

for $n = 1$ and 2 determine representations of the 1- and 2-dimensional delta functions.

Solution: The case $n = 1$ is simple:

$$\frac{1}{2\pi} \int_{-\infty}^{\infty} dk e^{-\epsilon |k|} e^{ikx} = \frac{1}{\pi} \int_{0}^{\infty} dk e^{-\epsilon k} \cos kx = \frac{1}{\pi} \frac{\epsilon}{\epsilon^2 + x^2}, \quad (\epsilon > 0).$$

We thus obtain the following representation for the delta function in one dimension:

$$\delta(x) = \lim_{\epsilon \to 0} \frac{1}{2\pi} \int_{-\infty}^{\infty} dk e^{-\epsilon |k|} e^{ikx} = \lim_{\epsilon \to 0} \frac{1}{\pi} \frac{\epsilon}{\epsilon^2 + x^2}.$$

Higher dimensional cases are not quite so easy. In two dimensions we have

$$\frac{1}{(2\pi)^2} \int dk_x dk_y e^{-\epsilon \sqrt{k_x^2 + k_y^2}} e^{i(k_x x + k_y y)} = \frac{1}{(2\pi)^2} \int_{0}^{\infty} k dk \int_{-\pi}^{\pi} d\theta e^{-\epsilon k} e^{ik|\mathbf{x}| \cos \theta}$$

$$= \frac{1}{(2\pi)^2} \int_{-\pi}^{\pi} d\theta \left(-\frac{d}{d\epsilon} \right) \frac{1}{\epsilon - i|\mathbf{x}| \cos \theta}.$$

The integral

$$\int_{-\pi}^{\pi} \frac{d\theta}{\epsilon - ix \cos \theta}$$

can be evaluated with the help of Tables of Integrals and with caution in the two cases $\epsilon^2 \lessgtr \mathbf{x}^2$. In both cases one obtains $2\pi/\sqrt{\epsilon^2 + \mathbf{x}^2}$. Hence the representation that follows is

$$\delta(\mathbf{x}) = \lim_{\epsilon \to 0} \frac{1}{(2\pi)^2} \int d^2 k e^{-\epsilon |\mathbf{k}|} e^{i\mathbf{k} \cdot \mathbf{x}} = \lim_{\epsilon \to 0} \frac{\epsilon}{2\pi (\epsilon^2 + \mathbf{x}^2)^{3/2}}.$$

Representations of the type discussed here are frequently useful in applications.

Example 2.6: Further representations of the delta function

Show that[*]

$$K(\mathbf{x}) = \lim_{x_0 \to 0} \tilde{K}(\mathbf{x}), \quad \tilde{K}(\mathbf{x}) = c \frac{x_0^n}{[x_0^2 + \sum_{j=1}^{n} x_j^2]^n}$$

is a delta function with coefficient 1 if the constant c is chosen suitably.

Solution: In order to establish the result, one has to show that (a) $K(\mathbf{x}) = 0$ for $x_j \neq 0$, and (b) $\int d\mathbf{x} K(\mathbf{x}) = 1$ and independent of x_0. We limit ourselves here exercise-wise to a few remarks. First of all we convince ourselves that a rescaling of $x_i \to x_i' = x_i/x_0$ verifies that the integral over K is independent of x_0:

$$\int \tilde{K}(\mathbf{x}) d\mathbf{x} = c \int \frac{x_0^n d\mathbf{x}}{[x_0^2 + \mathbf{x}^2]^n} = c \int \frac{d\mathbf{x}'}{[1 + \mathbf{x}'^2]^n}.$$

---

[*]See also E. Witten, *Adv. Theor. Math. Phys.* **2** (1998) 253; hep-th/9802150.

By going to polar coordinates it is not too difficult to verify that for example in the case $n = 2$: $c = 1/\pi$, in the case $n = 3 : c = 4/\pi^2$, and so on. Also the following representation can be verified:

$$\delta(\mathbf{x}) = \lim_{x_0 \to 0} \frac{(2h_+ - 1)! x_0^{-2h_-}}{\pi(2h_+ - 2)!} \left[\frac{x_0}{x_0^2 + x^2} \right]^{2h_+} \equiv \lim_{x_0 \to 0} K(\mathbf{x}), \quad \text{with } h_+ + h_- = 1.$$

For example we have for $n = 2$ (in this case $h_+ = 1, h_- = 0$):

$$\delta(\mathbf{x}) = \lim_{x_0 \to 0} \frac{x_0^2}{\pi(x_0^2 + \mathbf{x}^2)^2}.$$

We shall encounter in particular this case $n = 2$ in Chapter 21 in the regularisation of the field of a magnetic monopole.

Example 2.7: The Coulomb potential in higher dimensions
Determine the Coulomb potential ϕ in $N > 3$ space dimensions.

Solution: One way to solve this problem is the following. In 3 dimensions

$$\boldsymbol{\nabla} \cdot \mathbf{E} = \frac{\rho}{\epsilon_0}, \quad \boldsymbol{\nabla} = \left(\mathbf{e}_r \frac{\partial}{\partial r}, \mathbf{e}_\varphi \frac{1}{r \sin\theta} \frac{\partial}{\partial \varphi}, \mathbf{e}_\theta \frac{1}{r} \frac{\partial}{\partial \theta} \right).$$

For charge e at $\mathbf{r}' = \mathbf{r}$ we have (cf. Example 2.4)

$$\rho(\mathbf{r} - \mathbf{r}') = e\delta(\mathbf{r} - \mathbf{r}') = \frac{e}{r'^2} \delta(r - r')\delta(\cos\theta - \cos\theta')\delta(\varphi - \varphi').$$

Thus (since the field is radial),

$$E_r = \int^r \frac{\rho(r')}{\epsilon_0} dr'$$

and with further integrating out the angles

$$\int_0^{2\pi} d\varphi' \int_0^{-1} d\cos\theta' E_r = \int dr' \frac{e}{\epsilon_0 r'^2} \delta(r - r')\delta(\varphi - \varphi')\delta(\cos\theta - \cos\theta')d\varphi' d\cos\theta',$$

i.e.

$$-4\pi E_r = \frac{e}{\epsilon_0 r^2}, \quad \phi = -\frac{e}{4\pi\epsilon_0 r}.$$

For the case of N space dimensions we require the N-dimensional generalisation, in particular

$$dr = r^{N-1}(\sin\theta_1)^{N-2}(\sin\theta_2)^{N-3} \cdots (\sin\theta_{N-2})^2 dr d\theta_1 \cdots d\theta_{N-1} \equiv r^{N-1} dr d\Omega_{N-1}.$$

The solid angle Ω_{N-1} can be deduced from the equality of the following integrals

$$\int d^N x e^{-\frac{1}{2}\sum_{i=1}^N x_i^2} = \int_0^\infty dr r^{N-1} d\Omega_{N-1} e^{-\frac{1}{2}r^2},$$

and so from

$$(2\pi)^{N/2} = \frac{2^{N/2}}{N} \left(\frac{N}{2} \right)! \, \Omega_{N-1}.$$

Hence

$$\Omega_{N-1} = \frac{N\pi^{N/2}}{(N/2)!}.$$

For $N = 3$ with $(-1/2)! = \sqrt{\pi}$ (see gamma functions and factorials in Tables of Special Functions), we obtain $\Omega_2 = 4\pi$, as expected. Hence

$$E_r = -\frac{e}{\Omega_{N-1}\epsilon_0 r^{N-1}} \quad \text{and} \quad \phi(r) = -\frac{e}{(N-2)\Omega_{N-1}\epsilon_0 r^{N-2}}.$$

Example 2.8: The Coulomb potential in space dimensions 1 and 2
What is the space dependence of the Coulomb potential in a world with one space dimension?
What is it in the case of 2 space dimensions?

Solution: We have $\nabla \cdot \mathbf{E} = 4\pi k\rho$ with $\mathbf{E} = -\nabla\phi$; hence in the case of one space dimension

$$\frac{d^2\phi}{dx^2} = -4\pi k\rho = -\frac{\rho}{\epsilon_0}.$$

With the Green's function $G(x)$, the solution is

$$\phi(x) = \int dx' G(x - x')\left(-\frac{\rho(x')}{\epsilon_0}\right), \quad \frac{d^2 G(x)}{dx^2} = \delta(x).$$

The Green's function may be obtained by 'trial and error' or, better, by contour integration which will be explained later. The result is

$$G(x) = \frac{1}{2\pi} \int \frac{e^{ikx}}{-k^2} dk = \frac{1}{2}|x|,$$

as one can verify by differentiation. Thus it follows that in particular for $\rho(x') = e\delta(x' - x_0)$:

$$\phi(x) = -\frac{1}{\epsilon_0} \int dx'\rho(x')|x - x'|, \quad \text{or} \quad \phi(x) = -\frac{e}{\epsilon_0}|x - x_0|.$$

Such a linear potential is known as a *confinement potential*, since particles (like quarks) bound together by such a potential cannot be separated with finite energy. (See also Example 18.5).
In the case of 2 space dimensions one can go to polar coordinates r, φ, so that

$$\left(\frac{\partial^2}{\partial r^2} + \frac{1}{r}\frac{\partial}{\partial r}\right)\phi = -\frac{\rho(r)}{\epsilon_0}.$$

In a similar way one finds that $\phi = \text{const.} \ln r$.

2.6 Potential Energy of Charges in Electric Fields

Let a charge q be moved from A to B *against* the field $\mathbf{E}$. The work done in this process is

$$W = -\int_A^B q\mathbf{E}(\mathbf{r}) \cdot d\mathbf{r} = q \int_A^B \nabla\phi \cdot d\mathbf{r}$$

$$= q \int_A^B d\phi = q(\phi(B) - \phi(A)). \tag{2.14}$$

The minus sign indicates that the force exerted on the charge q is directed opposite to the field $\mathbf{E}$. The final result shows that the work done is independent of the path from A to B, that is, $q\mathbf{E}$ is a *conservative* field, i.e. $\nabla \times \mathbf{E} = 0$.

2.7   The Electric Field at Charged Surfaces

We now consider the behaviour of the electric field at charged surfaces or films.[†] Following Faraday it is very instructive to draw lines of force starting from positive charges and aiming at negative charges. Already in drawing these lines of force, as in Fig. 2.6, one observes that the electric field suffers a discontinuity at the charged surface, because, as the term "surface" implies (with the same medium on either side), its charge is the same on both sides, so that the vector $\mathbf{E}$ on one side is opposite to that on the other side.
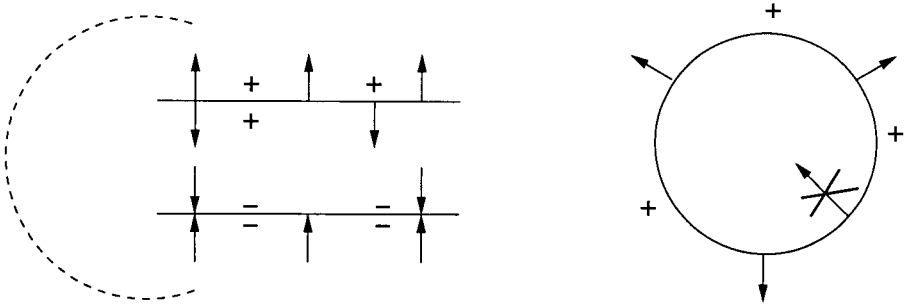
Fig. 2.6 Charged surfaces and lines of force.

Since we wish to consider charged surfaces, we first define as *charge per unit area* the expression $\sigma(\mathbf{r})$ in the following relation, in which q is the total charge of a surface area $\triangle F$, i.e.

$$q = \int_{\triangle F} \sigma(\mathbf{r})dF.$$

If $\mathbf{E}_1$ and $\mathbf{E}_2$ are as in Fig. 2.6 the vectors $\mathbf{E}$ at a point on the surface but on either sides and directed away as for a positive charge, then we write the difference of their components along the surface normals $E_{2n} - E_{1n}$ and the difference along a surface tangent $E_{2t} - E_{1t}$. We let $\mathbf{n}$ be a unit vector along the surface normal on the side with index "2" of the charged surface. In the first place we consider a surface of thickness d and volume $\triangle V = d\triangle F$. We apply the Gauss law to this volume and then allow d to approach zero, i.e.

[†]We are not yet considering macroscopic electrostatics and therefore have not yet introduced the dielectric displacement $\mathbf{D}$. Nonetheless a comment seems appropriate at this point. Since the expression for $\mathbf{E}$ contains the factor $1/\epsilon_0$ or $1/\epsilon$, the displacement $\mathbf{D}$ as the product $\epsilon\mathbf{E}$ is independent of the dielectric constant, and this implies the continuity of $\mathbf{D}$ at the interface between media. One should note the difference between this and the case to be considered here, i.e. that of a charged surface in the same medium.

we consider

$$\int_{F(\triangle V), d \to 0} \mathbf{E} \cdot d\mathbf{F} = 4k\pi q = 4k\pi \int_{\triangle V, d \to 0} \sigma(\mathbf{r})dF$$
$$= 4k\pi\sigma\triangle F. \tag{2.15}$$

For thickness $d \to 0$ we have

$$\int_{F(\triangle V), d \to 0} \mathbf{E} \cdot d\mathbf{F} = (\mathbf{E}_2 + \mathbf{E}_1) \cdot \triangle \mathbf{F} = (\mathbf{E}_2 - \mathbf{E}_1) \cdot \mathbf{n}\triangle F,$$

where $\mathbf{n}$ is a unit vector directed along the normal on side "2" of the surface. It follows that

$$(\mathbf{E}_2 - \mathbf{E}_1) \cdot \mathbf{n} = 4k\pi\sigma(\mathbf{r}). \tag{2.16}$$

Thus in passing through the charged surface of vanishing thickness the normal component of the electric field $\mathbf{E}$ jumps from one value to another depending on $\sigma(\mathbf{r})$.

We consider two particular cases. As the first we consider the case of a *large charged plate*, whose boundary effects (distortions of lines of force) may be neglected (hence the specification 'large'). In this case we have as depicted in Fig. 2.7,

$$\mathbf{E}_2 \cdot \mathbf{n} = E, \quad \mathbf{E}_1 \cdot \mathbf{n} = -E,$$

so that

$$2E = 4\pi k\sigma, \quad E = \frac{\sigma}{2\epsilon_0}.$$

Thus the field has the same value $2\pi k\sigma$ on both sides of the plate, but is in each case directed away from the plate; the difference is $4\pi k\sigma$.
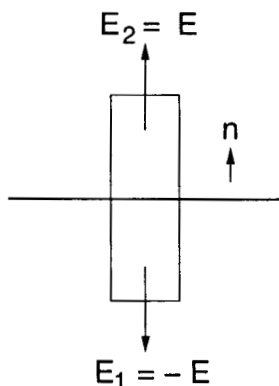
Fig. 2.7 Large charged plate.

These results, in particular formula (2.16), are of considerable importance and will be used in many examples. In the second particular case which we consider, we assume that the plate is made of an *ideal conducting material*,[‡] i.e. a material in which the electrons can move about freely without doing any work. We let the thickness of the plate be $d \neq 0$. In this case the electrostatic potential $\phi = $ const., so that inside the plate the electric field vanishes, i.e. $E_1 = 0$, and the field *above* the plate is $E = 4\pi k\sigma$.

The discontinuity formula is valid at every surface. If $\mathbf{E}$ has the same value on both sides of the surface, then $\sigma = 0$ on this surface. Conversely if $\mathbf{E}_1 \cdot \mathbf{n}, \mathbf{E}_2 \cdot \mathbf{n}$ are different on both sides, the surface carries a charge. This is generally valid for the refraction of light at an interface.

In the preceding we considered the fields $\mathbf{E}_1, \mathbf{E}_2$ as originating solely from the one and only charged surface considered. We can drop this restriction now and allow charges and surfaces anywhere in space. These have the effect of distorting the lines of force of $\mathbf{E}_1$ and $\mathbf{E}_2$, and to change their magnitudes. This does not change anything in the previous considerations since these involve only the difference

$$\mathbf{E}_1 \cdot \mathbf{n}_1 + \mathbf{E}_2 \cdot \mathbf{n}_2 = E_{1n} - E_{2n}.$$

Finally we consider the *tangential components*. Since $\mathbf{E}$ is a conservative field, it follows that $\oint \mathbf{E} \cdot dr = 0$ (along a closed path). If we choose a path along just above the surface to just below as in Fig. 2.8, then

$$(\mathbf{E}_{2t} - \mathbf{E}_{1t}) \cdot \triangle \mathbf{r} = 0,$$

i.e. the tangential components along an arbitrary direction remain unchanged:

$$\mathbf{E}_{2t} = \mathbf{E}_{1t}. \tag{2.17}$$
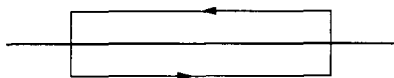
Fig. 2.8 The closed path above and below the surface.

Comment

The basis of our considerations so far was the experimentally determined Coulomb law. If one recalls that the Coulomb law has the same form as

---

[‡]Conductors will be treated in detail later.

Newton's gravitational law, one may enquire about further analogies between electrostatics and gravitation theory. Indeed a relation analogous to that for the electric field at a charged surface applies also for the gravitational field at a surface of mass. For applications we refer to the literature.[§]

2.8 Examples

The following examples are needed later. We therefore treat them in some detail. Most of these examples consider condensers which are combinations of conductors separated by an insulating material (i.e. a dielectric). However, the second conductor is not always apparent as such, e.g. the walls of a room in which an insulated conductor is placed (see e.g. Example 3.13).

Example 2.9: The parallel plate condenser

Two parallel plates made of conductor material and separated by a distance d are given, one with charge $+q$, the other with charge $-q$, as depicted in Fig. 2.9. Determine the capacity of the condenser.

Solution: Let σ be the charge per unit area, i.e. $\sigma = q/F$. We use the coordinate system $O(x, y)$ with origin at the centre of the condenser and the y-axis perpendicular to the plates. The lines of force (in the direction of $\mathbf{E}$) proceed from $+q$ to $-q$. We apply the Gauss law to the cylindrical volume with cross sectional area $\triangle F$, i.e.

$$\oint \mathbf{E} \cdot d\mathbf{F} = 4\pi k q.$$

Since the total charge contained in the cylindrical volume is zero, i.e. $\sigma \triangle F + (-\sigma)\triangle F = 0$, and since the tangential components of the field
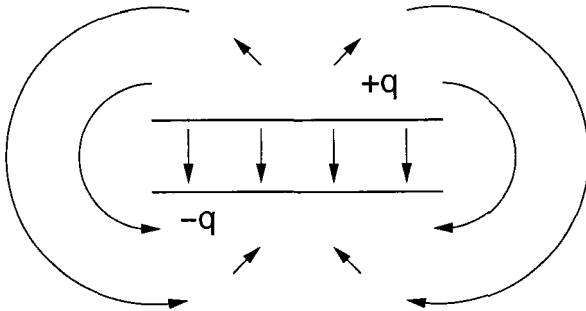
Fig. 2.9 The parallel plate condenser.

(parallel to the condenser plates) vanish, we obtain

$$E(y) - E(-y) = 0.$$

[§]S. K. Blau, E. I. Guendelman and A. H. Guth, *Phys. Rev.* **D35** (1987) 1747; see in particular Eqs. (3.3) to (3.4), (4.11) and (4.27).

This relation expresses the symmetry of the field on both sides above and below the condenser plates. In the case of a *finite condenser* it is difficult to compute the field in the external domains. The reason is the *deformation of the lines of force at the open ends* as indicated in Fig. 2.9. However, we can approximate this finite condenser by a segment of a large spherical condenser consisting of two concentric shells. As we shall see in Example 2.10 the field outside is zero. In some books it is simply assumed that the field outside is zero. We see that this is an approximation. We now apply Gauss' law to a single plate and make use of the approximations just explained (i.e. that the field outside is zero); then in Eq. (2.16) we have to put $\mathbf{E}_1 = 0$. Then it follows with $\mathbf{E}_2 \equiv \mathbf{E}$ that

$$(\mathbf{E} - 0) \cdot \mathbf{n} \triangle F = 4\pi k \sigma \triangle F,$$

i.e. the field in the condenser is given by

$$E = 4\pi k \sigma, \quad k = 1/4\pi\epsilon_0. \tag{2.18}$$

The difference V of the potentials betwen the two plates is defined as

$$V = \phi_2 - \phi_1 = \int_0^d \mathbf{E} \cdot d\mathbf{r} = Ed = 4\pi k \sigma d.$$

The *capacity* (American usage: capacitance) C of a condenser is defined as

$$C = \frac{q}{V} = \frac{q}{4\pi k \sigma d} = \frac{F}{4\pi dk}. \tag{2.19}$$

Capacity is a measure of the ability of a system to store electric energy. We saw previously that in MKSA units the constant k has a complicated dimension. The unit of capacity in this system is the farad (F), i.e. 1 farad = 1 coulomb per volt. In the Gaussian system the capacity C is given in cm with 1 farad $\Leftrightarrow 9 \times 10^{11}$ cm. Since a farad is a huge quantity, it is customary to employ micro-farads with 1 μF $= 10^{-6}$ F.

Example 2.10: The spherical condenser

The condenser consists of two conducting, concentric, spherical surfaces with internal radius r_1 and external radius r_2 and charge $+q$ on the inner shell and $-q$ on the outer shell. Determine the capacity of the condenser.

Solution: In view of the spherical symmetry of the condenser, the field components $\mathbf{E}_\theta, \mathbf{E}_\phi = 0$.
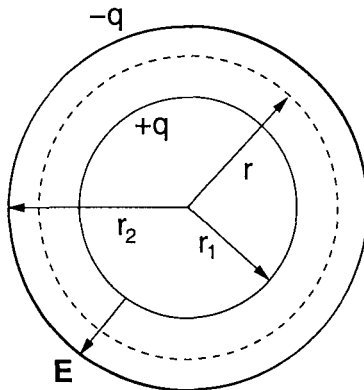
Fig. 2.10 The spherical condenser.

We consider a spherical surface of radius $r_1 - \epsilon, \epsilon > 0$. Then, according to the Gauss law,

$$\int_F \mathbf{E} \cdot d\mathbf{F} = 4\pi k \int_{V(F)} \rho(\mathbf{r}')d\mathbf{r}' = 0,$$

since the volume $V(r_1 - \epsilon)$ does not enclose a charge. Hence

$$\int_F \mathbf{E} \cdot d\mathbf{F} = 4\pi r_1^2 E_{r_1} = 0, \quad E_{r_1} = 0.$$

This means that the field in the interior of a charged closed surface is zero, i.e. this is field-free.[¶] Thus the field of the condenser originates from the inner shell. We now consider a sphere of radius $r > r_1$:

$$\int \mathbf{E} \cdot d\mathbf{F} = 4\pi r^2 E_r = 4\pi k q,$$

so that

$$E_r = E = k\frac{q}{r^2}$$

as for a point charge $+q$ at the origin. Outside of r_2, i.e. for $r > r_2$, the field again vanishes, because there the field contributions of the two equal and opposite charge distributions cancel each other (one could say, the field there corresponds to the sum of the fields of point charges $+q$ and $-q$ at the origin). With this we obtain for the potential V of the condenser

$$V = \phi_2 - \phi_1 = qk \int_{r_1}^{r_2} \frac{dr}{r^2}$$
$$= kq\left(\frac{1}{r_1} - \frac{1}{r_2}\right),$$

and for the capacity

$$C = \frac{q}{V} = \frac{r_1 r_2}{k(r_2 - r_1)}.$$

With reference to a *single* spherical shell we can define its capacity as

$$C = \text{charge/potential},$$

so that in the case of the inner sphere

$$C = \frac{q}{\phi_1} = \frac{qr_1}{kq}, \quad \text{i.e.} \quad C = 4\pi\epsilon_0 r_1.$$

Example 2.11: The van de Graaff generator

A van de Graaff generator has a spherically shaped dome made of conducting material which receives charges at the rate of a current of 1 μA. The dome has a radius of 1 m. At a critical electric field strength of 3×10^6 V/m the surrounding air produces sparks. How many seconds after the beginning of the charge transfer to the dome does this effect appear? $(4\pi\epsilon_0 = 10^7/c^2)$.

Solution: In the van de Graaff generator charge is sprayed on a transmission belt made of insulator material. The belt transports the charge to the conducting dome of radius R as depicted in Fig. 2.11. For a single spherical shell the capacity C is defined by the ratio of charge to potential, as we saw previously, so that with potential $\phi = kq/R$ the capacity is given by

$$C = \frac{q}{\phi} = \frac{qR}{kq} = 4\pi\epsilon_0 R$$

[¶]This was the observation of Cavendish in his experiment which we discussed earlier.
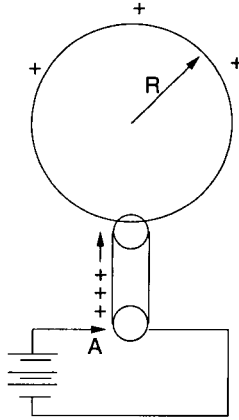
Fig. 2.11   The van de Graaff generator.

for a charge q. If charge is continuously added, the potential ϕ changes accordingly, i.e.

$$C\frac{d\phi}{dt} = \frac{dq}{dt} = I,$$

where I is the current. We thus have the relation

$$\frac{d\phi}{dt} = \frac{I}{C}, \quad \phi = \frac{I}{C}t$$

with $\phi = 0$ at time $t = 0$. The potential or voltage increases until the corresponding field $\mathbf{E}$ around the sphere is strong enough to knock electrons out of the surrounding air — the corresponding atomic transitions then become visible in the form of sparks. In air this stage is attained when $E_{\text{crit}} = 3 \times 10^6$ V/m. The field $\mathbf{E}$ in the neighbourhood of the dome is given by

$$E = k\frac{q}{r^2}, \quad r > R.$$

For $r \sim R$:

$$E = k\frac{q}{R^2} = \frac{\phi}{R},$$

or $\phi = ER$, so that $\phi_{\text{crit}} = E_{\text{crit}}R$. For $\phi_{\text{crit}} = 3 \times 10^6$ V, we have $R = 1$m and

$$C = 4\pi\epsilon_0 R = \frac{10^6}{c^2} \times 1 = \frac{10^7}{(3 \times 10^8)^2} \quad \text{farad.}$$

This implies that

$$C = \frac{1}{9} \text{ nF.}$$

Thus the appropriate current I is: $I = \phi_{\text{crit}}C/t$, i.e.

$$I = \frac{3 \times 10^6 V}{9 \times 10^9 \text{F}} \frac{1}{t} \text{ A} = \frac{0.3 \times 10^{-3}}{t} \text{ A.}$$

For $I = 1 \times 10^{-6}$ A:

$$1 \times 10^{-6} \text{ A} = \frac{0.3 \times 10^{-3}}{t} \text{ A}$$

we obtain $t = 0.3 \times 10^3$ seconds. Thus if the dome is charged at the rate of a current of 1 micro-ampere, it takes about 300 seconds for the potential to reach its maximal value.

Example 2.12: The cylindrical condenser

This condenser consists of two coaxial cylinders of height h and radii $r_1, r_2, r_2 > r_1$. The inner cylinder carries charge q, the outer cylinder charge $-q$. What is the capacity of the condenser?

Solution: Clearly one uses cylindrical coordinates θ, ρ, z in order to exploit the cylindrical symmetry.
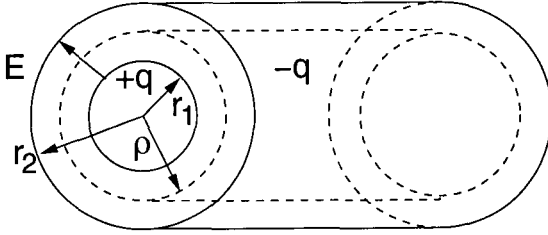
Fig. 2.12 The cylindrical condenser.

We consider a cylindrical surface of radius ρ where $r_2 > \rho > r_1$. In view of the cylindrical symmetry $E_\theta = E_z = 0, E_\rho = E$. The Gauss law implies

$$\oint \mathbf{E} \cdot d\mathbf{F} = E_\rho \cdot 2\pi\rho h = 4\pi k q,$$

so that

$$E_\rho = \frac{2qk}{\rho h} \frac{q}{2\pi\epsilon_0 h\rho}.$$

We thus obtain for the potential V:

$$V = \int_{r_1}^{r_2} E_\rho d\rho = \frac{2qk}{h} \int_{r_1}^{r_2} \frac{d\rho}{\rho} = \frac{2qk}{h} \ln \frac{r_2}{r_1}$$

and for the capacity

$$C = \frac{q}{V} = \frac{h/k}{2\ln(r_2/r_1)}.$$

The finite cylindrical condenser can be approximated by a segment of a toroidal condenser, like the parallel plate condenser can be approximated by a segment of the spherical condenser.

Example 2.13: The dipole

A dipole is generally defined as a system of two spatially separated charges $+q, -q$. Obviously an infinitesimally imagined parallel plate condenser is of this form. We now want to obtain the potential of such an arrangement of charges, and we shall define in this context the important concept of a dipole moment.

Solution: We consider first two infinitesimal parallel plates with surface areas dF', dF'', and charge densities $+\sigma(\mathbf{r}'), -\sigma(\mathbf{r}')$, which are elements of surfaces S', S'' with charges $+\sigma, -\sigma$ as depicted in Fig. 2.13. The potential at the point P is given by

$$\phi(\mathbf{r}) = k \int \frac{\sigma(\mathbf{r}')dF'}{|\mathbf{r} - \mathbf{r}'|} - k \int \frac{\sigma(\mathbf{r}')dF'}{|\mathbf{r} - \mathbf{r}' + d\mathbf{r}'|}.$$

In the following one should not confuse the line element dr with the volume element $d\mathbf{r}$ (the distinction should be clear from the context). We have

$$\frac{1}{|\mathbf{r} - \mathbf{r}' + dr'|} = \frac{1}{\{(\mathbf{r} - \mathbf{r}')^2 + 2dr' \cdot (\mathbf{r} - \mathbf{r}') + (dr')^2\}^{1/2}}$$

$$\simeq \frac{1}{|\mathbf{r} - \mathbf{r}'|}\left(1 - \frac{(\mathbf{r} - \mathbf{r}') \cdot dr'}{(\mathbf{r} - \mathbf{r}')^2} + \cdots\right).$$
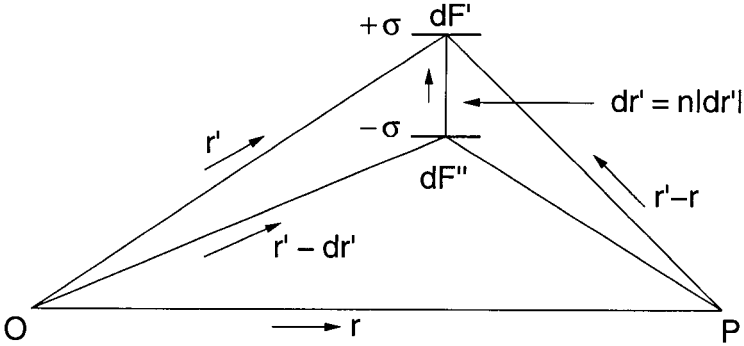
Fig. 2.13 The dipole as an infinitesimal condenser.

It follows that

$$\phi(\mathbf{r}) \simeq k \int \frac{\sigma(\mathbf{r}')(\mathbf{r} - \mathbf{r}') \cdot dr'}{(\mathbf{r} - \mathbf{r}')^3} dF'.$$

We put

$$D(\mathbf{r}') \equiv \lim_{dr' \to 0} \sigma(\mathbf{r}')|dr'|.$$

The quantity D is called *dipole density* (i.e. we consider a case in which the potential between the surface elements dF', dF'' remains constant, when $dr' \to 0$ and $\sigma(\mathbf{r}') \to \infty$). An *infinitesimal dipole surface* (area $\triangle F$) in the limit $\to 0$ is sometimes also described as a *point dipole*. Then (in the limit $dr' \to 0$)

$$\phi(\mathbf{r}) \simeq k \int \frac{D(\mathbf{r}')\mathbf{n} \cdot (\mathbf{r} - \mathbf{r}')}{(\mathbf{r} - \mathbf{r}')^3} dF'.$$

Let $d\Omega$ be the solid angle subtended by the area dF' at P. Let $\mathbf{n}$ be a unit vector perpendicular to dF' and directed to the outside (i.e. $\mathbf{n}$ parallel to dr'). It then follows from Fig. 2.13 that

$$\frac{\mathbf{n} \cdot (\mathbf{r}' - \mathbf{r})dF'}{(\mathbf{r} - \mathbf{r}')^3} = d\Omega$$

and hence

$$\phi(\mathbf{r}) = -\int_F D(\mathbf{r}')d\Omega$$

(the minus sign appears because $d\Omega$ involves the vector $(\mathbf{r}' - \mathbf{r})$ and not $(\mathbf{r} - \mathbf{r}')$ like ϕ). Replacing

F by $\triangle F$ we obtain *the potential of a point dipole*:

$$\phi(\mathbf{r}') = -k \int_{\triangle F} D(\mathbf{r}') d\Omega$$

$$= -k \int_{\triangle F} \frac{\sigma(\mathbf{r}')|d\mathbf{r}'|dF'\mathbf{n} \cdot (\mathbf{r}' - \mathbf{r})}{(\mathbf{r} - \mathbf{r}')^3}$$

$$= -k \frac{\mathbf{n} \cdot (\mathbf{r}' - \mathbf{r})}{(\mathbf{r} - \mathbf{r}')^3} |d\mathbf{r}'| \int_{\triangle F} \sigma(\mathbf{r}') dF'$$

$$= -k\mathbf{p} \cdot \frac{(\mathbf{r}' - \mathbf{r})}{(\mathbf{r} - \mathbf{r}')^3}, \qquad (2.20)$$

where

$$\mathbf{p} = \mathbf{n}|d\mathbf{r}'|Q, \quad Q = \int_{\triangle F} \sigma(\mathbf{r}') dF'.$$

The expression $\mathbf{p}$ is called *dipole moment* of the point dipole; this is a vector quantity with *direction from the negative charge to the positive charge*. Q is the charge.

2.9   Conductors and Electrical Screening

A *conductor* is a macroscopic object in which charges are contained which can move about freely (i.e. without doing work). Inside the conductor the electric field $\mathbf{E}$ therefore vanishes, because only then a charge can move about freely in the conductor. It follows that if the conductor is given a charge, this charge must be located in a thin layer at its surface. This must be so because if we consider an arbitrary closed Gaussian surface wholly within the conductor material, this cannot contain charge since everywhere on the surface $\mathbf{E} = 0$. Thus inside the conductor the charge is zero at every point (the charges of the conductor material itself averaging out to zero). It follows that the surface of the conductor is an *equipotential surface*.

Example 2.14: The charged sphere
Let a spherically shaped conductor be given with charge q. Explain with the help of potential considerations that the charge accumulates on the outer surface of the conductor.

Solution: We consider a sphere with external radius b and an internal hollow space of radius $r_0 = a < b$. We subdivide the intermediate region into concentric spherical shells of radii $r_1, r_2, \ldots$. Each shell initially has charge zero. Since the sphere is a conductor, each of its parts has the same constant potential. If we give, e.g. the sphere with radius r_2 the charge q, the potential changes for all shells with radius $r_i, i > 3$. Thus the gradient of a potential arises. Obviously the energy of the system becomes minimal when the charge q is located on the outermost shell (recall that the modulus of the Coulomb potential decreases with increase of the distance from the charge).

The electric field inside arbitrarily shaped conducting bodies whose surfaces are charged but which do not contain any enclosed charges is always zero. This is so, because if we apply to the interior the integral form of the

law of Gauss, we obtain

$$\int_{F'} \mathbf{E} \cdot d\mathbf{F} = 0.$$

The right hand side is zero, since the surface F' does not enclose charges. If there were inside the given positively charged, closed surface a field $E \neq 0$ then the field would be directed from the surface towards the interior. All contributions to the integral would then be positive. The integral would be a sum of positive contributions. All these would therefore have to vanish, implying that the field inside is zero. This result is apparently also valid for homogeneous charge distributions on the surface. Such closed bodies which do not contain charges are called *Faraday cages* because Faraday crept into such a container in an attempt to measure the field therein.[||] Containers of this type are used when field-free regions are needed. That the field inside closed charged bodies has to be zero can also be seen by the impossibility to sketch lines of force in the interior which travel from positive charges to negative charges.

A conductor is said to be *earthed* if its potential is the same as at infinity, i.e. zero. This effect is achieved in practice by connection with the ground, i.e. the "Earth".

If a body, e.g. a solid sphere, is made of a nonconducting material, that is a dielectric medium, which is polarised in an applied electric field, then one part of the body can be positively charged and another negatively. These cases will be treated later.

Finally we consider some *systems of conductors and their electric screening*. Above we obtained the Poisson equation for the electric potential ϕ. Assuming also the *superposition principle*, we can treat the case of n conductors with potentials ϕ_i and charges $q_i, i = 1, 2, \ldots, n$, in vacuum or air by writing

$$\phi_i = \sum_{j=1}^{n} p_{ij} q_j$$

or inversely

$$q_i = \sum_{j=1}^{n} C_{ij} \phi_j.$$

The coefficients C_{ij} are described as *capacities*, and the coefficients p_{ij} as *induction coefficients*. Obviously these coefficients depend on the geometry of the conductors and their arrangement in space. In the special case of two

---

[||]See e.g. Zahn [8], pp. 77–78.

conductors with charges q and $-q$ (as in preceding examples), we obtain

$$\phi_1 = (p_{11} - p_{12})q, \quad \phi_2 = (p_{21} - p_{22})q,$$

or

$$\phi_1 - \phi_2 = (p_{11} - p_{12} - p_{21} + p_{22})q.$$

Hence

$$V = \phi_1 - \phi_2 = \frac{q}{C},$$

where

$$C = \frac{1}{(p_{11} - p_{12} - p_{21} + p_{22})}.$$

Expressed in terms of the coefficients C_{ij}, one obtains

$$C = \frac{C_{11} + C_{12} + C_{21} + C_{22}}{C_{11}C_{22} - C_{12}C_{21}}.$$

It can be shown that**

$$C_{ij} = C_{ji}, \quad p_{ij} = p_{ji}.$$

Electrical screening results if one conductor is completely surrounded by another conductor. Instead of considering the most general case, we consider here the example of 3 concentric spherical shells which provide such enclosures. We will find that, as illustrated in Fig. 2.14, conductor 1 is screened off by conductor 3 and other external conductors in the sense that the coefficient $C_{13} = 0$.

We consider the three concentric spherical conductor shells with charges q_i and radii $r_i, i = 1, 2, 3$ and $r_1 < r_2 < r_3$. We know that in the case of a single sphere with charge q and radius r_0:

$$\phi = k\frac{q}{r_0} \quad \text{for} \quad r < r_0 \text{ (there } E = 0\text{)},$$

$$\phi = k\frac{q}{r} \quad \text{for} \quad r > r_0 \text{ (there } E \neq 0\text{)}.$$

It follows that

$$\phi_1 = k\left[\frac{q_1}{r_1} + \frac{q_2}{r_2} + \frac{q_3}{r_3}\right],$$

$$\phi_2 = k\left[\frac{q_1}{r_2} + \frac{q_2}{r_2} + \frac{q_3}{r_3}\right],$$

$$\phi_3 = k\left[\frac{q_1}{r_3} + \frac{q_2}{r_3} + \frac{q_3}{r_3}\right]. \tag{2.21}$$

**See e.g. B. H. Chirgwin, C. Plumpton and C. M. Kilmister [9], pp. 82–84.

From these relations we deduce that if

$$\phi_i = \sum_j p_{ij} q_j,$$

we have

$$\frac{k}{r_2} = p_{12} = p_{21} = p_{22}, \qquad p_{13} = p_{31} = p_{23} = p_{33} = p_{32} = \frac{k}{r_3}.$$

The matrix (C_{ij}) is the inverse of the matrix (p_{ij}) (and one can show that the inverse exists, i.e. $\det(p_{ij}) \neq 0$). According to the rule for establishing the elements of an inverse matrix we have for example in the case of the (3×3)-matrix

$$A = \begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{pmatrix}$$
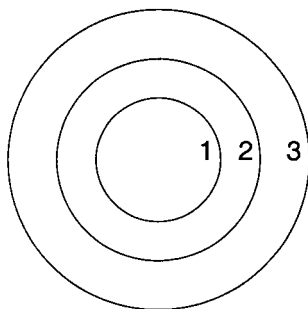
Fig. 2.14 Conducting shells screening the inner ones.

the inverse[tt]

$$A^{-1} = \frac{1}{\det A} \begin{pmatrix} |b_2 c_3| & -|b_1 c_3| & |b_1 c_2| \\ -|a_2 c_3| & |a_1 c_3| & -|a_1 c_2| \\ |a_2 b_3| & |a_1 b_3| & |a_1 b_2| \end{pmatrix},$$

where

$$|b_2 c_3| \equiv b_2 c_3 - b_3 c_2$$

and so on. Applied to the matrix (C_{ij}), this implies (corresponding to $|b_1 c_2|$) $C_{13} \propto |p_{12} p_{23}|$, i.e.

$$C_{13} \propto p_{12} p_{23} - p_{22} p_{13} = 0$$

[tt]See e.g. A. C. Aitken [10], p. 52, or any other book on determinants and matrices.

according to the above relations. This means that the charge q_1 remains unaffected by the field ϕ_3, or external conductors. This is what one calls electrical screening. In practice the conductor which is used for screening purposes is frequently "*earthed*", i.e. in the previous example one would have put $\phi_2 = 0$.

Example 2.15: The two hemispherical shells

A conductor has the shape of a spherical shell with radius a. The shell is divided diametrically into two hemispherical parts. Calculate the force which is needed to keep the hemispheres together if the sphere carried originally the total charge q ($q = 4\pi a^2 \sigma$).

Solution: The hemispheres are separated. Thus each has on both sides a field $\mathbf{E} \neq 0$. In this case (as we saw above) the electric field at the surface is given by

$$E = \frac{\sigma}{2\epsilon_0} = \frac{q}{\epsilon_0 8\pi a^2}.$$

We now calculate the force acting in the direction of z (see Fig. 2.15). To this end we compute the force acting on the surface charge σ in the general direction θ, take the component along z of this, (multiplication by $\cos\theta$), and add all force components in the direction of z (integration over θ for one hemisphere). We thus obtain for the force in the direction of z

$$
\begin{aligned}
F_z &= \int_0^{\pi/2} 2\pi a \sin\theta \, ad\theta \sigma \frac{q}{\epsilon_0 8\pi a^2} \cos\theta \\
&= 2\pi a^2 \frac{q}{4\pi a^2} \frac{q}{\epsilon_0 8\pi a^2} \int_0^{\pi/2} \sin\theta \cos\theta d\theta \\
&= \frac{q^2}{16\pi a^2 \epsilon_0} \int_0^{\pi/2} \sin\theta d(\sin\theta) \\
&\equiv \frac{q^2}{32\pi a^2 \epsilon_0} \equiv k\frac{q^2}{8a^2}.
\end{aligned}
\tag{2.22}
$$
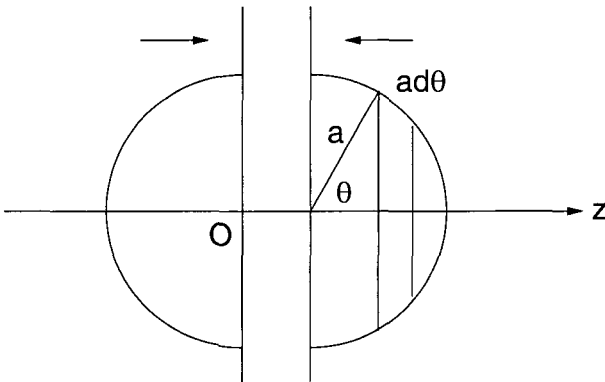
Fig. 2.15 The sphere divided into hemispheres.

If one considers a spherical condenser consisting of two concentric spherical shells with radii a and b ($a > b$) and one divides this diametrically into two parts, and if q and $-q$ are the charges on the

spheres originally, then the following force is required to keep the two parts together[*]

$$k\frac{q^2}{8}\left(\frac{1}{b^2} - \frac{1}{a^2}\right)$$

(difference of two forces so that the electric field remains between the two spheres).

Example 2.16: Two connected spherical shells

Two spherically shaped conductors with radii a and b and respectively charges q_1 and q_2 are separated from one another by a distance $r \gg a, b$. The conductors are connected by a thin conducting wire as indicated in Fig. 2.16. Calculate by what amount the energy of the system is thereby lowered.

Solution: Connecting the spheres by a wire alters the equipotential surfaces of the system. The charges rearrange themselves such that the energy is minimised. We first calculate the energy of the conductors when not connected. For this we calculate the potential of each of the conductors. These are in general determined by both charge distributions. The potential of the first conductor follows from a contribution $q_1/4\pi\epsilon_0 a$, originating from the charge of this conductor, and an additional contribution due to the charge q_2. Since $r \gg a, b$, we can approximate the latter by $q_2/4\pi\epsilon_0 r$. Hence the potential of the first conductor is

$$\phi_1 = \frac{1}{4\pi\epsilon_0}\left[\frac{q_1}{a} + \frac{q_2}{r}\right].$$

Analogously we obtain for the potential of the other conductor

$$\phi_2 = \frac{1}{4\pi\epsilon_0}\left[\frac{q_2}{b} + \frac{q_1}{r}\right].$$
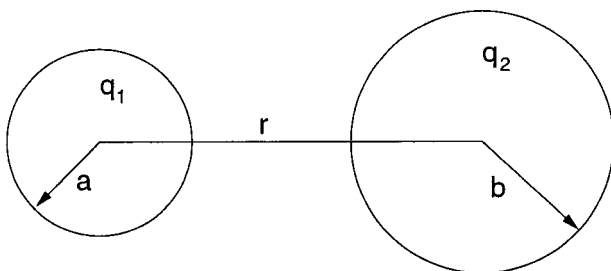
Fig. 2.16   Two spherical shells connected by a wire.

The energy of the system of separated conductors is therefore

$$W = q_1\phi_1 + q_2\phi_2$$
$$= \frac{q_1}{4\pi\epsilon_0}\left[\frac{q_1}{a} + \frac{q_2}{r}\right] + \frac{q_2}{4\pi\epsilon_0}\left[\frac{q_2}{b} + \frac{q_1}{r}\right]$$
$$= \frac{1}{4\pi\epsilon_0}\left[\frac{q_1^2}{a} + \frac{q_2^2}{b} + \frac{2q_1q_2}{r}\right].$$

---

[*]See problem 1, p. 54, of L. G. Chambers [11].

When the conductors are connected by a conducting wire they form an equipotential surface. The charges redistribute themselves to minimise the potential. Let q_1', q_2' be the charges of the conductors in the new state of equilibrium. Since no charge is lost, we must have

$$(a) \qquad q_1' + q_2' = q_1 + q_2.$$

Since both spheres now have the same potential we have

$$(b) \quad \phi_1' = \frac{1}{4\pi\epsilon_0} \left[\frac{q_1'}{a} + \frac{q_2'}{r} \right] = \frac{1}{4\pi\epsilon_0} \left[\frac{q_2'}{b} + \frac{q_1'}{r} \right] = \phi_2'.$$

This implies that

$$\frac{q_1'}{a} + \frac{q_2'}{r} = \frac{q_2'}{b} + \frac{q_1'}{r},$$

i.e.

$$q_1' = q_2' R, \quad q_2' = \frac{q_1'}{R}, \quad \text{where} \quad R = \left(\frac{\frac{1}{b} - \frac{1}{r}}{\frac{1}{a} - \frac{1}{r}} \right).$$

With (a) we have

$$q_1' = (q_1 + q_2 - q_1')R$$

and so

$$q_1' = \frac{(q_1 + q_2)R}{1 + R}.$$

Similarly

$$q_2' = \frac{(q_1 + q_2)\frac{1}{R}}{1 + \frac{1}{R}}.$$

With these expressions we obtain the new energy of the system as

$$W' = q_1'\phi_1' + q_2'\phi_2',$$

i.e.

$$W' = \frac{1}{4\pi\epsilon_0} \left[\frac{q_1'^2}{a} + \frac{q_2'^2}{b} + \frac{2q_1'q_2'}{r} \right].$$

The amount by which the *energy is lowered* is therefore $W - W'$; for $a = b$, this is

$$(W - W')\bigg|_{a=b} = \frac{(q_1 - q_2)^2}{8\pi\epsilon_0} \left[\frac{1}{a} - \frac{1}{r} \right].$$

This is the result.[†]

Example 2.17: The charges of connected spherical shells

Two conducting spherical shells have radii R_1 and R_2. Their centres are a distance D, $D \gg R_1 + R_2$ apart. The sphere with radius R_1 is given a charge Q_1 and the sphere with radius R_2 the charge Q_2. A thin conducting wire is now added to connect the two spheres. Calculate the resulting surface charge densities on both spheres, as well as the electric field strengths. Without further calculations discuss the effect that appears when the system with very small radius R_2 is subjected to a high voltage (i.e. why high voltage systems avoid sharp points like that of the small sphere).

Solution: It is easiest to obtain the potential of the spheres from the Gauss law,

$$\int \mathbf{E} \cdot d\mathbf{F} = 4\pi k \int \rho(\mathbf{r}')d\mathbf{r}' = 4\pi kQ,$$

[†]See also LI. G. Chambers [11], problem 8, p. 55.

i.e. for $r > R$:

$$4\pi r^2 E_r = 4\pi kQ, \qquad E_r = k\frac{Q}{r^2}. \tag{2.23}$$

In general the charge of the distant sphere contributes to the potential of the other sphere (ignoring the former is not always a good approximation). We can assume for simplicity here that at the other sphere the potential of the distant sphere does not vary much with the radius of the sphere considered, so that

$$\phi_1 = k\left[\frac{Q_1}{R_1} + \frac{Q_2}{D}\right] \quad \text{and} \quad \phi_2 = k\left[\frac{Q_2}{R_2} + \frac{Q_1}{D}\right]. \tag{2.24}$$

The wire connecting the spheres ensures that both spheres together form an equipotential surface with potential

$$\phi_0 = k\left[\frac{Q_1'}{R_1} + \frac{Q_2'}{D}\right] = k\left[\frac{Q_2'}{R_2} + \frac{Q_1'}{D}\right], \tag{2.25}$$

where Q_1' and Q_2' are the new charges, in other words the charges rearrange themselves on the spheres in such a way that both spheres have the same potential ϕ_0. Since in this process charge is neither created nor annihilated, we must have

$$Q_1 + Q_2 = Q_1' + Q_2'.$$

From this and Eq. (2.25) we can calculate Q_1' and Q_2':

$$Q_1' = \frac{R_1}{R_2}Q_2'\left[1 + O\left(\frac{1}{D}\right)\right]$$

for D large. By neglecting contributions of the order of $1/D$ we obtain

$$Q_1 + Q_2 \simeq \frac{R_1}{R_2}Q_2' + Q_2',$$

i.e.

$$Q_2' \simeq \frac{R_2}{R_1 + R_2}(Q_1 + Q_2), \qquad Q_1' \simeq \frac{R_1}{R_1 + R_2}(Q_1 + Q_2).$$

The potential ϕ_0 is therefore

$$\phi_0 = \frac{Q_1 + Q_2}{R_1 + R_2}\left[1 + O\left(\frac{1}{D}\right)\right].$$

Now, surface charge density = charge/area, so that

$$\sigma_1 \simeq \frac{Q_1'}{4\pi R_1^2} = \frac{Q_1 + Q_2}{4\pi R_1(R_1 + R_2)}$$

and

$$\sigma_2 \simeq \frac{Q_2'}{4\pi R_2^2} = \frac{Q_1 + Q_2}{4\pi R_2(R_1 + R_2)}.$$

According to Eq. (2.23) the corresponding field strengths are given by

$$E_1 \simeq k\frac{Q_1'}{R_1^2}, \qquad E_2 \simeq k\frac{Q_2'}{R_2^2},$$

i.e.

$$E_1 \simeq 4\pi k\sigma_1, \qquad E_2 \simeq 4\pi k\sigma_2.$$

When R_2 is very small, E_2 becomes very large. If a high voltage is applied, for $E_2 > E_{2,\text{critical}}$ this electric field is so large that it can eject electrons from atoms of the surrounding air. The resulting reordering of atomic states leads to the emission of light which can be observed in the form of sparks.

2.10 Energy of Charge Distributions

The *potential energy of a point charge* q_i at the point $\mathbf{r}_i$ in the electric field $\mathbf{E} = -\boldsymbol{\nabla}\phi$ of a point charge q_k at point $\mathbf{r}_k$ is the Coulomb potential

$$k\frac{q_i q_k}{|\mathbf{r}_i - \mathbf{r}_k|}, \quad i \neq k.$$

The potential energy of N charges $q_1, q_2, \ldots, q_N$ is correspondingly

$$W = \frac{1}{2}k \sum_i \sum_{k \neq i} \frac{q_i q_k}{|\mathbf{r}_i - \mathbf{r}_k|},$$

since $\sum_i \sum_{k \neq i}$ sums over i and k, i.e. *twice*. For *continuous charge distributions* $\rho(\mathbf{r})$ we have analogously

$$W = \frac{1}{2}k \int dV \int dV' \frac{\rho(\mathbf{r})\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} = \frac{1}{2}\int dV \rho(\mathbf{r})\phi(\mathbf{r}),$$

$$\phi(\mathbf{r}) = k \int \frac{dV'\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|}.$$

This expression for W does not explicitly
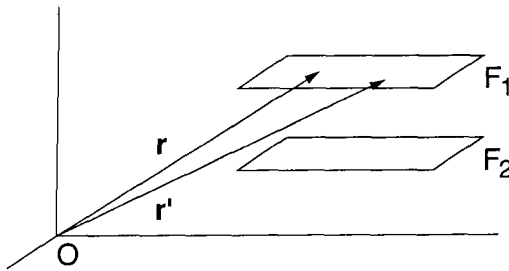
Fig. 2.17 Parts of a single surface.

exclude the point where $\mathbf{r} - \mathbf{r}' = 0$, so that it is possible that the expression involves divergent contributions. This is a problem connected with the Coulomb potential and requires a correction which we cannot enter into here (it is concerned with the problem dealt with under the term *"self-energy"*), which, however, is to be understood as a screening effect and replaces effectively the Coulomb potential $\sim 1/r$ by a type of Yukawa potential $\sim e^{-\mu r}/r$. In the case of surface charges we have correspondingly

$$W = \frac{1}{2}k \int_F dF(\mathbf{r}) \int_F dF(\mathbf{r}') \frac{\sigma(\mathbf{r})\sigma(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|}.$$

For example, F can consist of two condenser plates with areas F_1 and F_2 as illustrated in Fig. 2.17.

We can also express W in terms of $\mathbf{E}$: We saw that $\rho(\mathbf{r})$ is the source function of the potential ϕ, i.e.

$$\nabla^2 \phi \equiv \triangle \phi = -4\pi k \rho(\mathbf{r}).$$

Hence

$$W = -\frac{1}{8\pi k} \int \phi \nabla^2 \phi dV = -\frac{1}{8\pi k} \int \phi \triangle \phi dV.$$

However,

$$\nabla \cdot (\phi \nabla \phi) = \phi \triangle \phi + (\nabla \phi)^2,$$

so that

$$W = -\frac{1}{8\pi k} \int dV [\nabla \cdot \{\phi(\nabla \phi)\} - (\nabla \phi)^2].$$

But

$$\int dV \nabla \cdot (\phi \nabla \phi) = \int d\mathbf{F} \cdot (\phi \nabla \phi).$$

Now, $\phi \sim 1/r, \nabla \phi \sim 1/r^2$. If we integrate over an infinitely large surface (i.e. with $r \to \infty$), the integral vanishes, since $dF \propto r^2 d\Omega$ which implies

$$\int d\mathbf{F} \cdot (\phi \nabla \phi) \sim \frac{1}{r} \int d\Omega \to 0$$

for $r \to \infty$. It follows that

$$W = \frac{1}{8\pi k} \int_{V_\infty} dV \mathbf{E}^2. \tag{2.26}$$

Thus the expression $\mathbf{E}^2/8\pi k$ represents the *energy density of the electric field*. One should observe: For large values of r the integrand of W behaves as dr/r^2, but from this we cannot conclude that it vanishes, since the integral is the sum of many nonvanishing positive contributions. On the contrary, inserting the Coulomb potential for a single charge at $r = 0$ one obtains $W = \infty$.

Example 2.18: [‡] The gravitational potential

Consider a static particle with charge Q. With the help of Einstein's formula *energy = mass·c²* calculate the mass density which is equivalent to the energy density of the electric field of the charge, and then insert this together with the contribution of a "bare" (i.e. fieldless) mass m_0 of the particle into the equation for Newton's gravitational potential ϕ, i.e.

$$\triangle \phi = 4\pi G \rho_{\text{total}}.$$

[‡]See M. Visser, *Phys. Lett.* **A139** (1989) 99.

Taking into account an additional contribution to ρ, which makes ϕ its own source (similar to the case in electrostatics), solve the equation with the help of the ansatz

$$\psi(\phi) = e^{\frac{\phi}{2c^2}}$$

and the substitution $u = 1/r$. Then calculate m_0. What is the significance of the result?

Solution: At a distance r from the charge Q the electric field is ($k = 1/4\pi\epsilon_0$):

$$\mathbf{E} = \frac{Q}{4\pi\epsilon_0} \frac{\mathbf{r}}{r^3} .$$

The corresponding electromagnetic energy density is (here only the electric part):

$$\rho_{\text{elm}} = \frac{1}{8\pi k} |\mathbf{E}|^2 = \frac{1}{8\pi} \frac{1}{4\pi\epsilon_0} \frac{Q^2}{r^4} .$$

According to Einstein's energy-mass relation the electric energy density corresponds to a mass density

$$\rho = \frac{1}{8\pi k c^2} |\mathbf{E}|^2 .$$

Let m_0 be the so-called bare mass of the particle (like the electron) with charge (say) Q (i.e. not taking into account the radiation field of the particle). Then the total mass density is ($\int d\mathbf{r}\,\delta(\mathbf{r}) = 1$)

$$\rho = m_0\delta(\mathbf{r}) + \frac{1}{8\pi k c^2} |\mathbf{E}|^2 .$$

This mass density contributes to the gravitational potential in space (like any other mass). Newton's gravitational potential ϕ is the solution of the equation

$$\triangle\phi = 4\pi G \rho_{\text{total}}$$

(here G is the gravitational constant). The density ρ_{total} is the total mass density contributing to the gravitational potential which now includes the electromagnetic energy density (in this way gravity couples to any kind of energy density). One contribution is apparently the contribution ρ. Another contribution has its origin in the gravitational field itself. Like the electromagnetic energy density ρ_{elm} the (Newtonian) gravitational field also possesses an energy density which is given by the analogous expression (with minus sign in view of the attractive nature of the gravitational field)

$$\rho_{\text{Grav}} = -\frac{1}{8\pi G c^2} |\nabla\phi|^2 .$$

Adding this contribution also to ρ we obtain a nonlinear self-interaction of the gravitational field which could be looked at as a neo-Newtonian version which incorporates model-like the nonlinear contributions in Einstein's theory. We obtain therefore

$$\triangle\phi = 4\pi G \left[m_0\delta(\mathbf{r}) + \frac{1}{8\pi c^2} \frac{1}{4\pi\epsilon_0} \frac{Q^2}{r^4} \right] - \frac{1}{2c^2} |\nabla\phi|^2 . \tag{2.27}$$

In solving the equation we set

$$\psi(\phi) = e^{\frac{\phi}{2c^2}} ,$$

so that

$$\nabla\psi(\phi) = \frac{1}{2c^2} (\nabla\phi)\psi(\phi) .$$

and

$$\triangle\psi = \boldsymbol{\nabla} \cdot \boldsymbol{\nabla}\psi(\phi)$$
$$= \frac{1}{2c^2}\boldsymbol{\nabla} \cdot \left\{ (\boldsymbol{\nabla}\phi)\psi(\phi) \right\}$$
$$= \frac{1}{2c^2}\left\{ \boldsymbol{\nabla}\psi \cdot \boldsymbol{\nabla}\phi + \psi(\phi)\triangle\phi \right\}$$
$$= \frac{1}{2c^2}\left\{ \frac{1}{2c^2}(\boldsymbol{\nabla}\phi)^2 + \triangle\phi \right\}\psi(\phi). \tag{2.28}$$

Replacing here $\triangle\phi$ by the above expression, we have

$$\triangle\psi = \left[\frac{2\pi G m_0}{c^2}\delta(\mathbf{r}) + \frac{G}{4c^2 4\pi\epsilon_0 c^2}\frac{Q^2}{r^4} \right]\psi.$$

The right side is spherically symmetric. This suggests the separation of $\triangle$ in spherical polar coordinates r, θ, ϕ. Then setting $u = 1/r$, one obtains

$$\triangle\psi = u^4 \frac{\partial^2}{\partial u^2}\psi,$$

i.e.

$$\frac{\partial^2}{\partial u^2}\psi = \frac{GQ^2}{4c^2 4\pi\epsilon_0 c^2}\psi, \qquad \frac{\partial^2\psi}{\partial u^2} - \frac{\tilde{Q}^2}{4}\psi = 0$$

with $\tilde{Q} = \kappa Q, \kappa = \sqrt{G}/\sqrt{4\pi\epsilon_0}c^2$. It should be noted that because $r^4\delta(\mathbf{r}) = 0$, the other source term has disappeared (but will be regained by consideration of $r \to 0$, which is here excluded). The resulting equation can now be solved easily provided we know the boundary condition that $\psi(\phi)$ ought to satisfy. Since ϕ is to vanish for $r \to \infty$, we require

$$r \to \infty: \quad \psi(\phi) \to 1,$$

and hence

$$\psi = \frac{\cosh(A - \frac{\tilde{Q}}{2r})}{\cosh A}, \quad A = \text{const.}$$

However, for $r \to \infty$,

$$\phi = 2c^2 \ln\psi = 2c^2 \ln\left\{ \frac{\cosh(A - \frac{\tilde{Q}}{2r})}{\cosh A} \right\}$$
$$= 2c^2 \ln\left(1 - \frac{1}{2r}\tilde{Q}\tanh(A) + \cdots \right)$$
$$= -\frac{1}{r}\tilde{Q}c^2 \tanh(A) + \cdots .$$

Since this expression represents the gravitational potential (for $r \to \infty$) the quantity $\tilde{Q}c^2 \tanh(A)$ has to be interpreted as gravitational mass. With the above expression for ψ we obtain in addition

$$\phi = 2c^2 \ln\left(\frac{\cosh(A - \frac{\tilde{Q}}{2r})}{\cosh A} \right),$$

and with this

$$\boldsymbol{\nabla}\phi = \tilde{Q}c^2 \tanh\left(A - \frac{\tilde{Q}}{2r} \right)\frac{\mathbf{r}}{r^3}$$

and (see below)

$$\triangle\phi = -4\pi\tilde{Q}c^2\delta(\mathbf{r}) + \frac{\tilde{Q}^2c^2}{2r^4\cosh^2(A - \frac{\tilde{Q}}{2r})}. \tag{2.29}$$

The *first contribution* on the right of Eq. (2.29) is obtained as follows. For $r \neq 0$ we have

$$\boldsymbol{\nabla}\cdot\left(\frac{\mathbf{r}}{r^3}\right) = \frac{3r^2 - 3r^2}{r^5} = 0.$$

For $r \to 0$ we have

$$\lim_{r\to 0}\tanh\left(A - \frac{\tilde{Q}}{2r}\right) = \tanh(-\infty) = -1.$$

Then, for $r \to 0$: $\boldsymbol{\nabla}\phi \simeq -\tilde{Q}c^2\mathbf{r}/r^3$ and

$$\int \boldsymbol{\nabla}\cdot\boldsymbol{\nabla}\phi\, d\mathbf{r} = -\tilde{Q}c^2\int_V \boldsymbol{\nabla}\cdot\left(\frac{\mathbf{r}}{r^3}\right) d\mathbf{r}$$

$$= -\tilde{Q}c^2\int_{F(V)} \frac{\mathbf{r}}{r^3}\cdot d\mathbf{F} = -\tilde{Q}c^2\int \frac{\mathbf{r}}{r^3}(d\Omega r^2)\frac{\mathbf{r}}{r} = -4\pi\tilde{Q}c^2 \equiv -4\pi\tilde{Q}c^2\int\delta(\mathbf{r})d\mathbf{r}.$$

Hence $\triangle\phi$ acquires the contribution $-4\pi\tilde{Q}c^2\delta(\mathbf{r})$. Comparison of Eqs. (2.27) and (2.29) then yields $m_0 = -\frac{\tilde{Q}c^2}{G} = -Q/\sqrt{4\pi\epsilon_0 G}$.

The *second contribution* on the right side of Eq. (2.29) follows from the sum of the other two contributions on the right side of Eq. (2.27) with $1/\cosh^2 x = 1 - \tanh^2 x$.

Example 2.19: The spherical condenser

Show that for the spherical condenser (radii a and b, $a < b$, with charges $q, -q$) the electrostatic energy is $W = q^2/2C$. Calculate the electrostatic pressure on the outer spherical shell and compare the result with the product of charge/area of the outer shell with the electric field strength at the latter.

Solution: The charge on the inner sphere (radius a, charge $+q$) acts like concentrated at the centre. Therefore the field E_r at a distance r away from O is

$$E_r = k\frac{q}{r^2}, \quad a < r < b.$$

Hence the total electrostatic energy is

$$W = \frac{1}{8\pi k}\int d\mathbf{r}\, E_r^2 = \frac{1}{8\pi k}\int 4\pi r^2 dr\, E_r^2 = \frac{kq^2}{2}\int_a^b \frac{dr}{r^2} = \frac{kq^2}{2}\left(\frac{1}{a} - \frac{1}{b}\right).$$

Since according to Example 2.10 the capacity C is

$$C = \frac{1}{k(\frac{1}{a} - \frac{1}{b})},$$

we obtain $W = q^2/2C$. The force acting on the external spherical shell is

$$-\frac{\partial W}{\partial b} = -\frac{kq^2}{2b^2},$$

so that the pressure on this shell, force/area, is

$$-\frac{kq^2}{2b^2 4\pi b^2} = -\frac{kq^2}{8\pi b^4}$$

(toward the centre). On the other hand, the charge density σ of the outer shell is $= -q/4\pi b^2$. Thus

$$\sigma E_{r=b} = -\frac{q}{4\pi b^2}\frac{kq}{b^2} = -\frac{kq^2}{4\pi b^4}.$$

This is not the pressure, which acts on the outer shell, since the surface charge on the rest of the outer shell exerts a force in the opposite direction and thus compensates partially the force due to the charge on the inner sphere.

Example 2.20: A charged particle inside a charge density

A spherically shaped volume is uniformly charged with charge density per unit volume q_v. A particle with mass m and charge $-q$ is placed into this spherical volume. Show that the mass executes harmonic oscillations and calculate their frequency.

Solution: We have here

$$\int \mathbf{E} \cdot d\mathbf{F} = 4\pi k \sum_i q_i, \quad \text{i.e.} \quad E_r 4\pi r^2 = 4\pi k \frac{4}{3}\pi r^3 q_v,$$

and hence

$$E_r = \frac{4}{3}\pi r q_v k,$$

so that the equation of motion of the particle is

$$m\ddot{r} = -\frac{4}{3}\pi r q q_v k \quad \text{or} \quad m\ddot{r} + \omega^2 r = 0$$

with

$$\omega^2 = \frac{4\pi k q q_v}{3m}, \quad T = \frac{2\pi}{\omega} = 2\pi\sqrt{\frac{3m}{4\pi q q_v k}}.$$

Example 2.21: A multivalued potential

Determine the equipotential surfaces of the potential

$$\phi = \tan^{-1}\frac{y}{x}.$$

What is the value of a contour integral $\int \boldsymbol{\nabla}\phi \cdot d\mathbf{l}$ along a closed path A, which does not encircle the z-axis, and what is the value of a contour integral along a closed path B, which does encircle the z-axis? (Remark: The function ϕ will be discussed further, in particular in Sec. 8.4).

Solution: Since $\phi = \tan^{-1}\frac{y}{x} = \theta \pm n\pi$, n a positive integer or zero, the equipotential surfaces are the radial planes with $\theta = $ const. These intersect along the z-axis. These planes therefore violate the physically expected single valuedness of the potential and the z-axis is the seat of singularities.
    The gradient is

$$\boldsymbol{\nabla}\phi = -\frac{y}{x^2+y^2}\mathbf{e}_x + \frac{y}{x^2+y^2}\mathbf{e}_y, \quad \text{so that} \quad \boldsymbol{\nabla}\phi \cdot d\mathbf{l} = \frac{x\,dy - y\,dx}{x^2+y^2}.$$

For a circle around the origin in the plane of x and y, and $a = $ const., one has $x = a\cos\theta, y = a\sin\theta, x\,dy - y\,dx = a^2 d\theta$ and thus $\int \boldsymbol{\nabla}\phi \cdot d\mathbf{l} = \int d\theta$. This expression is independent of a and is therefore valid for an arbitrary contour. Since the direction of $\boldsymbol{\nabla}\phi$ is that of $\mathbf{e}_\theta$, contour components along $\mathbf{e}_z$ and $\mathbf{e}_r$ do not contribute. In the case of contour A we clearly have $\int d\theta = [\theta - \theta] = 0$. Also curl grad $\phi = 0$. In the case of contour B, however, the value of the integral depends on the endpoint. The potential is multivalued, i.e. each time O is encircled its value changes by $\pm 2\pi$. Such a contour cannot be shrunk to zero without cutting the singularity at O.

Chapter 3

Applications of Electrostatics

3.1 Introductory Remarks

In the following we investigate further extensions of electrostatics as well as more complicated applications, thereby also emphasising the methodical differences. Green's theorems and Fourier transforms are introduced, and finally the multipole expansion of the potential of a charge distribution is derived.

3.2 Method I: The Gauss Law

The problems we considered thus far were mostly treated with the help of the integral form of the Gauss law, and so with

$$\int \mathbf{E} \cdot d\mathbf{F} = 4\pi k \sum_i q_i.$$

We now investigate other methods.

3.3 Method II: Poisson and Laplace Equations

In the preceding we encountered the Poisson equation as the equation

$$\nabla^2 \phi = -4\pi k \rho.$$

In solving this equation we require boundary conditions. These are:
(1) $\phi = $ const. *on conducting surfaces* (in an ideal conductor the electrons do not perform work) with the discontinuity for the derivatives as for the

electric field strength, i.e. (cf. Eq. (2.16))

$$-\left(\frac{\partial\phi}{\partial n}\right)_2 + \left(\frac{\partial\phi}{\partial n}\right)_1 = 4\pi k\sigma, \qquad (3.1)$$

(2) *otherwise ϕ continuous* (this is no contradiction with the difference of the potential on both sides of a dipole layer;[*] the potential of a dipole $\mathbf{p}$ is, as we shall see later, $\mathbf{p}\cdot\mathbf{r}/r^3$, and this is continuous in $\mathbf{r}$ (no step function $\theta(x)$). If ϕ were discontinuous anywhere, i.e. if ϕ contained a step-function $\theta(x-x_0)$, then at this point $\mathbf{E} = -\boldsymbol{\nabla}\phi$ would be proportional to $\delta(x-x_0)$, i.e. infinite, which would be meaningless and unphysical.

Example 3.1: The cylindrical condenser
Calculate the potential difference V of a cylindrical condenser.

Solution: We consider again the cylindrical condenser without boundary effects and charges q and $-q$ on the surfaces of radii r_1 and $r_2 > r_1$ and height h. In the region between the cylindrical surfaces we have no charge, so that $\nabla^2\phi = 0$. In view of the cylindrical symmetry $\partial\phi/\partial\theta = 0 = \partial\phi/\partial z$, so that separation of the Laplace operator yields

$$\frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial\phi}{\partial r}\right) = 0,$$

i.e. $\partial\phi/\partial r = \alpha/r, \alpha = $ const. Hence $E = -\partial\phi/\partial r = -\alpha/r$. At $r = r_1$ we have (see Eq. (2.18)) $E = 4\pi k\sigma = 4\pi kq/F = 4\pi kq/2\pi hr_1$, so that $\alpha = -2qk/h$, and for E and the potential difference V we obtain as before

$$E = \frac{2qk}{h}\frac{1}{r}, \qquad V = \int_{r_1}^{r_2} E dr = \frac{2qk}{h}\ln\frac{r_2}{r_1}.$$

The result is positive as a consequence of its physical interpretation as the work done *in the direction of* $\mathbf{E}$.

As a more advanced application we consider a proportional counter.[†]

Example 3.2: The proportional counter
Calculate the potential of a proportional counter. An important part of a proportional counter is its circular cylindrically shaped anode of radius a made of wire and a corresponding circular cylindrical cathode with radius $b > a$. The intermediate space is filled with an ionisable mixture of gases. In practice one frequently has the situation that the volume charge density ρ differs significantly from zero only along a length L (with $|z| \leq L/2$) of the cylinders. Show that the potential $\phi(r, z)$ in this intermediate space obtained with the following boundary conditions

$$\phi(a, z) = V, \quad \phi(b, z) = 0, \quad \rho = \begin{cases} \text{const.} & \text{for} \quad |z| \leq L/2, \\ 0 & \text{for} \quad |z| > L/2, \end{cases}$$

is given by

$$\phi(r, z) = \frac{V}{\ln(a/b)}$$
$$+ \rho\sum_n B_n[J_0(\mu_n r) + K_n Y_0(\mu_n r)]\begin{cases} 1 - e^{-\mu_n L/2}\cosh(\mu_n L) & \text{for} \quad |z| \leq L/2, \\ \sinh(\mu_n L/2)e^{-\mu_n|z|} & \text{for} \quad |z| > L/2, \end{cases}$$

---

[*]See also for instance Greiner [4], p. 29.

[†]H. Sipila, V. Vanha-Honko and J. Bergqvist, *Nucl. Instr. Methods* **176** (1980) 381.

where μ_n satisfies the following relation involving the special functions J and Y known as Bessel and Neumann functions:

$$J_0(\mu_n a)Y_0(\mu_n b) - J_0(\mu_n b)Y_0(\mu_n a) = 0.$$

Which condition determines the coefficients K_n, and which the coefficients B_n?

Solution: For the detailed solution we refer this time to the literature cited above.

3.4 Method III: Direct Integration

In this case one integrates either *vectorially*, as in

$$\mathbf{E} = k \int \frac{\mathbf{r}dq}{r^3}, \quad dq = \rho dV,$$

or *scalar-wise* as in

$$\phi = k \int dqr.$$

We illustrate the method again by application to some important examples.

Example 3.3: The finite, charged, thin rod
Calculate the field surrounding a charged, thin rod of finite length.

Solution: The word "thin" implies that the rod is to be considered as a line. Let q be its charge per unit length. We choose the coordinates as in Fig. 3.1.
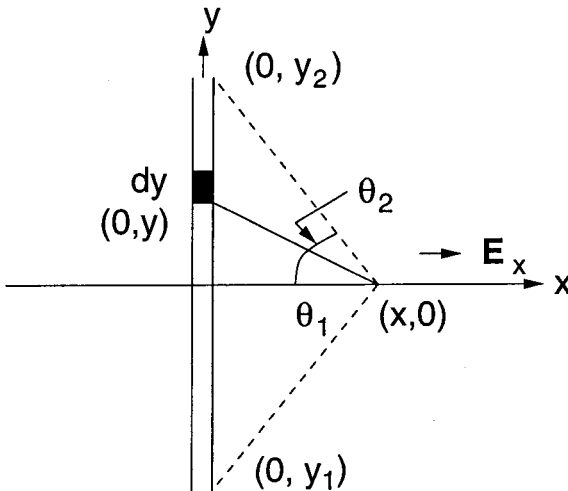
Fig. 3.1 The charged rod.

The potential energy at the point $(x, 0)$ in Fig. 3.1 is

$$\phi = kq \int_{y_1}^{y_2} \frac{dy}{\sqrt{x^2 + y^2}} = q\left[\sinh^{-1}\frac{y}{x}\right]_{y_1}^{y_2} = qk\left\{\sinh^{-1}\frac{y_2}{x} - \sinh^{-1}\frac{y_1}{x}\right\}. \tag{3.2}$$

Hence the electric field at the point $(x, 0)$ is

$$E_x = -\frac{\partial \phi}{\partial x} = \frac{qk}{x} \left\{ \frac{y_2}{\sqrt{y_2^2 + x^2}} - \frac{y_1}{\sqrt{y_1^2 + x^2}} \right\} = \frac{qk}{x} \left\{ \sin\theta_2 - (-\sin\theta_1) \right\}. \qquad (3.3)$$

In the limit $y_2, -y_1 \to \infty$: $\theta_1, \theta_2 \to \pi/2$, i.e. $E_x \to 2qk/x$ as for a point charge at the origin. One can show that the equipotential surfaces are ellipses with foci at the ends of the rod. In Example 3.3 we calculated as a matter of simplicity only the x-component of the field. We can deduce from Eq. (3.2) what the potential is outside the wire. To obtain this, we use the expansions

$$\sinh^{-1} x = \ln(2x) + \frac{1}{4x^2} + O\left(\frac{1}{x^4}\right) \quad \text{for } x > 1,$$

$$\sinh^{-1} x = -\ln|2x| - \frac{1}{4x^2} + O\left(\frac{1}{x^4}\right) \quad \text{for } x < -1.$$

Inserting the dominant contributions into Eq. (3.2) we obtain

$$\phi(x) \simeq qk\left\{ \ln\left(\frac{2y_2}{x}\right) + \ln\left|\frac{2y_1}{x}\right| \right\} = -2qk\ln x + qk\ln|4y_2y_1|, \quad \left(\frac{y_2}{x} > 1, \frac{y_1}{x} < 1\right).$$

Replacing x by r leads to the expression

$$\phi(r) = -\frac{q}{2\pi\epsilon_0} \ln r + \text{const.}, \quad r > 0,$$

in which the constant can be infinite. In the computation of the field strength, however, the contribution of the constant drops out. With the help of the last relation and the product formula

$$\sinh x = x \prod_{n=1}^{\infty} \left(1 + \frac{x^2}{n^2\pi^2}\right)$$

one can calculate the potential of a system of parallel wires, a distance b apart, and each carrying the same charge. One finds

$$\phi(r) = -\frac{q}{2\pi\epsilon_0} \ln\sinh\left(\frac{\pi r}{b}\right) + \text{const.}$$

Systems of this or a similar kind are used in particle detectors.[‡]

3.5   Method IV: Kelvin's Method of Image Charges

In an electrostatic problem the potential determines in a unique way quantities like $\mathbf{E}$ and σ. It is therefore possible to imagine certain potential distributions replaced by fictitious charges called "*image charges*" and to calculate the field of these. An example is shown in Fig. 3.2. The effect of the earthed conductor (i.e. with potential $\phi = 0$, this alone is our definition of "*earthed*") can also be achieved by a fictitious charge.

[‡]See e.g. T.J. Killian, *Nucl. Instr. Methods* **176** (1980) 355.
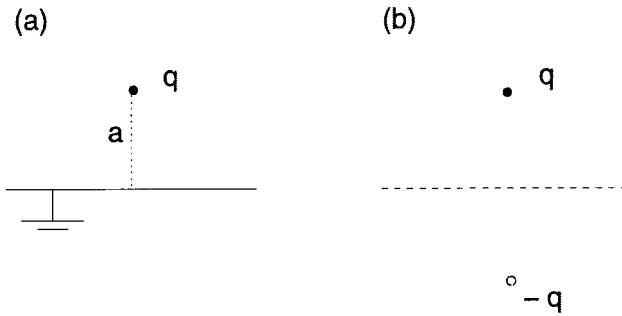
Fig. 3.2 (a) The earthed conductor, and (b) its fictitious replacement.

The earthed conductor implies a *boundary condition*: $\phi = 0 = \phi(r = \infty)$. The field **E** then results as the effect of both charges, $+q$ and $-q$, in the region *above* the conductor. A similar example is illustrated in Fig. 3.3; again the field of case (a) can be calculated from the situation of case (b).
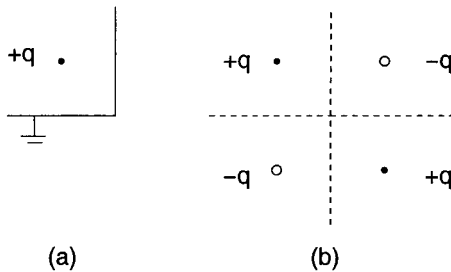
Fig. 3.3 (a) An earthed conductor, and (b) its fictitious replacement.

In the case of two earthed conducting plates and a single charge between them we have to introduce a large number of image charges as shown in Fig. 3.4.
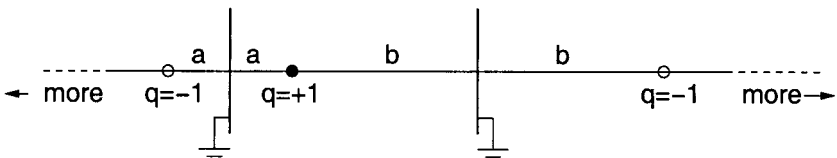
Fig. 3.4 Charge q between earthed conductor plates.

The potential at an arbitrary point between the plates enclosing the charge (in Fig. 3.4 with charge $+1$), is then the same as that of the given charge together with the image charges. An analogous situation is obtained for the

earthed conductor plates arranged as in Fig. 3.5.

Although the method of image charges is used in applications, it is not always treated in relevant texts. Some texts which include it are those of Stiefel [15] and Zahn [8]. The method of image charges finds application for instance in the calculation of the electric field in particle detectors.[§]
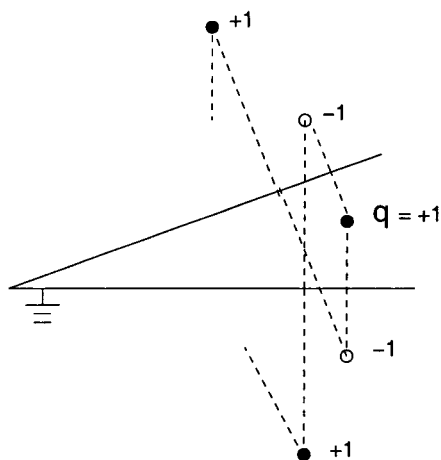
Fig. 3.5 Conductors with angular orientation.
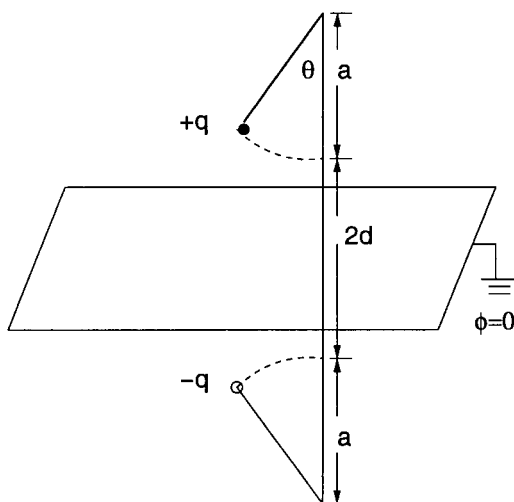
Fig. 3.6 The oscillating charge.

[§]See e.g. W. Weihs and G. Zech, *Numerical Computation of Electrostatic Fields in Multiwire Chambers*, Report Univ. Siegen, May 1989.

Example 3.4: The oscillating charge

A pointlike mass m with charge q at the end of a string of length a oscillates in the gravitational field of the Earth above an earthed horizontal metal plate. The mass oscillates like a pendulum in a plane. Determine the equation of motion of the pendulum expressed in terms of the angular deflection θ. What is the period in the case of small oscillations?

Solution: Using the method of image charges as sketched in Fig. 3.6, we place a charge $-q$ in the position of a mirror image of the charge $+q$, so that in the plane of the earthed conductor the potential ϕ is zero. In this way we can ignore the plate (or its boundary condition) in the evaluation of the electric field. It is easiest to use the principle of conservation of energy E. The total energy of the charge $+q$ is

$$E = \frac{1}{2}m(a\dot\theta)^2 + mga(1 - \cos\theta) - \frac{q^2}{(2d + 2a - 2a\cos\theta)},$$

where d is the distance shown in Fig. 3.6. With

$$\frac{d}{d\theta}\left(\frac{1}{2}\dot\theta^2\right) = \ddot\theta, \quad \frac{dE}{d\theta} = 0,$$

it follows that

$$ma^2\ddot\theta + mga\sin\theta + \frac{q^2 a\sin\theta}{(d + a - a\cos\theta)^2} = 0.$$

For θ small we have

$$\ddot\theta + \Omega^2\theta \simeq 0, \quad \Omega^2 = \left(\frac{mg + q^2/d^2}{ma}\right).$$

Thus the desired period is $T = 2\pi/\Omega$.

3.6 Theoretical Aspects of Image Charges

3.6.1 The Induced Charge

Before we investigate general theoretical aspects of the method, we consider an example: A charge q placed outside an earthed spherical shell of radius a. Different from the cases discussed above, in which we considered either point charges *or* charge distributions, we now have the example of a mixed case. *"Earthed"* means potential ϕ as on the ground, and this means as at $r = \infty$. Everywhere on the spherical surface the potential is to be $\phi = 0$. Apart from this, ϕ is a continuous function (so that $\mathbf{E} = -\nabla\phi$ is nowhere infinite).

The first question is therefore: Where do I have to put which charge so that at any point on the sphere one has $\phi = 0$? Inside the conductor, i.e. on the sphere, the charges (electrons) can move about freely (performing no work, since $\phi = 0$, also $\mathbf{E} = 0$). As indicated in Fig. 3.7 we put a fictitious charge $-\mu q$ into the interior of the sphere. In order that $\phi = 0$ at a point P on the shell, we must have

$$0 = \frac{q}{r_1} - \frac{\mu q}{r_2},$$

i.e. $r_2 = \mu r_1$. From the geometry of Fig. 3.7 we obtain

$$r_1^2 = d^2 + a^2 - 2ad\cos\theta = d^2\left(1 + \frac{a^2}{d^2} - 2\frac{a}{d}\cos\theta\right),$$

and

$$r_2^2 = a^2 + \nu^2 - 2a\nu\cos\theta = a^2\left(1 + \frac{\nu^2}{a^2} - 2\frac{\nu}{a}\cos\theta\right).$$

It follows that r_2 is proportional to r_1 if

$$\nu = \frac{a^2}{d}.$$

Then $r_2 = ar_1/d$, i.e. $\mu = a/d$. In other words, a fictitious charge $-aq/d$ has to be placed at a point, called the *"image point"*, a distance a^2/d away from O, so that $\phi = 0$ at all points on the sphere.
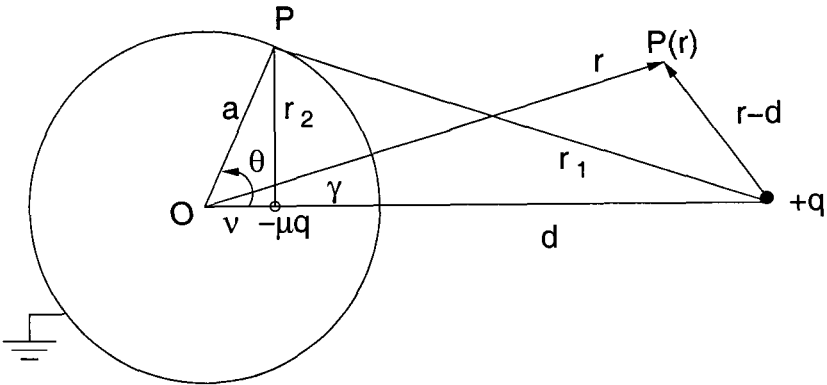
Fig. 3.7 The earthed spherical shell.

The second question: What is the potential $\phi(\mathbf{r})$ at an arbitrary point $P(\mathbf{r})$ outside the earthed sphere, where $r > a$? Outside the sphere

$$\mathbf{E} \equiv \mathbf{E}_2 = -\nabla\phi \neq 0.$$

At an arbitrary point $\mathbf{r}$ outside the sphere

$$4\pi\epsilon_0\phi(\mathbf{r}) = \left(\frac{q}{|\mathbf{r}-\mathbf{d}|} - \frac{qa}{d|\mathbf{r}-\boldsymbol{\nu}|}\right)$$

$$\underset{r\geq a}{=} \frac{q}{(r^2+d^2-2rd\cos\gamma)^{1/2}} - \frac{aq}{d(r^2+\nu^2-2r\nu\cos\gamma)^{1/2}},$$

and this vanishes for $r = a$. Hence inside the sphere $\mathbf{E} = 0$, i.e. ϕ constant. This constant is zero since on the sphere $\phi = 0$. (If the entire sphere is filled with conductor material, the potential is everywhere constant or zero as on the boundary surface). The relation we derived earlier for the passage of the electric field through a charged surface (now with $\mathbf{E}_1 = 0$) is

$$(\mathbf{E}_2 - \mathbf{E}_1) \cdot \mathbf{n} = 4\pi k\sigma, \quad \text{or} \quad -\left(\frac{\partial\phi}{\partial n}\right)_2 + \left(\frac{\partial\phi}{\partial n}\right)_1 = 4\pi k\sigma$$

where $\partial\phi/\partial n = (\partial\phi/\partial r) \cdot \mathbf{n}$. This allows us to compute the surface charge density σ. The corresponding charge on the sphere is called *induced charge*. This charge follows from the fact that the charge outside repels charges of the same polarity in the conductor and attracts those of opposite polarity.

The third question: What is the induced charge? The induced charge is a real charge which we can calculate. This charge is given by

$$Q = \int_S \sigma(\gamma) a^2 d\Omega, \quad d\Omega = 2\pi d(\cos\gamma)$$

(a circular annulus about the horizontal axis in Fig. 3.7 at angular height γ and of infinitesimal width $ad\gamma$ has area $(ad\gamma)2\pi(a\sin\gamma)$; hence the element of area under the integral).

We compute first σ. Since $\mathbf{E} = -\boldsymbol{\nabla}\phi = \mathbf{E}_2$, we obtain from the above formula

$$\mathbf{E}_2 \cdot \mathbf{n} = -\boldsymbol{\nabla}\phi \cdot \mathbf{n} = 4\pi k\sigma,$$

i.e.

$$-\frac{1}{4\pi}\frac{\partial\phi}{\partial n}\bigg|_{r=a} = -\frac{1}{4\pi}\frac{\partial\phi}{\partial r}\bigg|_{r=a} = k\sigma.$$

Since $S \equiv$ sphere, it follows that

$$4\pi\epsilon_0 \frac{\partial\phi}{\partial r} = -\frac{2q(r - d\cos\gamma)}{2(r^2 + d^2 - 2rd\cos\gamma)^{3/2}} + \frac{2aq(r - \nu\cos\gamma)}{2d(r^2 + \nu^2 - 2r\nu\cos\gamma)^{3/2}}.$$

For $\nu = a^2/d$ and $r = a$ we obtain

$$\underbrace{4\pi\epsilon_0}_{1/k}\frac{\partial\phi}{\partial r}\bigg|_{r=a} = -\frac{q(a - d\cos\gamma)}{(a^2 + d^2 - 2ad\cos\gamma)^{3/2}} + \frac{aq(a - \frac{a^2}{d}\cos\gamma)}{d(a^2 + \frac{a^4}{d^2} - 2\frac{a^3}{d}\cos\gamma)^{3/2}}$$

$$= \frac{q\frac{d^2}{a}(1 - \frac{a^2}{d^2})}{(a^2 + d^2 - 2ad\cos\gamma)^{3/2}}.$$

We note that $(\phi)_a = 0, (\partial\phi/\partial r)_a \neq 0$. Hence

$$k\sigma(\gamma) = -\frac{1}{4\pi}\frac{\partial\phi}{\partial r}\bigg|_{r=a} = -k\frac{q}{4\pi}\frac{d^2}{a}\frac{(1-\frac{a^2}{d^2})}{(a^2+d^2-2ad\cos\gamma)^{3/2}}$$

$$= -k\frac{q}{4\pi ad}\frac{(1-\frac{a^2}{d^2})}{(1+\frac{a^2}{d^2}-\frac{2a}{d}\cos\gamma)^{3/2}}.$$

The total charge on the sphere is therefore

$$Q = \int \sigma(\gamma)a^2 \cdot 2\pi d(\cos\gamma)$$

$$= -\frac{qa}{2d}(1-\frac{a^2}{d^2})\int_{-1}^{1}\frac{d(\cos\gamma)}{(1+\frac{a^2}{d^2}-\frac{2a}{d}\cos\gamma)^{3/2}}$$

$$= -\frac{qa}{2d}(1-\frac{a^2}{d^2})\left[\frac{(-2)(-d/2a)}{(1+\frac{a^2}{d^2}-\frac{2a}{d}\cos\gamma)^{1/2}}\right]_{-1}^{1}$$

$$= -\frac{qa}{2d}(1-\frac{a^2}{d^2})\frac{d}{a}\left[\frac{1}{1-\frac{a}{d}}-\frac{1}{1+\frac{a}{d}}\right] = -q\frac{a}{d}.$$

This is precisely the fictitious or image charge. Thus the fictitious charge ensures that $\phi = 0$ on the surface of the spherical conductor, but the conductor itself, that is the sphere, has total charge $-qa/d$. If we want to achieve that the (earthed or unearthed, insulated) sphere carries the total charge zero, we have to supply the sphere with extra charge $+qa/d$. Summarising, we can say: The charge induced on the sphere has the same effect as the charge $-qa/d$ at ν.
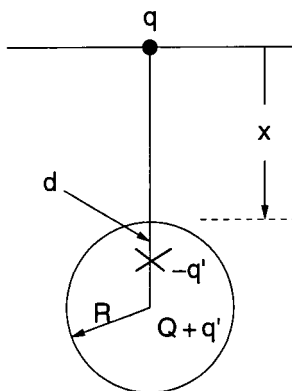
Fig. 3.8 The hanging spherical shell.

Example 3.5: The spherical shell hanging in the gravitational field

A conductor in the form of a homogeneous spherical shell of radius R, is attached to one end of an elastic string. With the other end, at which a charge $+q$ is placed, the sphere is hanging in the gravitational field of the Earth. Establish the equation from which the extension of the string can be calculated.

Solution: We let Q be the charge of the sphere (if not zero) and x_0 the unextended length of the string. The conductor with the constant potential can effectively be replaced by an image charge $-q'$, whereby q' and the position d, see Fig. 3.8, of the image charge assume the values we calculated previously. We then obtain for the equation of motion of the spherical shell of mass m (observe that the rigid body "spherical shell" can be treated like a pointlike particle)

$$m\ddot{x} = mg - k(x - x_0) - \frac{qq'}{(x_0 + d + (x - x_0))^2} + \frac{q(Q + q')}{|x + R|^2}.$$

As we explained previously: The unearthed sphere carries the total charge Q.

3.6.2 Green's Theorems

The previous example illustrates the point in the following formulation of a problem. We had the relation

$$\phi(\mathbf{r}) = k \int_V \frac{\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}',$$

i.e. if $\rho(\mathbf{r}')$ is given, then this relation yields the potential $\phi(\mathbf{r})$. In electrostatics one frequently encounters a different type of problem: $\phi(\mathbf{r})$ is known on certain surfaces, i.e. boundary conditions supplementing the Poisson equation are given, but not $\rho(\mathbf{r})$, and it is required to determine for instance $\mathbf{E}$. The solution of such boundary problems is simplified with the help of *Green's Theorems*:

Theorem (1)

$$\int_V \left[\varphi \triangle \psi + (\boldsymbol{\nabla} \varphi) \cdot (\boldsymbol{\nabla} \psi) \right] dV = \int_F \varphi(\boldsymbol{\nabla} \psi) \cdot d\mathbf{F},$$

Theorem (2)

$$\int_V \left[\varphi \triangle \psi - \psi \triangle \varphi \right] dV = \int_F \left[\varphi \boldsymbol{\nabla} \psi - \psi \boldsymbol{\nabla} \varphi \right] \cdot d\mathbf{F}.$$

We *verify* these theorems: Set

$$\mathbf{A} = \varphi \boldsymbol{\nabla} \psi.$$

Then

$$\boldsymbol{\nabla} \cdot \mathbf{A} = \boldsymbol{\nabla} \cdot (\varphi \boldsymbol{\nabla} \psi) = \varphi \triangle \psi + (\boldsymbol{\nabla} \varphi) \cdot (\boldsymbol{\nabla} \psi).$$

But with Gauss' divergence theorem

$$\int_V \boldsymbol{\nabla} \cdot \mathbf{A} dV = \int_F \mathbf{A} \cdot d\mathbf{F} = \int_F \mathbf{A} \cdot \mathbf{n} dF.$$

Moreover

$$\mathbf{A} \cdot \mathbf{n} = \varphi(\boldsymbol{\nabla}\psi) \cdot \mathbf{n} \equiv \varphi \frac{\partial \psi}{\partial n},$$

where $\mathbf{n}$ is a unit vector pointing vertically out of the surface F. We then have

$$\int_V \left[\varphi \triangle \psi + (\boldsymbol{\nabla}\varphi) \cdot (\boldsymbol{\nabla}\psi) \right] dV = \int_F \varphi(\boldsymbol{\nabla}\psi) \cdot d\mathbf{F},$$

which is the first theorem. If we interchange in this φ and ψ, we obtain

$$\int_V \left[\psi \triangle \varphi + (\boldsymbol{\nabla}\psi) \cdot (\boldsymbol{\nabla}\varphi) \right] dV = \int_F \psi(\boldsymbol{\nabla}\varphi) \cdot d\mathbf{F}.$$

Subtracting this equation from that of the first theorem, we obtain the second theorem, i.e.

$$\int_V \left[\varphi \triangle \psi - \psi \triangle \varphi \right] dV = \int_F \left[\varphi \boldsymbol{\nabla}\psi - \psi \boldsymbol{\nabla}\varphi \right] \cdot d\mathbf{F}.$$

An application which will be needed later is the case of

$$\psi = \frac{1}{|\mathbf{r} - \mathbf{r}'|}, \qquad \varphi = \phi,$$

so that

$$\triangle\psi = \triangle \frac{1}{|\mathbf{r} - \mathbf{r}'|} = -4\pi\delta(\mathbf{r} - \mathbf{r}'), \quad \triangle\varphi = \triangle\phi(\mathbf{r}) = -4\pi k\rho(\mathbf{r}).$$

The second of Green's theorems then implies

$$-4\pi \int_V \phi(\mathbf{r}')\delta(\mathbf{r} - \mathbf{r}')dV' + 4k\pi \int_V \frac{\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|}dV'$$

$$= \int_F \left[\phi\frac{\partial}{\partial n'}\left(\frac{1}{|\mathbf{r} - \mathbf{r}'|} \right) - \frac{1}{|\mathbf{r} - \mathbf{r}'|}\frac{\partial \phi}{\partial n'} \right] dF'. \tag{3.4}$$

In particular, *for $\mathbf{r}$ inside V*:

$$\phi(\mathbf{r}) = k \int_V \frac{\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|}dV' + \frac{1}{4\pi} \int_F \left[\frac{1}{|\mathbf{r} - \mathbf{r}'|}\frac{\partial \phi}{\partial n'} - \phi\frac{\partial}{\partial n'}\frac{1}{|\mathbf{r} - \mathbf{r}'|} \right] dF'. \tag{3.5}$$

Allowing here the area of integration to become infinite, we have

$$dF' \sim r'^2 d\Omega, \quad \phi \propto \frac{1}{r}, \quad \frac{\partial\phi}{\partial n} \propto \frac{1}{r^2}, \quad \frac{1}{|\mathbf{r} - \mathbf{r}'|}\frac{\partial\phi(\mathbf{r}')}{\partial n'} \sim \frac{1}{r'^3},$$

$$\frac{dF'}{r'^3} \sim \frac{d\Omega}{r'} \xrightarrow{r', F' \to \infty} 0.$$

We thus obtain the result familiar from Chapter 2, i.e.

$$\phi(\mathbf{r}) = k \int_{V_\infty} \frac{\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} dV'.$$

On the other hand, *if the volume of integration V does not enclose charges,* we have $\rho = 0$ and

$$\phi(\mathbf{r})\bigg|_{(\mathbf{r}\in V)} = \frac{1}{4\pi} \int_F \left[\frac{1}{|\mathbf{r} - \mathbf{r}'|}\frac{\partial\phi(\mathbf{r}')}{\partial n'} - \phi\frac{\partial}{\partial n'}\frac{1}{|\mathbf{r} - \mathbf{r}'|} \right] dF'. \tag{3.6}$$

Consequently the potential ϕ at $\mathbf{r}$ is determined by values and derivatives of ϕ on the boundary.

One describes as *Dirichlet boundary conditions* those that specify the potential on the boundary (not necessarily as zero!). One describes as *Neumann boundary conditions* those which specify the normal derivative on the boundary, i.e. $\partial\phi/\partial n$. By a specification of Dirichlet *or* Neumann boundary conditions *the potential is uniquely determined.* This can be seen as follows. Setting $u = \phi_1 - \phi_2$, with ϕ_1, ϕ_2 assumed to be two different solutions (but single-valued on the boundary, i.e. we do not assume multi-valued boundary conditions), then

$$\triangle u = 0, \quad \text{since} \quad \triangle\phi_{1,2} = -4\pi k\rho.$$

Setting in the first of Green's theorems $\varphi = \psi = u$, we have

$$\int_V (u\triangle u + \boldsymbol{\nabla} u \cdot \boldsymbol{\nabla} u) dV = \int_F u\boldsymbol{\nabla} u \cdot d\mathbf{F}.$$

In the case of Dirichlet boundary conditions u vanishes on the boundary of F (given uniquely), in the case of Neumann boundary conditions $\boldsymbol{\nabla} u$. Thus, in either of the two cases $\int_F = 0$ and hence

$$\int_V (\boldsymbol{\nabla} u)^2 dV = 0, \quad \text{i.e.} \quad \boldsymbol{\nabla} u = 0,$$

implying u inside $V = $ const. In the case of Dirichlet boundary conditions u is zero on the boundary; hence the constant is zero, i.e. $u = 0$. In the case of Neumann boundary conditions the constant can be $\neq 0$, and so ϕ_1 and ϕ_2 differ at most by an insignificant constant.

3.6.3   Green's Function and Image Potential

We saw earlier that

$$\nabla_{\mathbf{r}}^2 \frac{1}{|\mathbf{r} - \mathbf{r}'|} = -4\pi\delta(\mathbf{r} - \mathbf{r}'). \tag{3.7}$$

(The expression $1/|\mathbf{r} - \mathbf{r}'|$ is at $\mathbf{r}$ the potential of a point charge at $\mathbf{r}'$.)
The function $1/|\mathbf{r} - \mathbf{r}'|$, however, is *only one class* of functions G, for which
Eq. (3.7) is valid, i.e. we can have

$$\triangle G \equiv \nabla^2 G(\mathbf{r}, \mathbf{r}') = -4\pi\delta(\mathbf{r} - \mathbf{r}')$$

(for the moment we call G Green's function *in spite of* the factor -4π; here
this is only a matter of names!) with

$$G(\mathbf{r}, \mathbf{r}') = \frac{1}{|\mathbf{r} - \mathbf{r}'|} + F(\mathbf{r}, \mathbf{r}'), \quad \text{where } \triangle F(\mathbf{r}, \mathbf{r}') = 0. \tag{3.8}$$

The function F is called *image potential*. Solutions of the equation $\triangle F = 0$
are known as *harmonic functions*. Previously we obtained the expression
(3.5) as a general integral representation of the electrostatic potential. In
order to obtain an expression which takes also the image potential into ac-
count, we proceed as follows. In the second of Green's theorems we replace
φ by ϕ and ψ by $G(\mathbf{r}, \mathbf{r}')$:

$$\int_V \left(\phi \triangle G(\mathbf{r}, \mathbf{r}') - G(\mathbf{r}, \mathbf{r}') \triangle \phi \right) dV(\mathbf{r}')$$

$$= \int_F \left(\phi \frac{\partial G(\mathbf{r}, \mathbf{r}')}{\partial n'} - G(\mathbf{r}, \mathbf{r}') \frac{\partial \phi}{\partial n'} \right) dF(\mathbf{r}'),$$

and hence we obtain

$$\phi(\mathbf{r}) = k \int_V \rho(\mathbf{r}') G(\mathbf{r}, \mathbf{r}') dV'$$

$$+ \frac{1}{4\pi} \int_{F(V)} \left[G(\mathbf{r}, \mathbf{r}') \frac{\partial \phi(\mathbf{r}')}{\partial n'} - \phi(\mathbf{r}') \frac{\partial G(\mathbf{r}, \mathbf{r}')}{\partial n'} \right] dF' \tag{3.9}$$

(since $\triangle G(\mathbf{r}, \mathbf{r}') = -4\pi\delta(\mathbf{r} - \mathbf{r}')$, $\mathbf{r} \in V$, $\triangle\phi = -4k\pi\rho$). This expression still
involves both $\phi(F)$ and $(\partial\phi/\partial n)_F$ (i.e. the potential on the boundary and
its normal derivative on the boundary). With a suitable choice of $G(\mathbf{r}, \mathbf{r}')$,
i.e. $F(\mathbf{r}, \mathbf{r}')$, one or the other surface integral can be eliminated, so that
an expression for $\phi(\mathbf{r})$ results with *either Dirichlet or Neumann boundary
conditions*. In the case of a *Dirichlet boundary condition* we choose

$$G_D(\mathbf{r}, \mathbf{r}') \equiv G(\mathbf{r}, \mathbf{r}') \bigg|_{\mathbf{r}' \text{ on } F} = 0.$$

Then

$$\phi(\mathbf{r}) = k \int_V \rho(\mathbf{r}')G_D(\mathbf{r},\mathbf{r}')dV' - \frac{1}{4\pi}\int_F \phi(\mathbf{r}')\frac{\partial G(\mathbf{r},\mathbf{r}')}{\partial n'}dF'$$

$$= k \int_V \rho(\mathbf{r}')G_D(\mathbf{r},\mathbf{r}')dV', \quad \text{if} \quad \phi(F) = 0. \tag{3.10}$$

In Eq. (3.10) we make the following substitution for G (this is explained in more detail in Sec. 3.6.4):

$$G \rightarrow G_D(\mathbf{r},\mathbf{r}') = \frac{1}{|\mathbf{r}-\mathbf{r}'|} + F(\mathbf{r},\mathbf{r}').$$

Substituting this into Eq. (3.9), we obtain

$$\phi(\mathbf{r})\Big|_{(\mathbf{r}\in V)} = k \int_V \rho(\mathbf{r}')\left\{\frac{1}{|\mathbf{r}-\mathbf{r}'|} + F(\mathbf{r},\mathbf{r}')\right\}dV'$$

$$+ \frac{1}{4\pi}\int_F \left[\left\{\frac{1}{|\mathbf{r}-\mathbf{r}'|} + F(\mathbf{r},\mathbf{r}')\right\}\frac{\partial \phi(\mathbf{r}')}{\partial n'}\right.$$

$$\left. - \phi(\mathbf{r}')\frac{\partial}{\partial n'}\left\{\frac{1}{|\mathbf{r}-\mathbf{r}'|} + F(\mathbf{r},\mathbf{r}')\right\}\right]dF', \tag{3.11}$$

and using Eq. (3.5),

$$0 = k \int_V \rho(\mathbf{r}')F(\mathbf{r},\mathbf{r}')dV' + \frac{1}{4\pi}\int_F \left[F(\mathbf{r},\mathbf{r}')\frac{\partial \phi(\mathbf{r}')}{\partial n'} - \phi(\mathbf{r}')\frac{\partial F(\mathbf{r},\mathbf{r}')}{\partial n'}\right]dF'.$$

This equation represents a relation between the image potential $F(\mathbf{r},\mathbf{r}')$ and the boundary conditions. If $\phi(\mathbf{r}) = 0$ on the boundary of V (Dirichlet condition), then the following relation

$$0 = k \int_V \rho(\mathbf{r}')F(\mathbf{r},\mathbf{r}')dV' + \frac{1}{4\pi}\int_F F(\mathbf{r},\mathbf{r}')\frac{\partial \phi(\mathbf{r}')}{\partial n'}dF' \tag{3.12}$$

is the equation from which, in principle, the image potential $F(\mathbf{r},\mathbf{r}')$ is to be determined. The validity of this relation can be verified in the case of the spherically shaped conductor treated above, although such a check is nontrivial. In the case of Neumann boundary conditions, which we shall not pursue further here, one has to choose (see e.g. Greiner [4])

$$\partial G_N(\mathbf{r},\mathbf{r}')/\partial n' = -4\pi/F.$$

3.6.4   The Image Potential in an Example

We ask ourselves now: What do the preceding considerations imply in the concrete case of the earlier problem of a charge q outside an earthed sphere? We refer to Fig. 3.9. The Dirichlet boundary condition is

$$\phi(\mathbf{r} = \mathbf{a}) = \phi(\mathbf{r} = \infty) = 0.$$

The charge density (in V) is

$$\rho(\mathbf{r}') = q\delta(\mathbf{r}' - \mathbf{y}).$$

The function $G_D(\mathbf{r}, \mathbf{r}')$ as solution of $\triangle G = -4\pi\delta(\mathbf{r} - \mathbf{r}')$ with boundary condition $G = 0$ on the boundary, where $r' = a$, is given by the following expression, as we verify below:

$$G_D(\mathbf{r},\mathbf{r}') = \frac{1}{|\mathbf{r} - \mathbf{r}'|} + \frac{q'}{q|\mathbf{r} - \frac{a^2}{r'^2}\mathbf{r}'|}, \quad \text{i.e.} \quad F(\mathbf{r},\mathbf{r}') = \frac{q'}{q|\mathbf{r} - \frac{a^2}{r'^2}\mathbf{r}'|}, \quad (3.13)$$

where $q' = -qa/r'$. In Eq. (3.13), the first term is the inhomogeneous solution of the equation for G. We see immediately, that G_D vanishes for $r' = a$, since

$$G_D(\mathbf{r},\mathbf{r}')\bigg|_{r'=a} = \left[\frac{1}{|\mathbf{r} - \mathbf{r}'|} + \frac{q'}{q|\mathbf{r} - \frac{a^2}{r'^2}\mathbf{r}'|}\right]_{r'=a} = 0.$$
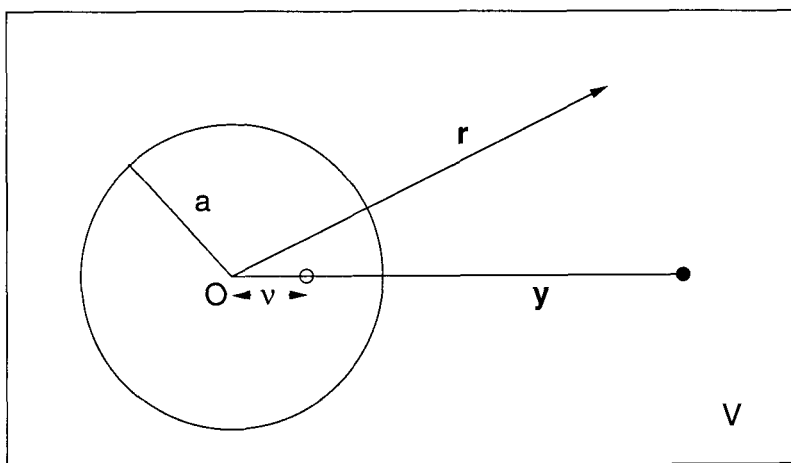
Fig. 3.9   The sphere as boundary surface.

One can easily verify that the expression G_D satisfies its differential equation. Assume $\mathbf{r}' \neq \mathbf{r}''$. Then

$$\triangle_r \left(\frac{a}{|\mathbf{r} - \mathbf{r}'|} + \frac{b}{|\mathbf{r} - \mathbf{r}''|} \right) = \frac{\partial^2}{\partial(\mathbf{r} - \mathbf{r}')^2} \left(\frac{a}{|\mathbf{r} - \mathbf{r}'|} + \frac{b}{|\mathbf{r} - \mathbf{r}' + \mathbf{r}' - \mathbf{r}''|} \right)$$
$$= -4\pi a \delta(\mathbf{r} - \mathbf{r}').$$

The following expression (with $\rho(\mathbf{r}') = q\delta(\mathbf{r}' - \mathbf{y})$) arises from F, i.e. this is the contribution to the potential at $\mathbf{r}$ arising from the image charge in our example above:

$$k \int_V \rho(\mathbf{r}')F(\mathbf{r}, \mathbf{r}')dV' = k \int_V q\delta(\mathbf{r}' - \mathbf{y})F(\mathbf{r}, \mathbf{r}')dV'$$
$$= kqF(\mathbf{r}, \mathbf{y}) = k\frac{q'(r' = y)}{|\mathbf{r} - \frac{a^2}{y^2}\mathbf{y}|}$$
$$= k\frac{q'(r' = y)}{|\mathbf{r} - \nu|},$$

where $q'(r' = y) = -aq/y$. The expression (see Eq. (3.12)):

$$-\frac{1}{4\pi} \int_{F(V)} F(\mathbf{r}, \mathbf{r}') \frac{\partial \phi(\mathbf{r}')}{\partial n'}dF'$$

represents the same, but calculated as the potential of the equivalent induced charge distribution on the sphere. In order to verify this, we would have to show that (with $F(\mathbf{r}, \mathbf{r}')|_{r'=a} = -1/|\mathbf{r} - \mathbf{r}'|_{r'=a}$)

$$\frac{1}{4\pi} \int_{r'=a} \frac{1}{|\mathbf{r} - \mathbf{r}'|} \frac{\partial \phi(\mathbf{r}')}{\partial r'}dF'$$
$$= \frac{1}{4\pi} \int_0^\pi \frac{2\pi a \sin \gamma' a d\gamma'}{\sqrt{r^2 + a^2 - 2ra\cos(\gamma' - \gamma)}} \cdot \frac{kq\frac{y^2}{a}(1 - \frac{a^2}{y^2})}{(\sqrt{a^2 + y^2 - 2ay\cos\gamma'})^3}$$

is equal to

$$\frac{kqa}{y\left|\mathbf{r} - \frac{a^2}{y}\frac{\mathbf{y}}{y}\right|} = \frac{kqa}{y\sqrt{r^2 + \frac{a^4}{y^2} - 2\frac{a^2r}{y}\cos\gamma}}.$$

For a point charge q at $\mathbf{r}' = \mathbf{y}$, we obtain from $\phi(\mathbf{r})$ (see Eq. (3.10))

$$\phi(\mathbf{r}) = qk\left[\frac{1}{|\mathbf{r} - \mathbf{y}|} + \frac{q'}{q|\mathbf{r} - \frac{a^2}{y^2}\mathbf{y}|} \right],$$

i.e. as obtained from the charge and the image charge.

We do not investigate further examples here. We only remind that in the case of a conducting sphere in a homogeneous electric field we replace the homogeneous field by two opposite but equal charges, initially with the same separations from the sphere, but finally removed to infinity.

Example 3.6: The image charge potential

Point charges q_i at (r_i, θ_i, ϕ_i) produce at point (r, θ, ϕ) the potential $\Phi(r, \theta, \phi)$. Show that the image charges $q_i' = \frac{a}{r_i} q_i$ of a conducting sphere of radius a at the points $(\frac{a^2}{r_i}, \theta_i, \phi_i)$ generate at (r, θ, ϕ) the potential

$$\Phi(r, \theta, \phi) = \frac{a}{r} \Phi\left(\frac{a^2}{r}, \theta, \phi \right).$$

Solution: The solution follows immediately from the preceding considerations (compare with the image potential contribution in "*The second question*" above; $\frac{a^2}{r}$ corresponds to ν there).

Example 3.7: A point charge below a spherical shell

A charge q with mass m is placed at a point distance z below the centre of a fixed, conducting and earthed spherical shell of radius a. Verify that the charge induced on the sphere can be replaced by an image charge $-qa/z$ at a distance a^2/z below the centre of the sphere (this part of the Example has been dealt with above). Taking into account also the gravitational field of the Earth calculate the potential energy of the charge. Finally assume that the charge q is allowed to fall (from rest) and obtain its velocity at a depth z_2 below the centre of the sphere (at $z = z_1 = 0$).

Solution: The attractive force between the charge q and the image charge $-qa/z$ is

$$q \frac{qa}{z} \frac{1}{(z - \frac{a^2}{z})^2} = \frac{q^2 az}{(z^2 - a^2)^2}.$$

The electrostatic potential energy is

$$\int \text{force} \cdot dz = q^2 a \int_\infty^z \frac{z \, dz}{(z^2 - a^2)^2} = \frac{1}{2} q^2 a \int_\infty^{z^2} \frac{d(z^2)}{(z^2 - a^2)^2} = -\frac{1}{2} \frac{q^2 a}{(z^2 - a^2)}.$$

The gravitational potential energy is $= -mgz$ ('minus', since measured downward). Hence energy conservation implies (energy at $z =$ energy at z_2)

$$\left[\frac{1}{2} mv^2 - mgz - \frac{1}{2} \frac{q^2 a}{(z^2 - a^2)} \right]_z^{z_2} = 0,$$

i.e.

$$v = \sqrt{ 2g(z_2 - z) \left\{ 1 - \frac{q^2 a(z_2 + z)}{2mg(z_2^2 - a^2)(z^2 - a^2)} \right\} }.$$

Example 3.8: A point charge in front of a conducting plate

A point charge $+q$ is placed in front of an infinitely large conducting plane at shortest separation d. The plane is earthed. The perpendicular from the charge to the plane meets the latter at a point O which is the centre of a circle in the plane of the conductor with radius R_0. Calculate the ratio Q_0/q, where Q_0 is the charge induced on the area of the circle with radius R_0.

Solution: The plane is a conductor which means that the electrons can move about freely without doing work. The potential ϕ on the entire conductor is therefore constant, i.e. zero. We therefore put $\phi = 0$ on the plane (earthed). We can achieve the same condition on the potential by placing a

charge $-q$ at a distance d on the other side of the plane, as depicted in Fig. 3.10. We now consider an arbitrary point $P(\mathbf{R})$ on the same side as charge $+q$. We use O as origin of coordinates and the line connecting the charges as x-axis. Let $\mathbf{R} = (X, Y, Z)$. The potential at P is then (with $k = 1/4\pi\epsilon_0$)

$$\phi(\mathbf{R}) = k\left[\frac{q}{R_+} - \frac{q}{R_-}\right] = qk\left[\frac{1}{\sqrt{(X-d)^2 + Y^2 + Z^2}} - \frac{1}{\sqrt{(X+d)^2 + Y^2 + Z^2}}\right].$$

We have

$$\frac{\partial \phi}{\partial X} = qk\left[\frac{-(X-d)}{[(X-d)^2 + Y^2 + Z^2]^{3/2}} - \frac{-(X+d)}{[(X+d)^2 + Y^2 + Z^2]^{3/2}}\right].$$

For $X \to 0$, it follows that

$$\left(\frac{\partial \phi}{\partial X}\right)_{X\to 0} = \frac{2qdk}{(d^2 + Y^2 + Z^2)^{3/2}}.$$
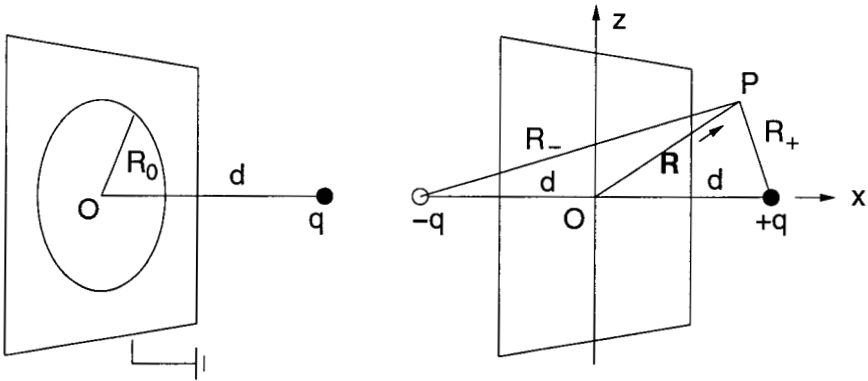
Fig. 3.10 Charge q in front of a conducting plane.

We apply the Gauss law to the volume V enclosed by the plane to the left, where $\mathbf{E} \to \mathbf{E}_1$, and obtain:

$$\int \mathbf{E} \cdot d\mathbf{F} = 4\pi k \sum(\text{charges}) = 0,$$

since no real charge is enclosed to the left. This means $E_{1x} = 0$, i.e. $E_x = -\partial\phi/\partial x = 0$ to the left of the infinitely large conducting plane, as well as inside. We obtain the charge density per unit area, σ, from the relation

$$(\mathbf{E}_2 - \mathbf{E}_1) \cdot \mathbf{n} = 4\pi k\sigma,$$

i.e. in the present case

$$E_{2x} = -\left(\frac{\partial \phi}{\partial x}\right)_{x\to 0} = -\frac{2qdk}{(d^2 + Y^2 + Z^2)^{3/2}}.$$

Hence

$$\sigma = -\frac{qd}{2\pi(d^2 + Y^2 + Z^2)^{3/2}}.$$

The desired ratio is therefore (with $r^2 = Y^2 + Z^2$)

$$\frac{Q_0}{q} = -d \int \frac{dF}{2\pi(d^2 + Y^2 + Z^2)^{3/2}} = -d \int_{\theta=0}^{2\pi} \int_{r=0}^{R_0} \frac{r\,d\theta\,dr}{2\pi(d^2 + r^2)^{3/2}}$$

$$= -d \int_0^{R_0} \frac{r\,dr}{(d^2 + r^2)^{3/2}} = d \left[\frac{1}{(r^2 + d^2)^{1/2}} \right]_0^{R_0}$$

$$= d \left[\frac{1}{\sqrt{R_0^2 + d^2}} - \frac{1}{d} \right].$$

3.7   Method V: Conjugate Functions

This method is applicable only in 2-dimensional problems and depends on 5 main points which we therefore consider first.

Let $W = U + iV$ be a function of the complex variable $z = x + iy$ with $U(x,y), V(x,y)$ real functions of the variables x and y. Assume also that $\lim_{\delta z \to 0} \delta W/\delta z = dW/dz$. Then one verifies readily one after the other the following relations:

(1).

$$\frac{\partial}{\partial y} = i \frac{\partial}{\partial x} \qquad \left(\text{from } \frac{\partial}{\partial y} = \frac{\partial}{\partial z} \frac{\partial z}{\partial y}, \ \frac{\partial}{\partial x} = \frac{\partial}{\partial z} \frac{\partial z}{\partial x} \right).$$

(2).

$$\frac{\partial V}{\partial y} = \frac{\partial U}{\partial x}, \quad \frac{\partial V}{\partial x} = -\frac{\partial U}{\partial y}.$$

(From (1) applied to $U + iV$; the equations are called *Cauchy–Riemann equations*).

(3). At points where they intersect, the lines $U = $ const. are perpendicular to lines $V = $ const. With the help of the previous relations, one can verify that

$$\nabla V \cdot \nabla U = 0, \quad \text{i.e. } \nabla V \perp \nabla U.$$

It follows that lines of constant U are always perpendicular to lines of constant V. In other words, lines of constant force are perpendicular to lines of constant potential (i.e. the equipotentials).

(4). U and V satisfy the 2-dimensional Laplace equation, i.e.

$$\nabla^2 U = 0, \quad \nabla^2 V = 0,$$

as follows from (2). One of the reasons for the importance of the Cauchy–Riemann equations in physics lies in the fact that any of their solutions are automatically harmonic functions, i.e. solutions of Laplace's equation.

(5). Since $(\nabla U) \cdot (\nabla U) = (\nabla V) \cdot (\nabla V)$, we have

$$|\nabla V| = |\nabla U|.$$

We now calculate the capacity of a 2-plate condenser in two dimensions with charge q on plate A. To do this, we assume the two plates A, A', of the condenser have unit length in the additional third dimension. The plates therefore lie along lines of constant potential V as in Fig. 3.11. One may note that the plates are not necessarily parallel (see examples below). Let δr be an element of length of plate A.
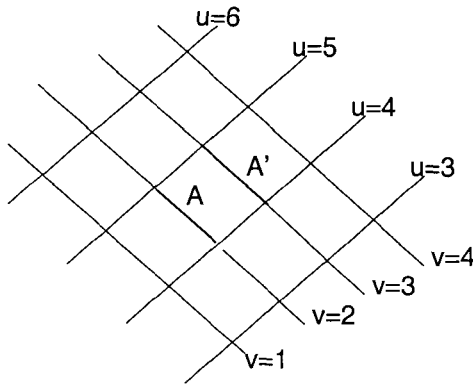
Fig. 3.11 The 2-plate condenser with plates A, A'.

Then, (always of unit length in the (here) irrelevant third dimension)

$$q = \int \sigma dr,$$

where σ is again the charge per unit area. Since the electric field strength E is again approximately zero on the backside of a plate, E is given by $E = 4\pi k \sigma$ (cf. Eq. (2.16)). Hence

$$q = \frac{1}{4\pi k} \int |E| dr = \frac{1}{4\pi k} \int |\nabla V| dr = \frac{1}{4\pi k} \int |\nabla U| dr = \frac{1}{4\pi k} \int \frac{\partial U}{\partial r} dr.$$

It follows that

$$q = \frac{1}{4\pi k} [U],$$

where $[\cdots]$ stands for the corresponding difference. For the capacity of the condenser we then obtain the formula:

$$C = \frac{q}{[V]} = \frac{1}{4\pi k} \frac{[U]}{[V]}.$$

Example 3.9: The simple parallel plate condenser
Calculate the capacity of the parallel plate condenser with the help of the preceding considerations.

Solution: We put the plates along the x-direction and assume they are a distance $y = d$ apart. We set $W = z$, so that $U = x, V = y$. Then the capacity is

$$C = \frac{1}{4\pi k}\frac{[U]}{[V]} = \frac{1}{4\pi k}\frac{[x]}{[y]} = \frac{1}{4\pi kd}$$

per unit area (again with unit length in the third direction which is normally called z, but here $z = x + iy$).
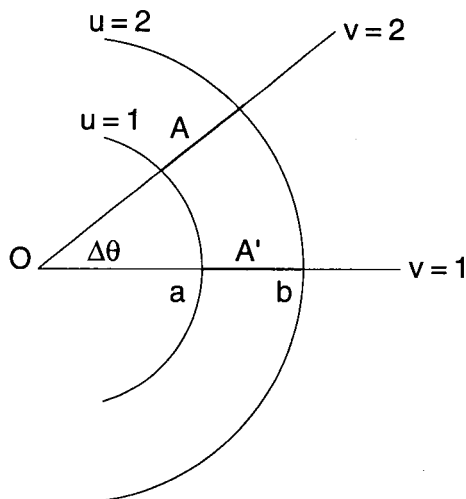
Fig. 3.12 Condenser plates A, A'.

Example 3.10: The condenser with radially assembled plates
Calculate the capacity of the condenser with angle $\triangle\theta$ between the plates.

Solution: In this case we choose $W = \ln z = \ln r + i\theta$, so that $U = \ln r, V = \theta$. Let the angle between the plates with radial lengths $b - a$ be the polar angle $\triangle\theta$ as in Fig. 3.12. In this case we have

$$C = \frac{1}{4\pi k}\frac{[U]}{[V]} = \frac{1}{4\pi k}\frac{[\ln r]}{[\theta]} = \frac{\ln(b/a)}{4\pi k\triangle\theta}.$$

Example 3.11: The cylindrical condenser
Calculate its capacity.

Solution: In the preceding example we reverse the roles of U and V, so that $U = \theta, V = \ln r$. We then obtain the condenser consisting of two coaxial cylinders of radii a and $b > a$. The capacity is then obtained as

$$C = \frac{1}{4\pi k}\frac{[U]}{[V]} = \frac{1}{4\pi k}\frac{[\theta]}{[\ln r]} = \frac{2\pi}{4\pi k \ln(b/a)} = \frac{1}{2k \ln(b/a)}.$$

With this method we can consider very different cases — e.g. the case of non-coaxial cylinders. To do this, we consider a new function $W(z)$. We

let this be $W(z) = V + iU$ (V is later identified with the potential) and $z = x + iy$. Let $P(x,y)$ be the point in the complex plane as indicated in Fig. 3.13, and let r_+, r_- be the distances from $P(x,y)$ to two points, which are along the x-axis a distance $2d$ apart. We choose the following function W:

$$W = \ln \frac{z_+}{z_-} = \ln \frac{r_+}{r_-} + i(\theta_+ - \theta_-) \equiv V + iU.$$

We have therefore

$$V = \ln \frac{r_+}{r_-}, \quad U = \theta_+ - \theta_-.$$
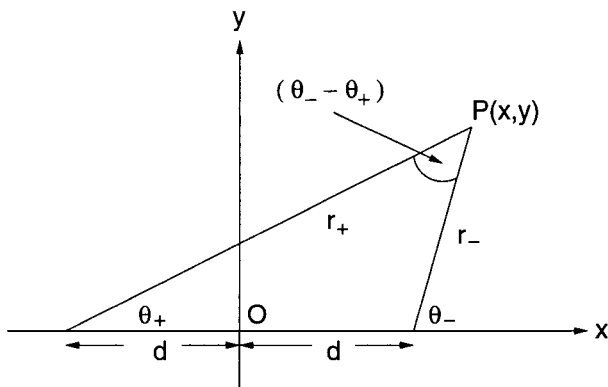
Fig. 3.13 Point P in the complex plane.

We can rewrite the first equation with the help of Fig. 3.13 as

$$e^{2V} = \frac{r_+^2}{r_-^2} = \frac{(x+d)^2 + y^2}{(x-d)^2 + y^2}.$$

With further algebraic manipulation of this equation we have

$$x^2(1 - e^{2V}) + d^2(1 - e^{2V}) + y^2(1 - e^{2V}) + 2xd(1 - e^{2V}) = 0,$$

from which the following equation of a circle results

$$y^2 + (x - d \coth V)^2 = \left(\frac{d}{\sinh V} \right)^2. \tag{3.14}$$

The equipotentials $\pm V = $ const. are circles (with change of sign of V also $\coth V$ and $\sinh V$ change signs). Let $V = V_+$ be the potential on one (cylindrical) plate of the condenser with radius $a_r = d/ \sinh V_+$ and with its centre at $b_+ = d \coth V_+$ on the x-axis. Then $b_+/a_r = \cosh V_+$ and so

$$V_+ = \cosh^{-1} \frac{b_+}{a_+}.$$

Let $V = -V_-$ on the other plate of the condenser. In this case we have

$$V_- = \cosh^{-1} \frac{b_-}{a_-},$$

where $b_- = d \coth V_-$ and $a_- = d/\sinh V_-$. The capacity between the cylinders is now

$$C = \frac{q}{[V]} = \frac{1}{4\pi k} \frac{[U]}{[V]} = \frac{1}{4\pi k} \frac{[\theta_+ - \theta_-]}{[V_+ + V_-]}.$$

The entire domain of $(\theta_+ - \theta_-)$ is 2π. It therefore follows that

$$C = \frac{1}{4\pi k} \frac{2\pi}{\left(\cosh^{-1} \frac{b_+}{a_+} + \cosh^{-1} \frac{b_-}{a_-}\right)}. \tag{3.15}$$

We now consider some applications.

Example 3.12: Parallel wires a distance D apart and with radii a
Again the problem is to calculate the capacity.

Solution: In this case $a_+ = a_- = a$ and $b_+ = b_- = D/2$, so that the capacity is given by

$$C = \frac{1}{4k \cosh^{-1} \frac{D}{2a}}.$$
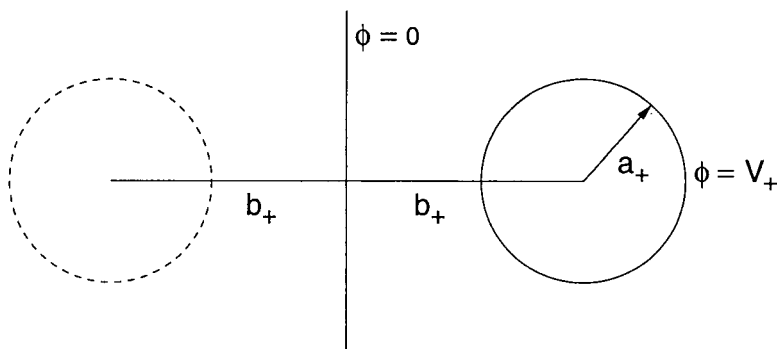
Fig. 3.14 The fictitious cylinder in Kelvin's method.

Example 3.13: The wire in air (equivalent to a cylinder and a plane)
Calculate its capacity.

Solution: In this case one of the cylinders is the surface of the Earth with $a_- = b_-$ and $b_- \to \infty$ (the circle with centre to $-\infty$ and radius to $+\infty$, becomes a straight line representing the surface of the Earth). This means $V_- = 0$. It then follows from Eq. (3.15) that

$$C = \frac{1}{2k \cosh^{-1}\left(\frac{b_+}{a_+}\right)}.$$

This result can also be obtained with Kelvin's method of image charges; cf. Fig. 3.14. The capacity of the system consisting of one cylinder and the parallel image cylinder with opposite charge is given by the expression (3.15) as

$$C_{\text{parallel cylinders}} = \frac{1}{4k \cosh^{-1}\left(\frac{b_+}{a_+}\right)},$$

where b_+ is the height of the actual cylinder above the ground (the Earth's surface where $V = 0$) and a_+ its radius. The difference of potentials of the two parallel cylinders is $[V] = V_+ - (-V_+) = 2V_+$. Since $C \propto 1/[V]$, and in the case of one cylinder above the surface at which $V = 0$, we have in the latter case $[V] = V_+ - 0 = V_+$, so that

$$C_{\text{cylinder to Earth}} = 2C_{\text{parallel cylinders}} = \frac{1}{2k \cosh^{-1}\left(\frac{b_+}{a_+}\right)}.$$

Example 3.14: Electrolytic determination of the capacity

Use the cylindrical condenser to suggest an experiment for the electrolytic determination of its capacity with the help of Eq. (2.16), i.e. the relation $\mathbf{E} \cdot \mathbf{n} = 4\pi k\sigma_{\text{charge}}$, where σ_{charge} is the charge per unit area on one cylindrical electrode, and $\mathbf{E} = 0$ inside the conducting material of the electrode. (Note: This Example assumes familiarity with Secs. 5.2 and 6.1).

Solution: As a simple application of the cylindrical condenser one can imagine — as indicated in Fig. 3.15 — two cylindrical electrodes in a container with a salt solution (cf. Sec. 6.1 where Ohm's law $V = IR$, current $I = \int \mathbf{j} \cdot d\mathbf{F}$, is written $j = \sigma E$ with conductivity $\sigma = 1/\rho$, where ρ is the resistivity; this is unfortunately standard notation, so that this ρ must not be confused with the charge density, nor conductivity σ with charge per unit area).
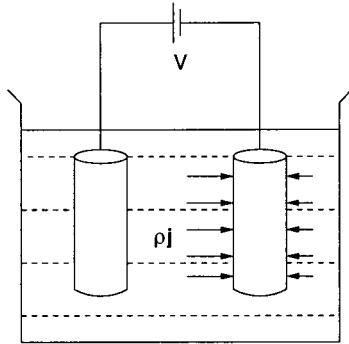
Fig. 3.15 The electrolytic tank containing two electrodes.

Consider the salt solution with resistivity ρ. Macroscopically positive and negative (ion) charges in this solution neutralise each other so that the macroscopic charge density ρ_{charge} in the solution is nought (hence the solution can be looked at as a dielectric medium). Thus the potential ϕ and hence the voltage V between the electrodes is given by the Laplace equation $\nabla^2\phi = 0$, i.e. with source term zero. One can use a probe connected to a voltmeter to determine the potential distribution which would have existed with the same conducting plates in a homogeneous dielectric. Let $\mathbf{j}$ be the current density in the solution (cf. Sec. 5.2) and $d\mathbf{F}$ an element of area of one of the electrodes. Since (cf. Sec. 6.1) $\mathbf{j} = \mathbf{E}/\rho$, the resistance R of the solution is given by Ohm's law as

$R = V/I$ with the current I entering each electrode given by $I = \int \mathbf{j} \cdot d\mathbf{F}$, so that

$$R = \frac{V}{\int \mathbf{j} \cdot d\mathbf{F}} = \frac{\rho V}{\int \mathbf{E} \cdot d\mathbf{F}}$$

$$= \frac{\rho V}{4\pi k \int \sigma_{\text{charge}} dF}.$$

The capacity between the electrodes is therefore, if σ_{charge} denotes the charge q of an electrode per unit area,

$$C = \frac{q}{V} = \frac{\int \sigma_{\text{charge}} dF}{V} = \frac{\rho}{4\pi k R},$$

as one would have written in the case of a dielectric. Thus the capacity of the system of cylinders can be determined experimentally from the known resistivity ρ of the solution into which the cylinders are immersed. R is the resistance of the electrolyte.

3.8   Orthogonal Functions

3.8.1   Orthogonal Functions in General

We begin with the expansion of arbitrary functions in terms of a complete set of basis functions.

Functions $U_i(x), i = 1, 2, \ldots, n$ (with n finite or infinite), are said to be *orthogonal* in $[a, b]$, if

$$\int_a^b U_n^\star(x) U_m(x) dx = s_n \delta_{nm}.$$

The nonnegative number

$$\|U\| = \sqrt{\int_a^b U^\star(x) U(x) dx}$$

is called *norm* of $U(x)$. U is said to be *normalised*, if

$$\int_a^b U^\star(x) U(x) dx = 1.$$

The *analogy with orthogonal vectors* is obvious. The *expansion of an arbitrary function $f(x)$ in terms of $\{U_n(x)\}$* is given by the sum

$$\sum_{n=1}^N a_n U_n(x).$$

It is necessary to establish whether the sum converges for $N \to \infty$ towards $f(x)$, in the sense of *convergence in the mean*, i.e. (definition) one demands

$$\lim_{N \to \infty} \int_a^b \left| f(x) - \sum_{n=1}^N a_n U_n(x) \right|^2 dx = 0,$$

so that

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

(implying uniform convergence), and what the coefficients are. We construct the *mean squared error*

$$M_N = \int_a^b |f(x) - \sum_{n=1}^{N} a_n U_n(x)|^2 dx$$

and determine a_n such that M_N is minimal. M_N is minimised for

$$\frac{\partial M_N}{\partial a_j} = 0, \quad \frac{\partial M_N}{\partial a_j^\star} = 0,$$

i.e.

$$\int_a^b U_j \left(f^\star(x) - \sum_{n=1}^{N} a_n^\star U_n^\star \right) dx = 0$$

and the complex conjugate. This implies for normalised $U_n(x)$ the relations

$$a_j^\star = \int_a^b f^\star(x) U_j dx, \quad a_j = \int_a^b f(x) U_j^\star dx.$$

Since

$$M_N = \int_a^b \left\{ f^\star(x) f(x) + \sum_{n,n'} a_n^\star a_{n'} U_n^\star(x) U_{n'}(x) \right.$$

$$\left. -f^\star(x) \sum_n a_n U_n(x) - f(x) \sum_n a_n^\star U_n^\star(x) \right\} dx$$

$$= \int_a^b dx\, f^\star(x) f(x) - \sum_n |a_n|^2,$$

it follows immediately from $M_N \geq 0$ that

$$\int_a^b |f(x)|^2 dx \geq \sum_{n=1}^{N} |a_n|^2.$$

This relation is known as *Parseval relation*. When the equality applies, the a_n describe functions $f(x)$ completely — the Parseval relation is then called *completeness relation*. Analogous considerations apply to functions in higher dimensions.

3.8.2   Fourier Series, Fourier Expansions

The functions

$$\varphi_n(x): \frac{1}{\sqrt{2\pi}}, \frac{\sin x}{\sqrt{\pi}}, \frac{\cos x}{\sqrt{\pi}}, \frac{\sin 2x}{\sqrt{\pi}}, \frac{\cos 2x}{\sqrt{\pi}}, \ldots$$

constitute in the interval $-\pi \le x \le \pi$ a complete orthonormal system. This follows from the integrals

$$\frac{1}{\pi} \int_{-\pi}^{\pi} \cos mx \cos nx dx = \delta_{mn},$$

$$\frac{1}{\pi} \int_{-\pi}^{\pi} \sin mx \sin nx dx = \delta_{mn},$$

$$\int_{-\pi}^{\pi} \cos mx \sin nx dx = 0.$$

Then in the interval $-\pi \le x \le \pi$

$$f(x) = \sum_{n=0}^{\infty} c_n \varphi_n(x), \quad c_n = \int_{-\pi}^{\pi} f(x)\varphi_n(x)dx.$$

In addition $f(x + 2\pi) = f(x)$, i.e. the function is periodic.

Complex Fourier Series
The functions

$$\varphi_n(x) = \frac{1}{\sqrt{2l}} e^{in\pi x/l}, \quad n = 0, \pm 1, \pm 2, \ldots$$

constitute in the interval $-l \le x \le l$ a complete orthonormal system with

$$\int_{-l}^{l} dx \varphi_n(x)\varphi_m^{\star}(x) = \delta_{mn},$$

so that the periodic function $f(x) = f(x + 2l)$ can be expanded as

$$f(x) = \sqrt{2l} \sum_{n=-\infty}^{\infty} c_n \varphi_n(x), \quad c_n = \frac{1}{2l} \int_{-l}^{l} f(x)e^{-in\pi x/l}dx.$$

Fourier Integrals
We now want to explore the *expansion of nonperiodic functions* by considering in the last relations the limit $l \to \infty$. For $l \to \infty$ the interval $i\pi x/l$ between the exponents of terms in

$$f(x) = \sum_{n=-\infty}^{\infty} c_n e^{in\pi x/l}$$

becomes increasingly smaller; in the limit, one has the continuum. With

$$k_n = \frac{n\pi}{l} \text{ (so that } k_{n+1} - k_n = \pi/l), \quad \frac{lc_n}{\pi} = g(k_n), \quad (3.16)$$

we have

$$f(x) = \pi \sum_{n=-\infty}^{\infty} e^{ik_n x} \frac{g(k_n)}{l} = \sum_{n=-\infty}^{\infty} e^{ik_n x} g(k_n)(k_{n+1} - k_n),$$

which in the limit $l \to \infty$ yields the representation

$$f(x) = \int_{-\infty}^{\infty} dk g(k) e^{ikx}$$

($k_{n+1} - k_n \sim dk = \pi/l$, and for l very large: $\pi/l \sim \pi\epsilon = \delta k$; for n finite we have $k_n = n\pi/l \sim n\pi\epsilon$, and for $n \to \pm\infty : k_n \to k_{\pm\infty} = \pm\infty$). Using Eq. (3.16), we have

$$c_n = \frac{1}{2l} \int_{-l}^{l} f(x) e^{-in\pi x/l} dx = \frac{\pi g(k_n)}{l},$$

and so obtain correspondingly

$$g(k_n) = \frac{1}{2\pi} \int_{-\infty}^{\infty} f(x) e^{-ik_n x} dx,$$

i.e.

$$g(k) = \frac{1}{2\pi} \int_{-\infty}^{\infty} f(x) e^{-ikx} dx. \quad (3.17)$$

This expression is called *Fourier transform* or *spectral function* of $f(x)$. We can easily obtain the inverse of Eq. (3.17) using

$$\delta(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{ikx} dk,$$

because then

$$\int_{-\infty}^{\infty} g(k) e^{iky} dk = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{iky} dk \int_{-\infty}^{\infty} f(x) e^{-ikx} dx$$

$$= \int_{-\infty}^{\infty} dx f(x) \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{ik(y-x)} dk$$

$$= \int_{-\infty}^{\infty} f(x) \delta(y - x) dx = f(y).$$

A necessary condition for the possibility to represent a function $f(x)$ as a Fourier integral is the convergence of the integral

$$\int_{-\infty}^{\infty} |f(x)| dx, \quad \text{i.e.} \quad \lim_{x \to \pm\infty} f(x) = 0.$$

3.8.3   Spherical Functions (Legendre, Associated Legendre, Hyperspherical Functions)

We encountered the following factor earlier (in the integral representation of the solution of Poisson's equation), and now consider it in more detail. Written as an expansion the factor is

$$\frac{1}{|\mathbf{r} - \mathbf{r}'|} = \frac{1}{\sqrt{r^2 + r'^2 + 2rr'\cos\alpha}}$$

$$= \frac{1}{r}\left\{1 + P_1(x)\frac{r'}{r} + P_2(x)\left(\frac{r'}{r}\right)^2 + \cdots\right\}\theta(r - r') + (r \leftrightarrow r')$$

$$= \frac{1}{r}\sum_{l=0}^{\infty} P_l(x)\left(\frac{r'}{r}\right)^l \theta(r - r') + (r \leftrightarrow r'), \qquad (3.18)$$

where with $x = \cos\alpha$

$$P_0(x) = 1, \quad P_1(x) = x, \quad P_2(x) = \frac{1}{2}(3x^2 - 1), \quad \ldots.$$

These polynomials are called *Legendre polynomials*. These are in general given by

$$P_l(x) = \frac{1}{2^l l!}\frac{d^l}{dx^l}(x^2 - 1)^l$$

and define a complete orthonormal system in the interval $x \in [1, -1], \alpha \in [0, \pi]$:

$$\int_{-1}^{1} P_l(x)P_m(x)dx = \frac{2}{2l + 1}\delta_{lm}.$$

A complete orthonormal system of functions which is defined on the *unit sphere* with $\theta \in [0, \pi], \varphi \in [0, 2\pi]$ is given by the *hyperspherical functions* Y_{lm} which are defined as

$$Y_{lm}(\theta, \varphi) = \left\{\frac{(2l + 1)(l - m)!}{4\pi(l + m)!}\right\}^{1/2} P_l^m(\cos\theta)e^{im\varphi},$$

where $-l \leq m \leq l$ and integral, and

$$P_l^m(x) = (-1)^m(1 - x^2)^{m/2}\frac{d^m}{dx^m}P_l(x)$$

are the *associated Legendre polynomials*. For instance

$$Y_{0,0} = \frac{1}{\sqrt{4\pi}}, \quad Y_{10} = \sqrt{\frac{3}{4\pi}}\cos\theta, \quad Y_{11} = -\sqrt{\frac{3}{8\pi}}\sin\theta e^{i\varphi}. \qquad (3.19)$$

The associated Legendre polynomials with negative indices m are defined by

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

General relations of these are

$$P_l^m(x) = \frac{(-1)^m}{2^l l!}(1-x^2)^{m/2}\frac{d^{l+m}}{dx^{l+m}}(x^2-1)^l,$$

$$\int_{-1}^1 P_{l'}^m(x)P_l^m(x)dx = \frac{2}{2l+1}\frac{(l+m)!}{(l-m)!}\delta_{ll'}$$

or

$$\int_0^{2\pi} d\varphi \int_0^\pi Y_{l'm'}^\star(\theta,\varphi)Y_{lm}(\theta,\varphi)\sin\theta d\theta = \delta_{ll'}\delta_{mm'}.$$

The *completeness relation* follows in analogy to the basis change of vectors, $\{\mathbf{e}_i\} \to \{\mathbf{E}_k\}$, from $\mathbf{e}_i \cdot \mathbf{e}_j = \delta_{ij}$ to $\sum_k(\mathbf{e}_i \cdot \mathbf{E}_k)(\mathbf{E}_k \cdot \mathbf{E}_j) = \delta_{ij}$, as the relation

$$\sum_{l=0}^\infty \sum_{m=-l}^l Y_{lm}^\star(\theta',\varphi')Y_{lm}(\theta,\varphi) = \delta(\varphi-\varphi')\delta(\cos\theta-\cos\theta')$$

(in the above analogy k corresponds to m,l, and i,j correspond to θ,φ). In books on Special Functions also the following relation called *addition theorem* can be found:

$$P_l(\cos\alpha) = \frac{4\pi}{2l+1}\sum_{m=-l}^l Y_{lm}^\star(\theta',\varphi')Y_{l,m}(\theta,\varphi), \qquad (3.20)$$

where (see below)

$$\cos\alpha = \sin\theta\sin\theta'\cos(\varphi-\varphi') + \cos\theta\cos\theta' = \cos(\theta-\theta'), \quad \text{if} \quad \varphi = \varphi'$$

(hence its name "addition theorem"), where θ and φ are the spherical coordinates of a vector $\mathbf{r}$, i.e.

$$\mathbf{r} = (r\sin\theta\cos\varphi, r\sin\theta\sin\varphi, r\cos\theta),$$

$$\mathbf{r}' = (r'\sin\theta'\cos\varphi', r'\sin\theta'\sin\varphi', r'\cos\theta'),$$

so that

$$\mathbf{r}\cdot\mathbf{r}' = rr'(\sin\theta\sin\theta'\cos(\varphi-\varphi') + \cos\theta\cos\theta').$$

We use the addition theorem in order to rewrite Eq. (3.18):

$$\frac{1}{|\mathbf{r} - \mathbf{r}'|} = \frac{1}{r} \sum_{l=0}^{\infty} P_l(x) \left(\frac{r'}{r}\right)^l \theta(r - r') + (r \leftrightarrow r')$$

$$= \sum_{l=0}^{\infty} \sum_{m=-l}^{l} \frac{4\pi}{2l + 1} \frac{r'^l}{r^{l+1}} Y_{lm}^\star(\theta', \varphi') Y_{lm}(\theta, \varphi) \theta(r - r')$$

$$+ (r \leftrightarrow r'). \tag{3.21}$$

The hyperspherical functions are also obtained as the solutions of Laplace's equation in spherical polar coordinates.

3.9   The Multipole Expansion

We consider a *charge distribution* (density ρ) in a domain $|\mathbf{r}| < a$. Then the potential is for

$$r > a: \quad \phi(\mathbf{r}) = k \int \frac{\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} dV' \tag{3.22}$$

(the origin of coordinates for instance in the charge centre of mass), provided the volume of integration extended to infinity does not contain conducting surfaces (equipotentials specified by boundary conditions). Thus, only the expression (3.22) remains of the general relation (3.5). This remaining part (3.22) collects contributions only from points $\mathbf{r}'$ in space at which $\rho(\mathbf{r}') \neq 0$. Since $r' < a < r$, this means, that we can expand $1/|\mathbf{r} - \mathbf{r}'|$ in powers of r'/r. This allows us to substitute the expression (3.21) in Eq. (3.22):

$$\phi(\mathbf{r}) = k \sum_{l=0}^{\infty} \sum_{m=-l}^{l} \frac{4\pi}{2l + 1} \frac{1}{r^{l+1}} Y_{lm}(\theta, \varphi) \int \rho(\mathbf{r}') r'^l Y_{lm}^\star(\theta', \varphi') dV'. \tag{3.23}$$

The expressions

$$q_{lm}^\star = k \int Y_{lm}^\star(\theta', \varphi') r'^l \rho(\mathbf{r}') dV' \tag{3.24}$$

are called *multipole moments of the charge distribution* $\rho(\mathbf{r}')$: Specifically the case $l = 0$ is called *monopole moment*, the case $l = 1$ *dipole moment*, the case $l = 2$ *quadrupole moment*, and so on. Since $l = 0, 1, 2, \ldots$, and $m = -l, -l + 1, \ldots, l$, the $2l + 1$ components of q_{lm} form a quantity known as a *spherical tensor of the l-th rank*. With

$$Y_{l,-m} = (-1)^m Y_{lm}^\star$$

it follows that
$$q_{l,-m}^{\star} = (-1)^m q_{l,m}.$$

The multipole moments describe how the charge is distributed:

(1).
$$q_{00} = k \int Y_{00}^{\star}(\theta', \varphi')\rho(\mathbf{r}')dV' = \frac{k}{\sqrt{4\pi}} \int \rho(\mathbf{r}')dV' = \frac{kQ}{\sqrt{4\pi}},$$

(2).
$$q_{11}^{\star} = k \int \rho(\mathbf{r}')r'Y_{11}^{\star}(\theta', \varphi')dV' = -k\sqrt{\frac{3}{8\pi}} \int \rho(\mathbf{r}')r' \sin\theta' e^{-i\varphi'} dV'$$

$$= -k\sqrt{\frac{3}{8\pi}} \int \rho(\mathbf{r}')r'(\sin\theta' \cos\varphi' - i\sin\theta' \sin\varphi')dV',$$

$$q_{11}^{\star} = -k\sqrt{\frac{3}{8\pi}} \int \rho(\mathbf{r}')(x' - iy')dV',$$

or, with $p_{x,y} = \int \rho(\mathbf{r}')(x', y')dV'$,

$$q_{11}^{\star} = -k\sqrt{\frac{3}{8\pi}}(p_x + ip_y),$$

(3).
$$q_{1-1} = -q_{11}^{\star} = k\sqrt{\frac{3}{8\pi}}(p_x - ip_y),$$

(4).
$$q_{10} = k \int \rho(\mathbf{r}')r'Y_{10}(\theta', \varphi')dV' = k\sqrt{\frac{3}{4\pi}} \int \rho(\mathbf{r}')r' \cos\theta' dV' = k\sqrt{\frac{3}{4\pi}}p_z.$$

Here p_x, p_y, p_z are the Cartesian components of the *dipole moment*.

We thus obtain (to be verified below) for $\phi(\mathbf{r})$ the following series expansion (we leave the verification of the quadrupole term as an exercise):

$$\phi(\mathbf{r}) = \frac{kQ}{r} + \frac{k\mathbf{p} \cdot \mathbf{r}}{r^3} + \frac{1}{6} \sum_{i,j} \frac{Q_{ij}(3x_ix_j - r^2\delta_{ij})}{r^5} + \cdots . \qquad (3.25)$$

This expansion is obtained since:

$$\phi(\mathbf{r}) = \frac{4k\pi}{r}Y_{00}(\theta, \varphi) \int Y_{00}(\theta', \varphi')\rho(\mathbf{r}')dV'$$

$$+ \frac{4k\pi}{3r^2}Y_{11}(\theta, \varphi) \int Y_{11}^{\star}(\theta', \varphi')r'\rho(\mathbf{r}')dV'$$

$$+ \frac{4k\pi}{3r^2}Y_{10}(\theta, \varphi) \int Y_{10}(\theta', \varphi')r'\rho(\mathbf{r}')dV'$$

$$+ \frac{4k\pi}{3r^2}Y_{1,-1}(\theta, \varphi) \int Y_{1,-1}(\theta', \varphi')r'\rho(\mathbf{r}')dV' + \cdots ,$$

which reduces to

$$
\phi(\mathbf{r}) = \frac{kQ}{r} + \frac{4k\pi}{3r^2}\left(-\sqrt{\frac{3}{8\pi}}\sin\theta e^{i\varphi}\right)\left(-\sqrt{\frac{3}{8\pi}}(p_x - ip_y)\right)
$$
$$
+\frac{4k\pi}{3r^2}\left(\sqrt{\frac{3}{4\pi}}\cos\theta\right)\left(\sqrt{\frac{3}{4\pi}}p_z\right)
$$
$$
+\frac{4k\pi}{3r^2}\left(\sqrt{\frac{3}{8\pi}}\sin\theta e^{-i\varphi}\right)\left(\sqrt{\frac{3}{8\pi}}(p_x + ip_y)\right) + \cdots
$$
$$
= \frac{kQ}{r} + \frac{k}{2r^2}\sin\theta e^{-i\varphi}(p_x - ip_y) + \frac{k\cos\theta}{r^2}p_z
$$
$$
+\frac{k\sin\theta}{2r^2}\sin\theta e^{-i\varphi}(p_x + ip_y) + \cdots ,
$$

and so to

$$
\phi(\mathbf{r}) = \frac{kQ}{r} + \frac{kr\cos\theta}{r^3}p_z + \frac{kr\sin\theta\cos\varphi}{r^3}p_x + \frac{r\sin\theta\sin\varphi}{r^3}p_y + \cdots
$$
$$
= \frac{kQ}{r} + \frac{k\mathbf{p}\cdot\mathbf{r}}{r^3} + \cdots .
$$

The first term kQ/r is the potential of a point charge Q at the origin, seen at $\mathbf{r}$. The potential of a dipole $\mathbf{p}$ at the origin and observed at $\mathbf{r}$ is $k\mathbf{p}\cdot\mathbf{r}/r^3$, and so on. For the electric field strength we then obtain

$$
\mathbf{E} = -\nabla\phi(\mathbf{r}) = -\nabla\left[k\frac{Q}{|\mathbf{r}|} - k\mathbf{p}\cdot\nabla\frac{1}{|\mathbf{r}|} + \cdots\right]. \tag{3.26}
$$

This expression will be needed later. The potential of a dipole $\mathbf{p}$ at the origin, but observed at $\mathbf{r}$ is therefore

$$
\phi_d(\mathbf{r}) = k\frac{\mathbf{p}\cdot\mathbf{r}}{r^3}. \tag{3.27}
$$

Its field strength $\mathbf{E}_d$ is

$$
\mathbf{E}_d = -\nabla\phi_d(\mathbf{r}) = -k\frac{\mathbf{p} - 3\mathbf{n}(\mathbf{n}\cdot\mathbf{p})}{r^3}, \tag{3.28}
$$

where $\mathbf{n} = \mathbf{r}/r$. The expansion (3.26) tells us that, as we approach the individual charges by going to smaller values of r, we see more and more of the structure of the charge distribution.

We close this consideration with a remark on the *interaction of 2 dipoles*. What we mean by this interaction is the energy of one dipole in the field of the other. This interaction is defined by the expression
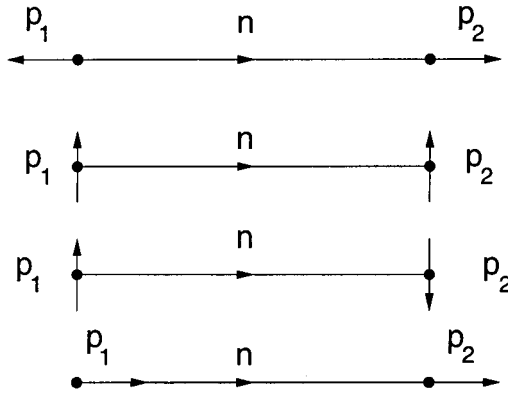
Fig. 3.16 Relative orientations of dipole pairs.

$$W_{12} = -\mathbf{p}_2 \cdot \mathbf{E}_1 = -\mathbf{p}_2 k \left[- \frac{\mathbf{p}_1 - 3\mathbf{n}(\mathbf{n} \cdot \mathbf{p}_1)}{r^3} \right]$$

$$= \frac{k}{r^3} \left[\mathbf{p}_1 \cdot \mathbf{p}_2 - 3(\mathbf{n} \cdot \mathbf{p}_1)(\mathbf{n} \cdot \mathbf{p}_2) \right]. \tag{3.29}$$

We see that this energy and hence force depends also on the orientation of the dipoles to the straight line connecting them. We set

$$W \equiv k \frac{p_1 p_2}{r^3}.$$

In the case of the orientations depicted in Fig. 3.16 we then have the following energy: (a) $W_{12} = 2W$, (b) $W_{12} = W$, (c) $W_{12} = -W$, (d) $W_{12} = -2W$ (minimum of W_{12}). We make an important and fundamental observation: The minimum of the energy ensues for the *same orientation* of the dipoles. This is a fundamental phenomenon, which is also seen to hold for instance for atomic spin orientations. In quantum statistics it is shown that this behaviour is reached at the zero point of the absolute temperature scale.

Example 3.15: Dipole and quadrupole distributions
For the dipole and quadrupole distributions shown in Fig. 3.17 calculate the electric field at the point P.

Solution: In the first case the electric field at the point P in Fig. 3.17 is the sum of a repulsive contribution and an attractive contribution with the resultant as shown antiparallel to the dipole:

$$E_P = \frac{(+q)\cos\theta}{4\pi\epsilon_0[r^2 + d^2/4]} - \frac{(-q)\cos\theta}{4\pi\epsilon_0[r^2 + d^2/4]},$$

where $\cos\theta = (d/2)/\sqrt{r^2 + d^2/4}$, so that

$$E_P = \frac{qd}{4\pi\epsilon_0[r^2 + d^2/4]^{3/2}} \simeq \frac{p}{4\pi\epsilon_0 r^3},$$
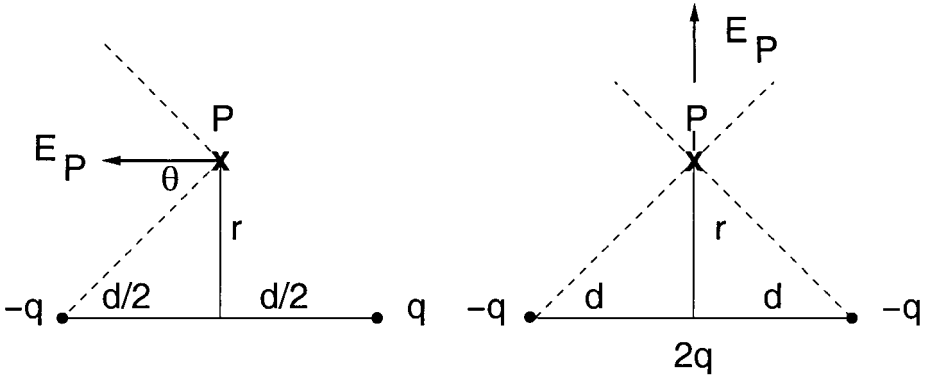
Fig. 3.17 Dipole and quadrupole distributions.

for $r \gg d$, where $p = qd$ is the dipole moment.

In the second case the resultant field is directed along r and away from the dipole:

$$
\begin{aligned}
E_P &= \frac{2q}{4\pi\epsilon_0 r^2} - 2\frac{qr}{4\pi\epsilon_0 (r^2 + d^2)^{3/2}} \\
&= \frac{2q}{4\pi\epsilon_0}\left[\frac{1}{r^2} - \frac{r}{(r^2 + d^2)^{3/2}}\right] \\
&\simeq \frac{3pd}{4\pi\epsilon_0 r^4} \quad \text{for } r \gg d.
\end{aligned}
$$

Example 3.16: Dipole-dipole interaction

Consider two atoms as harmonic oscillators (spring constant $k = \omega_0^2$) and add to these the dipole-dipole interaction (separation of the atoms: R). What is then the energy of the system? In a similar way one can consider two electric LC circuits (a distance R apart) with eigenfrequency $\omega_0 = 1/\sqrt{LC}$, in which the condensers are treated like dipoles.

Solution: Without the dipole-dipole interaction the energy of the system is

$$
E_0 = \frac{1}{2}m(\dot{x}_1^2 + \dot{x}_2^2) + \frac{1}{2}\omega_0^2(x_1^2 + x_2^2),
$$

and quantum mechanically

$$
E_0 = \left(N_1 + \frac{1}{2}\right)\hbar\omega_0 + \left(N_2 + \frac{1}{2}\right)\hbar\omega_0,
$$

where $N_1, N_2 = 0, 1, 2, \ldots$. With attractive dipole-dipole interaction

$$
-ax_1 x_2, \quad a \propto \frac{1}{R^3}
$$

and $k \equiv \omega_0^2$, the energy becomes

$$
\begin{aligned}
E &= \frac{1}{2}m(\dot{x}_1^2 + \dot{x}_2^2) + \frac{1}{2}k(x_1^2 + x_2^2) - ax_1 x_2 \\
&= \frac{1}{2}m(\dot{x}_1^2 + \dot{x}_2^2) + \frac{1}{2}(k - a)(x_1^2 + x_2^2) + \frac{1}{2}a(x_1 - x_2)^2.
\end{aligned}
$$

We set

$$y_1 = x_1 + x_2, \quad y_2 = x_1 - x_2,$$

so that

$$x_1 = \frac{1}{2}(y_1 + y_2), \quad x_2 = \frac{1}{2}(y_1 - y_2),$$

and E is then diagonalised, i.e.

$$E = \frac{1}{2}m\left(\frac{\dot{y}_1}{\sqrt{2}}\right)^2 + \frac{1}{2}m\left(\frac{\dot{y}_2}{\sqrt{2}}\right)^2 + \frac{1}{2}(k-a)\frac{1}{2}m\left(\frac{y_1}{\sqrt{2}}\right)^2 + \frac{1}{2}(k+a)\frac{1}{2}m\left(\frac{y_2}{\sqrt{2}}\right)^2.$$

Quantum mechanically this is

$$E = \left(N_1 + \frac{1}{2}\right)\hbar\sqrt{k-a} + \left(N_2 + \frac{1}{2}\right)\hbar\sqrt{k+a}.$$

For the difference with E_0, i.e. with no dipole-dipole interacion, we obtain for the ground state of the system $(N_1 = 0, N_2 = 0)$

$$\Delta E = E - E_0\Big|_{N_1, N_2 = 0} = \frac{1}{2}\hbar(\sqrt{k-a} - \sqrt{k}) + \frac{1}{2}\hbar(\sqrt{k+a} - \sqrt{k})$$

$$= \frac{1}{2}\hbar\omega_0\left(\sqrt{1 - \frac{a}{\omega_0^2}} + \sqrt{1 + \frac{a}{\omega_0^2}} - 2\right)$$

$$\simeq \frac{1}{2}\hbar\omega_0\left(-\frac{a^2}{4\omega_0^4}\right) \propto \frac{1}{R^6}.$$

This is an energy similar to that of a van der Waals force. One can look at this calculation as a rough physical model of the Casimir effect in producing the potential $\propto 1/R^6$, at least this was suggested in the literature.¶ Actually the Casimir effect is a boundary effect, as we shall see in Chapter 20.

¶See D. Kleppner, *Physics Today* (Oct. 1990) 9; cf. also the Casimir effect in Chapter 20.

This page intentionally left blank

Chapter 4

Macroscopic Electrostatics

4.1 Introductory Remarks

In this chapter we extend the previous considerations to dielectric media (generally with negligible conductivity). This leads to the consideration of the polarisation of these media and their macroscopic effects, the most prominent concept introduced here being that of the dielectric displacement **D**.

4.2 Dielectrics and Dielectric Displacement

So far we considered point charges, for which the following equations hold:

$$\boldsymbol{\nabla} \cdot \mathbf{E} = 4k\pi\rho, \quad \boldsymbol{\nabla} \times \mathbf{E} = 0. \tag{4.1}$$

Here, as in our previous considerations, **E** and ρ are constant in time. In macroscopic dimensions, yet in a small volume, the number of charge-carrying particles is of the order of (say) 10^{18}, most of which are in microscopic motion. In the context of macroscopic electrostatics we consider only such cases which are macroscopically static, i.e. without an observable change in time. This means we assume that all fluctuations in space and time taking place microscopically average out such that they leave no effect macroscopically. Thus we again assume the charge distributions and fields to be time independent. Later it will be a decisive step to assume that the equations of electrostatics hold also for time dependent cases. We define the mean values

$$\mathbf{r} \notin V' : \langle \mathbf{E}(\mathbf{r}) \rangle := \frac{1}{V'} \int_{V'} \mathbf{E}(\mathbf{r} + \mathbf{r}')dV',$$

$$\mathbf{r}'' \notin V' : \langle \rho(\mathbf{r}'') \rangle := \frac{1}{V'} \int_{V'} \rho(\mathbf{r}'' + \mathbf{r}')dV'. \tag{4.2}$$

Here V' is as shown in Fig. 4.1 the volume containing the molecules.
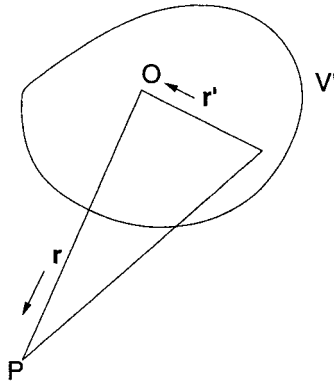
Fig. 4.1 The macroscopic volume V'.

We imagine the field $\mathbf{E}$ with definite direction directed from a charge $+q$ to a charge $-q$. If we place charges or charged particles into this field the positive charges will be attracted by $-q$ and repelled by q, and corresponding effects apply to the negative charges. In this way a separation of the charges, contained in the medium or material arises and is described as its *polarisation.* This is indicated in Fig. 4.2.
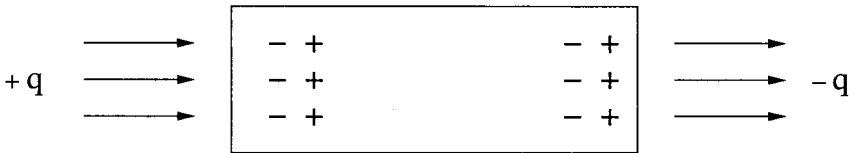
Fig. 4.2   Polarisation of a dielectric in an external field.

Thus, in an externally applied electric field the dielectric substance is polarised with the face at one end exhibiting a positive induced charge, and that at the other the equal and opposite charge. The object can then be considered as an arrangement with alternately positively charged and negatively charged cross sections resulting from a corresponding alignment of its constituent atoms or molecules as dipoles. Thus every such constituent particle behaves like an oriented dipole. The induced charges are described as *polarisation charges.* The quadrupole moments of molecules are usually negligibly small. We define as *(charge) centre of mass* of a molecule j with charge $q_j = \int \rho_j(\mathbf{r}')dV'$ the expression

$$\mathbf{r}_j = \frac{\int \mathbf{r}'\rho_j(\mathbf{r}')dV'}{\int \rho_j(\mathbf{r}')dV'} = \frac{\int \mathbf{r}'\rho_j(\mathbf{r}')dV'}{q_j}$$

(integration over the volume of the molecule). Then the electric field strength $\mathbf{E}$ at a point $\mathbf{r}$ outside the volume containing n molecules is, with

$$\frac{\mathbf{r} - \mathbf{r}'}{|\mathbf{r} - \mathbf{r}'|^3} = -\nabla_{\mathbf{r}} \frac{1}{|\mathbf{r} - \mathbf{r}'|} = \nabla_{\mathbf{r}'} \frac{1}{|\mathbf{r}' - \mathbf{r}|},$$

given by

$$\mathbf{E}(\mathbf{r}) = -\nabla \sum_{j=1}^{n} \left[\frac{kq_j}{|\mathbf{r} - \mathbf{r}_j|} + k\mathbf{p}_j \cdot \nabla_j \frac{1}{|\mathbf{r} - \mathbf{r}_j|} \right]. \tag{4.3}$$

There are two ways to interpret this equation:

(a) Given the charge distribution with corresponding dipole moments (e.g. polar molecules), then $\mathbf{E}(\mathbf{r})$ is the electric field strength produced by these.

(b) Given electrically neutral atoms, which are placed in an externally applied field $\mathbf{E}$ (this is frequently the case, since dipole moments can average out as a result of their motion), then this external field produces the polarisation with induced charges and dipole moments $\mathbf{p}_j$ of the molecules.

With the help of delta distributions which are integrated over we can rewrite the last expression as

$$\mathbf{E}(\mathbf{r}) = -\nabla \sum_{j} \int dV'' \left[\frac{kq_j \delta(\mathbf{r}_j - \mathbf{r}'')}{|\mathbf{r} - \mathbf{r}''|} + k\delta(\mathbf{r}_j - \mathbf{r}'')\mathbf{p}_j \cdot \nabla_{\mathbf{r}''} \frac{1}{|\mathbf{r} - \mathbf{r}''|} \right].$$

It follows that ($V = V'$ being the volume containing the molecules)

$$
\begin{aligned}
\langle \mathbf{E}(\mathbf{r}) \rangle \quad &= \quad \frac{1}{V'} \int_{V'} E(\mathbf{r} + \mathbf{r}')dV' \\
&= \quad -\frac{1}{V} \sum_{j} \int_{V} dV' \nabla \int_{V} dV'' \left[\frac{kq_j \delta(\mathbf{r}_j - \mathbf{r}'')}{|\mathbf{r} + \mathbf{r}' - \mathbf{r}''|} \right. \\
&\qquad\qquad \left. + k\delta(\mathbf{r}_j - \mathbf{r}'')\mathbf{p}_j \cdot \nabla'' \frac{1}{|\mathbf{r} + \mathbf{r}' - \mathbf{r}''|} \right] \\
&\overset{\mathbf{s}=\mathbf{r}''-\mathbf{r}'}{=} \quad -\frac{1}{V} \sum_{j} \int dV' \nabla \int_{V} d\mathbf{s} \left[\frac{kq_j \delta(\mathbf{r}_j - \mathbf{s} - \mathbf{r}')}{|\mathbf{r} - \mathbf{s}|} \right. \\
&\qquad\qquad \left. + k\delta(\mathbf{r}_j - \mathbf{s} - \mathbf{r}')\mathbf{p}_j \cdot \nabla_{\mathbf{s}} \frac{1}{|\mathbf{r} - \mathbf{s}|} \right]. \tag{4.4}
\end{aligned}
$$

Integrating over $\mathbf{s}$ reproduces the first part of the earlier expression (where the integrand is the expression (4.3) with $\mathbf{r} \to \mathbf{r} + \mathbf{r}'$); the variables $\mathbf{s}$ and $\mathbf{r}'$ must therefore be considered as independent variables of integration, and

we obtain

$$\langle \mathbf{E}(\mathbf{r}) \rangle = -\boldsymbol{\nabla} \int_V ds \left[\frac{1}{|\mathbf{r} - \mathbf{s}|} \frac{1}{V} \int_V dV' \sum_j q_j k \delta(\mathbf{r}_j - \mathbf{s} - \mathbf{r}') \right.$$

$$\left. + \left(\boldsymbol{\nabla}_s \frac{1}{|\mathbf{r} - \mathbf{s}|} \right) \frac{k}{V} \int_V dV' \sum_j \mathbf{p}_j \delta(\mathbf{r}_j - \mathbf{s} - \mathbf{r}') \right] \qquad (4.5)$$

(interchanging averaging and taking of gradient). We now set

$$\frac{1}{V} \int_V dV' \sum_j q_j \delta(\mathbf{r}_j - \mathbf{s} - \mathbf{r}') \equiv N(\mathbf{s}) \langle q_{\mathrm{mol}}(\mathbf{s}) \rangle \equiv \rho(\mathbf{s}),$$

$$\frac{1}{V} \int_V dV' \sum_j \mathbf{p}_j \delta(\mathbf{r}_j - \mathbf{s} - \mathbf{r}') \equiv N(\mathbf{s}) \langle \mathbf{p}_{\mathrm{mol}}(\mathbf{s}) \rangle \equiv \mathbf{P}(\mathbf{s}), \qquad (4.6)$$

where we define as

$N(\mathbf{s})$: *the number of molecules per unit volume* at $\mathbf{s}$,
$\langle q_{\mathrm{mol}}(\mathbf{s}) \rangle$: *average charge per molecule* at $\mathbf{s}$,
$\rho(\mathbf{s})$: *macroscopic charge density* at $\mathbf{s}$,
$\mathbf{P}(\mathbf{s})$: *polarisation vector (= dipole moment per unit volume)* at $\mathbf{s}$,
$\langle \mathbf{p}_{\mathrm{mol}}(\mathbf{s}) \rangle$: *average dipole moment per molecule* at $\mathbf{s}$.

We then have (with $\mathbf{s}$ replaced by $\mathbf{r}'$)

$$\mathbf{E}^*(\mathbf{r}) \equiv \langle \mathbf{E}(\mathbf{r}) \rangle = -\boldsymbol{\nabla} \int_V dV' \left[\frac{k\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} + k\mathbf{P}(\mathbf{r}') \cdot \boldsymbol{\nabla}' \frac{1}{|\mathbf{r} - \mathbf{r}'|} \right]$$

$$\equiv -\boldsymbol{\nabla} \phi_\rho - \boldsymbol{\nabla} \phi^P.$$

In view of continuity for $\mathbf{r} \neq \mathbf{r}'$ it is permissible to write the divergence (see explanations below)

$$\boldsymbol{\nabla} \cdot \mathbf{E}^*(\mathbf{r}) = -\int_V dV' \left[k\rho(\mathbf{r}')\boldsymbol{\nabla}^2 \frac{1}{|\mathbf{r} - \mathbf{r}'|} + k\mathbf{P}(\mathbf{r}') \cdot \boldsymbol{\nabla}' \left\{ \boldsymbol{\nabla}^2 \frac{1}{|\mathbf{r} - \mathbf{r}'|} \right\} \right]$$

$$= 4\pi \int_V dV' \left[k\rho(\mathbf{r}')\delta(\mathbf{r} - \mathbf{r}') + k\mathbf{P}(\mathbf{r}') \cdot \boldsymbol{\nabla}'\delta(\mathbf{r} - \mathbf{r}') \right]$$

$$= 4\pi k\rho(\mathbf{r}) - 4\pi k\boldsymbol{\nabla} \cdot \mathbf{P}(\mathbf{r})$$

$$= \frac{1}{\epsilon_0} \left[\rho(\mathbf{r}) - \boldsymbol{\nabla} \cdot \mathbf{P}(\mathbf{r}) \right], \qquad (4.7)$$

where in the step from the first line to the second we used the replacement

$$\boldsymbol{\nabla}^2 \frac{1}{|\mathbf{r} - \mathbf{r}'|} \to -4\pi\delta(\mathbf{r} - \mathbf{r}'),$$

and in the step from the second line to the third

$$\nabla' \to -\nabla,$$

and then ∇ was put in front of the integral. We now set

$$\mathbf{D} \equiv \epsilon_0 \mathbf{E}^* + \mathbf{P}. \tag{4.8}$$

The vector $\mathbf{D}$ is called *dielectric displacement.* As a consequence we have the important relation

$$\nabla \cdot \mathbf{D} = \rho. \tag{4.9}$$

Here, we observe that the polarisation charges (cf. ρ_P below) are not sources of $\mathbf{D}$. Since $\mathbf{E}^* = -\nabla$ of something, but always *curl grad* $= 0$, it follows that

$$\nabla \times \mathbf{E}^* = 0. \tag{4.10}$$

Equations (4.9) and (4.10) are the macroscopic equivalents of the microscopic equations $\nabla \cdot \mathbf{E} = 4\pi k \rho$ and $\nabla \times \mathbf{E} = 0$.

We observed at the beginning, that in the MKSA-system of units the factor $k = 1/4\pi\epsilon_0$ is not dimensionless. As a consequence $\mathbf{D}$ has a dimension different from that of $\mathbf{E}$. In the MKSA-system the unit of dielectric displacement $\mathbf{D}$ is *one coulomb* (C) *per square meter*; in Gaussian units

$$1 \ \mathrm{C \ m^{-2}} = 12\pi \times 10^5 \ \text{statvolt } \mathrm{cm^{-1}} \ (\text{statcoulomb } \mathrm{cm^{-2}}).$$

We deduce from the above relation

$$\mathbf{D} = \epsilon_0 \mathbf{E}^* + \mathbf{P}$$

that the polarisation $\mathbf{P}$ is measured in the MKSA-system in the same units as $\mathbf{D}$.

We distinguish between the two ways to see this:

(a) Given a charge distribution in space (like, for example, of molecules, also polar molecules) whose electrostatic potential contains contributions of dipole moments (higher multipole moments being negligible), then $\mathbf{E}(\mathbf{r})$ is the strength of the electric field of this charge distribution.

(b) Given electrically neutral atoms which macroscopically form a neutral dielectric, and that these are placed in an externally applied electric field $\mathbf{E}(\mathbf{r})$. Then this external field generates the polarisation of the dielectric with induced charges or induced dipole moments which are then defined by the contributions in the above expression for $\mathbf{E}(\mathbf{r})$. (In the case of a neutral atom, which is considered classically, we can have a dipole moment; quantum

mechanically, however, where, for instance, the electron of the hydrogen atom would be found only with a certain probability somewhere near the proton, hence also its charge, the dipole moment of such an atom averages out).

Consider an element of length l of a dielectric as illustrated in Fig. 4.3. Here $\triangle F$ is the cross sectional area with charge Q or charge per unit area $\sigma = Q/\triangle F$. The dipole moment is charge $\times$ distance between the charges, i.e. if $\mathbf{P}$ is the density of dipole moments,

$$V|\mathbf{P}| = Ql = \sigma_P \triangle Fl = \sigma_P V.$$

It follows that

$$|\mathbf{P}| = \sigma_P \equiv \mathbf{P} \cdot \mathbf{n}. \tag{4.11}$$
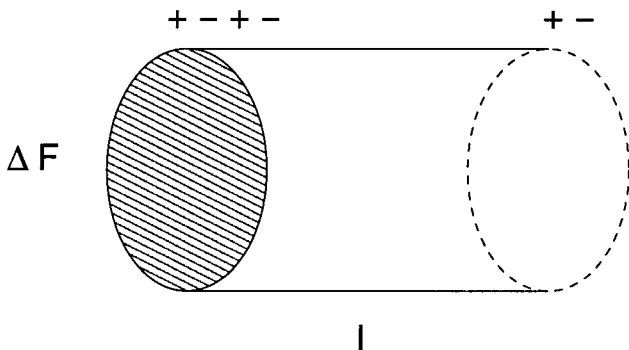
Fig. 4.3 An element of a dielectric.

We now look in more detail at the expression ϕ^P introduced above and contained in $\mathbf{E}^*(\mathbf{r})$. We have

$$
\begin{aligned}
\phi^P(\mathbf{r}) &= k \int_V dV' \mathbf{P}(\mathbf{r}') \cdot \nabla' \frac{1}{|\mathbf{r} - \mathbf{r}'|} \\
&= k \int_V dV' \left[\nabla' \cdot \left(\mathbf{P}(\mathbf{r}') \frac{1}{|\mathbf{r} - \mathbf{r}'|} \right) - \frac{1}{|\mathbf{r} - \mathbf{r}'|} \nabla' \cdot \mathbf{P}(\mathbf{r}') \right] \\
&= k \int_{F(V)} d\mathbf{F} \cdot \mathbf{P}(\mathbf{r}') \frac{1}{|\mathbf{r} - \mathbf{r}'|} - k \int dV' \frac{\nabla' \cdot \mathbf{P}(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} \\
&\equiv \phi_\sigma^P + \phi_\rho^P.
\end{aligned}
\tag{4.12}
$$

We let σ_P be the induced charge density per unit area (induced through polarisation by the external field) defined by

$$\sigma_P = \mathbf{P}(\mathbf{r}) \cdot \mathbf{n}, \tag{4.13}$$

and we let ρ_P be the similarly induced charge density per unit volume defined by

$$\nabla \cdot \mathbf{P} = -\rho_P. \tag{4.14}$$

Without the external field $\mathbf{E}$, the polarisation $\mathbf{P}$ would be zero (because the dipoles would average out). For $|\mathbf{E}|$ not too large (and not for ferroelectric substances, i.e. not universally) we have — from here on we replace $\mathbf{E}^*$ by $\mathbf{E}$ — the linear relation

$$\mathbf{P} = \chi_e \epsilon_0 \mathbf{E}, \tag{4.15}$$

in which the constant χ_e is the so-called *electric susceptibility*. For *isotropic dielectrics* $\mathbf{P} \propto \mathbf{E}$:

$$P_i(\mathbf{E}) = P_i(0) + \sum_j a_{ij} E_j + \cdots , \quad P_i(0) = 0, \quad a_{ij} \propto \delta_{ij}.$$

Since $\mathbf{D} = \epsilon_0 \mathbf{E} + \mathbf{P}$, it follows that

$$\mathbf{D} = (\epsilon_0 + \epsilon_0 \chi_e)\mathbf{E} \equiv \epsilon \mathbf{E}, \quad \epsilon \equiv \epsilon_0 + \epsilon_0 \chi_e, \quad \chi_e = \frac{1}{\epsilon_0}(\epsilon - \epsilon_0) \geq 1. \tag{4.16}$$

The quantity ϵ is called *dielectric constant*. The ratio ϵ/ϵ_0 for the case of the vacuum is 1 and for example for air 1.0005, for glass 5 up to 8 and for water 81. Above we had $\nabla \cdot \mathbf{D} = \rho$, so that, if ϵ is constant,

$$\nabla \cdot \mathbf{E} = \frac{\rho}{\epsilon}.$$

One defines as *relative permittivity* the expression

$$\kappa = \frac{\epsilon}{\epsilon_0} = 1 + \chi_e, \quad E = \frac{D}{\kappa \epsilon_0}.$$

We imagine the formation of polarisation charges as explained above: The electrically neutral atom when placed in an electric field assumes the shape of a dipole as a result of the deformation polarisation (classically the shifting of positive charge to one side, and negative charge to the other). *Polar molecules* are those which possess their own dipole moment. In many other cases rotation, vibration etc. average the dipole moments out to zero in the laboratory frame of reference.

It remains to consider the modification of the Gauss law. We had

$$\mathbf{D} = \epsilon_0 \mathbf{E} + \mathbf{P}.$$

It follows that

$$\int dV' \nabla \cdot \mathbf{D} = \epsilon_0 \int dV' \nabla \cdot \mathbf{E} + \int dV' \nabla \cdot \mathbf{P},$$

i.e.

$$\int dV' \rho = \epsilon_0 \int \mathbf{E} \cdot d\mathbf{F} + \int dV'(-\rho_P),$$

or

$$Q = \epsilon_0 \int \mathbf{E} \cdot d\mathbf{F} - Q_P$$

and hence

$$\epsilon_0 \int \mathbf{E} \cdot d\mathbf{F} = Q + Q_P. \tag{4.17}$$

This is the *modified Gauss law*, in which Q is the *true macroscopic charge* and Q_P the *induced charge*. One should note that Q_P does not have to be zero; in the case of unpolarised dielectrics the sum of the charges vanishes for every volume element, i.e. $\rho = 0$. In the case of polarisation in general $\rho_P dV \neq 0$, in view of a different distribution (cf. a volume element of the box in Fig. 4.2). See, however, Sec. 4.5, where $\rho_P = 0$.

Example 4.1: Relative permittivity*

An insulating medium fills the space between two parallel plates with charge densities σ and $-\sigma$ coulomb/meter^2. Determine the relative permittivity of the dielectric, i.e. κ (with $E = D/\kappa\epsilon_0$), expressed in terms of the induced charges ($\pm\sigma_P$ coulomb/meter^2) at the end-faces.

Solution: We start from the standard equations, i.e.

$$\epsilon_0 \mathbf{E} = \mathbf{D} - \mathbf{P}, \quad \text{with} \quad \boldsymbol{\nabla} \cdot \mathbf{D} = \rho, \quad \boldsymbol{\nabla} \cdot \mathbf{P} = -\rho_P.$$

Polarisation charges are not sources of $\mathbf{D}$. The decisive point here is to distinguish clearly between the real applied charges on the plates at either end ($\sigma, -\sigma$ per unit area), and the charges induced by them in the medium, i.e. $-\sigma_P$ at the medium end next to $+\sigma$ and $+\sigma_P$ at the medium end next to $-\sigma$. Hence we have

$$\int \boldsymbol{\nabla} \cdot \mathbf{D} dV = \int \mathbf{D} \cdot d\mathbf{F} = \int \rho dV = Q,$$

so that for an element $\triangle F$ of area of a plate

$$\mathbf{D} \cdot \triangle \mathbf{F} = Q, \quad D = \frac{Q}{\triangle F} = \sigma,$$

with $\mathbf{D}$ parallel to $\triangle \mathbf{F}$. However, in

$$\int \boldsymbol{\nabla} \cdot \mathbf{P} dV = \int \mathbf{P} \cdot d\mathbf{F} = -\int \rho_P dV = -Q_P,$$

$\mathbf{P}$ is parallel to $-\mathbf{D}$ and hence

$$-\mathbf{P} \cdot \triangle \mathbf{F} = -Q_P, \quad P = \frac{Q_P}{\triangle F} = \sigma_P.$$

Thus

$$E = \frac{D - P}{\epsilon_0} = \frac{\sigma - \sigma_P}{\epsilon_0} = \frac{\sigma(\sigma - \sigma_P)}{\epsilon_o \sigma} = \frac{D\ (\sigma - \sigma_P)}{\epsilon_0 \quad \sigma} \equiv \frac{D}{\epsilon_0 \kappa},$$

where

$$\kappa = \frac{\sigma}{\sigma - \sigma_P}.$$

---

*See also *The Electromagnetic Problem Solver* [21], p. 213.

4.3   The Behaviour of D at an Interface

We are now interested in the behaviour of **D** on both sides of a macroscopically thin, charged surface, in much the same way as in the case of **E**. We have

$$\nabla \cdot \mathbf{D} = \rho$$

(the polarisation charges are not sources of **D**). Thus

$$\int \nabla \cdot \mathbf{D} = \int \mathbf{D} \cdot d\mathbf{F} = \int dV \rho = Q,$$

and hence for an infinitesimal surface element as indicated in Fig. 4.4 we have
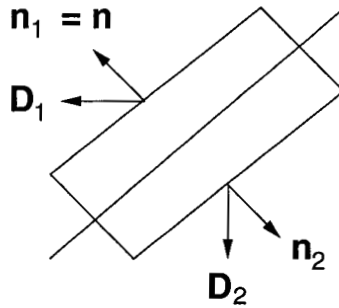
Fig. 4.4   Element of a charged boundary surface.

$$\mathbf{D}_1 \cdot \mathbf{n}_1 + \mathbf{D}_2 \cdot \mathbf{n}_2 = \frac{Q}{\Delta F} = \sigma,$$

or with **n** as a unit vector in the direction of $\mathbf{n}_1$:

$$(\mathbf{D}_1 - \mathbf{D}_2) \cdot \mathbf{n} = \sigma. \qquad (4.18)$$

But since

$$\epsilon_0 \int \mathbf{E} \cdot d\mathbf{F} = (Q + Q_P),$$

we obtain for **E** instead of **D**:

$$\epsilon_0 (\mathbf{E}_1 - \mathbf{E}_2) \cdot \mathbf{n} = (\sigma + \sigma_P). \qquad (4.19)$$

On the other hand, since $\nabla \times \mathbf{E} = 0$, one obtains for a closed contour from just above to just below the interface that

$$\oint \mathbf{E} \cdot d\mathbf{s} = \int \nabla \times \mathbf{E} \cdot d\mathbf{F} = 0,$$

(a) $\epsilon_1 > \epsilon_2$ (b) $\epsilon_2 > \epsilon_1$
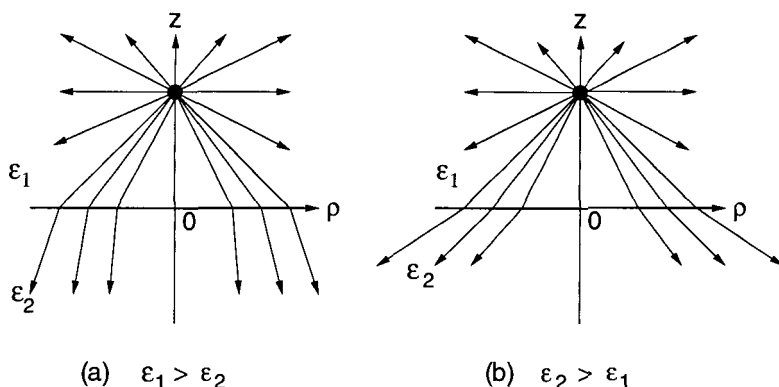
Fig. 4.5 Lines of constant $\mathbf{E}$ for (a) $\epsilon_1 > \epsilon_2$ and (b) $\epsilon_2 > \epsilon_1$.

so that the tangential components of $\mathbf{E}$, i.e. those in the plane of the interface surface, are equal, i.e.

$$E_{1t} - E_{2t} = 0. \qquad (4.20)$$

Since $\nabla \times \mathbf{D} \neq 0$ for $\mathbf{P} \neq 0$, there is no corresponding relation for $\mathbf{D}$.

4.4   Examples

Example 4.2: Lines of E at interfaces

A charge q is placed at $z = 1$ along the z-axis in a medium with dielectric constant ϵ_1 in the region $z > 0$. The medium in the region $z < 0$ has dielectric constant ϵ_2. Sketch the paths of lines of $\mathbf{E}$ in the cases $\epsilon_1 > \epsilon_2$ and $\epsilon_2 > \epsilon_1$.

Solution: With $\rho = \sqrt{x^2 + y^2}$ the potential in the first medium is given by

$$\phi = \frac{q}{4\pi\epsilon_1 \sqrt{\rho^2 + (z-1)^2}},$$

so that

$$\text{for } z > 0: \quad \mathbf{E}^{(1)} = -\nabla\phi = \frac{q}{4\pi\epsilon_1} \left(\frac{\rho\mathbf{e}_\rho + (z-1)\mathbf{e}_z}{[\rho^2 + (z-1)^2]^{3/2}} \right).$$

At the interface to the second medium along the interface normals we have ("± 0" meaning $\pm\epsilon, \epsilon > 0$ and small)

$$(\epsilon_1 E_z^{(1)})_{z=+0} = (\epsilon_2 E_z^{(2)})_{z=-0}$$

(no charge given on the interface). Moreover tangentially

$$[E_\rho^{(1)}]_{z=+0} = [E_\rho^{(2)}]_{z=-0}.$$

Hence

$$(E_z^{(2)})_{z=-0} = \frac{\epsilon_1}{\epsilon_2}(E_z^{(1)})_{z=+0} = -\frac{q}{4\pi\epsilon_2[\rho^2 + 1]^{3/2}}$$

and

$$(E_\rho^{(2)})_{z=-0} = (E_\rho^{(1)})_{z=+0} = \frac{q\rho}{4\pi\epsilon_1[\rho^2 + 1]^{3/2}}.$$

The trajectories of lines of $\mathbf{E}$ are shown in Figs. 4.5 (a) and (b). For $\epsilon_2 \gg \epsilon_1$, i.e. $\epsilon_2 \to \infty$, the second medium behaves approximately like a conductor (no field).

Example 4.3: The spherical condenser

A condenser consists of two concentric, conducting spherical shells with inner charge q and radii a and b, $a < b$. The space between the spheres from a to r_0 is filled with a dielectric with dielectric constant ϵ_1, and from r_0 to b with a different medium of dielectric constant ϵ_2 (both linear). The outer sphere is earthed. Calculate the capacity of the condenser.

Solution: We have spherical symmetry with

$$\int \mathbf{D} \cdot d\mathbf{F} = D_r \cdot 4\pi r^2 = q, \quad D_r = \frac{q}{4\pi r^2}$$

for all r. Moreover, $E_r = -\partial\phi/\partial r$ for potential ϕ. We assume the linearity $\mathbf{D} = \epsilon\mathbf{E}$. Then:
- for $a \leq r < r_0$: $E_r = D_r/\epsilon_1 = q/4\pi\epsilon_1 r^2$, so that $\phi = -\int E_r dr = q/4\pi\epsilon_1 r + C_1$;
- for $r_0 < r \leq b$: $E_r = D_r/\epsilon_2 = q/4\pi\epsilon_2 r^2$, so that $\phi = q/4\pi\epsilon_2 r + C_2$.

At $r = a$: $\phi(r = a) = q/4\pi\epsilon_1 a + C_2$.
Continuity of ϕ at $r = r_0$:

$$\frac{q}{4\pi\epsilon_1 r_0} + C_1 = \frac{q}{4\pi\epsilon_2 r_0} + C_2.$$

Earthing at $r = b$:

$$\phi(r = b) = 0, \quad \frac{q}{4\pi\epsilon_2 b} + C_2 = 0.$$

Thus

$$\phi(r = a) = \frac{q}{4\pi\epsilon_1 a} + \frac{q}{4\pi\epsilon_2 r_0} - \frac{q}{4\pi\epsilon_1 r_0} - \frac{q}{4\pi\epsilon_2 b}.$$

Hence the capacity C is

$$C = \frac{q}{\phi(r = a)} = \frac{4\pi}{\left[\frac{1}{\epsilon_1 a} - \frac{1}{\epsilon_2 b} + \frac{1}{r_0}\left(\frac{1}{\epsilon_2} - \frac{1}{\epsilon_1}\right)\right]}.$$

Example 4.4: The coaxial cable

A coaxial cable consists of a conducting wire with circular cross section of radius a and charge e per unit length, which is surrounded by an insulating layer of thickness d and a thin, hollow, conducting and earthed cylinder of radius $c \gg a + d$ as indicated in Fig. 4.6. Determine the capacity of the cable per unit length.

Solution: We have cylindrical symmetry. The outer cylinder is earthed, and the dielectric is assumed to be linear, i.e. $D \propto E, D = \epsilon E$. We use

$$\int_F \mathbf{D} \cdot d\mathbf{F} = \text{charge enclosed}$$

(i.e. e per unit length with the integral over the cylindrical surface). We then have

$$0 < r < a: \quad E = 0 \quad \text{(conductor)}.$$

According to our earlier treatment of the cylindrical condenser we have for

$$a < r < a + d: \quad E = \frac{e}{2\pi\epsilon r},$$

and for

$$a + d < r < c: \quad E = \frac{e}{2\pi\epsilon_0 r}.$$

Integrating with the help of $E = -d\phi/dr$:

$$0 < r < a: \quad \phi = \phi_1 = \text{const.},$$

$$a < r < a + d: \quad \phi = -\frac{e}{2\pi\epsilon}\ln r + C_1,$$

$$a + d < r < c: \quad \phi = -\frac{e}{2\pi\epsilon_0}\ln r + C_2.$$

The boundary conditions are

$$r = a: \quad \phi = \phi_1 = -\frac{e}{2\pi\epsilon}\ln a + C_1,$$

$$r = a + d \text{ and } \phi \text{ continuous}: \quad -\frac{e}{2\pi\epsilon}\ln(a+d) + C_1 = -\frac{e}{2\pi\epsilon_0}\ln(a+d) + C_2,$$

$$r = c: \quad \phi_2 = 0 \text{ (earthed)}, \quad -\frac{e}{2\pi\epsilon_0}\ln c + C_2 = 0.$$

Inserting the latter two expressions into the
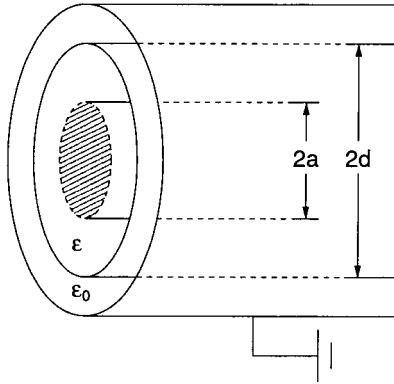
Fig. 4.6 The coaxial cable.

first for the constants, we obtain

$$\phi_1 = -\frac{e}{2\pi\epsilon}\ln a + C_1 = -\frac{e}{2\pi\epsilon}\ln a - \frac{e}{2\pi\epsilon_0}\ln(a+d) + \frac{e}{2\pi\epsilon_0}\ln C + \frac{e}{2\pi\epsilon}\ln(a+d)$$

$$= \frac{e}{2\pi\epsilon_0}\left[\frac{\epsilon_0}{\epsilon}\ln\left(\frac{a+d}{a}\right) + \ln\left(\frac{c}{a+d}\right)\right].$$

Thus, for the capacity C follows

$$C = \frac{e}{\phi_1 - \phi_2} = \frac{2\pi\epsilon_o}{\left[\frac{\epsilon_0}{\epsilon}\ln\left(\frac{a+d}{a}\right) + \ln\left(\frac{c}{a+d}\right)\right]}.$$

4.5 Polarisation of a Sphere in an Electric Field $\mathbf{E}_0$

The sphere[†] is electrically neutral (i.e. originally uncharged, $\rho = 0$) and is assumed to be made of a material with dielectric constant ϵ_2. The externally applied electric field $\mathbf{E}_0$ is taken to be homogeneous and parallel to the unit vector $\mathbf{e}_z$ in a medium with dielectric constant ϵ_1. We take the centre of the sphere as the origin of the frame of coordinates as shown in Fig. 4.7.
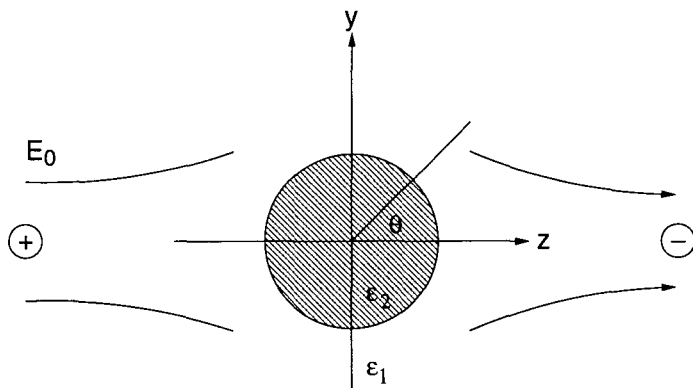
Fig. 4.7   A spherical dielectric in the field $\mathbf{E}_0$.

The polarisation charges on the surface give the sphere the appearance of a dipole placed at the origin and pointing in the direction of z (we shall see that the induced volume charge density in the sphere, i.e. ρ_P, vanishes). Let $\mathbf{E}_i$ be the field $\mathbf{E}$ inside the sphere and $\mathbf{E}_a$ the field outside.

We begin by writing down the equations of (now macroscopic) electrostatics for inside *and* outside the sphere and homogeneous, linearly isotropic dielectrics ($\rho = 0$ implies no "*true*" charges):

$$\nabla \cdot \mathbf{D} = 0, \qquad \mathbf{D} = \epsilon_0 \mathbf{E} + \mathbf{P} = \epsilon \mathbf{E},$$
$$\mathbf{P} \cdot \mathbf{n} = \sigma_P, \qquad \nabla \cdot \mathbf{P} = -\rho_P \tag{4.21}$$

and

$$\nabla \times \mathbf{E} = 0, \qquad \mathbf{E} = -\nabla \varphi. \tag{4.22}$$

From these equations we obtain

$$\epsilon_0 \nabla \cdot \mathbf{E} + \nabla \cdot \mathbf{P} = 0,$$

[†]The following considerations correspond to those of a sphere. In the case of a cylinder some differences arise as can be seen e.g. in the steps from Eq. (4.24) to Eq. (4.25). The case of the cylinder is treated in Example 4.5.

and so

$$\epsilon_0 \triangle \varphi = -\rho_P.$$

On the other hand

$$0 = \boldsymbol{\nabla} \cdot \mathbf{D} = \epsilon \boldsymbol{\nabla} \cdot \mathbf{E} = -\epsilon \triangle \varphi.$$

Hence, (for $\epsilon \neq 0$) $\triangle \varphi = 0$ and so $\rho_P = 0$ for the assumed linear dielectric, i.e. for $\mathbf{P}$ parallel to $\mathbf{E}$ (not for other cases).

At the interface the boundary conditions derived in Sec. 4.3 have to be satisfied. This means,

(1). $E_{1t} = E_{2t}$, and so

$$E_{1\theta}\bigg|_{r=a} = E_{2\theta}\bigg|_{r=a},$$

or

$$\frac{\partial \varphi_i}{\partial \theta}\bigg|_{r=a} = \frac{\partial \varphi_a}{\partial \theta}\bigg|_{r=a},$$

(2).

$$(\mathbf{D}_1 - \mathbf{D}_2) \cdot \mathbf{n} = \sigma = 0,$$

or

$$(\mathbf{D}_1 \cdot \mathbf{n}) = (\mathbf{D}_2 \cdot \mathbf{n}),$$

(where σ, the "*true*" surface charge density, is zero). Thus (since $\mathbf{D} = \epsilon \mathbf{E} = -\epsilon \boldsymbol{\nabla} \varphi$)

$$-\epsilon_1 \frac{\partial \varphi_a}{\partial r}\bigg|_{r=a} = -\epsilon_2 \frac{\partial \varphi_i}{\partial r}\bigg|_{r=a}.$$

In order to be able to solve these two equations, we make a plausible ansatz for E_i, taking into account $\boldsymbol{\nabla} \times \mathbf{E} = 0$. We make the *ansatz* of a homogeneous field $E_i \mathbf{e}_z$ in the interior of the sphere (and demonstrate that this ansatz together with φ_a satisfies the above equations or boundary conditions). The ansatz E_i implies

$$\varphi_i = -E_i z, \quad z = r \cos \theta. \tag{4.23}$$

From outside the sphere appears as a dipole along $\mathbf{e}_z$ with appropriate dipole potential, i.e. (taking into account, that $\mathbf{p}_z \cdot \mathbf{r} = p_z r \cos \theta = p_z z$)

$$\varphi_a = -E_0 z + k\frac{\mathbf{p}_z \cdot \mathbf{r}}{r^3} \tag{4.24}$$

(so that $E_\infty = -\partial \varphi_a / \partial z = E_0$). The second term on the right really results from solution of the Poisson equation $\triangle \varphi_a = 0$ with *Kelvin's theorem*, which

says that if $\varphi_a \propto r^n$ is a solution, then there is another solution $\propto 1/r^{n+1}$ for the same angular dependence. In the present case of a sphere

$$r = (x^2 + y^2 + z^2)^{1/2}, \quad dr/dy = y/r,$$

so that

$$\frac{d\varphi_a}{dy} = -k\frac{3p_z zy}{r^5}, \quad E_{ay} = k\frac{3p_z zy}{r^5}. \tag{4.25}$$

Analogously, we obtain

$$E_{az} = E_0 - k\frac{p_z}{r^3} + k\frac{3p_z z^2}{r^5}. \tag{4.26}$$

Thus outside the sphere and for r near a the field components E_{ay}, E_{ax}, E_z, are different from zero and lead to a curving of the lines of constant E or D (see below). The condition (1), in which we can replace $\partial/\partial\theta$ by $\partial/\partial\cos\theta$, yields

$$E_i = E_0 - k\frac{p_z}{a^3}. \tag{4.27}$$

We note in passing, that this equation follows also from the continuity of the potential at the boundary of the sphere, i.e. at $r = a$. Analogously, condition (2) yields the equation

$$\epsilon_1\left(E_0 + k\frac{2p_z}{a^3}\right) = \epsilon_2 E_i. \tag{4.28}$$

From Eqs. (4.27) and (4.28) we obtain

$$E_i - E_0 = -k\frac{p_z}{a^3} = \frac{\epsilon_1 E_0 - \epsilon_2 E_i}{2\epsilon_1},$$

i.e.

$$E_i = \frac{\frac{3}{2}E_0}{1 + \frac{\epsilon_2}{2\epsilon_1}} = \frac{3\epsilon_1}{\epsilon_2 + 2\epsilon_1}E_0 = E_i.$$

(a) (b)
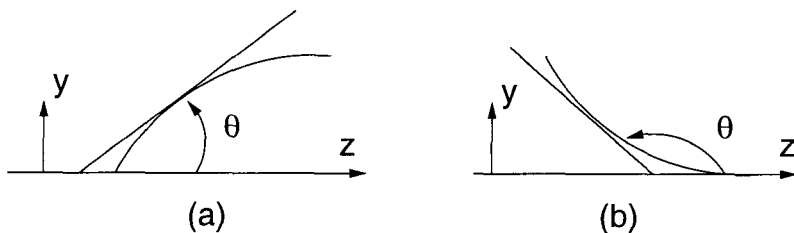
Fig. 4.8 Bending of lines of $\mathbf{E}$.

In Example 4.5 the corresponding result is derived for a cylinder. From Eq. (4.28) we also obtain

$$kp_z = (E_0 - E_i)a^3 = E_0\left(1 - \frac{3\epsilon_1}{\epsilon_2 + e\epsilon_1}\right)a^3,$$

and hence

$$kp_z = \frac{\epsilon_2 - \epsilon_1}{\epsilon_2 + 2\epsilon_1}a^3 E_0. \tag{4.29}$$

It follows, that $p_z > $ (or $<$) 0, if $\epsilon_2 > $ (or $<$) ϵ_1. Substituting the expression (4.29) into E_{ay} for p_z allows us to determine the distortion of the field lines in the neighbourhood of the sphere. Let $\tan\theta$ be the gradient of the $\mathbf{E}$ field line of force at a point. Then, we have

$$\tan\theta = \frac{E_{ay}}{E_{az}} \sim \frac{E_{ay}}{E_0} \begin{cases} > 0 & \text{for} \quad \epsilon_2 > \epsilon_1, \\ < 0 & \text{for} \quad \epsilon_2 < \epsilon_1. \end{cases}$$

The correspondingly distorted lines of $\mathbf{E}$ are shown in Fig. 4.8. Figure 4.8(a) shows $\tan\theta > 0$ for $z > 0$ and (b) $\tan\theta < 0$ for $z > 0$. The next figure, Fig. 4.9, shows the paths of lines of D for (a) $\epsilon_2 > \epsilon_1$ and (b) $\epsilon_2 < \epsilon_1$.
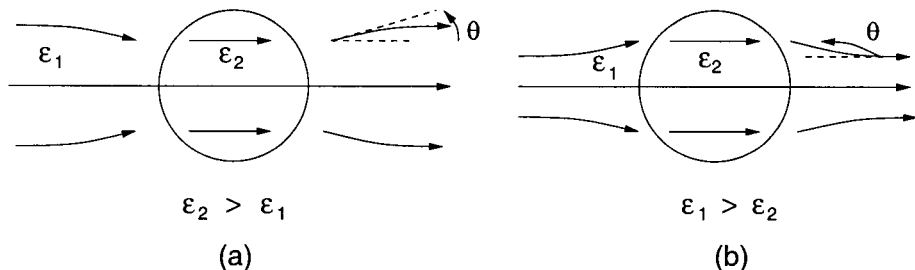
Fig. 4.9 Lines of **D**.

The limiting case of $\epsilon_2 \to \infty$: This is the case of the conducting sphere: $E_i = 0$ (i.e. potential $\varphi = $ const.) in the interior. This case is depicted in Fig. 4.10.
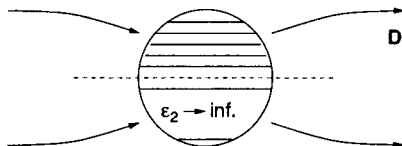
Fig. 4.10 Lines of **D**.

One should note that in Fig. 4.9(a) the behaviour of the D-lines in the interior: $\epsilon \to \infty$, $E_i \to 0$, but $D = \epsilon E_i =$ finite.

The limiting case $\epsilon_2 \lll \epsilon_1$: In this case we have

$$E_i = \frac{3\epsilon_1}{\epsilon_2 + 2\epsilon_1} E_0 \sim \frac{3}{2} E_0,$$

but $D = \epsilon E$, so that $D_2 = \epsilon_2 E_2 = \frac{3}{2}\epsilon_2 E_0 \sim$ is small for $\epsilon_2 \to 0$. In this case the D-field is expelled from the sphere, as indicated in Fig. 4.11. This is the electric analogy to a superconductor in magnetism.

Finally we ask: *How do lines of* $\mathbf{E}$ *differ from lines of* $\mathbf{D}$? D-lines cannot be created at the boundary since $\mathbf{\nabla} \cdot \mathbf{D} = 0$ (no external charges). In the interior as
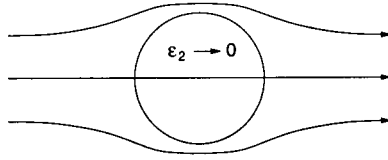
Fig. 4.11 Expulsion of lines of $\mathbf{D}$ for $\epsilon_2 \to 0$.

well as the exterior of the sphere, $\mathbf{P}$ is parallel to $\mathbf{E}$, if the dielectric is isotropic, i.e. $\mathbf{P} = \chi \mathbf{E}$. In the *nonisotropic case* this does not apply. Then the polarisation charges ρ_P are additional sources for lines of $\mathbf{E}$ as indicated in Fig. 4.12, because then

$$0 = \mathbf{\nabla} \cdot \mathbf{D} = \epsilon_0 \mathbf{\nabla} \cdot \mathbf{E} + \mathbf{\nabla} \cdot \mathbf{P} = \epsilon_0 \mathbf{\nabla} \cdot \mathbf{E} - \rho_P,$$

i.e.

$$\mathbf{\nabla} \cdot \mathbf{E} = \frac{1}{\epsilon_0} \rho_P.$$
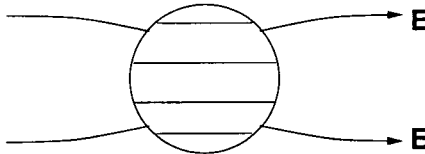
Fig. 4.12 *Additional* lines of $\mathbf{E}$ in the presence of polarisation charges.

Example 4.5: Polarisation of a cylinder in a homogeneous field E_0
Determine the field inside a cylinder of radius $r = a$ and with axis along the x-axis in analogy to the case of the sphere above.

Solution: The potential φ outside and far away ($r \gg a$) is

$$\varphi_a = -E_0 z, \qquad z = r\cos\theta,$$

and we have $\triangle\varphi = 0$, except where there are charges. For the potentials outside and inside we set respectively (see the comment on *Kelvin's theorem* after Eq. (4.24))

$$\varphi_a = -\left(E_0 r + \frac{B}{r}\right)\cos\theta,$$

and

$$\varphi_i = -E_i z.$$

At $r = a$ we must have $\varphi_a = \varphi_i$, i.e.

$$E_0 a + \frac{B}{a} = E_i a,$$

and continuity of $\mathbf{D}\cdot\mathbf{n} = \epsilon E_r = -\epsilon\partial V/\partial r$. Hence

$$\epsilon_1\left(E_0 - \frac{B}{a^2}\right) = \epsilon_2 E_i.$$

The two equations can be solved for B and E_i giving

$$B = \left(\frac{\epsilon_1 - \epsilon_e}{\epsilon_1 + \epsilon_2}\right)a^2 E_0, \qquad E_i = \left(\frac{2\epsilon_1}{\epsilon_1 + \epsilon_2}\right)E_0.$$

An application of this result for E_i can be found in Example 4.6.

4.6   Energy of the Electric Field

4.6.1   The Energy Density Formula

In accordance with the definition of the potential, the potential *energy* of a charge q in the potential ϕ is $W = q\phi$. In the case of a number of charges the energy is, as we saw earlier,

$$W = \frac{1}{2}\sum_i q_i\phi_i, \qquad \phi_i = \sum_{q_k\neq q_i} k\frac{q_k}{|\mathbf{r}_i - \mathbf{r}_k|}.$$

In the case of continuous charge distributions we have

$$W = \frac{1}{2}\int \phi(\mathbf{r})\rho(\mathbf{r})d\mathbf{r}.$$

Since $\nabla\cdot\mathbf{D} = \rho$, we have

$$\begin{aligned}
W &= \frac{1}{2}\int \phi(\mathbf{r})\nabla\cdot\mathbf{D}d\mathbf{r} \\
&= \frac{1}{2}\left\{\int \nabla\cdot(\phi\mathbf{D})d\mathbf{r} - \int \mathbf{D}\cdot\nabla\phi d\mathbf{r}\right\} \\
&= \frac{1}{2}\left\{\int_F \phi\mathbf{D}\cdot d\mathbf{F} + \int \mathbf{D}\cdot\mathbf{E}d\mathbf{r}\right\} \\
&= \frac{1}{2}\int \mathbf{D}\cdot\mathbf{E}d\mathbf{r}.
\end{aligned} \qquad (4.30)$$

To obtain the total energy, and in view of the concept of a field, we have to integrate over all space. Then the surface integral yields a contribution which tends to zero (as we explained earlier).

We now want to show that the total energy of the electrostatic field is

$$\frac{1}{2} \int \mathbf{D} \cdot \mathbf{E} dr,$$

even if the space encloses conductors, provided the volume of integration is only that of the dielectric. This result can be deduced as follows. The potential energy of a dielectric with enclosed conductors of potentials ϕ_n enclosing charges q_n is, with

$$q_n = \int_{V_n} \boldsymbol{\nabla} \cdot \mathbf{D} = \int_{\mathbf{F}_n^*} \mathbf{D}_n \cdot d\mathbf{F}^*$$

and

$$d\mathbf{F}^*{}_n = -d\mathbf{F}_n,$$

as can be seen from Fig. 4.13, given by

$$W = \frac{1}{2} \int_{V_{\text{diel}}} \phi \rho dV + \frac{1}{2} \sum_n \phi_n q_n$$

$$= \frac{1}{2} \int_{V_{\text{diel}}} \phi \boldsymbol{\nabla} \cdot \mathbf{D} dV + \frac{1}{2} \sum_n \phi_n \int_{\mathbf{F}_n^*} \mathbf{D}_n \cdot d\mathbf{F}_n^*$$

$$= \frac{1}{2} \left[\int_{V_{\text{diel}}} \boldsymbol{\nabla} \cdot (\phi \mathbf{D}) dV - \int_{V_{\text{diel}}} \mathbf{D} \cdot \boldsymbol{\nabla} \phi dV + \sum_n \phi_n \int_{\mathbf{F}_n^*} \mathbf{D}_n \cdot d\mathbf{F}^*{}_n \right].$$

This can now be rewritten as (note in the second step the contributions of
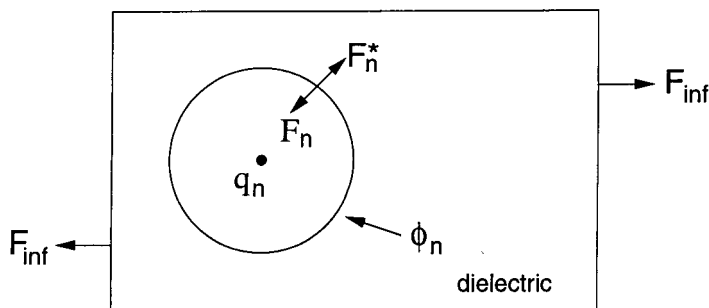
Fig. 4.13 A dielectric with conductor cavities.

F_{diel} from inside and outside boundaries of the dielectric)

$$
\begin{aligned}
W &= \frac{1}{2}\left[\int_{F_{\text{diel}}} \phi \mathbf{D} \cdot d\mathbf{F} + \int_{V_{\text{diel}}} \mathbf{D} \cdot \mathbf{E} dV - \sum_n \phi_n \int_{F_n} \mathbf{D}_n \cdot d\mathbf{F}_n\right] \\
&= \frac{1}{2}\left[\int_{F_\infty} \phi \mathbf{D} \cdot d\mathbf{F} + \sum_n \int_{F_n} \phi_n \mathbf{D}_n \cdot d\mathbf{F}_n \right. \\
&\quad \left. + \int_{V_{\text{diel}}} \mathbf{D} \cdot \mathbf{E} dV - \sum_n \phi_n \int_{F_n} \mathbf{D}_n \cdot d\mathbf{F}_n\right] \\
&= \frac{1}{2}\int_{V_{\text{diel}}} \mathbf{D} \cdot \mathbf{E} dV.
\end{aligned}
\tag{4.31}
$$

4.6.2   Polarisation Energy

Finally we consider the change in energy when a vacuum is replaced by a medium. Let $\triangle_0 W$ be this change from vacuum values $\mathbf{D}_0 = \epsilon_0 \mathbf{E}_0$ and $\mathbf{E}_0$ to medium values $\mathbf{D} = \epsilon_0 \mathbf{E} + \mathbf{P}$ and $\mathbf{E}$. Then

$$
\triangle_0 W = \frac{1}{2}\int d\mathbf{r}[\mathbf{D} \cdot \mathbf{E} - \mathbf{D}_0 \cdot \mathbf{E}_0].
$$

We set $\mathbf{D} = \mathbf{D}_0 + \delta \mathbf{D}$, so that

$$
\triangle_0 W = \frac{1}{2}\int d\mathbf{r}[\mathbf{D}_0 \cdot \mathbf{E} - \mathbf{D} \cdot \mathbf{E}_0] + \frac{1}{2}\int d\mathbf{r}(\mathbf{E} + \mathbf{E}_0) \cdot \delta \mathbf{D}.
$$

Here $(\mathbf{E} + \mathbf{E}_0) = -\boldsymbol{\nabla}\phi'$, where ϕ' is a potential. Using the relation

"div(scalar $\times$ vector) = grad(scalar)·(vector)+(scalar) div(vector)",

we have

$$
-\boldsymbol{\nabla}\phi' \cdot \delta\mathbf{D} = -\boldsymbol{\nabla} \cdot (\phi'\delta\mathbf{D}) + \phi' \underbrace{\boldsymbol{\nabla} \cdot \delta\mathbf{D}}_{\delta(\boldsymbol{\nabla}\cdot\mathbf{D}=\rho)}.
$$

With Gauss' divergence theorem the first term integrated over the volume becomes a surface integral which vanishes. The second contribution vanishes if no new charge $\delta\rho$ is introduced. Hence

$$
\begin{aligned}
\triangle_0 W &= \frac{1}{2}\int d\mathbf{r}[\mathbf{D}_0 \cdot \mathbf{E} - \mathbf{D} \cdot \mathbf{E}_0] \\
&= \frac{1}{2}\int d\mathbf{r}[\epsilon_0 \mathbf{E}_0 \cdot \mathbf{E} - (\epsilon_0 \mathbf{E} + \mathbf{P}) \cdot \mathbf{E}_0] \\
&= -\frac{1}{2}\int d\mathbf{r}[\mathbf{P} \cdot \mathbf{E}_0].
\end{aligned}
\tag{4.32}
$$

This result is used in the following example.

Example 4.6: Measurement of electric susceptibility

Consider a cylindrical U-tube filled with some liquid (l) to the same level in both arms in the original equilibrium position with vapour (v) above. A homogeneous electric field E_0 is applied to one arm of the U-tube. In the presence of the field the liquid rises to a height h above the equilibrium position. Considering this arrangement as a cylindrical dielectric exposed to an external field E_0, find an expression from which the electric susceptibility of the liquid ($\chi_l = \epsilon_l/\epsilon_0 - 1$) can be obtained if that of the vapour (χ_v) is known.
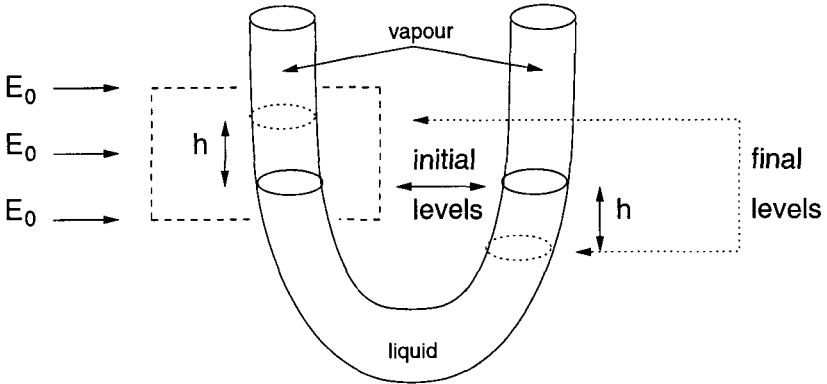
Fig. 4.14 The U-tube with one arm in field E_0.

Solution: Let F be the constant cross section of the tube and ρ the density of the liquid, and g the acceleration due to gravity. The hydrostatic force producing the rise of the liquid to height h above the original equilibrium position against the vapour is

$$\rho F(2h)g,$$

and the corresponding energy is

$$\int_0^h \rho F(2h)g\,dh = \rho g F h^2.$$

The change in electrostatic energy is — using Eq. (4.32) — the change in polarisation energy

$$\triangle_0 W_l - \triangle_0 W_v = -\frac{1}{2}\int d\mathbf{r}(P_l - P_v)E_0$$

$$= -\frac{1}{2}\int_0^h (Fdh)(P_l - P_v)E_0$$

$$= -\frac{1}{2}Fh(P_l - P_v)E_0.$$

The total change in energy is therefore

$$W_{\text{total}} = \rho g F h^2 - \frac{1}{2}Fh(P_l - P_v)E_0.$$

The condition for equilibrium is $\partial W_{\text{total}}/\partial h = 0$, i.e.

$$h = \frac{(P_l - P_v)E_0}{4\rho g}.$$

Now, for $i = l, v$ we have $P_i = \epsilon_0 \chi_i E_i$, $\chi_i = \epsilon_i/\epsilon_0 - 1$, where E_i is the field in the medium $i = l, v$. From Example 4.5 we obtain for the cylindrical media l, v

$$E_l = \frac{2\epsilon_0}{\epsilon_0 + \epsilon_l} E_o, \qquad E_v = \frac{2\epsilon_0}{\epsilon_0 + \epsilon_v} E_o.$$

Then

$$P_i = \frac{2\epsilon_0(\epsilon_i - \epsilon_0)}{\epsilon_0 + \epsilon_i} E_0 = 2\epsilon_0\left(1 - \frac{2\epsilon_0}{\epsilon_0 + \epsilon_i}\right) E_0.$$

Thus

$$h = \frac{2\epsilon_0 E_0^2}{4\rho g}\left[-\frac{2\epsilon_0}{\epsilon_0 + \epsilon_l} + \frac{2\epsilon_0}{\epsilon_0 + \epsilon_v}\right] = \frac{(\epsilon_0 E_0)^2}{\rho g}\frac{\epsilon_l - \epsilon_v}{(\epsilon_0 + \epsilon_l)(\epsilon_0 + \epsilon_v)},$$

which can be reexpressed as

$$h = \frac{\epsilon_0(\chi_l - \chi_v)}{\rho g(\chi_l + 2)(\chi_v + 2)} E_0^2.$$

From this we obtain

$$\chi_l = \frac{\epsilon_0 E_0^2 \chi_v + 2h\rho g(\chi_v + 2)}{\epsilon_0 E_0^2 - h\rho g(\chi_v + 2)}.$$

Hence by measuring h, and since the other quantities appearing on the right are assumed to be known, the susceptibility χ_l of the liquid can be determined.

4.7   Summary of Formulae of Electrostatics

We summarise here the most important formulae of electrostatics.

(1). **Multipole expansion**:

$$\phi(\mathbf{r}) = k\frac{Q}{r} + k\frac{\mathbf{p} \cdot \mathbf{r}}{r^3} + \cdots,$$

$$\mathbf{E} = -\boldsymbol{\nabla}\phi = -k\boldsymbol{\nabla}\left[\frac{Q}{r} - \mathbf{p} \cdot \boldsymbol{\nabla}\frac{1}{|\mathbf{r} - \mathbf{r}_j|} + \cdots\right],$$

but (note the signs of the dipole contributions)

$$\mathbf{E}(\mathbf{r}) = -k\boldsymbol{\nabla}\sum_j\left[\frac{q_j}{|\mathbf{r} - \mathbf{r}_j|} + \mathbf{p_j} \cdot \boldsymbol{\nabla}_j\frac{1}{|\mathbf{r}|} + \cdots\right],$$

since

$$\boldsymbol{\nabla}_{\mathbf{r}}\frac{1}{|\mathbf{r} - \mathbf{r}'|} = -\frac{\mathbf{r} - \mathbf{r}'}{|\mathbf{r} - \mathbf{r}'|^3}, \qquad \frac{\mathbf{r}}{r^3} = -\boldsymbol{\nabla}\frac{1}{|\mathbf{r}|}$$

and

$$\boldsymbol{\nabla}_{\mathbf{r}'}\frac{1}{|\mathbf{r}' - \mathbf{r}|} = -\frac{\mathbf{r}' - \mathbf{r}}{|\mathbf{r} - \mathbf{r}'|^3} = -\boldsymbol{\nabla}_{\mathbf{r}}\frac{1}{|\mathbf{r} - \mathbf{r}'|}.$$

(2). **Macroscopic electrostatics**:

$$\boldsymbol{\nabla} \cdot \mathbf{D} = \rho, \qquad \mathbf{D} = \epsilon_0\mathbf{E} + \mathbf{P},$$

$$\mathbf{P} \cdot \mathbf{n} = \sigma_P, \quad \nabla \cdot \mathbf{P} = -\rho_P,$$
$$\nabla \times \mathbf{E} = 0, \quad (\mathbf{E} = -\nabla\phi).$$

Note: We have to distinguish between ρ_P, σ_P and ρ, σ. For homogeneous, linearly isotropic dielectrics the following **dielectric material equations** can be used:

$$\mathbf{D} = \epsilon\mathbf{E}, \quad \mathbf{P} = \epsilon_0\chi\mathbf{E}.$$

(3). The behaviour at boundary surfaces:

$$(\mathbf{E}_1 - \mathbf{E}_2) \cdot \mathbf{n} = 4\pi k(\sigma + \sigma_P), \qquad (\mathbf{D}_1 - \mathbf{D}_2) \cdot \mathbf{n} = \sigma.$$

But always for tangential components of $\mathbf{E}$:

$$E_{1t} - E_{2t} = 0.$$

For $\mathbf{P} \neq 0$ there is no corresponding relation for $\mathbf{D}$.

(4). Energy of the electric field:

The total energy of the electrostatic field is

$$W = \frac{1}{2} \int \mathbf{D} \cdot \mathbf{E} dV.$$

This page intentionally left blank

Chapter 5

Magnetostatics

5.1 Introductory Remarks

Magnetostatics deals with *stationary currents*. So far we considered only static charges. A current, however, consists of moving charges. The fundamental observation, that magnetic fields exist in the neighbourhood of currents, and hence of moving charges, was made by Oersted in 1819. Oersted had observed that a magnetic needle aligns itself in the vicinity of a current carrying wire perpendicular to this wire. Within a very short time after this discovery (i.e. within a few years) Ampère published his results of a series of experiments which established the law of the force today named after him, and Biot and Savart observed the corresponding law (for a current element) that carries their names today. We are concerned with these laws in this chapter.

5.2 Continuity Equation and Stationary Current

Here by current density we mean only that of the conduction electrons; we do not mean the current density of (classically considered) circular currents of molecular electrons (which, by the way, are not even sufficiently well known). First we consider the case without matter effects. A precisely known current density $\mathbf{j}$ is defined by the charge density of positive charges multiplied by their velocity, i.e.

$$\mathbf{j} = \rho \frac{d\mathbf{s}}{dt} \tag{5.1}$$

in ampere/m^2 or C/m^2 s. We use $d\mathbf{s}$ for an element of the trajectory or path as indicated in Fig. 5.1, in order not to confuse this with the element of

some distance, dr, or the volume element $d\mathbf{r}$. As usual t represents the time variable.
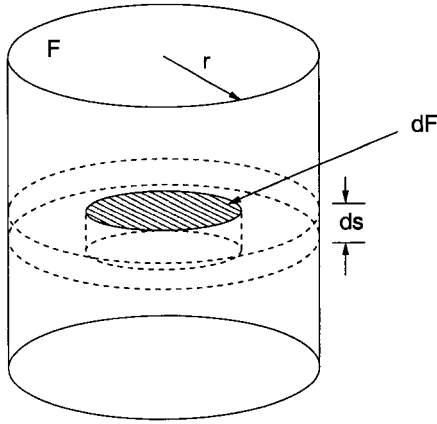
Fig. 5.1 The current element.

The *current I* is defined as that amount of positive* charge which passes through an area F per unit time, i.e.

$$I = \int_F \mathbf{j} \cdot d\mathbf{F} = jF \qquad (\mathrm{A} = \mathrm{C/s})$$

$$= \int_F \rho \frac{d\mathbf{s} \cdot d\mathbf{F}}{dt}$$

$$= \rho \frac{d\mathbf{s}}{dt} \cdot \mathbf{F}$$

$$= \rho \frac{dV}{dt} = \frac{dq}{dt}, \qquad (5.2)$$

if

$$\frac{d\rho}{dt} = 0, \quad \text{and} \quad q = \rho V \qquad (5.3)$$

(the latter is, in fact, the integrated form of the equation of continuity to be discussed below). With the help of the previous equation we also have

$$I d\mathbf{s} = \rho \frac{dV}{dt} d\mathbf{s} = \mathbf{j} dV$$

$$= (\rho dV) \frac{d\mathbf{s}}{dt} = dq \frac{d\mathbf{s}}{dt}. \qquad (5.4)$$

---

*For an explanation see for instance also Jackson [3], p. 169.

In the following we also write $\dot{s} = ds/dt$. In addition we have — these manipulations serve the purpose of making the definition of a *surface current* plausible —

$$dq\frac{ds}{dt} = \frac{dq/dt}{2\pi r}2\pi rds = \frac{I(2\pi rds)}{2\pi r} = \frac{I}{2\pi r}dF_{\text{cyl}}, \tag{5.5}$$

where

$$\mathbf{K} \equiv \frac{I}{2\pi r} = \frac{dq}{dF_{\text{cyl}}}\dot{s} = \sigma\mathbf{v} \tag{5.6}$$

is a *surface current*, also called layer of current, and σ charge per unit area. Hence we can write

$$\mathbf{I}ds = \mathbf{j}dV = \mathbf{K}dF, \tag{5.7}$$

where $\mathbf{I}$ represents the *line or linear current*, $\mathbf{j}$ *the volume current* density and $\mathbf{K}$ the *surface current* density. We now construct the divergence:

$$\boldsymbol{\nabla} \cdot \mathbf{j} = \boldsymbol{\nabla} \cdot (\rho\dot{\mathbf{s}}) = \rho\boldsymbol{\nabla} \cdot \dot{\mathbf{s}} + (\boldsymbol{\nabla}\rho) \cdot \dot{\mathbf{s}}.$$

But

$$\boldsymbol{\nabla} \cdot \dot{\mathbf{s}}\bigg|_{\mathbf{s}\equiv\mathbf{r}} = \boldsymbol{\nabla} \cdot \dot{\mathbf{r}} = \frac{d}{dt}\boldsymbol{\nabla} \cdot \mathbf{r} = \frac{d}{dt}3 = 0.$$

This condition, the vanishing of the divergence of a velocity, is known as condition of incompressibility (as in hydrodynamics). We obtain with this

$$\boldsymbol{\nabla} \cdot \mathbf{j} = (\boldsymbol{\nabla}\rho) \cdot \dot{\mathbf{s}}.$$

In general one has

$$\frac{d\rho(\mathbf{s}, t)}{dt} = \frac{\partial\rho}{\partial t} + (\boldsymbol{\nabla}\rho) \cdot \dot{\mathbf{s}}$$

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

Since (see above) $d\rho/dt = 0$, it follows that

$$0 = \frac{\partial\rho}{\partial t} + \boldsymbol{\nabla} \cdot \mathbf{j}. \tag{5.9}$$

This important equation is known as *equation of continuity*. The equation plays an important role in many branches of physics, including quantum mechanics and statistical mechanics. The condition $d\rho/dt = 0$ implies, that in every time interval the number of charges in every unit volume remains constant. This again implies, since charges move in and out, *that no charges are created or annihilated*. The significance of the equation of continuity is

therefore that as much charge as flows into a volume element per unit time, also flows out of this volume element again per unit time. Thus the amount of charge in every volume element remains constant, or — as one says — is conserved. Assume now that in addition $\partial\rho/\partial t = 0$. This condition says, that the density does not change explicitly with time. This condition is described as the condition of *stationarity* or (specifically in statistical mechanics) of *equilibrium*. We thus have as condition for a *stationary* or *steady current*

$$\boldsymbol{\nabla} \cdot \mathbf{j} = 0. \tag{5.10}$$

Stationarity is attained only after some time interval, e.g. after a condenser has been charged. Thus $\partial\rho/\partial t = 0$ does not apply immediately at the time of charging or discharging of a condenser.

Example 5.1: The relaxation time of a dielectric

The homogeneous dielectric of a condenser has a weak conductivity σ. Calculate with the help of the equation of continuity and Ohm's law in the form $\mathbf{j} = \sigma\mathbf{E}$ (cf. Sec. 6.1) the relaxation time of the dielectric (dielectric constant ϵ), that is, the time interval in the course of a discharging of the condenser, in which the charge density drops to the fraction $1/e$ of its original value.

Solution: We have

$$\boldsymbol{\nabla} \cdot \mathbf{j} + \frac{\partial\rho}{\partial t} = 0, \quad \mathbf{j} = \sigma\mathbf{E},$$

from which we obtain

$$\sigma\boldsymbol{\nabla} \cdot \mathbf{E} + \frac{\partial\rho}{\partial t} = 0.$$

Since the dielectric is homogeneous, i.e. isotropic, we have $\mathbf{D} = \epsilon\mathbf{E}$, $\epsilon = $ const., and so the equation implies $\boldsymbol{\nabla} \cdot \mathbf{D} = \rho$, $\boldsymbol{\nabla} \cdot \mathbf{E} = \rho/\epsilon$. With this and the preceding equation we obtain

$$\frac{\partial\rho}{\partial t} = -\frac{\sigma\rho}{\epsilon},$$

i.e.

$$\rho = \rho_0 e^{-\sigma t/\epsilon},$$

where at time $t = 0$ we have $\rho = \rho_0$. The relaxation time is therefore ϵ/σ. This expression is large for small σ. We observe here incidentally that in the Gaussian system of units with $\epsilon_0 = 1$ the conductivity is given in seconds^{-1}.

In much the same way as one can produce a steady flow of water with the help of a pump, one can use a battery in a closed circuit to produce a steady or stationary current. The area enclosed by this circuit is of particular significance for its magnetic properties. We can imagine the addition of further conducting wires as illustrated in Fig. 5.2 until the entire area is filled with a net of these covering the entire area of the original circuit which then appears as the boundary of the net. The stationary current is flowing in this boundary whereas currents in the net segments inside all cancel each other. It is because the internal currents all cancel each other (no matter

what the size of an element of the net may be), that the net has the same magnetic properties as the original circuit.
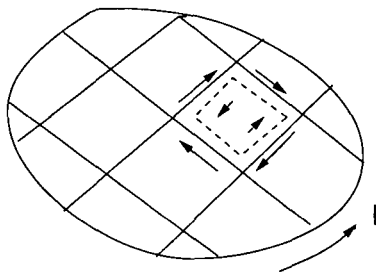
Fig. 5.2 The net of conducting wires.

5.3 Ampère's Experiments and the Law of Biot and Savart

In electrostatics we saw that Coulomb and Cavendish had discovered the law of force $\propto 1/r^2$ experimentally, but that the proportionality of the force to the charges was more or less guessed. In the case of magnetostatics Ampère discovered that the force acting between current carrying circuits which had been arranged in a particularly symmetric fashion, is similarly proportional to $1/r^2$, where r is the distance between the circuits, and he deduced from this that this behaviour is of a general kind.
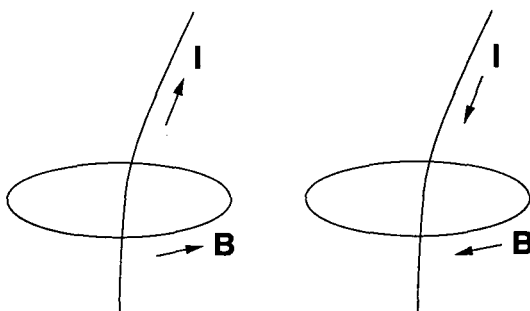
Fig. 5.3 The first experiment of Ampère.

Ampère was largely concerned with the study of *forces* due to current carrying wires. *In one experiment* he demonstrated that two wires close to each other and carrying equal but opposite currents do not exert a force on some

distant conducting wire. Today we know that the magnetic fields annul each other in this case, as illustrated in Fig. 5.3. In a *second experiment* Ampère demonstrated, as far as the effect on other current carrying conductors was concerned, that arrangements of conductors as in Fig. 5.4 have the same effect (i.e. the same field). In a *third experiment* Ampère showed, that the forces stemming from magnets or currents which act on a wire having the shape of a circular arc carrying the current I do not move this wire and hence must be vertical to this wire ($\mathbf{F} = \int I d\mathbf{s} \times \mathbf{B}(\mathbf{r})$). In this experiment Ampère used the subtle arrangement illustrated schematically in Fig. 5.5. The two mercury columns provided the necessary electric contact with the element of wire, which could move freely in its plane around a vertical.
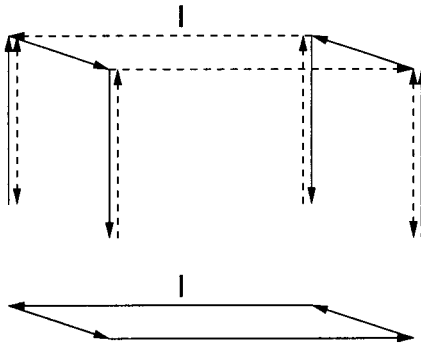
Fig. 5.4 The second experiment of Ampère.

Finally in a *fourth experiment* Ampère showed, that the force acting between current carrying loops is again proportional to $1/r^2$, r their separation.
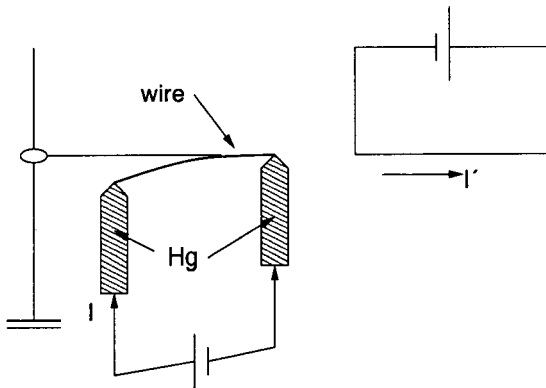
Fig. 5.5 The third experiment of Ampère.

To this end he constructed a symmetric arrangement of three circular coils carrying current I, whose radii were in the ratio $n^0(=1) : n^1 : n^2$. Let the distances of the first and the last coils from the second coil be a and b as indicated in Fig. 5.6. The central coil was then adjusted in such a way, that the forces due to the currents balance each other out. If, as stated, the radii of the coils are respectively $1, n, n^2$, then equilibrium (no motion) is obtained for (this was Ampère's experimental observation)

$$b = na.$$

If the forces are proportional to the circumferences or equi-angular arcs, then this means (with $f(x)$ for the dependence on the distance) for equilibrium at $b = na$:

$$2\pi \cdot 2\pi n \cdot f(a) = 2\pi n \cdot 2\pi n^2 \cdot f(b)\Big|_{b=na}$$

(ignoring the effects of the outer coils on each other), i.e.

$$f(a) = n^2 f(na),$$

i.e.

$$f(a) = \frac{\text{const.}}{a^2}$$

(so that $\text{const.}/a^2 = n^2 \text{const.}/n^2 a^2$). This is the $1/r^2$ law for currents.

The fundamental formula for the field outside a current carrying wire is today named after Biot and Savart, who had also discovered the law $\propto 1/r^2$ shortly before Ampère (1820).
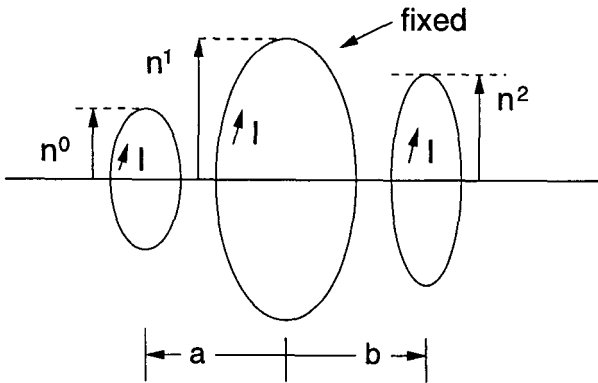
Fig. 5.6 The fourth experiment of Ampère.

The experiments of Ampère (1820–1825), however, were considerably more rigorous and were praised by Maxwell[†] as *"one of the most brilliant achievements in science"*.

The actual postulate or the law of magnetostatics derived from observation is *Ampère's law*, which (like that of the Coulomb force between charges in electrostatics) gives the force $\mathbf{F}$ between two current carrying string-like conductors as illustrated in Fig. 5.7, and can be written
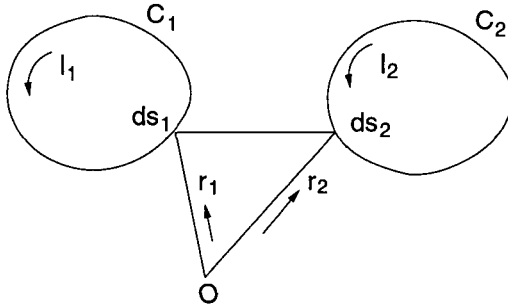
Fig. 5.7 Two conductor elements.

$$\mathbf{F} = \frac{\mu_0}{4\pi} \oint_{C_1} \oint_{C_2} \frac{I_1 d\mathbf{s}_1 \times \{I_2 d\mathbf{s}_2 \times (\mathbf{r}_1 - \mathbf{r}_2)\}}{|\mathbf{r}_1 - \mathbf{r}_2|^3}. \tag{5.11}$$

Using

$$d\mathbf{s}_1 \times \{d\mathbf{s}_2 \times (\mathbf{r}_1 - \mathbf{r}_2)\}$$

$$= \{d\mathbf{s}_1 \cdot (\mathbf{r}_1 - \mathbf{r}_2)\}d\mathbf{s}_2 - (d\mathbf{s}_1 \cdot d\mathbf{s}_2)(\mathbf{r}_1 - \mathbf{r}_2)$$

and Stokes' theorem, with which

$$\oint_{C_1} d\mathbf{s}_1 \cdot \frac{\mathbf{r}_1 - \mathbf{r}_2}{|\mathbf{r}_1 - \mathbf{r}_2|^3} = -\oint_{C_1} d\mathbf{s}_1 \cdot \boldsymbol{\nabla}_{\mathbf{r}_1} \frac{1}{|\mathbf{r}_1 - \mathbf{r}_2|}$$

$$= -\int_{F_{C_1}} d\mathbf{F} \cdot \boldsymbol{\nabla} \times \boldsymbol{\nabla} \frac{1}{|\mathbf{r}_1 - \mathbf{r}_2|},$$

$$= 0,$$

the force $\mathbf{F}$ is also given by

$$\mathbf{F} = -\frac{\mu_0}{4\pi} \oint_{C_1} \oint_{C_2} I_1 d\mathbf{s}_1 \cdot I_2 d\mathbf{s}_2 \frac{(\mathbf{r}_1 - \mathbf{r}_2)}{|\mathbf{r}_1 - \mathbf{r}_2|^3}. \tag{5.12}$$

[†]See e.g. Harnwell [14], p. 298.

The *law of Biot–Savart* thus abstracts the observations of Biot, Savart and Ampère by specifying an expression for the magnetic field **B**, called *magnetic induction* or *magnetic flux density*,[‡] but frequently also simply magnetic field, which a conductor with current I, or the conductor element of length ds — see Fig. 5.8 — generates at a point **r** outside the conductor, and is defined by

$$d\mathbf{B}(\mathbf{r}) := k' \frac{I(\mathbf{r}')d\mathbf{s}(\mathbf{r}') \times (\mathbf{r} - \mathbf{r}')}{|\mathbf{r}_1 - \mathbf{r}_2|^3}, \quad k' = \frac{\mu_0}{4\pi}, \tag{5.13}$$

or

$$\mathbf{B}(\mathbf{r}) = k' \int_{\text{conductor}} \frac{I(\mathbf{r}')d\mathbf{s}(\mathbf{r}') \times (\mathbf{r} - \mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|^3}$$
$$= k' \int_{\text{conductor}} d\mathbf{r}' \frac{\mathbf{j}(\mathbf{r}') \times (\mathbf{r} - \mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|^3}. \tag{5.14}$$

In MKSA-units the constant k' is given by

$$k' = \frac{\mu_0}{4\pi}, \quad \mu_0 = 4\pi \times 10^{-7} = 12.566 \times 10^{-7} \text{ N A}^{-2}.$$

The magnetic induction **B** is then given in units called tesla (T) (named after the engineer of this name); 1 tesla $= 10^4$ gauss. In Gaussian units the constant is $k' = 1/c$.
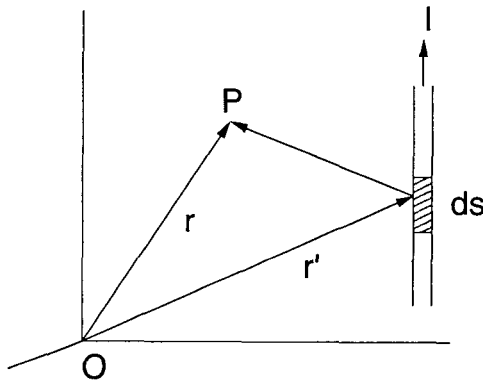
Fig. 5.8 The conductor element.

As mentioned above, Ampère was mainly concerned with the investigation of the force that one current exerts on another current carrying conductor. Consequently it became customary to describe as *Ampère's law* the

---

[‡]We begin here with the induction, although we have not yet considered media. Some authors therefore start with the magnetic field strength **H**, which we introduce later.

expression for the force, that a field **B** exerts on a conductor with current I. This means the expression obtained by inserting the above expression for $d\mathbf{B}$ into that for **F**, i.e.

$$\mathbf{F} = \int I(\mathbf{r}')ds(\mathbf{r}') \times \mathbf{B}(\mathbf{r}'). \tag{5.15}$$

In the MKSA-system of units the proportionality constant appearing in this expression is 1. This constant follows by demanding that force equals time derivative of momentum; indeed the logical derivation of this expression is to go to the relativistic theory, obtain there the four-momentum of a charge in the gauge field, and then to differentiate the latter's space-like components with respect to time. Some authors, e.g. Greiner [4], say simply, the constant k' *"is obtained experimentally as 1/c"* in the Gaussian system.[§]

With the relation $Ids = \mathbf{j}dV$ it follows that

$$\mathbf{F} = \int_{\text{conductor}} \mathbf{j}(\mathbf{r}') \times \mathbf{B}(\mathbf{r}')dr'. \tag{5.16}$$

For a point charge q moving with velocity **v**, we can write

$$\mathbf{j} = q\mathbf{v}\delta(\mathbf{r} - \mathbf{r}') \tag{5.17}$$

(a precisely known current density), so that

$$\mathbf{F} = q\mathbf{v}(\mathbf{r}) \times \mathbf{B}(\mathbf{r}). \tag{5.18}$$

This expression is known as *Lorentz force* (cf. Jackson [3], p. 238, where this expression is initially written down with little justification). A version of this equation which is of practical importance, and is also described as *Ampère's law*, is obtained by writing it, e.g. in the simple perpendicular case, as

$$dF = IBds, \quad I = dq/dt.$$

In the case of a coil with N turns in the field B, the torque (moment) of this force is then $dFdl = INBdA$, where dA is the area $dsdl$.

The expression for **B** should be compared with that for **E**, i.e. the formula

$$\mathbf{E}(\mathbf{r}) = k \int dr' \frac{\rho(\mathbf{r}')(\mathbf{r} - \mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|^3}, \quad k = \frac{1}{4\pi\epsilon_0}.$$

The field **B** follows for moving charges (current density **j**) in a similar way as the field **E** for static charges (charge density ρ).

[§]One way of determining experimentally the constant relating electric units to magnetic units (which turns out to be c) is discussed briefly in Appendix B in relation to the experiment of Rosa and Dorsey.

5.4   Examples

In the following we consider examples which are important in later contexts, and therefore do not simply serve the purpose of illustration or exercise.

Example 5.2: The field B of a long, straight conducting wire

Calculate the field **B** of an infinitely long, straight conductor carrying current I (see Fig. 5.9).

Solution: From the Biot–Savart law we obtain (with $k' = \mu_0/4\pi$):

$$\mathbf{B}(\mathbf{r}) = Ik' \int_{-\infty}^{\infty} \frac{d\mathbf{s} \times (\mathbf{r} - \mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|^3}.$$

With the geometry of Fig. 5.9 this can be written

$$B(\mathbf{r}) = Ik' \int_{-\infty}^{\infty} \frac{ds|\mathbf{r} - \mathbf{r}'|\sin\theta}{|\mathbf{r} - \mathbf{r}'|^3},$$
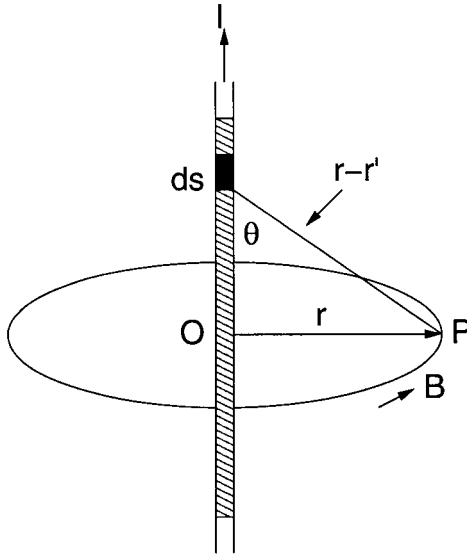
Fig. 5.9 The long, straight conducting wire.

so that

$$\begin{aligned}
B(\mathbf{r}) &= Ik' \int_{-\infty}^{\infty} \frac{r\,ds}{|\mathbf{r} - \mathbf{r}'|^3} = Irk' \int_{-\infty}^{\infty} \frac{ds}{(r^2 + s^2)^{3/2}} \\
&= \frac{Irk'}{r^2} \left[\frac{s}{\sqrt{r^2 + s^2}} \right]_{-\infty}^{\infty} = \frac{2Irk'}{r^2} \\
&= \frac{\mu_0}{2\pi} \frac{I}{r}.
\end{aligned} \tag{5.19}$$

Example 5.3: The force between two long, thin, parallel conducting wires

Calculate the force which two infinitely long, parallel conductors exert on each other.

Solution: We consider the case as depicted in Fig. 5.10. Conductor 2 generates the field $\mathbf{B}_2$ at conductor 1 which follows from Eq. (5.19):

$$B_2 = \frac{2I_2}{d} k',$$

where d is the distance between the conductors. The field $\mathbf{B}_2$ is perpendicular to the currents I_1, I_2. Hence the force, that $\mathbf{B}_2$ exerts on conductor 1, is (cf. Eq. (5.16))

$$F = k' \int I_1 ds(\mathbf{r}') \frac{2I_2}{d} = k' \frac{2I_1 I_2}{d} \int ds(\mathbf{r}'). \tag{5.20}$$
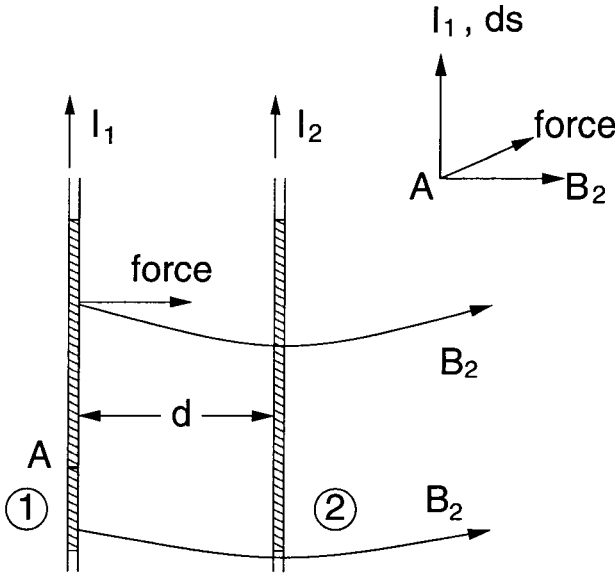
Fig. 5.10 Two long, thin conductors.

The force per unit length of the conductors is therefore:

$$k' \frac{2I_1 I_2}{d} = \mu_0 \frac{I_1 I_2}{2\pi d}.$$

For I_1, I_2 parallel we have the situation illustrated in Fig. 5.10.

Example 5.4: The circular or ring conductor

As another important example for the calculation of the field we consider the circular or ring-shaped conductor shown in Fig. 5.11 with current I and radius a.

Solution: The components of the field $\mathbf{B}$ perpendicular to the z-axis cancel each other along this axis. The field $d\mathbf{B}$ at $\mathbf{r}$ along the axis due to the element ds of the conductor is according to the Biot–Savart law given by:

$$d\mathbf{B}(\mathbf{r}) = Ik' \frac{ds(\mathbf{r}') \times (\mathbf{r} - \mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|^3}.$$

Along the z-axis $d\mathbf{B}$ is

$$dB(\mathbf{r})\cos\theta = \frac{dB(\mathbf{r})a}{|\mathbf{r} - \mathbf{r}'|}.$$

Hence the magnetic induction B at point P along the z-axis is:
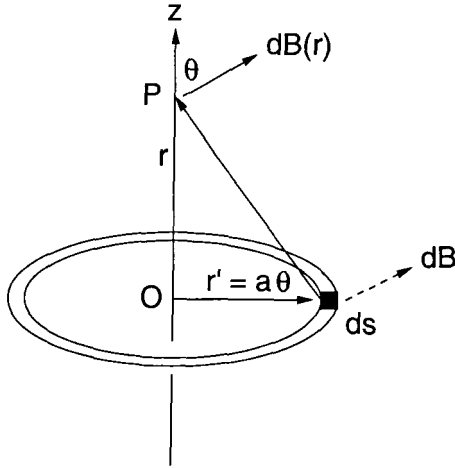
Fig. 5.11 The ring-shaped conductor.

$$
\begin{aligned}
B &= \int_{conductor} \frac{a\,dB(\mathbf{r})}{|\mathbf{r} - \mathbf{r}'|} \\
&= k' \int \frac{I|d\mathbf{s} \times (\mathbf{r} - \mathbf{r}')|}{|\mathbf{r} - \mathbf{r}'|^3} \frac{a}{|\mathbf{r} - \mathbf{r}'|} \\
&= Iak' \int \frac{a\,d\theta'}{|\mathbf{r} - \mathbf{r}'|^3} \quad \text{(since } d\mathbf{s} \perp \mathbf{r} - \mathbf{r}',\ ds = a\,d\theta') \\
&= Iak' \frac{2\pi a}{(r^2 + a^2)^{3/2}} \quad \text{(since } \mathbf{r}' \perp \mathbf{r}),
\end{aligned}
\tag{5.21}
$$

i.e. for $r \gg a$:

$$B \simeq 2I\pi a^2 \frac{k'}{r^3} = \frac{\mu_0}{2}\frac{Ia^2}{r^3}. \tag{5.22}$$

We will deal with this problem again later (see Example 5.7), because one expects the field B at a point $\mathbf{r}$ to be related to the solid angle which the ring-shaped conductor subtends at this point.

Example 5.5: The magnetic moment
The magnetic moment of a circuit is defined as 1/2 the volume integral of the moment of the current density, i.e. as

$$\mathbf{m} := \frac{1}{2} \int d\mathbf{r}[\mathbf{r} \times \mathbf{j}(\mathbf{r})] = \frac{1}{2} \int \mathbf{r} \times Id\mathbf{s}(\mathbf{r}) = I \int d\mathbf{F}, \tag{5.23}$$

where, see Fig. 5.12 (a), $d\mathbf{F} = \frac{1}{2}\mathbf{r} \times d\mathbf{s}$ is an element of area. Calculate the field.

Solution: In the special case of the ring-shaped conductor considered above we have

$$m = I\pi a^2,$$

a relation which is also known as *Ampère's dipole law.* By definition the current I has the direction of positive charge; hence **m** is opposite to the direction of negative charge. In terms of the expression for m we can rewrite B of Eq. (5.22) as

$$B = k' \frac{2m}{r^3}.$$

It is instructive to compare this result with the field of an electric dipole with moment **p** as illustrated in Fig. 5.12(b). The electric field at point P follows, as we saw, from

$$\mathbf{E} = -\boldsymbol{\nabla}\phi_{\text{dipole}}, \quad \phi_{\text{dipole}} = k\frac{\mathbf{p}\cdot\mathbf{r}}{r^3} \overset{\text{here}}{=} k\frac{pr}{r^3} = k\frac{p}{r^2},$$

so that

$$E = k\frac{2p}{r^3}.$$

(a)   **(b)**
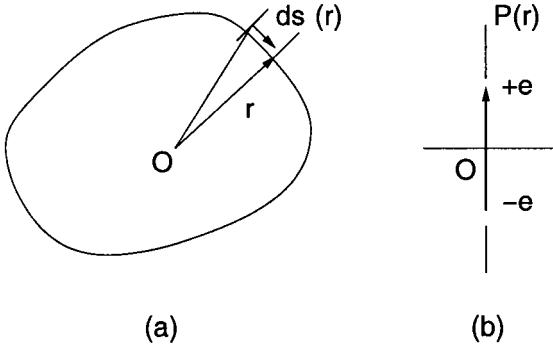
Fig. 5.12   (a) Circuit element, (b) dipole.
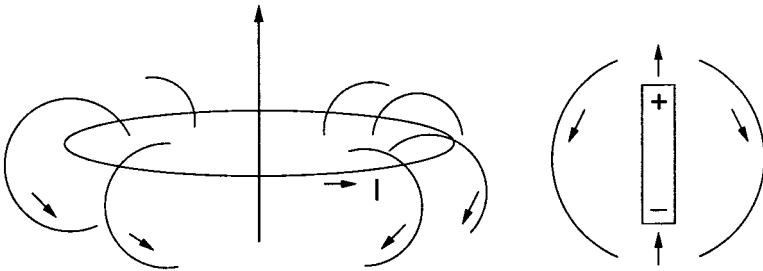
Fig. 5.13 Analogy between the fields of a circular current and a magnet.

The general expression for **E**, i.e. at a point P in the direction of the polar angle θ as seen from the origin, can immediately be written down (or see e.g. [4], p. 36)

$$\mathbf{E} = k\frac{2p\cos\theta}{r^3}\mathbf{e}_r + k\frac{p\sin\theta}{r^3}\mathbf{e}_\theta, \quad k = \frac{1}{4\pi\epsilon_0}, \tag{5.24}$$

in agreement with the above expression for polar angle $\theta = 0$. Taking into account that (see Chapter 1 or later)

$$k'/k = 1/c^2, \quad B = \frac{k}{c^2}\frac{2m}{r^3} = \frac{E}{p}\frac{m}{c^2},$$

we obtain the correspondence

$$B \leftrightarrow \frac{1}{c}E, \quad m \leftrightarrow cp.$$

This means, the magnetic moment **m** generates the field **B** in complete analogy to the generation of the field **E** by an electric dipole, and

$$\begin{aligned}
\mathbf{B}(\mathbf{r}) &= k'\frac{2m\cos\theta}{r^3}\mathbf{e}_r + k'\frac{m\sin\theta}{r^3}\mathbf{e}_\theta \\
&= k'\frac{3\mathbf{r}(\mathbf{m}\cdot\mathbf{r}) - r^2\mathbf{m}}{r^5} \\
&= -k'\boldsymbol{\nabla}\frac{\mathbf{m}\cdot\mathbf{r}}{r^3}.
\end{aligned} \tag{5.25}$$

We see therefore that a circular current generates an induction field **B**, which sufficiently far away looks like that of a dipole, as indicated in Fig. 5.13.

5.5 The Electromagnetic Vector Potential A

We return to the law of Biot–Savart. With

$$\frac{\mathbf{r} - \mathbf{r}'}{|\mathbf{r} - \mathbf{r}'|^3} = -\boldsymbol{\nabla}_\mathbf{r}\frac{1}{|\mathbf{r} - \mathbf{r}'|} = \boldsymbol{\nabla}'\frac{1}{|\mathbf{r} - \mathbf{r}'|}$$

we have

$$\begin{aligned}
\mathbf{B}(\mathbf{r}) &= -k'\int d\mathbf{r}'\mathbf{j}(\mathbf{r}') \times \boldsymbol{\nabla}_\mathbf{r}\frac{1}{|\mathbf{r} - \mathbf{r}'|} \\
&= \boldsymbol{\nabla}_\mathbf{r} \times k'\int\frac{\mathbf{j}(\mathbf{r}')d\mathbf{r}'}{|\mathbf{r} - \mathbf{r}'|}
\end{aligned} \tag{5.26}$$

(where $k' = \mu_0/4\pi$). But "*div curl*" is always zero. Hence

$$\boldsymbol{\nabla} \cdot \mathbf{B}(\mathbf{r}) = 0. \tag{5.27}$$

This equation has *no source term* on the right hand side. This means in physical terms: There are *no single magnetic poles*, i.e. *monopoles*. How did we arrive at this result: The empirically obtained law of Biot–Savart gives the field **B** which surrounds a conductor with current I; we saw that this field corresponds to that of a magnetic moment, a magnetic moment, however, corresponds to a pairing of magnetic poles — hence no single magnetic poles enter the consideration. It should be noted that we are (apparently) dealing with macroscopic considerations (the Biot–Savart law applies to macroscopic

conductors); these can, however, also be interpreted microscopically (the classically considered electron encircling a nucleus represents a current). The effect of the magnetic induction $\mathbf{B}$ on magnetisable matter will later be considered separately, in analogy with macroscopic electrostatics.

We now define the *electromagnetic vector potential* $\mathbf{A}$ by

$$\mathbf{A}(\mathbf{r}) = k' \int \frac{\mathbf{j}(\mathbf{r}')d\mathbf{r}'}{|\mathbf{r} - \mathbf{r}'|}, \tag{5.28}$$

so that

$$\mathbf{B} = \boldsymbol{\nabla} \times \mathbf{A} \tag{5.29}$$

($\mathbf{A}$ is not unique, since $\mathbf{A} + \boldsymbol{\nabla}\phi$, for ϕ a scalar function, is also possible; see gauge transformations below). We next evaluate the expression $\boldsymbol{\nabla} \times \mathbf{B}$, in which we use first the formula

$$\boldsymbol{\nabla} \times (\boldsymbol{\nabla} \times \mathbf{A}) = \boldsymbol{\nabla}(\boldsymbol{\nabla} \cdot \mathbf{A}) - \triangle\mathbf{A} \tag{5.30}$$

(*"curl curl = grad div − div grad "*), then

$$\triangle_{\mathbf{r}} \frac{1}{|\mathbf{r} - \mathbf{r}'|} = -4\pi\delta(\mathbf{r} - \mathbf{r}')$$

(and integrating over the delta function), and finally use Gauss' divergence theorem. We then have

$$\boldsymbol{\nabla} \times \mathbf{B} = \boldsymbol{\nabla} \times \left(\boldsymbol{\nabla}_{\mathbf{r}} \times k' \int \frac{\mathbf{j}(\mathbf{r}')d\mathbf{r}'}{|\mathbf{r} - \mathbf{r}'|} \right)$$

$$= k'\boldsymbol{\nabla} \int \boldsymbol{\nabla}_{\mathbf{r}} \cdot \frac{\mathbf{j}(\mathbf{r}')d\mathbf{r}'}{|\mathbf{r} - \mathbf{r}'|} - k' \int \triangle_{\mathbf{r}} \frac{\mathbf{j}(\mathbf{r}')d\mathbf{r}'}{|\mathbf{r} - \mathbf{r}'|}$$

$$= k'\boldsymbol{\nabla} \int \left(\boldsymbol{\nabla}_{\mathbf{r}} \cdot \frac{\mathbf{j}(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} \right) d\mathbf{r}' + 4\pi k'\mathbf{j}(\mathbf{r}).$$

Since $\boldsymbol{\nabla}_r$ acts only on the denominator this can be rewritten as

$$\boldsymbol{\nabla} \times \mathbf{B} = -k'\boldsymbol{\nabla} \int \mathbf{j}(\mathbf{r}') \cdot \left(\boldsymbol{\nabla}_{\mathbf{r}'} \frac{1}{|\mathbf{r} - \mathbf{r}'|} \right) d\mathbf{r}' + 4\pi k'\mathbf{j}(\mathbf{r})$$

$$= -k'\boldsymbol{\nabla} \int d\mathbf{r}' \left[\boldsymbol{\nabla}' \cdot \left(\frac{\mathbf{j}(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} \right) - \frac{1}{|\mathbf{r} - \mathbf{r}'|}\boldsymbol{\nabla}' \cdot \mathbf{j}(\mathbf{r}') \right] + 4\pi k'\mathbf{j}(\mathbf{r})$$

$$= -k'\boldsymbol{\nabla} \int_{F_\infty} d\mathbf{F} \cdot \frac{\mathbf{j}(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|}$$

$$+ k'\boldsymbol{\nabla} \int \frac{\boldsymbol{\nabla}' \cdot \mathbf{j}(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}' + 4\pi k'\mathbf{j}(\mathbf{r}). \tag{5.31}$$

Here the first term vanishes, if we extend the area of integration F to infinity and remember, that $\mathbf{j} = 0$ at $r = \infty$, since the conductor and hence $\mathbf{j}$ are located in a finite part of space. The second term also vanishes, since in magnetostatics $\nabla \cdot \mathbf{j} = 0$, which is the condition for stationary currents. Thus there remains *in magnetostatics*

$$\nabla \times \mathbf{B} = \mu_0 \mathbf{j}(\mathbf{r}). \tag{5.32}$$

As a *subsidiary result* we obtain:

$$
\begin{aligned}
\nabla \cdot \mathbf{A}(\mathbf{r}) &= k' \nabla_{\mathbf{r}} \cdot \int \left(\frac{\mathbf{j}(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} \right) d\mathbf{r}' \\
&= -k' \int \mathbf{j}(\mathbf{r}') \cdot \left(\nabla_{\mathbf{r}'} \frac{1}{|\mathbf{r} - \mathbf{r}'|} \right) d\mathbf{r}' \\
&= -k' \int d\mathbf{r}' \left[\nabla' \cdot \left(\frac{\mathbf{j}(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} \right) - \frac{1}{|\mathbf{r} - \mathbf{r}'|} \nabla' \cdot \mathbf{j}(\mathbf{r}') \right] \\
&= -k' \int_{F_\infty} d\mathbf{F} \cdot \frac{\mathbf{j}(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} + k' \int d\mathbf{r}' \frac{\nabla' \cdot \mathbf{j}(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} \\
&= 0
\end{aligned}
\tag{5.33}
$$

for reasons explained above. Thus

$$\nabla \cdot \mathbf{A}(\mathbf{r}) = 0. \tag{5.34}$$

This relation is known as *Coulomb gauge*. It can be shown,[¶] that this condition has the physical meaning that only two components of the vector potential **A** are independent, and that the vector **A** is orthogonal to the propagation vector **k** of the free electromagnetic wave, and hence implies the *transversality* of the wave (fields **E**, **H** perpendicular to the direction of propagation).

One should note: Here we arrived at the relation $\nabla \cdot \mathbf{A} = 0$ only as a consequence of $\nabla \cdot \mathbf{j} = 0$, i.e. in the case of stationary currents. A condition of this kind, however, is always required for the vector potential, as a consequence of the so-called *gauge invariance of the theory*. By this one means (put simply) the *invariance* of the field equations or equations of motion, here the Maxwell equations, under transformations (here for magnetostatics) called *gauge transformations*, i.e.

$$\mathbf{A} \to \mathbf{A}' = \mathbf{A} + \nabla \chi, \quad \phi \to \phi' = \phi,$$

[¶]See Chapter 18.

(more generally $\phi' = \phi - \partial\chi/\partial t$), where $\triangle\chi = 0$. The invariance of the equations obtained so far follows immediately:

$$\mathbf{E} \to \mathbf{E}' = -\boldsymbol{\nabla}\phi' = -\boldsymbol{\nabla}\phi = \mathbf{E},$$

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

$$\boldsymbol{\nabla} \cdot \mathbf{A} = 0 \to \boldsymbol{\nabla} \cdot \mathbf{A}' = \boldsymbol{\nabla} \cdot (\mathbf{A} + \boldsymbol{\nabla}\chi) = \boldsymbol{\nabla} \cdot \mathbf{A} = 0 \ \text{ (since } \triangle\chi = 0).$$

Again one should note: The invariant fields $\mathbf{E}$ and $\mathbf{B}$ are macroscopically observable quantities, the nonunique gauge potential, of course, is not.

5.6   Integral Form of Ampère's Law

We return to our earlier considerations. We know that

$$\triangle\frac{1}{|\mathbf{r} - \mathbf{r}'|} = -4\pi\delta(\mathbf{r} - \mathbf{r}').$$

Hence we have with Eq. (5.28):

$$\triangle\mathbf{A}(\mathbf{r}) = -4\pi k' \int \delta(\mathbf{r} - \mathbf{r}')\mathbf{j}(\mathbf{r}')d\mathbf{r}',$$

i.e.

$$\triangle\mathbf{A}(\mathbf{r}) = -\mu_0\mathbf{j}(\mathbf{r}). \tag{5.35}$$

Thus the *current density* $\mathbf{j}$ *is the source of the vector potential* $\mathbf{A}$. This equation should be compared with the Poisson equation $\triangle\phi = -4\pi k\rho = \rho/\epsilon_0$ in electrostatics, in which the charge density ρ appears as the source function of the potential ϕ.
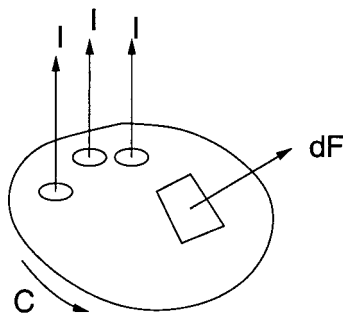
Fig. 5.14   Currents through a surface.

We now take a closer look at Eq. (5.32). This equation is also described as *Ampère's law*; in fact it is the differential form of the Biot–Savart law.

(Since for a gauge transformation $\mathbf{\nabla} \times \mathbf{B} = \mathbf{\nabla} \times \mathbf{B}'$, the invariance of this equation requires that $\mathbf{j} = \mathbf{j}'$). We apply to Eq.(5.32) the *theorem of Stokes* and obtain

$$\int \mathbf{\nabla} \times \mathbf{B} \cdot d\mathbf{F} = \oint_C \mathbf{B} \cdot d\mathbf{s} = \mu_0 \int_{F(C)} \mathbf{j} \cdot d\mathbf{F}, \qquad (5.36)$$

where $d\mathbf{s}$ is an element of the path C enclosing the area $F(C)$. Let us assume that n conductors are enclosed by C as indicated in Fig. 5.14 (or one with several turns or wiggles passing through the area enclosed by C, some also in opposite directions). Let I be the current flowing n times through the surface enclosed by C in one direction (or the net number, if currents are also going in the opposite direction). Then

$$\mu_0 \int_{F(C)} \mathbf{j} \cdot d\mathbf{F} = \mu_0 n I,$$

and hence

$$\oint_C \mathbf{B} \cdot d\mathbf{s} = \mu_0 n I. \qquad (5.37)$$

This important formula is the *integral form of Ampère's law*. In view of Eqs. (5.32) and (5.35) we have

$$\mathbf{\nabla} \times \mathbf{B} + \triangle \mathbf{A} = 0.$$

But

$$\mathbf{\nabla} \times (\mathbf{\nabla} \times \mathbf{A}) = \mathbf{\nabla}(\mathbf{\nabla} \cdot \mathbf{A}) - \triangle \mathbf{A} = 0 - \triangle \mathbf{A}$$

and always $\mathbf{\nabla} \times \mathbf{\nabla} = 0$. Hence

$$\mathbf{B} = \mathbf{\nabla} \times \mathbf{A} + \mathbf{\nabla}\phi_0.$$

Since, as we saw, $\mathbf{\nabla} \cdot \mathbf{B} = 0$ and always "*div curl* $= 0$", it follows that ϕ_0 is solution of $\triangle \phi_0 = 0$. If $\mathbf{A} = 0$, then it follows that $\mathbf{B}$ can always be expressed as a gradient (the so-called Coulomb law of magnetostatics for hypothetical monopoles).

5.7 Further Examples

Example 5.6: The magnetic field of a long solenoid with current I

Since the solenoid is stated to be *long*, the magnetic field in its interior is practically homogeneous and outside practically zero, as we can see immediately by drawing lines of constant $\mathbf{B}$ into the diagram as in Fig. 5.15. One observes that the lines crowd together in the inside region and thin out outside. The field inside is to be calculated.

Solution: In Fig. 5.15 a path C is shown, which encloses N turns of the solenoid. Hence

$$\oint_C \mathbf{B} \cdot d\mathbf{s} = IN\mu_0,$$

where IN is the current flowing through the surface
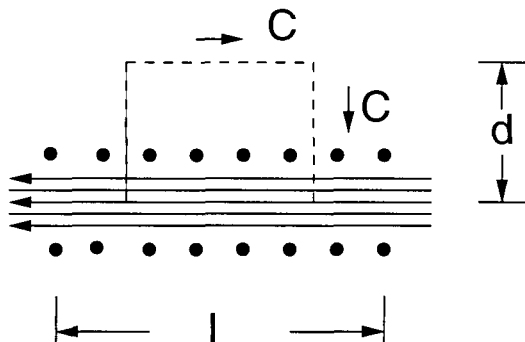
Fig. 5.15 Longitudinal cross section of the solenoid.

enclosed by C. Since the field $\mathbf{B}$ outside the solenoid is practically zero, and $\mathbf{B}$ along the two transverse sections of length d in Fig. 5.15 is also practically zero, and inside the solenoid the field is practically homogeneous, it follows that

$$Bl = IN\mu_0, \quad \text{i.e.} \quad B = In\mu_0, \tag{5.38}$$

where $n = N/l$ is the number of turns per unit length of the solenoid.
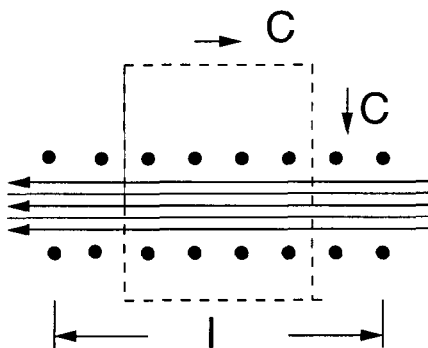
Fig. 5.16 A different closed path C.

For a better understanding we choose another closed path C as indicated in Fig. 5.16. In this case the field along the path is everywhere zero, i.e. $\oint_C \mathbf{B} \cdot d\mathbf{s} = 0$. On the other hand (since the path this time encloses ingoing as well as outgoing currents)

$$\int_C \mathbf{j} \cdot d\mathbf{F} = 2I - 2I = 0.$$

Thus here the path was selected in such a way that one is unable to determine the field B.

Example 5.7: Significance of solid angle subtended by a current loop

(a) Show that the field $\mathbf{B}(\mathbf{r})$ at point $\mathbf{r}$ due to a closed current carrying loop (current I) is given by

$$\mathbf{B}(\mathbf{r}) = -\frac{\mu_0}{4\pi}\boldsymbol{\nabla}(I\Omega),$$

where Ω is the solid angle subtended by the current loop at the point $\mathbf{r}$.

(b) With the help of the expression obtained under (a) for the field $\mathbf{B}(\mathbf{r})$ calculate the field $B(z)$ of a ring-shaped conductor of radius a at a point $P(z)$, distance z above the centre O of the ring-shaped conductor.

(c) A solenoid of radius a, length l, with current I, consists of n uniform turns per unit length. Determine the component of the field $\mathbf{B}$ in axial direction at a point a distance x away from one end of the solenoid (e.g. from the origin O there on the left side of the axis).

Solution:

(a) We consider, as indicated in Fig. 5.17(a), a displacement of the point P by $d\mathbf{r}$, which is equivalent to a displacement $-d\mathbf{r}'$ of the element of area $d\mathbf{F}' = d\mathbf{r}' \times d\mathbf{s}(\mathbf{r}')$, i.e. $d\mathbf{r} = -d\mathbf{r}'$. $(d(\mathbf{r} - \mathbf{r}') = -d(\mathbf{r}' - \mathbf{r}))$.
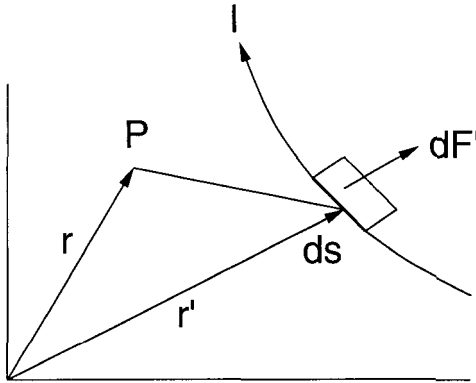
Fig. 5.17 (a) The element of area.

Starting from the formula

$$\mathbf{B}(\mathbf{r}) = \frac{\mu_0 I}{4\pi}\oint \frac{d\mathbf{s}(\mathbf{r}') \times (\mathbf{r} - \mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|^3},$$

we then have

$$\mathbf{B}(\mathbf{r}) \cdot d\mathbf{r} = -\frac{\mu_0 I}{4\pi}\oint d\mathbf{r}' \cdot \frac{d\mathbf{s}(\mathbf{r}') \times (\mathbf{r} - \mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|^3} = \frac{\mu_0 I}{4\pi}\oint \frac{d\mathbf{r}' \times d\mathbf{s}(\mathbf{r}') \cdot (\mathbf{r} - \mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|^3}$$

$$= \frac{\mu_0 I}{4\pi}\oint \frac{d\mathbf{F}' \cdot (\mathbf{r} - \mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|^3}$$

$$\equiv -k'I d\omega', \quad k' = \frac{\mu_0}{4\pi},$$

where $d\omega'$ is the solid angle subtended at P by the entire ring-shaped area, that arises in the shift of the loop. Thus

$$\mathbf{B}(\mathbf{r}) \cdot d\mathbf{r} = -k'I d\omega' = -k'd(I\Omega),$$

where Ω is the change of the solid angle of the conductor loop at P. The solution of the last equation is therefore

$$\mathbf{B}(\mathbf{r}) = -k'\boldsymbol{\nabla}(I\Omega). \tag{5.39}$$

(b) We now consider a circular conductor of radius a. The problem is to determine the solid angle subtended by the circular conductor at the point P in Fig. 5.17(b). We determine this by calculating the area F_S of the spherical shell (centre P, radius $\sqrt{a^2 + z^2}$), which has the conductor as boundary. We obtain this area from the corresponding part of the area of the corresponding cylindrical envelope dashed in Fig. 5.17(b). It can be shown that these two areas are equal,[||]
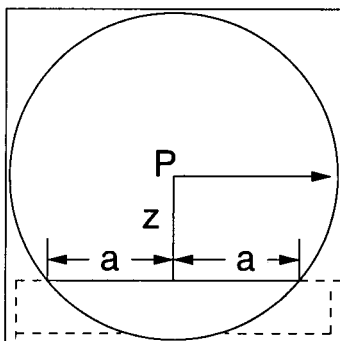
Fig. 5.17 (b) The spherical shell and its cylindrical envelope (dashed).

i.e.

$$F_S = (2\pi \times \text{cylindrical radius}) \times \text{cylinder height},$$

implies

$$F_S = 2\pi\sqrt{z^2 + a^2}(\sqrt{z^2 + a^2} - z).$$

It follows that

$$\Omega = \frac{\text{area}}{(\text{radius})^2} = \frac{F_S}{(\sqrt{z^2 + a^2})^2} = 2\pi\left(1 - \frac{z}{\sqrt{z^2 + a^2}}\right).$$

Hence

$$B_z = -k'I\frac{\partial}{\partial z}\Omega = k'I2\pi\left\{\frac{1}{\sqrt{z^2 + a^2}} - \frac{z^2}{(z^2 + a^2)^{3/2}}\right\}$$

$$= \frac{2\pi k'Ia^2}{(z^2 + a^2)^{3/2}}$$

in agreement with Eq. (5.21).

(c) Next we consider a cylindrical solenoid of radius a and length l as indicated in Fig. 5.17(c). The problem is to determine B in axial direction at point P, distance x from O. First we consider only one turn at y and so ndy turns in the immediate neighbourhood. Hence we calculate the solid angle subtended by the turns at y at the point P. This solid angle is, again calculated as above

---

[||]To this end one considers the ring-shaped area of the spherical shell of radius a (in Fig. 5.17(b) this radius is $\sqrt{a^2 + z^2}$) with ring radius $2\pi a \sin\theta$ and ring width ds. This radius multiplied by the arc length ds gives the area $2\pi a \sin\theta \cdot ds = 2\pi adx$; the latter expression is the area of an element of the enveloping cylinder. This relation can also be found in the literature, e.g. in K.E. Bullen, *An Introduction to the Theory of Mechanics* (Science Press, Sydney, 1951), p. 171.

with the help of the enveloping cylinder (observe $\sqrt{a^2 + (x-y)^2}$ is the radius of the sphere which is equal to the radius of the enveloping cylinder):

$$\Omega = \frac{\text{area}}{(\text{radius})^2} = \frac{2\pi\sqrt{a^2 + (x-y)^2}\{\sqrt{a^2 + (x-y)^2} - (x-y)\}}{(\sqrt{a^2 + (x-y)^2})^2}$$

$$= 2\pi\left\{1 - \frac{x-y}{\sqrt{a^2 + (x-y)^2}}\right\}.$$
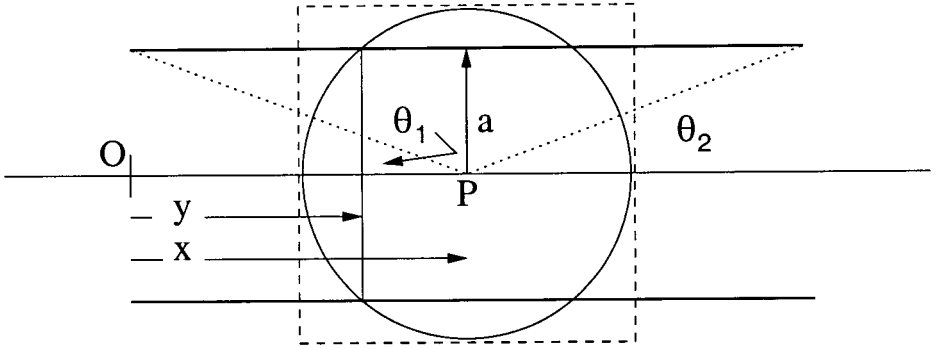
Fig. 5.17 (c) The solenoid.

Hence B_x:

$$B_x^{(1)} = -k'I\frac{\partial}{\partial x}\left\{2\pi\left(1 - \frac{x-y}{\sqrt{a^2 + (x-y)^2}}\right)\right\}.$$

We have to sum over the field contributions due to all turns. The interval dy at y contains ndy turns, so that its field contribution is

$$dB_x = -k'I\frac{\partial}{\partial x}\left\{2\pi\left(1 + \frac{y-x}{\sqrt{a^2 + (x-y)^2}}\right)\right\}ndy,$$

and altogether

$$B_x = -k'I\frac{\partial}{\partial x}2\pi n\int_0^l dy\left(1 + \frac{y-x}{\sqrt{a^2 + (y-x)^2}}\right)$$

$$= -k'I2\pi n\frac{\partial}{\partial x}\left\{l + \left[\sqrt{a^2 + (y-x)^2}\right]_0^l\right\}$$

$$= -k'I2\pi n\frac{\partial}{\partial x}\{l + \sqrt{a^2 + (l-x)^2} - \sqrt{a^2 + x^2}\}.$$

$$= k'In2\pi\left[\frac{l-x}{\sqrt{a^2 + (l-x)^2}} + \frac{x}{\sqrt{a^2 + x^2}}\right].$$

Hence

$$B_x = k'In2\pi\left[\cos\theta_2 + \cos\theta_1\right]$$

$$\overset{l\to\infty}{\longrightarrow} 4\pi k'In = \mu_0 In$$

in agreement with Eq. (5.38).

Example 5.8: The quadratic current loop

A current I flows in a planar, quadratic loop of wire with sides of length l. Determine the field $\mathbf{B}$ at the centre of the square.

Solution: We consider first the field at a point P, originating from one side of the quadratic loop. According to the Biot–Savart law we have, see Fig. 5.18 (a),

$$\mathbf{dB} = \frac{\mu_0}{4\pi} \frac{I ds(\mathbf{r}') \times (\mathbf{r} - \mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|^3},$$

i.e. (see Fig. 5.18)

$$dB = \frac{\mu_0}{4\pi} \frac{I dx \sin \alpha PQ}{PQ^3}.$$

We have

$$PQ = \frac{r}{\sin \alpha}, \quad \tan \alpha = -\frac{r}{x}, \quad -x = r \cot \alpha, \quad dx = \frac{r d\alpha}{\sin^2 \alpha},$$

so that

$$dB = \frac{\mu_0}{4\pi} \frac{I r d\alpha (\sin \alpha)^3}{(\sin \alpha)^2 r^2} = \frac{\mu_0}{4\pi} \frac{I}{r} \sin \alpha d\alpha,$$

and hence

$$B = \frac{\mu_0}{4\pi} \frac{I}{r} (\cos \alpha_1 - \cos \alpha_2).$$
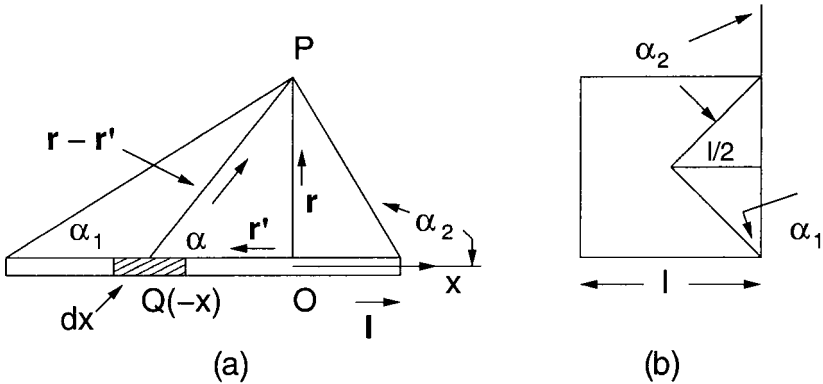
Fig. 5.18 (a) The field point P, and (b) the quadratic loop.

The requested field at P is therefore

$$B = 4 \frac{\mu_0}{4\pi} \frac{I}{l/2} (\cos 45° - \cos 135°) = \frac{\mu_0}{4\pi} \frac{4I \cdot 2}{l} \frac{2}{\sqrt{2}} = \frac{\mu_0}{4\pi} \frac{8\sqrt{2}I}{l}.$$

Example 5.9: An electron in a vector potential

An electron at rest (initial velocity zero!) is exposed to electromagnetic radiation given by

$$\mathbf{A} = (0, A(x - ct), 0).$$

Formulate the (nonrelativistic) equation of motion of the electron and derive, one after the other, the z-, y-, x-components $(\dot{z}, \dot{y}, \dot{x})$ of the velocity of the electron. Assume that $A(u)$ tends to zero for $u \to \infty$.

Solution: The force acting on the electron is the Lorentz force. Hence the equation of motion is

$$\dot{\mathbf{p}} = e(\mathbf{E} + \mathbf{v} \times \mathbf{B}), \tag{5.40}$$

where $\mathbf{v} = (\dot{x}, \dot{y}, \dot{z})$ (with the electron initially at the origin). Here $\mathbf{E} = -\partial \mathbf{A}/\partial t$ (the contribution $-\nabla\phi$ drops out with the given condition), and $\mathbf{B} = \nabla \times \mathbf{A}$. Hence in our case

$$\mathbf{E} = -\left(0, \frac{\partial A(x - ct)}{\partial t}, 0\right) = -(0, -cA'(x - ct), 0), \tag{5.41}$$

where $A' = \partial A(x - ct)/\partial x$. Moreover,

$$\mathbf{B} = \nabla \times \mathbf{A} = \begin{vmatrix} \mathbf{e}_x & \mathbf{e}_y & \mathbf{e}_z \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ 0 & A(x - ct) & 0 \end{vmatrix} = (0, 0, A'),$$

i.e. $\mathbf{B} = (0, 0, A')$. Hence

$$\mathbf{v} \times \mathbf{B} = (\dot{x}, \dot{y}, \dot{z}) \times (0, 0, A') = (\dot{y}A', -\dot{x}A', 0). \tag{5.42}$$

The expressions (5.41) and (5.42) now have to be inserted into (5.40) and yield

$$\dot{\mathbf{p}} = e(\dot{y}A', (c - \dot{x})A', 0). \tag{5.43}$$

Thus $p_z = 0$, i.e. $\dot{z} = 0$. For the y-component we have

$$m\ddot{y} = e(c - \dot{x})A', \quad c - \dot{x} = -\frac{d}{dt}(x - ct), \quad \therefore (c - \dot{x})A' = -A'\frac{d}{dt}(x - ct) = -\dot{A},$$

$$\therefore m\ddot{y} = -e\dot{A}, \quad \therefore \dot{y} = -\frac{e}{m}A + \text{const.}, \quad \text{const.} = 0 \quad \text{with} \quad \dot{y} = 0, A = 0 \quad \text{at} \quad t = 0.$$

Finally the x-component is given by

$$m\ddot{x} = e\dot{y}A', \quad \therefore \ddot{x} = -\frac{e^2}{m^2}AA' = -\frac{e^2}{m^2}\frac{1}{2(\dot{x} - c)}\frac{d}{dt}A^2.$$

Thus

$$2\ddot{x}(\dot{x} - c) = -\frac{e^2}{m^2}\frac{d}{dt}A^2,$$

so that

$$\dot{x}^2 - 2c\dot{x} = -\frac{e^2}{m^2}A^2 + \underbrace{\text{const.}}_{0}, \quad \dot{x} = c \pm \sqrt{c^2 - \frac{e^2}{m^2}A^2},$$

i.e.

$$\dot{x} = c - \sqrt{c^2 - \frac{e^2}{m^2}A^2} \simeq \frac{1}{2c}\left(\frac{eA}{m}\right)^2.$$

This page intentionally left blank

Chapter 6

Macroscopic Magnetostatics

6.1 Introductory Remarks

Up to this point we have only considered currents I or precisely known current densities $\mathbf{j}$ which describe a stationary microscopic flow of charge. In macroscopic problems the total current density is not known precisely. The atomic or molecular ring-like currents (and spins) of the electrons in matter determine the latter's magnetic properties, whose effective magnetic moments contribute to the vector potential $\mathbf{A}(\mathbf{r})$, as also the conduction electrons of the macroscopic transport of charge. For macroscopic effects again only an averaging over a macroscopic volume is meaningful. We can keep the microscopic current density $\mathbf{j}$ considered thus far as the density of macroscopic charge transport; this has to be supplemented, however, by the contributions of microscopic molecular currents. Hence we write

$$\text{total current density} = \mathbf{j} + \mathbf{j}_{\text{mol}}.$$

We are interested here in *currents in conductors*. We defined the ideal conductor electrostatically as an *equipotential domain* in which the electrons can move about without doing work. If a *potential difference* V is applied, and hence a field $\mathbf{E}$,

$$\phi \equiv V = \int \mathbf{E} \cdot d\mathbf{s},$$

e.g. $V = V_1 - V_2 = EL$ (L = length of the conductor), then the conduction electrons move in the direction determined by $\mathbf{E}$. One thus has a current I and a current density $\mathbf{j}$. However, in general these conduction electrons can not move about completely freely. The nuclei of the atoms of the conductor occupy points in the lattice of the rigid body structure of the metal, and this lattice interacts with the conduction electrons, although relatively weakly.

Something similar applies in the case of an externally applied magnetic field. The *resistance of the lattice* is evident in *Ohm's law*, i.e. in the relation

$$\mathbf{j} = \sigma \mathbf{E}, \tag{6.1}$$

where $\sigma = 1/\rho$ is the conductivity (gases, for instance, are not linear conductors, i.e. in their case Eq. (6.1) does not apply, instead $\mathbf{j} = f(\mathbf{E})$). Hence, with F the appropriate cross sectional area ($I = \int \mathbf{j} \cdot d\mathbf{F}$),

$$I = Fj = F\sigma E,$$

and so (see above)

$$\frac{V}{I} = \frac{EL}{F\sigma E} = \frac{L}{F\sigma} \equiv \rho \frac{L}{F} \equiv R.$$

Here ρ is the *resistivity*, also called specific resistance, and R the resistance, and we see that this depends on the geometry of the conductor. Here we usually avoid the use of ρ for resistivity since this can be confused with ρ for charge density. Thus the charge density appears in the product $\mathbf{E} \cdot \mathbf{j} = \rho \mathbf{E} \cdot d\mathbf{s}/dt$ which (since it represents work per volume and time) is the energy transferred per unit volume and per unit time from the conduction electrons to the lattice.

6.2   Macroscopic Magnetisation

We return to our original considerations. In keeping with the subdivision of the entire current density we write the total vector potential correspondingly

$$\mathbf{A} = \mathbf{A}_{\mathrm{macr}} + \mathbf{A}_{\mathrm{mol}} \tag{6.2}$$

or

$$\mathbf{A} = k' \int dr' \frac{\mathbf{j}(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} + k' \int dr' \frac{\mathbf{j}_{\mathrm{mol}}(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|}, \tag{6.3}$$

where $\mathbf{j}$ is the conduction current density, $\mathbf{j}_{\mathrm{mol}}$ the contribution of the molecular ring currents and $k' = \mu_0/4\pi$. On the other hand, we saw in the preceding chapter that the field $\mathbf{B}$ of a magnetic dipole moment $\mathbf{m}$ is given by

$$\mathbf{B} = k' \frac{3\mathbf{n}(\mathbf{n} \cdot \mathbf{m}) - \mathbf{m}}{r^3} \quad \text{with} \quad \mathbf{n} = \frac{\mathbf{r}}{r} \quad \text{and} \quad \mathbf{B} = \nabla \times \mathbf{A}.$$

We verify first that the vector potential of this expression for $\mathbf{B}$ is given by

$$\mathbf{A} = k' \frac{\mathbf{m} \times \mathbf{r}}{r^3}, \tag{6.4}$$

and our intention is to express A_{mol} in this form.

Example 6.1: Vector potential A of a magnetic dipole moment m
Verify Eq. (6.4).

Solution: We use the formula

$$\nabla \times (\mathbf{a} \times \mathbf{b}) = \mathbf{a}(\nabla \cdot \mathbf{b}) - \mathbf{b}(\nabla \cdot \mathbf{a}) + (\mathbf{b} \cdot \nabla)\mathbf{a} - (\mathbf{a} \cdot \nabla)\mathbf{b},$$

which implies

$$\nabla \times \left(\frac{\mathbf{m} \times \mathbf{r}}{r^3}\right) = \nabla \times \left(\mathbf{m} \times \frac{\mathbf{r}}{r^3}\right)$$

$$= \mathbf{m}\left(\nabla \cdot \frac{\mathbf{r}}{r^3}\right) - \frac{\mathbf{r}}{r^3}(\nabla \cdot \mathbf{m}) + \left(\frac{\mathbf{r}}{r^3} \cdot \nabla\right)\mathbf{m} - (\mathbf{m} \cdot \nabla)\frac{\mathbf{r}}{r^3}$$

$$= \mathbf{m}\left(\nabla \cdot \frac{\mathbf{r}}{r^3}\right) - (\mathbf{m} \cdot \nabla)\frac{\mathbf{r}}{r^3},$$

since $\mathbf{m} = $ const. Moreover, since,

$$\frac{\partial r}{\partial x} = \frac{\partial \sqrt{x^2 + y^2 + z^2}}{\partial x} = \frac{x}{r}, \quad \text{or} \quad \nabla r = \frac{\mathbf{r}}{r},$$

and

$$\frac{\partial}{\partial x}\left(\frac{x}{r^3}\right) = \frac{1}{r^3} - \frac{3x}{r^4}\frac{\partial r}{\partial x} = \frac{1}{r^3} - \frac{3x}{r^4}\frac{x}{r},$$

so that

$$\left(\nabla \cdot \frac{\mathbf{r}}{r^3}\right) = \frac{3}{r^3} - \frac{3}{r^5}r^2 = 0,$$

it follows that

$$\nabla \times \left(\frac{\mathbf{m} \times \mathbf{r}}{r^3}\right) = -(\mathbf{m} \cdot \nabla)\frac{\mathbf{r}}{r^3}.$$

However

$$(\mathbf{m} \cdot \nabla)\frac{\mathbf{r}}{r^3} = m_x \frac{\partial}{\partial x}\left(\frac{\mathbf{r}}{r^3}\right) + \cdots = m_x\left(\frac{\mathbf{e}_x}{r^3} - \frac{3}{r^4}\mathbf{r}\frac{x}{r}\right) + \cdots$$

$$= \frac{\mathbf{m}r^2}{r^5} - \frac{3\mathbf{r}(\mathbf{m} \cdot \mathbf{r})}{r^5}.$$

It follows that

$$\nabla \times \left(\frac{\mathbf{m} \times \mathbf{r}}{r^3}\right)k' = k'\frac{3\mathbf{r}(\mathbf{m} \cdot \mathbf{r}) - \mathbf{m}r^2}{r^5}$$

in agreement with the above claim.

Thus we can write for the vector potential at the point $\mathbf{r}$, which is due to molecule "i" ($\mathbf{m}_i$ the latter's total magnetic moment)

$$\mathbf{A}_{\text{mol},i}(\mathbf{r}) = k'\frac{\mathbf{m}_i \times (\mathbf{r} - \mathbf{r}_i)}{|\mathbf{r} - \mathbf{r}_i|^3}.$$

Hence

$$\mathbf{A}(\mathbf{r}) = k'\int d\mathbf{r}'\frac{\mathbf{j}(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} + k'\sum_i \frac{\mathbf{m}_i \times (\mathbf{r} - \mathbf{r}_i)}{|\mathbf{r} - \mathbf{r}_i|^3}. \tag{6.5}$$

We now define as

$N(\mathbf{r})$: the *average number of molecules per unit volume* at $\mathbf{r}$,

$\langle \mathbf{m}(\mathbf{r}) \rangle$: the *average molecular magnetic moment* at $\mathbf{r}$,

$\mathbf{M}{:=}N\langle \mathbf{m} \rangle$: the *macroscopic magnetisation*, i.e. *magnetic dipole density*.

Hence we can write

$$\Sigma_i(\text{contribution})_i = \int (\text{density of contributions})dr',$$

and thus Eq. (6.5)

$$\mathbf{A}(\mathbf{r}) = k' \int d\mathbf{r}' \frac{\mathbf{j}(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} + k' \int d\mathbf{r}' \frac{\mathbf{M}(\mathbf{r}') \times (\mathbf{r} - \mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|^3}$$

$$= k' \int d\mathbf{r}' \frac{\mathbf{j}(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} + k' \int d\mathbf{r}' \mathbf{M}(\mathbf{r}') \times \nabla_{\mathbf{r}'} \frac{1}{|\mathbf{r} - \mathbf{r}'|}. \quad (6.6)$$

For the curl of the product of a scalar function ϕ multiplied by a vector function $\mathbf{M}$ the following formula can be found in Appendix C or in Tables of Formulae:

$$\nabla \times (\phi \mathbf{M}) = (\nabla \phi) \times \mathbf{M} + \phi \nabla \times \mathbf{M}, \quad (6.7)$$

or for the first term on the right side of this equation

$$\mathbf{M} \times (\nabla \phi) = -\nabla \times (\phi \mathbf{M}) + \phi \nabla \times \mathbf{M}.$$

With $\phi = 1/|\mathbf{r} - \mathbf{r}'|$ we obtain for the second term on the right side of Eq. (6.6):

$$\int d\mathbf{r}' \mathbf{M}(\mathbf{r}') \times \nabla_{\mathbf{r}'} \frac{1}{|\mathbf{r} - \mathbf{r}'|}$$

$$= -\int d\mathbf{r}' \nabla' \times \left(\frac{\mathbf{M}}{|\mathbf{r} - \mathbf{r}'|} \right) + \int d\mathbf{r}' \frac{1}{|\mathbf{r} - \mathbf{r}'|} \nabla' \times \mathbf{M}. \quad (6.8)$$

But, as we shall verify in a moment,

$$\int d\mathbf{r} \nabla \times \mathbf{B} = \int d\mathbf{F} \times \mathbf{B}. \quad (6.9)$$

Example 6.2: Verification of Eq. (6.9)

Verify Eq. (6.9) by scalar multiplication by a constant vector $\mathbf{C}$.

Solution: For the verification we multiply the right hand side by an arbitrary vector $\mathbf{C} = \text{const.}$ and consider

$$\mathbf{C} \cdot \int d\mathbf{F} \times \mathbf{B} \equiv \int \mathbf{C} \cdot (\mathbf{n} \times \mathbf{B})dF.$$

For the scalar triple product we have

$$(\mathbf{C} \times \mathbf{n}) \cdot \mathbf{B} = \mathbf{B} \cdot (\mathbf{C} \times \mathbf{n}) = (\mathbf{B} \times \mathbf{C}) \cdot \mathbf{n}.$$

Hence

$$\mathbf{C} \cdot \int d\mathbf{F} \times \mathbf{B} = \int (\mathbf{B} \times \mathbf{C}) \cdot d\mathbf{F}$$

$$\overset{\text{Gauss}}{=} \int d\mathbf{r} \boldsymbol{\nabla} \cdot (\mathbf{B} \times \mathbf{C}) \overset{\text{see below}}{=} \int d\mathbf{r} \mathbf{C} \cdot (\boldsymbol{\nabla} \times \mathbf{B}),$$

where in the last step we used for $\mathbf{u} = \mathbf{B}$ and $\mathbf{v} = \mathbf{C}$ and $\mathbf{C} = $ const. the formula (see Appendix C):

$$\boldsymbol{\nabla} \cdot (\mathbf{u} \times \mathbf{v}) = \mathbf{v} \cdot (\boldsymbol{\nabla} \times \mathbf{u}) - \mathbf{u} \cdot (\boldsymbol{\nabla} \times \mathbf{v}).$$

Hence Eq. (6.9) applied to the first integral on the right hand side of Eq. (6.8) yields

$$-\int d\mathbf{r}' \boldsymbol{\nabla}' \times \left(\frac{\mathbf{M}}{|\mathbf{r} - \mathbf{r}'|} \right) = -\int d\mathbf{F}' \times \frac{\mathbf{M}}{|\mathbf{r} - \mathbf{r}'|}.$$

Here (see above) $\mathbf{M}(\mathbf{r}') = N(\mathbf{r}')\langle \mathbf{m} \rangle$. The integration extends over the entire volume of the macroscopic body. Assuming localisation, there is no molecule outside and hence the surface integral there vanishes. Then

$$\int_F d\mathbf{F}' \times \frac{\mathbf{M}(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} = 0,$$

and with Eqs. (6.6) and (6.8) we are left with

$$\mathbf{A}(\mathbf{r}) = k' \int d\mathbf{r}' \frac{\mathbf{j}(\mathbf{r}') + \boldsymbol{\nabla}' \times \mathbf{M}(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} \tag{6.10}$$

for the vector potential at a point $\mathbf{r}$ outside the matter of the macroscopic body. Thus the macroscopic magnetisation $\mathbf{M}$ implies a *magnetisation current*

$$\mathbf{j}_M = \boldsymbol{\nabla} \times \mathbf{M}. \tag{6.11}$$

We therefore obtain according to our earlier considerations

$$\boldsymbol{\nabla} \times \mathbf{B} = \mu_0(\mathbf{j}(\mathbf{r}) + \mathbf{j}_M(\mathbf{r})) = \mu_0 \mathbf{j}_{\text{total}}, \tag{6.12}$$

where $\mathbf{j}_{\text{total}}$ is the total current density. With Eq. (6.11) we then have

$$\boldsymbol{\nabla} \times \mathbf{B} = \mu_0(\mathbf{j} + \boldsymbol{\nabla} \times \mathbf{M}), \tag{6.13}$$

or

$$\boldsymbol{\nabla} \times (\mathbf{B} - \mu_0 \mathbf{M}) = \mu_0 \mathbf{j},$$

or, introducing the vector $\mathbf{H}$,

$$\nabla \times \mathbf{H} = \mathbf{j} \tag{6.14}$$

with

$$\mathbf{H} = \frac{1}{\mu_0}\mathbf{B} - \mathbf{M}, \quad \mathbf{B} = \mu_0(\mathbf{H} + \mathbf{M}). \tag{6.15}$$

The vector $\mathbf{H}$ is called *magnetic field strength*. $\mathbf{H}$ does not depend on the molecular magnetic moments. (Note the analogy with electrostatics: There, since $\nabla \cdot \mathbf{D} = \rho$, the polarisation charges with density ρ_P are not sources of $\mathbf{D}$). This is the reason why in many texts $\mathbf{H}$ is introduced first, and not $\mathbf{B}$, since the current is first considered without the effect of magnetisation. This means the $\mathbf{B}$ we introduced originally is $\mu_0\mathbf{H}$. In vacuum $\mathbf{B} = \mu_0\mathbf{H}$.

The integral form of Eq. (6.14) is very important in applications, that is

$$\int_F (\nabla \times \mathbf{H}) \cdot d\mathbf{F} \stackrel{\text{Stokes}}{=} \oint_{C(F)} \mathbf{H} \cdot dl = \int_F \mathbf{j} \cdot d\mathbf{F}. \tag{6.16}$$

On the right hand side we do not write I, because in general (i.e. depending on the problem) the current can cross the area F several times and in different directions, and then the right hand side is an integral multiple of I. The formula (6.16) will be an important starting point in problems to be discussed later.

For many cases, i.e. types of matter, one observes the linearity

$$\mathbf{M} = \kappa\mathbf{H}, \quad \kappa = \text{const.} \tag{6.17}$$

In these *linear* cases

$$\mathbf{B} = \mu\mathbf{H}, \quad \mu = \mu_0(1 + \kappa). \tag{6.18}$$

Today the unit of *magnetic field strength* is ampere/meter, written A/m, and sometimes ampere-revolutions/meter (the obsolete unit is oersted (Oe)). Thus here in magnetostatics we have

$$\mathbf{A}(\mathbf{r}) = k' \int dr' \frac{\mathbf{j}_{\text{total}}(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|},$$

where

$$\begin{aligned}
\mathbf{j}_{\text{total}} &= \mathbf{j} + \mathbf{j}_M = \mathbf{j} + \nabla \times \mathbf{M} \\
&= \mathbf{j} + \nabla \times (\kappa\mathbf{H}) = \mathbf{j} + \kappa\nabla \times \mathbf{H} \\
&= \mathbf{j} + \kappa\mathbf{j} \equiv \frac{\mu}{\mu_0}\mathbf{j},
\end{aligned}$$

and $\mu = \mu_0(1 + \kappa)$. Thus we have

$$\mathbf{A}(\mathbf{r}) = k'\frac{\mu}{\mu_0} \int d\mathbf{r}' \frac{\mathbf{j}(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|}, \tag{6.19}$$

where

$$k'\frac{\mu}{\mu_0} = \frac{\mu_0}{4\pi}\frac{\mu}{\mu_0} = \frac{\mu}{4\pi}.$$

Thus one has almost the same formula (in these linear cases) as in the microscopic case, except that $\mu_0 \to \mu$, with μ incorporating matter effects.

6.3   Magnetic Properties of Matter

We saw above that

$$\frac{\mu}{\mu_0} = 1 + \kappa.$$

Here μ/μ_0 is the *relative permeability of the medium* and κ the *magnetic susceptibility*. One distinguishes between diamagnetic, paramagnetic and ferromagnetic materials. *Diamagnetic materials* are those consisting of atoms, for which the sum of the individual orbital angular momenta of the electrons is zero. Such atoms therefore do not possess a permanent total magnetic moment. However, when placed in an external field $\mathbf{B}$, they develop an induced magnetic moment on account of their motion in this field $\mathbf{B}$ (in accordance with the law of induction to be discussed later which implies an induced electromotive force which accelerates the electrons). According to the law of induction (cf. later: the Lenz rule) the induced magnetic moment or rather its field has the direction opposite to that of the applied field $\mathbf{B}$, so that

$$\kappa < 0, \quad \mu < 1.$$

The effect is independent of temperature (as the explanation implies).

Paramagnetic materials are those whose atoms possess permanent magnetic moments (in fact, their electrons outside closed electronic shells), which favour alignment with the applied field, so that for these materials

$$\kappa > 0, \quad \mu > 1.$$

(In the case of these the effect of induced magnetic moments is much smaller).

Ferromagnetic materials are those, for which

$$\kappa > 0, \quad \text{and} \quad \mu = \mu(\mathbf{H}) \gg 1$$

(see discussion of *hysteresis curves* (B versus H) in appropriate literature). In the case of these materials the connection between $\mathbf{B}$ and $\mathbf{H}$ depends

on the past of the system; the connection is therefore not unique. In addition the connection depends on the temperature (above the so-called Curie-temperature ferromagnetic materials become paramagnetic).

6.4   Energy of the Magnetic Field

For the *magnetic energy* one obtains, as we shall show later, an expression very similar to that for electric energy. We cite the result at this point for reasons of completeness:

$$W = \frac{1}{2} \int d\mathbf{r} \, \mathbf{B} \cdot \mathbf{H}. \tag{6.20}$$

6.5   Behaviour of B and H at Boundary Surfaces

In the case of the field $\mathbf{E}$ we saw that *without considerations of properties of matter* at a thin charged surface

$$(\mathbf{E}_2 - \mathbf{E}_1) \cdot \mathbf{n} = 4\pi k\sigma, \quad k = \frac{1}{4\pi\epsilon_0},$$

or *in the case of macroscopic electrostatics*, which takes the polarisation of the dielectric into account,

$$(\mathbf{E}_2 - \mathbf{E}_1) \cdot \mathbf{n} = 4\pi k(\sigma + \sigma_P),$$

and

$$(\mathbf{D}_2 - \mathbf{D}_1) \cdot \mathbf{n} = 4\pi k\sigma.$$

In the case of the field $\mathbf{B}$ we have $\nabla \cdot \mathbf{B} = 0$, i.e.

$$0 = \int_V dV \nabla \cdot \mathbf{B} = \int_{F(V)} \mathbf{B} \cdot d\mathbf{F},$$

so that

$$(\mathbf{B}_2 - \mathbf{B}_1) \cdot \mathbf{n} \triangle F = 0.$$

The zero on the right side implies that no single magnetic poles exist. Hence the normal component of the field at the surface is continuous, i.e.

$$B_{2n} = B_{1n}. \tag{6.21}$$

In the case of the field $\mathbf{E}$ we obtained from $\nabla \times \mathbf{E} = 0$ the relation

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

In the case of **H** we have
$$\nabla \times \mathbf{H} = \mathbf{j},$$

i.e.
$$\int_F \nabla \times \mathbf{H} \cdot d\mathbf{F} = \oint_C \mathbf{H} \cdot d\mathbf{s} = \int_F \mathbf{j} \cdot d\mathbf{F} = I. \qquad (6.22)$$

We now consider the contribution on the right side in relation to an element of area $d\mathbf{F}$ perpendicular to the interface between two media as in Fig. 6.1. Then $d\mathbf{F}$ is a vector in the plane of the interface. Applied to this element of area we can write Eq. (6.22):

$$\int_{\Delta F} \mathbf{j} \cdot d\mathbf{F} = \mathbf{j} \cdot \Delta \mathbf{F} = jld,$$
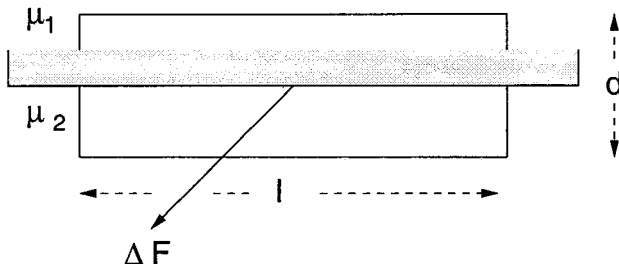
Fig. 6.1 Element of area ΔF.

where l is the length and d the height of the element of area. For conductors we have Ohm's law $\mathbf{j} = \sigma \mathbf{E}$ or $j = \sigma E$. In ideal conductors $\sigma = \infty$ with $\mathbf{E} = -\nabla \phi = 0$. It is then possible that

$$jld = \sigma Eld \neq 0 \quad \text{and} \quad \text{finite},$$

and one defines

$$\lim_{\sigma \to \infty, d \to 0} \sigma Ed := K, \quad \int \mathbf{j} \cdot d\mathbf{F} = Kl,$$

where K is a surface current density,[*] $\mathbf{j}$ is the volume current density. K is also called *linear current density*.[†] If the conductivity σ is finite, we have $K = 0$; i.e. in this case there is actually no surface current.[‡]

---

[*]See Jackson [3], p. 19.

[†]See Lim [2], p. 42.

[‡]See Jackson [3], p. 336. Nonetheless an effective surface current will be defined later in the treatment of wave guides, in fact with the help of $\mathbf{j}$.

The continuity condition for $\mathbf{H}$ can now be written (using $\oint_C \mathbf{H} \cdot d\mathbf{s}$ as in the case of $\mathbf{E}$)

$$H_{2t} - H_{1t} = K \quad (= 0 \text{ for } \sigma \text{ finite}), \tag{6.23}$$

i.e.

$$\mathbf{n} \times (\mathbf{H}^{(2)} - \mathbf{H}^{(1)}) = \mathbf{K}. \tag{6.24}$$

The surface current introduced here is the same as the surface current K referred to earlier. The name "linear current density" has its origin in the fact that in Eq. (6.22),

$$H_{2t} - H_{1t} = \frac{I}{2l},$$

the current appears linearly. This result is of considerable importance and, like its counterpart (2.16) in electrostatics, will be used in examples, as in Example 14.4.

6.6   Current Circuits Compared with Flux Circuits

It is interesting and instructive to enquire about analogies between electric and magnetic properties. In an electric circuit, i.e. *current carrying circuit*, with stationary current $I = jF$ defined by the condition $\boldsymbol{\nabla} \cdot \mathbf{j} = 0$, the current I is the constant quantity. The potential V, or the corresponding *electromotive force* of the circuit, is with $\mathbf{j} = \sigma \mathbf{E}, \mathbf{I} = \mathbf{j}F$ (see Eq. (5.2)) given by

$$\begin{aligned} V = \oint \mathbf{E} \cdot d\mathbf{l} &= \oint \frac{\mathbf{j} \cdot d\mathbf{l}}{\sigma} \\ &= \oint \frac{\mathbf{I} \cdot d\mathbf{l}}{\sigma F} \\ &= I \oint \frac{d\mathbf{l}}{\sigma F} = IR, \end{aligned} \tag{6.25}$$

where R is the *electric resistance* of the circuit.

The magnetic analogy is (see Eq. (6.16)) the *magnetomotive force*

$$\oint \mathbf{H} \cdot d\mathbf{l} = \int_F \mathbf{j} \cdot d\mathbf{F} = IN,$$

where now $\boldsymbol{\nabla} \cdot \mathbf{B} = 0$ with constant *magnetic flux* $\Phi = \int \mathbf{B} \cdot d\mathbf{F_0}$. N is the effective number of currents I passing through the surface $C(F)$ (i.e. the number of currents I coming out minus the number going in, i.e. in the opposite direction). The relation $B_{1n} = B_{2n} = \cdots$ tells us that the magnetic

flux is conserved, i.e. the lines of flux are continuous at a boundary surface F_0 (note the difference between the areas $F(C)$ and F_0). It follows that

$$\int_F \mathbf{j} \cdot d\mathbf{F} = \oint \frac{\mathbf{B} \cdot d\mathbf{l}}{\mu} = \oint \frac{F_0 \mathbf{B} \cdot d\mathbf{l}}{F_0 \mu} = \oint \frac{(B_n F_0) dl}{F_0 \mu} = \Phi \oint \frac{dl}{\mu F_0}. \quad (6.26)$$

If the magnetic flux Φ corresponds in the electric case to the current I, then the expression

$$\oint \frac{dl}{\mu F_0}$$

may be interpreted as *magnetic resistance*. It follows that we can calculate the magnetic flux with the relation (6.26), i.e. with

$$\Phi = \frac{\int_F \mathbf{j} \cdot d\mathbf{F}}{\oint \frac{dl}{\mu F_0}} = \frac{IN}{\oint \frac{dl}{\mu F_0}}. \quad (6.27)$$

This result will be used in Example 6.4 below.

6.7 Examples

Example 6.3: The electromagnet
Determine the magnetic field strength H_2 in the gap of the circular electromagnet shown in Fig. 6.2 (assuming constant and homogeneous fields).
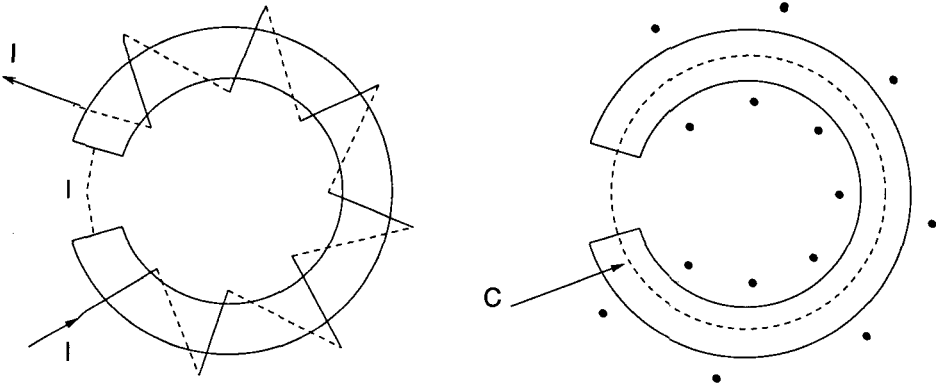
Fig. 6.2 The electromagnet.

Solution: The ring-shaped magnetic material with circular cross-sectional area F_0, permeability μ and mean circumference $L - l$, where l is the length of the gap (of air or vacuum), has N turns of a wire-like conductor carrying current I wound around it. We use the integral form of Eq. (6.14), i.e.

$$\oint_C \mathbf{H} \cdot d\mathbf{l} = \int_{F(C)} \mathbf{j} \cdot d\mathbf{F}, \quad (6.28)$$

where C is the path shown in Fig. 6.2 and $F(C)$ the area enclosed by it. Now $\mathbf{B} = \mu \mathbf{H}$, and at a boundary surface, which means at the two transverse areas, where the gap begins, as we saw, $B_{1n} = B_{2n}$. We set $B = B_1$ in the magnetic material. Then

$$H_1 \mu = H_2 \mu_0. \tag{6.29}$$

On the other hand, from Eq.(6.16) (since the wire crosses the area $F(C)$ a total of N times in the same direction)

$$H_1 L + H_2 l = IN. \tag{6.30}$$

From Eqs. (6.29) and (6.30) we eliminate H_1 and obtain

$$H_2 = \frac{IN}{l + \frac{\mu_0 L}{\mu}} \tag{6.31}$$

in ampere-turns/meter. The same result follows immediately from application of the flux formula (6.27), i.e. with

$$\Phi = \int_F \frac{\mathbf{j} \cdot d\mathbf{F}}{\oint \frac{dl}{\mu F_0}}.$$

Inserting the appropriate quantities, we obtain

$$\mu_0 H_2 F_0 = \frac{IN}{\frac{1}{F_0}\left(\frac{L}{\mu} + \frac{l}{\mu_0}\right)},$$

in agreement with the result above.

Example 6.4: The electromagnet with tapered poles towards the gap

We consider a toroidal electromagnet with circular cross section and radius R (of the cross-sectional area) along the torus part of length L (as in Example 6.3). Towards the gap (of length h), at both ends along a length l (always along the circle which coincides with the axis of the torus), the radius decreases from R to r at the poles. Determine the magnetic field strength H_2 in the gap.
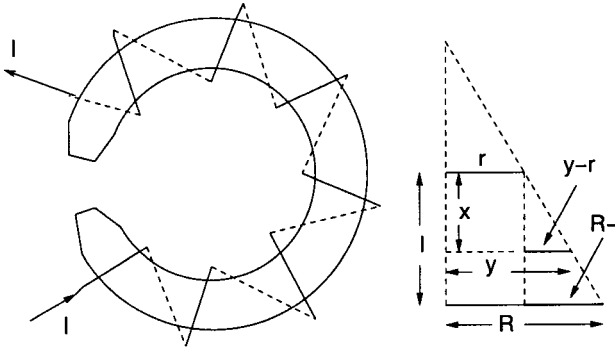
Fig. 6.3 The electromagnet with tapered poles.

Solution: We use Eq. (6.27). For this we have to evaluate

$$\underbrace{H_2 \mu_0 \pi r^2}_{\text{flux in the gap}} = \frac{IN}{\oint \frac{dl}{\mu F_0}}.$$

Here the line integrals along the gap and the main part of the torus follow as in Example 6.3, the respective cross-sectional areas being πr^2 and πR^2. Along each of the two tapered ends the radius y of the cross section of the electromagnet at a distance x measured from the small end is given by — as one can calculate with the method of congruent triangles (see Fig. 6.3) —

$$\frac{x}{l} = \frac{y-r}{R-r}, \quad \text{i.e.} \quad y = r + \frac{x}{l}(R-r)$$

(verification: for $x = 0$ this is r and for $x = l$ this is R), so that for each of the two pole pieces

$$\int \frac{dl}{\mu_1 F_0} = \frac{1}{\mu_1} \int_0^l dx \frac{1}{\pi\left(r + \frac{x}{l}(R-r)\right)^2}$$

$$= -\frac{l}{\pi\mu_1(R-r)}\left[\frac{1}{r + \frac{x}{l}(R-r)}\right]_0^l$$

$$= \frac{l}{\pi\mu_1 rR}.$$

Altogether we have

$$H_2\mu_0\pi r^2 = \frac{IN}{\frac{h}{\mu_0\pi r^2} + \frac{L}{\mu_1\pi R^2} + \frac{2l}{\pi\mu_1 rR}},$$

i.e.

$$H_2 = \frac{IN}{h + \frac{\mu_0 L r^2}{\mu_1 R^2} + \frac{\mu_0 2lr}{\mu_1 R}}.$$

We see that for $l \to 0, r \to R$ the result reduces to that of Example 6.3.

Example 6.5: The vector potential along a thin, circular wire[§]
Calculate the vector potential along a current carrying wire of tiny cross section and circular shape.

Solution: We use Eq. (6.19). With $ds = rd\phi$, where r is the radius formed by the wire, we have

$$\mathbf{A}(\mathbf{r}) = \frac{\mu}{4\pi}\int \frac{Ids(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}$$

$$= \frac{\mu}{4\pi}\int_{\phi=0}^{2\pi} \frac{\mathbf{I}(r'd\phi)}{r'\sqrt{2-2\cos\phi}}$$

$$= \mathbf{I}\frac{\mu}{4\pi\sqrt{2}}\int_0^{2\pi} \frac{d\phi}{\sqrt{2}\sin(\phi/2)} = \mathbf{I}\frac{\mu}{4\pi}\int_0^{\pi} \frac{dx}{\sin x}$$

$$= \mathbf{I}\frac{\mu}{4\pi}\left[\ln\left|\tan\frac{x}{2}\right|\right]_0^{\pi}$$

$$= \mathbf{I}\frac{\mu}{4\pi}\left[\ln\left|\frac{x}{2}\right| + \frac{1}{3}\left(\frac{x}{2}\right)^2 + \cdots\right]_{\delta\phi=2/\Lambda r\to 0}^{\pi-\delta\phi}$$

$$\sim \mathbf{I}\frac{\mu}{4\pi}\ln(\Lambda r), \tag{6.32}$$

where we used a not so familiar expansion from a Table of Formulae.[¶] The divergence that we encounter here (for $\Lambda \to 0$) appears because we assumed that the wire has (effectively) vanishing cross section ($r = r'$). For small thickness $1/\Lambda$ the dominant contribution is proportional to $\ln(\Lambda r)$,

---

[§]See also E. Witten, *Nucl. Phys.* **B249** (1985) 557.

[¶]I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series and Products* (Academic, 1965), formula 1.518(3), p. 46.

Λ having the dimension of a reciprocal length. Since the current is constant along the string, the same is true for $\mathbf{A}(\mathbf{r})$.

Example 6.6: Vector potential of a long, straight wire with current I
Determine the vector potential of the wire.

Solution: We start with the same relation for the vector potential as in Example 6.5, i.e.

$$\mathbf{A}(\mathbf{r}) = \frac{\mu}{4\pi} \int \frac{I ds(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|}.$$

We take the wire as lying along the y-direction. Then the formula implies that the components A_x, A_z are zero, and for the y-component we obtain for the field at a point x along the x-axis

$$A_y(x) = \frac{\mu I}{4\pi} \int \frac{dy}{\sqrt{x^2 + y^2}}$$

$$= \frac{\mu I}{4\pi} \left\{ \sinh^{-1} \frac{y_2}{x} - \sinh^{-1} \frac{y_1}{x} \right\}.$$

The last expression follows in a similar way as in the case of the electrostatic potential we encountered in Example 3.3.

Example 6.7: Energy of the magnetic moment m in the field B
Determine the energy of the magnetic moment $\mathbf{m}$.

Solution: We consider first the case of an electric dipole $\mathbf{p}$ in an applied field $\mathbf{E}$. The energy U of the dipole is the sum of the energies of the two charges q and $-q$ at points $\mathbf{r}_+$ and $\mathbf{r}_-$, a small distance d apart, i.e.

$$U = q\phi(\mathbf{r}_+) - q\phi(\mathbf{r}_-) = q\left[\phi(\mathbf{r}_- + \mathbf{d}) - \phi(\mathbf{r}_-)\right] = q\mathbf{d} \cdot \nabla\phi = -\mathbf{p} \cdot \mathbf{E},$$

with $\mathbf{p} = q\mathbf{d}$. We define the force $\mathbf{F}$ by the expression

$$U = -\int \mathbf{F} \cdot d\mathbf{s} = -\mathbf{p} \cdot \mathbf{E}.$$

Then

$$\mathbf{F} = -\nabla U = \nabla(\mathbf{p} \cdot \mathbf{E}).$$

In the magnetic case we have to remember first of all that $\mathbf{B} = \nabla \times \mathbf{A}$, i.e. $\mathbf{B}$ is in general not expressible as a gradient. However, we now consider an *external field*, i.e. a field $\mathbf{B}$, where the current density is zero. Then $\nabla \times \mathbf{B} = 0$, and $\mathbf{B}$ can be written as a gradient, i.e. derivable from a potential ϕ. It follows that we can use the relations of above and we can write

$$U = -\mathbf{m} \cdot \mathbf{B}, \qquad \mathbf{F} = \nabla(\mathbf{m} \cdot \mathbf{B}). \tag{6.33}$$

Example 6.8: Vector potential of a solid cylinder
An infinitely long solid cylinder of radius R (made of material with relative permeability $\mu/\mu_0 > 1$) carries the constant current density $\mathbf{j}_0$ in the direction of its axis. Calculate (a) the vector potential and (b) the magnetic field strength within and without the conductor by solving the Poisson equation for the vector potential.

Solution:
(a) Since

$$\mathbf{A}(\mathbf{r}) = \frac{\mu}{4\pi} \int d\mathbf{r'} \, \frac{\mathbf{j}(\mathbf{r'})}{|\mathbf{r} - \mathbf{r'}|},$$

the vector $\mathbf{A}$ is parallel to $\mathbf{j}$ and parallel to $\mathbf{e}_z$. The Poisson equation is $\triangle \mathbf{A} = -\mu \mathbf{j}, \mathbf{j} = j_0 \mathbf{e}_z \theta(R - \rho), j_0 = I/\pi R^2$, with cylindrical coordinates ρ, φ, z. In view of the cylindrical symmetry we have $A = A_z(\rho)$ and the Poisson equation is

$$\frac{1}{\rho} \frac{\partial}{\partial \rho} \left(\rho \frac{\partial}{\partial \rho} A_z(\rho) \right) = -\mu j_0 \theta(R - \rho).$$

$$(i) \ \rho > R : \ \rho \frac{\partial}{\partial \rho} A_z(\rho) = c \ (\text{const.}); \quad (ii) \ \rho < R : \ \frac{1}{\rho} \frac{\partial}{\partial \rho} \left(\rho \frac{\partial}{\partial \rho} A_z(\rho) \right) = -\mu j_0.$$

In case (i) we have

$$A_z(\rho) = c \ln R + A_{z0}.$$

In case (ii) we have

$$\frac{\partial}{\partial \rho} \left(\rho \frac{\partial}{\partial \rho} A_z(\rho) \right) = -\mu j_0 \rho,$$

and hence

$$A_z(\rho) = -\frac{\mu j_0}{4} \rho^2 + c_1 \ln \rho + c_2.$$

We choose $c_2 = 0$. Regularity at $\rho = 0$ requires $c_1 = 0$. Continuity at $\rho = R$ implies:

$$c \ln R + A_{z0} = -\frac{1}{4} \mu j_0 R^2.$$

We obtain therefore:

$$A_z(\rho) = \begin{cases} -\frac{1}{4} \mu j_0 R^2 & \text{for } \rho \leq R, \\ c \ln \frac{\rho}{R} - \frac{1}{4} \mu j_0 R^2 & \text{for } \rho \geq R. \end{cases}$$

(b) Considering the magnetic field strength, we have

$$\mu \mathbf{H} = \nabla \times \mathbf{A} = \mu H_\varphi(\rho) \mathbf{e}_\varphi = \left(\frac{\partial}{\partial z} A_\rho - \frac{\partial}{\partial \rho} A_z \right) \mathbf{e}_\varphi = -\frac{\partial}{\partial \rho} A_z \mathbf{e}_\varphi.$$

$$\therefore H_\varphi(\rho) = -\frac{1}{\mu} \frac{\partial}{\partial \rho} A_z(\rho) = \begin{cases} \frac{1}{2} j_0 \rho & \text{for } \rho \leq R, \\ -\frac{c}{\mu \rho} & \text{for } \rho \geq R. \end{cases}$$

Continuity at $\rho = R$ implies:

$$-\frac{c}{\mu R} = \frac{1}{2} j_0 R, -c = \frac{1}{2} j_0 \mu R^2.$$

$$\therefore H_\varphi(\rho) = \begin{cases} j_0 \rho & \text{for } \rho \leq R, \\ \frac{j_0 \mu R^2}{2\rho} & \text{for } \rho \geq R. \end{cases}$$

6.8   Summary of Formulae of Magnetostatics

We collect here for reference purposes the most important results of magnetostatics:

(1). **Current, equation of continuity and stationarity:**

$$I = \frac{dq}{dt}, \quad dq = \rho dV, \quad \frac{\partial \rho}{\partial t} + \nabla \cdot \mathbf{j} = 0, \quad \text{stationarity} \quad \nabla \cdot \mathbf{j} = 0.$$

(2). **Magnetic field strength:**

$$\nabla \times \mathbf{H} = \mathbf{j}, \quad \text{or} \quad \oint_C \mathbf{H} \cdot d\mathbf{l} = \int_{F(C)} \mathbf{j} \cdot d\mathbf{F},$$

$$\mathbf{j} = \rho \frac{d\mathbf{s}}{dt}, \quad I = \int \mathbf{j} \cdot d\mathbf{F}.$$

(3). **Macroscopic magnetostatics:**[||]

$$\mathbf{j}_{\text{total}} = \mathbf{j} + \mathbf{j}_M,$$

Magnetisation current density $\mathbf{j}_M = \nabla \times \mathbf{M}$,

$$\mathbf{B} = \mu_0(\mathbf{H} + \mathbf{M}).$$

(4). **Magnetic material equations:**

$$\mathbf{M} = \kappa \mathbf{H}, \quad \mathbf{B} = \mu \mathbf{H}, \quad \frac{\mu}{\mu_0} = 1 + \kappa.$$

(5). **Biot–Savart law:**

$$\mathbf{B}(\mathbf{r}) = \frac{\mu}{4\pi} \int \frac{I(\mathbf{r}')d\mathbf{s}(\mathbf{r}') \times (\mathbf{r} - \mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|^3}.$$

(6). **Ampère's law and Lorentz-force:**

$$\mathbf{F} = \int I(\mathbf{r}')d\mathbf{s}(\mathbf{r}') \times \mathbf{B}(\mathbf{r}'), \quad \mathbf{F} = q\mathbf{v}(\mathbf{r}) \times \mathbf{B}(\mathbf{r}).$$

(7). **No magnetic monopoles:**

$$\nabla \cdot \mathbf{B} = 0.$$

(8). **Behaviour at boundary surfaces:**

$$B_{2n} = B_{1n}, \quad H_{2t} - H_{1t} = K$$

(Surface current $K = 0$ for finite conductivity σ).

(9). **B-field energy:**

$$W = \frac{1}{2} \int \mathbf{B} \cdot \mathbf{H} dV.$$

[||] Some authors relate **B** to **H** by $\mathbf{B} = \mu_0\mathbf{H} + \mathbf{M}'$, in which case **M**' has units of **B** (weber/meter^2) instead of **H** (ampere/meter). See e.g. *The Electromagnetic Problem Solver* [21], Sec. II, p. 1-1.

Chapter 7

The Maxwell Equations

7.1 Preliminary Remarks

Now that we have dealt with electrostatics and magnetostatics also for macroscopic objects, the next step is to introduce time dependence. Proceeding in our phenomenological and historical approach we are led to consider next Faraday's law of induction. With this we can complete the equations of macroscopic electrodynamics with the addition of *Maxwell's displacement current*. The result is the full set of Maxwell's equations.

7.2 Time-Dependent Fields and Faraday's Law of Induction

Faraday discovered in 1831 *that an electric current arises in a closed wire loop if the wire is moved through a magnetic field*, in other words when the position or orientation of the wire with respect to the magnetic field is changed, or if the magnetic field varies with time. We consider two situations in which Faraday's observation applies. In the first case the field $\mathbf{B}$ is maintained constant in time.

(a) In an electric field $\mathbf{E}$ the charge q experiences the force

$$\mathbf{F} = q\mathbf{E}, \quad \frac{d\mathbf{F}}{dq} = \mathbf{E}. \tag{7.1}$$

This force $\mathbf{F}$ results from a nonvanishing potential difference in the conductor. On the other hand (cf. Lorentz-force), the field $\mathbf{B}$ acting on a charge dq moving with velocity $\mathbf{v} = d\mathbf{l}/dt$ implies that the latter experiences the force

$d\mathbf{F}$ given by

$$dF = dq\frac{d\mathbf{l}}{dt} \times \mathbf{B},$$

so that

$$\frac{d\mathbf{F}}{dq} = \frac{d\mathbf{l}}{dt} \times \mathbf{B}. \tag{7.2}$$

Identifying the forces of Eqs. (7.1) and (7.2) in order to arrive at an explanation of Faraday's observation, we obtain

$$dq\mathbf{E} = \frac{d\mathbf{l}}{dt} \times dq\mathbf{B}, \tag{7.3}$$

i.e.
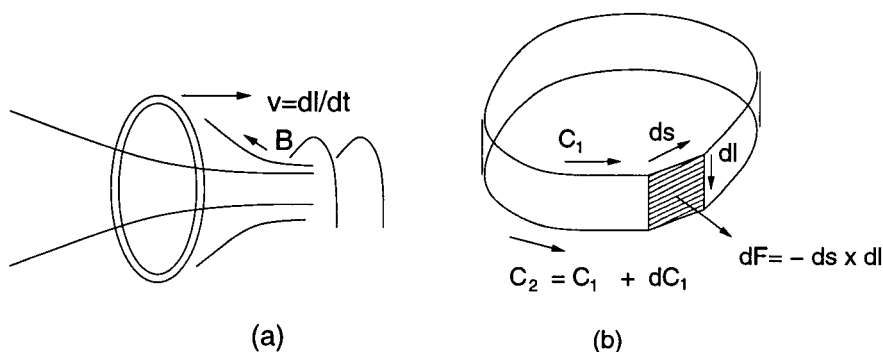
Fig. 7.1 (a), (b) The moving current loop.

$$\mathbf{E} = \frac{d\mathbf{l}}{dt} \times \mathbf{B},$$

provided the right side (or one component) is parallel to $\mathbf{E}$. This means, the electric force acting on the charge dq is equal to the force which the field $\mathbf{B}$ exerts on the charge dq moving with velocity $\mathbf{v}$. Put differently: The right side of Eq. (7.3), i.e. the force that $\mathbf{B}$ exerts on dq, *induces* the electric force $\mathbf{E}dq$, i.e. the Lorentz-force, acts on the electrons in the conductor, and hence a current in the conductor is observed, and so a potential difference or induced voltage. We assume this here as an *empirical finding*.[*]

We let C_1 be the initial position of the conducting loop and $C_2 = C_1 + \delta C_1$ its position a time interval δt later (as indicated in Fig. 7.1(b)). δC_1

---

[*]The induced voltage or potential difference V is also called *induced electromotive force*. V has the dimension of energy. The term *"force"* has a historical origin; in the 18th century, as also at the beginning of the 19th century, various quantities which today represent energy were described as "force", e.g. also in writings of Kant.

represents the displacement of the coordinates of the loop or, alternatively, the deformation of the loop. The potential difference V induced in the loop C_1 is according to its definition and after inserting the expression for $\mathbf{E}$ from above given by

$$
\begin{aligned}
V = \oint_{C_1} \mathbf{E} \cdot d\mathbf{s} &= \oint_{C_1} \left(\frac{d\mathbf{l}}{dt} \times \mathbf{B} \right) \cdot d\mathbf{s} \\
&= \int_{C_1} \left(d\mathbf{s} \times \frac{d\mathbf{l}}{dt} \right) \cdot \mathbf{B} = - \oint_{\text{displacement area}} \frac{d\mathbf{F} \cdot \mathbf{B}}{dt},
\end{aligned}
\tag{7.4}
$$

where the integration is over the area swept out by C_1 in its displacement to C_2. However, we always have $\nabla \cdot \mathbf{B} = 0$, i.e. integrating over the displacement-volume between C_1 and C_2

$$
\begin{aligned}
0 &= \int_{V(\text{displacement})} d\mathbf{r}\, \nabla \cdot \mathbf{B} \\
&= \int_{F(V)} \mathbf{B} \cdot d\mathbf{F} \\
&= \left\{ \int_{\text{displacement}} + \int_{F(C_1)} + \int_{F(C_1 + \delta C_1)} \right\} \mathbf{B} \cdot d\mathbf{F}.
\end{aligned}
$$

Here $F(C_1)$ is the area enclosed by C_1, and correspondingly $F(C_2)$ the area enclosed by C_2, and $F(\text{displacement})$ is as before the area swept out by the boundary. Then

$$
\begin{aligned}
- \int_{\text{boundary}} \mathbf{B} \cdot d\mathbf{F} &= \int_{F(C_1)} \mathbf{B} \cdot d\mathbf{F}_1 + \int_{F(C_1 + \delta C_1)} \mathbf{B} \cdot d\mathbf{F}_2 \\
&= - \left[\int_{F(C_1 + \delta C_1)} \mathbf{B} \cdot d\mathbf{F}_1 - \int_{F(C_1)} \mathbf{B} \cdot d\mathbf{F}_1 \right] \\
&= - \delta \int_{F(C_1)} \mathbf{B} \cdot d\mathbf{F}_1.
\end{aligned}
$$

It then follows with Eq. (7.4) that

$$
V = \oint_{C_1} \mathbf{E} \cdot d\mathbf{s} \stackrel{\text{Stokes}}{\equiv} \int_{F(C_1)} \nabla \times \mathbf{E} \cdot d\mathbf{F} = - \frac{d}{dt} \int_{F(C_1)} \mathbf{B} \cdot d\mathbf{F}_1.
$$

Thus in this case with the field $\mathbf{B}$ constant in time:

$$
\int_{F(C_1)} \nabla \times \mathbf{E} \cdot d\mathbf{F} = - \frac{d}{dt} \int_{F(C_1(t))} \mathbf{B} \cdot d\mathbf{F}_1.
\tag{7.5}
$$

In the next case we consider C_1 as fixed and $\mathbf{B}$ as varying with time.

(b) We have seen that at a sufficient distance away from a current carrying loop the *magnetic properties of the loop* can be considered as those of a magnetic dipole with moment $\mathbf{m}$ given by

$$\mathbf{m} = I \int_{C_1 \text{ fixed}} d\mathbf{F}.$$

The energy of this dipole in an external field $\mathbf{B}$ is (see Eq. (6.30))

$$U = -\mathbf{m} \cdot \mathbf{B} = -I \int_{C_1 \text{ fixed}} \mathbf{B} \cdot d\mathbf{F}.$$

We now assume: In the time interval δt the field $\mathbf{B}$ changes by an amount $\delta \mathbf{B}$ in such a way, that the amount of magnetic energy thereby gained per second is exactly compensated by the loss of a corresponding amount of energy as heat whilst *the current I is kept constant*. Then the energy lost as heat per second is

$$-\frac{dqV}{dt} = -IV = -I \oint \mathbf{E} \cdot d\mathbf{l}.$$

On the other hand, the magnetic energy gained per second is

$$+\frac{\partial U}{\partial t} = -I \int_{C_1 \text{ fixed}} \frac{\partial \mathbf{B}}{\partial t} \cdot d\mathbf{F}.$$

Hence the sum "loss + gain = 0" yields the relation

$$\oint \mathbf{E} \cdot d\mathbf{l} = -\int \frac{\partial \mathbf{B}}{\partial t} \cdot d\mathbf{F}. \tag{7.6}$$

Let Φ be the *magnetic flux* defined by

$$\Phi = \int_F \mathbf{B}(t) \cdot d\mathbf{F}.$$

The MKSA-unit of magnetic flux is the weber (Wb) (for conventions of notation see Appendix B). Equation (7.6) can then be written

$$IV = -I \frac{d\Phi}{dt}. \tag{7.7}$$

In this formulation Faraday's law is also known as *Neumann's law*. The minus sign in Eq. (7.7) expresses effectively what the so-called *Lenz rule* says (see below).

Recalling — as an instructive analogy — the following result from calculus [†]

$$\frac{d}{dt} \int_{a(t)}^{b(t)} f(x,t)dx = f(b(t),t)\frac{db}{dt} - f(a(t),t)\frac{da}{dt} + \int_{a(t)}^{b(t)} \frac{\partial f(b(t),t)}{\partial t}dx,$$

we see that the results (7.4), (7.5) can be combined in the following form (where instead of boundary values we have the path of integration $C_1(t)$):

$$\oint_{C_1(t)} \mathbf{E} \cdot d\mathbf{l} = \int_{F(C_1(t))} \boldsymbol{\nabla} \times \mathbf{E} \cdot d\mathbf{F} = -\frac{d}{dt} \int_{F(C_1(t))} \mathbf{B}(\mathbf{r},t) \cdot d\mathbf{F}. \qquad (7.8)$$

Here we are at present not so much interested in the variation of the positions of current circuits with time (but see Example 7.3), as in time-dependent fields $\mathbf{B}$. We assume therefore, that the form of C_1 remains unchanged and that C_1 is fixed in space. Then we have the case of Eq. (7.7), i.e.

$$\int_{F(C_1)} \boldsymbol{\nabla} \times \mathbf{E} \cdot d\mathbf{F} = - \int_{F(C_1)} \frac{\partial \mathbf{B}}{\partial t} \cdot d\mathbf{F}.$$

Since C_1 is arbitrary, we can equate the integrands and obtain

$$\boldsymbol{\nabla} \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}. \qquad (7.9)$$

This is the differential form of Faraday's law.
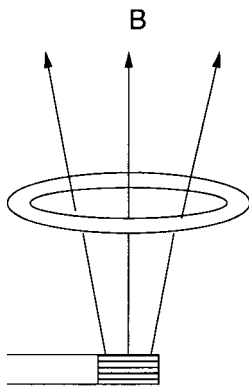
Fig. 7.2 The Lenz rule.

Finally we comment briefly on the so-called *Lenz rule*. This rule is useful in practical applications. In other respects there is no necessity to use it.

[†]See e.g. Whittaker and Watson [7], p. 67.

What the Lenz rule says is contained in Maxwell's equations. In the literature the Lenz rule is frequently — at least on a first reading — difficult to follow, e.g. with a statement like: "The Lenz rule says that the induced current and the associated magnetic flux have directions such that they act against the change of the external flux." Let us consider a conducting ring which falls in a field **B** as depicted in Fig. 7.2 (i.e. the ring maintains its shape but the flux through it changes). The change of flux $d\Phi$ or change in the number of lines of force cut by the falling ring in time dt is given by $d\Phi/dt$. According to Eq. (7.7) the magnitude of this rate gives the voltage V induced in the ring. Lenz's rule determines the sign in Eq. (7.7) by calling $V\,dt$ now induced flux (i.e. magnetic flux associated with the induced current) and saying, its sign is such that it cancels the corresponding external flux $d\Phi$.

At this point the natural question arises: Is there also a magnetic field strength **H** which is induced by time-variation of **E**? The answer is yes, and is provided by Maxwell's displacement current (see Example 7.1). Faraday had asked himself this question and performed experiments, but with no success.

7.3   Energy of the Magnetic Field

Starting from the expression

$$\mathbf{m} = I\mathbf{F}$$

for the magnetic moment of a circuit or network with current I and area F, and also from the expression for the potential energy U of the moment **m** in the field **B**, i.e.

$$U = -\mathbf{m} \cdot \mathbf{B} = -I\mathbf{B} \cdot \mathbf{F} \equiv -I \int \mathbf{B} \cdot d\mathbf{F},$$

we have for the variation of the energy of the field (where we have to reexpress the current I in terms of the field)

$$-\delta U \equiv \delta W = I\delta \int_{F \text{ fixed}} \mathbf{B} \cdot d\mathbf{F} = I \int_F (\nabla \times \delta \mathbf{A}) \cdot d\mathbf{F}$$

$$\stackrel{\text{Stokes}}{=} I \oint \delta \mathbf{A} \cdot d\mathbf{s} = \int d\mathbf{r}\delta \mathbf{A} \cdot \mathbf{j},$$

i.e.

$$\delta W = \int d\mathbf{r}\delta \mathbf{A} \cdot \mathbf{j}. \tag{7.10}$$

This expression has to be compared with the corresponding expression in electrostatics, i.e.

$$\delta W = \int d\mathbf{r}\delta\phi(\mathbf{r})\rho(\mathbf{r})$$

(observe that one can define an expression $\delta W' = \int d\mathbf{r}\delta\rho(\mathbf{r})\phi(\mathbf{r})$). Inserting into Eq. (7.10) the expression for $\mathbf{j}$ which we obtain from $\nabla \times \mathbf{H} = \mathbf{j}$, and using

$$\nabla \cdot (\mathbf{a} \times \mathbf{b}) = \mathbf{b} \cdot (\nabla \times \mathbf{a}) - \mathbf{a} \cdot (\nabla \times \mathbf{b}),$$

it follows that

$$\delta W = \int_{V_\infty} d\mathbf{r}\delta\mathbf{A} \cdot (\nabla \times \mathbf{H})$$

$$= \int d\mathbf{r}[\nabla \cdot (\mathbf{H} \times \delta\mathbf{A}) + \mathbf{H} \cdot (\nabla \times \delta\mathbf{A})]$$

$$= 0 + \int d\mathbf{r}\mathbf{H} \cdot \delta\mathbf{B},$$

where the first term was converted into a surface integral at infinity with the help of Gauss' divergence theorem and therefore vanishes. For $\mathbf{B} = \mu\mathbf{H}$ we have $\delta(\mathbf{H} \cdot \mathbf{B}) = 2\mathbf{H} \cdot \delta\mathbf{B}$ and hence

$$W = \frac{1}{2} \int d\mathbf{r}\ \mathbf{H} \cdot \mathbf{B}. \tag{7.11}$$

This result should be compared with the corresponding energy in electrostatics.

7.4 The Generalised Ohm's Law

In the case of a simple circuit with resistance R, generator voltage V_e, Ohm's law is $V_e = IR$, where I is the conduction current. Next we consider the case of a coil with n turns per unit length. In this case one has a magnetic field along the axis of the coil (as we saw), and so

$$H = nI, \quad B = \mu H.$$

If H is changed by changing I, also the flux $\Phi = n\times$ area of one turn $\times B$ changes and this induces in the conductor of the circuit an *induced potential*

$$V_i = -\frac{d\Phi}{dt}.$$

Hence the *generalised Ohm's law* is

$$IR = V = V_e + V_i = V_e - \frac{d\Phi}{dt}, \tag{7.12}$$

where V_e is the potential provided by an external source.

7.5   E and B with Time-Dependent Potentials

We saw: Since $\nabla \cdot \mathbf{B} = 0$ and *"div curl = 0 "*, the field $\mathbf{B}$ can always be written as $\mathbf{B} = \nabla \times \mathbf{A}$. According to Faraday's law we then have

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}$$

$$= -\nabla \times \frac{\partial \mathbf{A}}{\partial t},$$

so that $\mathbf{E} = -\partial \mathbf{A}/\partial t + $ something, whose curl is zero, i.e.

$$\mathbf{E} = -\frac{\partial \mathbf{A}}{\partial t} - \nabla \phi. \tag{7.13}$$

This equation differs in the first term from the equation obtained in electrostatics.

7.6   Displacement Current and Maxwell Equations

We obtained above the following equation

$$\nabla \times \mathbf{H} = \mathbf{j}. \tag{7.14}$$

Taking the divergence of both sides, we have

$$\nabla \cdot (\nabla \times \mathbf{H}) = \nabla \cdot \mathbf{j}.$$

The left side is always zero, since always *"div curl"* equals zero. But $\nabla \cdot \mathbf{j}$ is not zero except in magnetostatics. Hence, here (this means not in magnetostatics, where $\partial \rho/\partial t = 0$) we have

$$\nabla \cdot \mathbf{j} + \frac{\partial \rho}{\partial t} = 0,$$

or with

$$\nabla \cdot \mathbf{D} = \rho,$$

where $\mathbf{D}$ is the dielectric displacement,

$$\nabla \cdot \mathbf{j} + \nabla \cdot \frac{\partial \mathbf{D}}{\partial t} = 0.$$

Thus, Eq. (7.14) would have to be

$$\nabla \times \mathbf{H} = \mathbf{j} + \frac{\partial \mathbf{D}}{\partial t}. \tag{7.15}$$

The expression on the right of this equation is called *Maxwell current density* and the term proportional to $\partial \mathbf{D}/\partial t$ is called the density of *Maxwell's displacement current*. This displacement current puzzled people considerably in the early days of Maxwell theory. A detailed discussion can be found in the book of O'Rahilly.[‡] On p. 89, O'Rahilly remarks: *"In a flash of mathematical insight he (Maxwell) saw, that the addition of an extra term to one of the electromagnetic equations would make an immense difference"*. What irritated people at that time was that **j** can be clearly visualised as a current (density), but $\partial \mathbf{D}/\partial t$? Pohl (cf. O'Rahilly, p. 100) writes in his book: *"By the term 'displacement current' we denote an alteration of the electric field in time, i.e. the appearance and disappearance of lines of force. The term 'current' has no doubt been historically adopted from the analogy with water. In the conduction current atoms of electricity really do move or flow. In the term 'displacement current' there remains no trace of the original meaning of the word 'current' .... We shall see later, however, that the idea of the displacement current reveals unsuspected relationships to us and vastly extends our physical conceptions of the world. Here we shall content ourselves with mentioning that the light which reaches our eyes from any source is, from the point of view of physics, nothing but a displacement current."* Citing J. J. Thomson, O'Rahilly (p. 101) says: *"We may cheerfully confess that it (i.e. the displacement current) was his (Maxwell's) 'greatest contribution to physics', i.e. as an analytical formula."*

Today we are accustomed to the concept of fields already intuitively, so that the puzzles of long-ago lose their significance. In the book of Sommerfeld[§] one reads correspondingly: *"The vector field strength* **E** *is associated with a second electric vector* **D**. *We call this electric 'excitation', but will frequently ... adhere to the customary term of 'dielectric displacement' (Maxwell's 'displacement') .... We assumed here the difference between conductor and nonconductor (dielectric) as being self-evident. In reality there is no perfect dielectric, since even the best nonconducting material can become weakly conducting as e.g. under the influence of cosmic radiation. Maxwell therefore augments the displacement current to the total current* $\mathbf{C} = \dot{\mathbf{D}} + \mathbf{j}$ *.... This step of attributing equal importance to* $\dot{\mathbf{D}}$ *and* **j** *was a fundamentally new idea of Maxwell's and was a crucial prerequisite in enabling a complete formulation of electrodynamical phenomena. Similarly for a metallic conductor he supplements the conduction current* **j** *by the addition of a hypothetical displacement current* $\dot{\mathbf{D}}$, *although the first is much more important than the second except in the case of rapidly changing fields."*

[‡]O'Rahilly [6], Vol. 1, see pp. 95–101 and pp. 81–101, see also p. 232.

[§]Sommerfeld [1], pp. 8–9.

The displacement current $\partial \mathbf{D}/\partial t$ is in general of importance only,

(a) if $\mathbf{j} = 0$, thus, for instance, in the case of propagation of light through a dielectric, or

(b) in the case of high frequencies, when $\partial \mathbf{D}/\partial t \propto$ frequency, e.g. in the case of propagation of light through a metal.

We have thus completed *Maxwell's equations*. We summarise these now. First in *differential form*.

Maxwell's equations in differential form are:

$$\nabla \cdot \mathbf{D} = \rho,$$

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

$$\nabla \cdot \mathbf{B} = 0,$$

$$\nabla \times \mathbf{H} = \mathbf{j} + \frac{\partial \mathbf{D}}{\partial t}. \tag{7.16}$$

These are supplemented by the *equation of continuity*:

$$\nabla \cdot \mathbf{j} + \frac{\partial \rho}{\partial t} = 0, \tag{7.17}$$

and the *generalised Lorentz-force* per unit volume:

$$\mathbf{F} = \rho \mathbf{E} + \mathbf{j} \times \mathbf{B}. \tag{7.18}$$

Maxwell's equations are completed by the *connecting relations*:

$$\mathbf{D} = \epsilon_0 \mathbf{E} + \mathbf{P},$$

$$\mathbf{B} = \mu_0 (\mathbf{H} + \mathbf{M}). \tag{7.19}$$

The *matter equations* apply to special cases

$$\mathbf{D} = \epsilon \mathbf{E}, \quad \mathbf{B} = \mu \mathbf{H}. \tag{7.20}$$

The *generalised Ohm's law* is the relation

$$\mathbf{j} = \sigma \mathbf{E}. \tag{7.21}$$

Maxwell's equations in integral form are:

$$\oint \mathbf{D} \cdot d\mathbf{F} = Q,$$

$$V = \oint \mathbf{E} \cdot dl = -\frac{\partial \Phi}{\partial t}, \quad \Phi = \int_F \mathbf{B} \cdot d\mathbf{F},$$

$$\int_F \mathbf{B} \cdot d\mathbf{F} = 0,$$

$$\oint \mathbf{H} \cdot dl = \int_{F \text{ fixed}} \mathbf{j} \cdot d\mathbf{F} + \frac{d}{dt} \int_{F \text{ fixed}} \mathbf{D} \cdot d\mathbf{F}. \tag{7.22}$$

The *equation of continuity* follows from

$$\frac{d}{dt} \int_V \rho dV = - \int_F \mathbf{j} \cdot d\mathbf{F}. \tag{7.23}$$

The *main assumptions*, we made in our historical and phenomenological approach in order to arrive at Maxwell's equations were the empirically discovered laws of Coulomb, Ampère/Biot–Savart and Faraday (Coulomb in the case of electric charges, Ampère and Biot and Savart for magnetic fields of currents, and Faraday for the induced electromotive force). We also assumed, without saying so, that the equations obtained originally for static systems are also valid in the case of dynamic (i.e. time-dependent) systems.
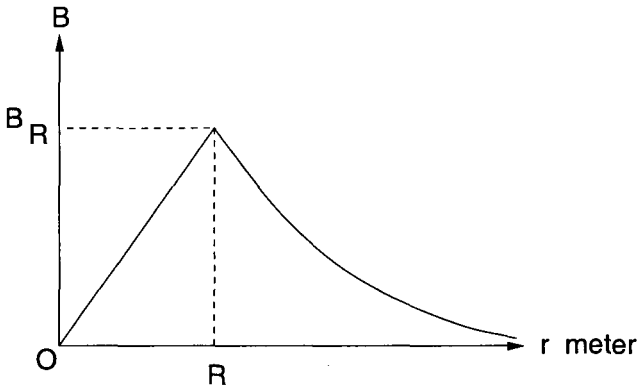
Fig. 7.3 Variation of the induced B with radius r.

Example 7.1: The displacement current in a condenser

A parallel plate condenser consists of two circularly shaped parallel plates of radius $R = 0.1$ meter. The plates are charged such that the electric field E between them is subjected to a constant time variation $dE/dt = 10^{13}$ volt meter^{-1}second^{-1}. The following relations apply: $B = \mu H, D = \epsilon E$.
(a) Calculate the displacement current between the plates of the condenser.
(b) Determine the induced magnetic field B between the plates as a function of the distance from the centre of the cylindrically symmetric condenser in a direction parallel to the plates. What is the value of B at $r = R$? (Note: $\epsilon_0 \approx 8.9 \times 10^{-12}$ ampere second/(volt meter), $\mu_0 = 4\pi \times 10^{-7}$ volt second/(ampere meter)).

Solution:
(a) The current I is given by $I = \int \mathbf{j} \cdot d\mathbf{F}$ ampere, where $\mathbf{j}$ is the current density. The density of the displacement current is $\partial \mathbf{D}/\partial t$, so that the displacement current itself is given by

$$I_D = \int_{F=\pi R^2} \frac{\partial \mathbf{D}}{\partial t} \cdot d\mathbf{F} = \epsilon_0 \frac{dE}{dt} \pi R^2 = 8.9 \times 10^{-12} \times 10^{13} \times 3.14 \times (0.1)^2 \text{ ampere} = 2.8 \text{ ampere.}$$

(b) We start with the Ampère–Maxwell-law $\boldsymbol{\nabla} \times \mathbf{H} = \mathbf{j} + \partial \mathbf{D}/\partial t$ with conduction current density $\mathbf{j} = 0$. Multiplying by μ_0 we obtain

$$\mu_0 \boldsymbol{\nabla} \times \mathbf{H} = \boldsymbol{\nabla} \times \mathbf{B} = \mu_0 \epsilon_0 \frac{\partial \mathbf{E}}{\partial t}.$$

We integrate this equation over a cross-sectional area and use Stokes' theorem. Then

$$\int_{F(C)} \boldsymbol{\nabla} \times \mathbf{B} \cdot d\mathbf{F} = \oint_{C(F)} \mathbf{B} \cdot d\mathbf{l} = \mu_0 \epsilon_0 \int_{F(C)} \frac{\partial \mathbf{E}}{\partial t} \cdot d\mathbf{F}.$$

For $r \leq R$, and C a circular path with radius r we obtain

$$B(2\pi r) = \mu_0 \epsilon_0 \frac{dE}{dt} \pi r^2, \quad \text{or} \quad B = \frac{1}{2} \mu_0 \epsilon_0 r \frac{dE}{dt}.$$

For $r \geq R$, on the other hand, we obtain, since the electric field is nonvanishing only between the plates (ignoring boundary effects),

$$B(2\pi r) = \mu_0 \epsilon_0 \frac{dE}{dt} \pi R^2, \quad \text{i.e.} \quad B = \frac{1}{2} \mu_0 \epsilon_0 \frac{R^2}{r} \frac{dE}{dt}.$$

The variation of this induced magnetic field with r is shown in Fig. 7.3.

$$B_R = \frac{1}{2} \mu_0 \epsilon_0 R \frac{dE}{dt} = \frac{1}{2} \times 4\pi \times 10^{-7} \times 8.9 \times 10^{-12} \times 0.1 \times 10^{13} = 5.6 \times 10^{-6} \text{ weber meter}^{-2}.$$

Although the displacement current is comparatively large, the magnetic field is rather small.

7.7 Poynting Vector and Conservation of Energy

The work W_0 done by charges per unit time in the presence of an electro-magnetic field is the *power*

$$\frac{dW_0}{dt} \equiv \frac{d \text{ work}}{dt}$$

$$= \sum_i (\text{force acting on charge } i) \cdot (\text{velocity of charge } i)$$

$$= \int \mathbf{v} \cdot (\rho \mathbf{E} + \mathbf{j} \times \mathbf{B}) dV. \tag{7.24}$$

Since $\mathbf{j} = \rho \mathbf{v}$, we have $\mathbf{v} \cdot \mathbf{j} \times \mathbf{B} = 0$. This implies (using $\mathbf{D} = \epsilon \mathbf{E}, \mathbf{B} = \mu \mathbf{H}$ in the second-last step)

$$\frac{dW_0}{dt} = \int \rho \mathbf{v} \cdot \mathbf{E} dV \equiv \int \mathbf{j} \cdot \mathbf{E} dV = \int dV \mathbf{E} \cdot \left(\boldsymbol{\nabla} \times \mathbf{H} - \frac{\partial \mathbf{D}}{\partial t} \right)$$

$$= \int dV (\mathbf{E} \cdot \boldsymbol{\nabla} \times \mathbf{H} - \mathbf{H} \cdot \boldsymbol{\nabla} \times \mathbf{E})$$

$$- \int dV (\mathbf{E} \cdot \frac{\partial \mathbf{D}}{\partial t} - \mathbf{H} \cdot \underbrace{\boldsymbol{\nabla} \times \mathbf{E}}_{-\frac{\partial \mathbf{B}}{\partial t}})$$

$$= \int dV \boldsymbol{\nabla} \cdot (\mathbf{E} \times \mathbf{H}) - \frac{1}{2} \frac{\partial}{\partial t} \int dV (\mathbf{E} \cdot \mathbf{D} + \mathbf{H} \cdot \mathbf{B})$$

$$= - \int (\mathbf{E} \times \mathbf{H}) \cdot d\mathbf{F} - \frac{\partial W}{\partial t}, \tag{7.25}$$

where we assumed that $\mathbf{D} \propto \mathbf{E}$ and $\mathbf{B} \propto \mathbf{H}$ in time dependence. The vector

$$\mathbf{S} = \mathbf{E} \times \mathbf{H} \quad \text{watt meter}^{-2} \tag{7.26}$$

is called *Poynting vector*. The expression for W is the *electromagnetic energy*,

$$W = \frac{1}{2} \int_V dV (\mathbf{E} \cdot \mathbf{D} + \mathbf{H} \cdot \mathbf{B}). \tag{7.27}$$

We have therefore:

$$\underbrace{\frac{d \text{ work}}{dt} = \int dV \mathbf{j} \cdot \mathbf{E}}_{\text{Ohmic power}} = - \underbrace{\int_{F(V)} \mathbf{S} \cdot d\mathbf{F}}_{d\mathbf{F} \text{ directed to outside}} \quad - \frac{\partial W}{\partial t}. \tag{7.28}$$

Thus *the work done per unit time is equal to the loss of electromagnetic energy per unit time* $(\partial W/\partial t)$ *plus gain or loss of energy through the walls of the system* $(\int \mathbf{S} \cdot d\mathbf{F})$. It should be noted that only the integral over the Poynting vector is a measurable quantity, i.e. the integral over a closed surface, not the transport of energy through some part of the volume, since only the total energy is a conserved quantity. Furthermore one should note that the definition of the Poynting vector involves some arbitrariness because we can always add to $\mathbf{S}$ a vector $\mathbf{T}$, whose divergence is zero, so that

$$\int_F \mathbf{T} \cdot d\mathbf{F} = \int_{V(F)} dV \boldsymbol{\nabla} \cdot \mathbf{T} = 0.$$

In the case of the "ideal conductor" the conduction electrons do no work, i.e. $\sigma = \infty$ and so $R = 0$, and $\mathbf{E} = -\boldsymbol{\nabla}\phi = 0$, so that $\mathbf{j} = \sigma\mathbf{E} = \text{finite}$, but $\int \mathbf{j} \cdot \mathbf{E} dV = 0$. (This is the reverse case to that of dielectrics, for which in the ideal case $\sigma = 0$).

In the case *with* resistance $dW_0/dt \neq 0$. In the "ideal conductor", we have $\int \mathbf{j} \cdot \mathbf{E} dV = 0$, since there $\mathbf{E} = 0$, i.e.

$$\int_{F(V)} \mathbf{S} \cdot d\mathbf{F} = -\frac{\partial W}{\partial t},$$

which implies: energy transmitted through the walls per unit time = loss of electromagnetic energy per unit time. The (somewhat paradox) example usually given to the above is that of *transport of energy in a conduction wire*.

7.7.1   Application: The Conduction Wire

The conduction wire we consider now is not made of ideal conductor material, since otherwise we would have to have $V = IR = 0$. Thus the conduction electrons suffer resistance and hence perform work $\neq 0$. The diagram in Fig. 7.4 shows, that the electric field $\mathbf{E}$ here points in the direction of the current and the vector $\mathbf{H}$ to its right into the paper.
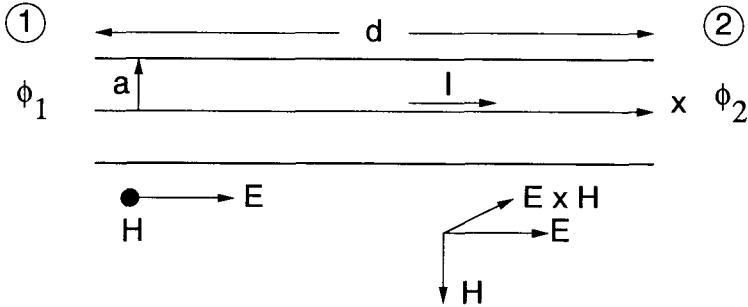
Fig. 7.4 The conduction wire.

We can obtain the direction of the Poynting vector with the "thumb rule" ($\mathbf{E}$ along the index finger, $\mathbf{H}$ along the middle finger and so $\mathbf{E} \times \mathbf{H}$ parallel to the thumb), as indicated in the figure. We have *in the wire*:

$$\mathbf{E} = -\boldsymbol{\nabla}\phi, \quad E_{\text{tang}} \text{ continuous,}$$

and

$$\int_1^2 \mathbf{E} \cdot dl = -\int_1^2 \boldsymbol{\nabla}\phi \cdot dl$$
$$= -\int_1^2 d\phi = -(\phi_2 - \phi_1) = \phi_1 - \phi_2$$
$$\equiv V.$$

But $E_{\text{tangential}}$ is continuous, as we saw. Hence also on the surface from outside: $E = V/d$. On the other hand, from

$$\boldsymbol{\nabla} \times \mathbf{H} = \mathbf{j} + \frac{\partial \mathbf{D}}{\partial t},$$

follows or (neglecting $\partial \mathbf{D}/\partial t$) from

$$\oint \mathbf{H} \cdot dl = \int \mathbf{j} \cdot d\mathbf{F} = I,$$

that

$$H = \frac{I}{2\pi a}.$$

Since $\mathbf{E} \perp \mathbf{H}$,

$$\mathbf{E} \times \mathbf{H} = \frac{V}{d}\frac{I}{2\pi a} = \frac{VI}{2\pi ad}.$$

Thus the amount of energy transported *into* the wire per unit time is:

$$-\int (\mathbf{E} \times \mathbf{H}) \cdot d\mathbf{F} = \frac{IV}{2\pi ad}2\pi ad = IV$$

= the amount of energy, that the wire loses per unit time in the form of heat. We see therefore: Energy is radiated away, but at the same time absorbed, *in such a way that the current is kept constant* (this latter point is not always given sufficient emphasis, and then makes the problem appear paradoxical).

The fields $\mathbf{E}$ and $\mathbf{H}$ are a consequence of the current I or the potential difference V; it is *not* possible to have I and V without $\mathbf{E}$ and $\mathbf{H}$, or the converse. It is thus a question of one's point of view whether one says, the power $-IV$ or the energy transported per unit time IV are a consequence of I and V or a consequence of the fields $\mathbf{E}$ and $\mathbf{H}$. The easiest way to understand this problem is probably as an example of the conservation of energy. For a fixed current I and thus in a static case $(\partial W/\partial t = 0)$ the loss of energy as a result of the resistance R is compensated by radiation *into* the conductor, i.e.

$$\int \mathbf{S} \cdot (-d\mathbf{F}) = IV = I(\phi_1 - \phi_2).$$

We can see this equivalence as follows:

$$\int_1^2 \mathbf{E} \cdot \mathbf{j} dV = -\int \mathbf{j} \cdot \nabla \phi dV = -\int_1^2 \mathbf{j} \cdot \left(\mathbf{e}_x \frac{d\phi}{dx} \underbrace{dx \, d\mathbf{F}}_{dV} \right)$$

$$= -\int_1^2 \int (\mathbf{j} \cdot d\mathbf{F}) d\phi = -\int_1^2 I d\phi = -I(\phi_2 - \phi_1)$$

$$= I(\phi_1 - \phi_2) = IV.$$

7.7.2 The Field Momentum Density

Let a volume V be given with charge distribution (i.e. density) ρ. The *force that the field exerts on the charges*, is the external force in Newton's equation

for the charged particle, i.e.

$$\frac{d\,(\text{mechanical momentum})}{dt} = \text{external force}$$

$$= \int \left(\mathbf{E} + \mathbf{v} \times \mathbf{B}\right)\rho dV = \int \left(\underbrace{\rho}_{\boldsymbol{\nabla} \cdot \mathbf{D}}\,\mathbf{E} + \underbrace{\mathbf{j}}_{\boldsymbol{\nabla} \times \mathbf{H} - \frac{\partial \mathbf{D}}{\partial t}} \times \mathbf{B}\right)dV$$

$$= \int \left[\mathbf{E}\boldsymbol{\nabla} \cdot \mathbf{D} + \left(\boldsymbol{\nabla} \times \mathbf{H} - \frac{\partial \mathbf{D}}{\partial t}\right) \times \mathbf{B}\right]dV$$

$$= -\int \left[\frac{\partial \mathbf{D}}{\partial t} \times \mathbf{B} + \mathbf{D} \times \frac{\partial \mathbf{B}}{\partial t}\right]dV$$

$$+ \int [\mathbf{E}\boldsymbol{\nabla} \cdot \mathbf{D} + \mathbf{H}\underbrace{\boldsymbol{\nabla} \cdot \mathbf{B}}_{0} + (\boldsymbol{\nabla} \times \mathbf{H}) \times \mathbf{B} + (\underbrace{\boldsymbol{\nabla} \times \mathbf{E}}_{-\frac{\partial \mathbf{B}}{\partial t}}) \times \mathbf{D}]dV.$$

It can be shown that the remaining terms can be combined in the following way:

$$\left\{\frac{d(\text{mechanical momentum})}{dt}\right\}_j = -\frac{\partial}{\partial t}\int (\mathbf{D} \times \mathbf{B})_j dV + \int \frac{\partial}{\partial x_i} T_{ij} dV,$$

where

$$T_{ij} = (D_i E_j + B_i H_j) - \delta_{ij}\frac{1}{2}(\mathbf{D} \cdot \mathbf{E} + \mathbf{B} \cdot \mathbf{H}).$$

This expression is equal to T_{ji} for $D = \epsilon E, B = \mu H$. The expression T_{ij} is called *Maxwell stress tensor*. In a self-evident way we have (on application of Gauss' divergence theorem)

$$\frac{d}{dt}(\mathbf{p}^{\text{mech}} + \mathbf{P}^{\text{field}})_j - \int_F T_{ij}(d\mathbf{F})_i = 0. \tag{7.29}$$

The term on the right represents the *flow* of momentum through the surface F. If this surface is shifted to infinity, so that the term on the right is zero, we have

$$\mathbf{p}^{\text{mech}} + \mathbf{P}^{\text{field}} = \text{constant},$$

where $\mathbf{P}^{\text{field}} = \int dV \mathbf{p}^{\text{field}}$, $\mathbf{p}^{\text{field}} = \mathbf{D} \times \mathbf{B} = $ volume density of the field momentum $\mathbf{P}^{\text{field}}$. This is the *law of conservation of momentum in elec-trodynamics*. Note that the field momentum density is not given by the Poynting vector! The expression

$$K_j = T_{ij}(d\mathbf{F})_i$$

is the j-component of a force acting on the area $d\mathbf{F}$. Since pressure = force/area, we see that

$$-\frac{K_j n_j}{dF} = -T_{ij} n_i n_j$$

(n_i: i-component of the unit vector) is the pressure of the radiation, i.e. the *radiation pressure*. The tails of comets are observed to point away from the sun; this is a consequence of the pressure of the radiation from the sun. We have here really been considering the microscopic field momentum. Considerations of the macroscopic bulk field momentum lead to difficulties which are pointed out in Jackson [3] and discussed in more detail in literature cited there.

Example 7.2: Separating a charge-anticharge pair

A charge q, originally together with a charge $-q$ at rest at some point, is moved away from the charge $-q$. Show that the mechanical momentum $\mathbf{p}^{\text{mech}}$ of the charge q is exactly cancelled by the field momentum.

Solution: The Lorentz force acting on a charge q in a flux density $\mathbf{B}$ (due to some current or moving charge) is $\mathbf{F}_q = q\mathbf{v} \times \mathbf{B}$. We give the charge q a mechanical momentum $\mathbf{p}^{\text{mech}}$ by pulling it away from $-q$ with a force equal and opposite to that of the Lorentz force. The charge $-q$, moving relatively in the opposite direction, provides the flux density $\mathbf{B}$. In fact from the Biot–Savart law we see that the fields $\mathbf{B}$ generated by the motion of the two opposite charges moving in opposite directions are opposite, i.e. $\mathbf{B}_q = -\mathbf{B}_{-q}$. In moving the charge it generates field momentum which has to be such that the total momentum is conserved, i.e. equal to that initially, i.e. zero. Thus from $\mathbf{F}_q$ and Newton's equation ($\mathbf{v} = d\mathbf{x}/dt$)

$$\mathbf{p}^{\text{mech}} = -q \int d\mathbf{x} \times \mathbf{B}_{-q}.$$

From the Gauss law we obtain, now with only the one charge q to be considered, $q = \int \mathbf{D} \cdot d\mathbf{a} = \int D_l da_l$, so that with $dx_j da_l = \delta_{jl} dV$ for element of area da, we have

$$p_i^{\text{mech}} = -q \int \epsilon_{ijk} dx_j B_k = -\int \epsilon_{ijk} dx_j B_k D_l da_l$$

$$= -\int \epsilon_{ijk} B_k D_l \delta_{jl} dV = -\int \epsilon_{ijk} B_k D_j dV$$

$$= -\int dV (\mathbf{D} \times \mathbf{B})_i.$$

Hence $\mathbf{p}^{\text{mech}} + \mathbf{P}^{\text{field}} = 0$. The components of the field momentum perpendicular to the line joining the charges cancel and leave only the component along that line.

7.8 Further Examples

Example 7.3: The variable conductor loop

A rectangular current loop with constant width $y = b$ consists of three fixed conducting rods and one movable conducting rod which allows to change the length x of the loop. This movable rod is moved with constant velocity $\mathbf{v}$ in the direction of length of the rectangular loop (i.e. in the direction of positive x). The time-dependent field of induction $\mathbf{B} = \mathbf{e}_z B_0 \cos \omega t$ acts in the overall space perpendicular to the plane of the loop.

(a) Calculate the total induced electromotive force V in the loop.

(b) What is obtained for V, if the movable conducting rod is held fixed at a distance x from the opposite side of the rectangle, but the entire current loop moves with velocity $\mathbf{v} = v\mathbf{e}_x$?

Solution:

(a) We have here a case, in which the area of the loop, as well as the field $\mathbf{B}$ vary with time. The variations of both contribute to the induced potential V. We use Faraday's law

$$\int_{F(C)} \boldsymbol{\nabla} \times \mathbf{E} \cdot d\mathbf{F} = -\frac{d}{dt} \int_{F(C)} \mathbf{B} \cdot d\mathbf{F},$$

i.e.

$$V = \oint_C \mathbf{E} \cdot d\mathbf{l} = -\frac{d}{dt} \int_{F(C)} \mathbf{B} \cdot d\mathbf{F}.$$

If the area F to be integrated over varies with time, we have to differentiate with respect to time here. We assume first that $\mathbf{B}$ is constant with respect to time. (Here only so that we do not have to drag its time derivative all the way along!). We then have

$$\frac{d}{dt} \int_{\mathbf{F}(t)} \mathbf{B} \cdot d\mathbf{F} = \lim_{\delta t \to 0} \left[\frac{\int_{\mathbf{F}(t+\delta t)} \mathbf{B} \cdot d\mathbf{F} - \int_{\mathbf{F}(t)} \mathbf{B} \cdot d\mathbf{F}}{\delta t} \right]$$

$$= \lim_{\delta t \to 0} \frac{1}{\delta t} \left[\int_{\mathbf{F}(t)+\delta t \frac{d\mathbf{F}}{dt}} \mathbf{B} \cdot d\mathbf{F} - \int_{\mathbf{F}(t)} \mathbf{B} \cdot d\mathbf{F} \right]$$

$$= \lim_{\delta t \to 0} \left[\frac{1}{\delta t} \delta t \frac{d\mathbf{F}}{dt} \cdot \frac{d}{d\mathbf{F}} \int \mathbf{B} \cdot d\mathbf{F} \Big|_{\mathbf{F}(t)} \right]$$

$$= (\mathbf{B})_{\mathbf{F}} \cdot \frac{d\mathbf{F}}{dt}. \tag{7.30}$$
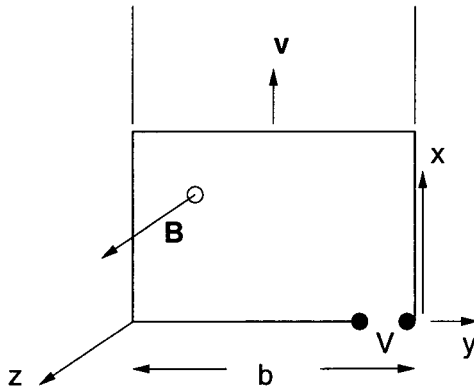
Fig. 7.5   The conducting loop with movable rod.

Including now also the time dependence of $\mathbf{B}$, we have

$$\frac{d}{dt} \int_{\mathbf{F}(t)} \mathbf{B} \cdot d\mathbf{F} = \int_F \frac{\partial \mathbf{B}(t)}{\partial t} \cdot d\mathbf{F} + (\mathbf{B})_{\mathbf{F}(t)} \cdot \frac{d\mathbf{F}}{dt}.$$

In the case of our example we have

$$V = \oint_{C(F)} \mathbf{E} \cdot d\mathbf{l} = -\frac{d}{dt} \int_{F(C)} \mathbf{B} \cdot d\mathbf{F} = -\mathbf{B} \cdot \frac{d\mathbf{F}(C)}{dt} - \int \frac{\partial \mathbf{B}}{\partial t} \cdot d\mathbf{F}.$$

We choose the area $\mathbf{F}$ to point in the direction of z, as indicated in Fig. 7.5 ($\mathbf{F} \parallel \mathbf{e}_z$). The voltage V is given by:

$$V = -\mathbf{B}bv \cdot (\mathbf{e}_y \times \mathbf{e}_x) + \int_0^x \omega B_0 \sin \omega t \underbrace{b\, dx}_{dF}$$

and hence by

$$\begin{aligned}
V &= -B_0 vb \cos \omega t + \omega B_0 bx \sin \omega t \\
&= B_0 b\{x\omega \sin \omega t - v \cos \omega t\} \\
&= B_0 b\sqrt{v^2 + x^2\omega^2} \sin(\omega t - \delta),
\end{aligned} \tag{7.31}$$

where $\delta = \tan^{-1}(v/x\omega)$, i.e. $\cos \delta = x\omega/\sqrt{v^2 + x^2\omega^2}$ and correspondingly for $\sin \delta$.

(b) We now assume, that the movable part of the rectangular loop is fixed to the other part of the loop (at position x) so that the area of the loop remains constant in time but the entire current loop moves with velocity $\mathbf{v}$ in the direction of $\mathbf{e}_x$. Then

$$\begin{aligned}
V &= -\frac{d}{dt} \int_{F(C)} \mathbf{B} \cdot d\mathbf{F} = -\int_{F(C)} \frac{d\mathbf{B}}{dt} \cdot d\mathbf{F} \\
&= -\int_{F(C)} \left[\frac{\partial \mathbf{B}}{\partial t} + (\mathbf{v} \cdot \boldsymbol{\nabla})\mathbf{B}\right] \cdot d\mathbf{F},
\end{aligned} \tag{7.32}$$

where $\mathbf{v}$ represents the velocity of the charges in $\mathbf{B}$. But now (for $\mathbf{v}$ constant with $\boldsymbol{\nabla} \cdot \mathbf{B} = 0$)

$$\boldsymbol{\nabla} \times (\mathbf{B} \times \mathbf{v}) = (\mathbf{v} \cdot \boldsymbol{\nabla})\mathbf{B} - (\mathbf{B} \cdot \boldsymbol{\nabla})\mathbf{v} + \mathbf{B}(\boldsymbol{\nabla} \cdot \mathbf{v}) - \mathbf{v}(\boldsymbol{\nabla} \cdot \mathbf{B}) = (\mathbf{v} \cdot \boldsymbol{\nabla})\mathbf{B}.$$

For given $\mathbf{B} = B\mathbf{e}_z$, we have $\boldsymbol{\nabla}B_z = 0$. Although we know that this contribution vanishes, we drag it along, so that:

$$\begin{aligned}
V &= -\int_{F(C)} \left[\frac{\partial \mathbf{B}}{\partial t} + \boldsymbol{\nabla} \times (\mathbf{B} \times \mathbf{v})\right] \cdot d\mathbf{F} \\
&= -\int_{F(C)} \frac{\partial \mathbf{B}}{\partial t} \cdot d\mathbf{F} - \oint_{C(F)} (\mathbf{B} \times \mathbf{v}) \cdot d\mathbf{l} \\
&= -\int_{F(C)} \frac{\partial \mathbf{B}}{\partial t} \cdot d\mathbf{F} + \oint_{C(F)} (\mathbf{v} \times \mathbf{B}) \cdot d\mathbf{l}.
\end{aligned} \tag{7.33}$$

In our case

$$\mathbf{v} \times \mathbf{B} = vB\mathbf{e}_x \times \mathbf{e}_z = vB\mathbf{e}_y,$$

$$\oint_{C(F)} (\mathbf{v} \times \mathbf{B}) \cdot d\mathbf{l} = vB \oint \mathbf{e}_y \cdot d\mathbf{l} = vB(b - b) = 0$$

(since the field $\mathbf{B}$ is constant, it has the same value along b as along $-b$; for this reason these contributions cancel). As expected, also the contribution of $\mathbf{v} \times \mathbf{B}$ is zero, and

$$V = -\int \frac{\partial \mathbf{B}}{\partial t} \cdot d\mathbf{F} = \omega B_0 bx \sin \omega t = B_0 b(\omega x) \sin \omega t.$$

This means, the induced potential is independent of the velocity. This is, what one expects, because in the rest frame of the conductor one observes — in the case of the constant field $\mathbf{B}$ — at every instant of time the same situation.
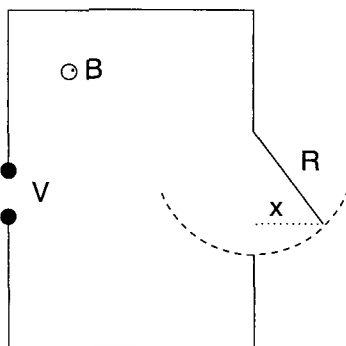
Fig. 7.6 The wire pendulum.

Example 7.4: The wire pendulum

A wire pendulum (of length R) oscillates with velocity $\dot{x} = \omega d \cos \omega t$, where d is the largest horizontal deflection of the pendulum as shown in Fig. 7.6. The constant magnetic field of B weber/meter^2 points out of the plane of the pendulum. Determine the induced voltage V.

Solution: We start with Faraday's law

$$V = -\frac{d}{dt} \int \mathbf{B} \cdot d\mathbf{F} = \frac{d}{dt}[B \times \text{area}(t)]$$

$$= \frac{d}{dt}[B(\text{constant area} + \text{varying area})].$$

Set $\sin \theta = x/R, \sin \theta_2 = d/R$, so that

$$\frac{\sin \theta}{\sin \theta_2} = \frac{x}{d}, \quad x(t) = \frac{d \sin \theta}{\sin \theta_2} = \frac{d \sin \omega t}{\sin \theta_2}.$$

The variable part of the area is then $\simeq \frac{1}{2}Rx(t)$. It follows that the requested voltage V is

$$V = \frac{d}{dt}\left[\frac{BR}{2}\frac{d \sin \omega t}{\sin \theta_2}\right] = \frac{1}{2}B\omega R^2 \cos \omega t.$$

Example 7.5: The revolving coil

In Example 6.3 the field in the gap of a circularly shaped electromagnet was calculated. A small circularly shaped coil of radius b and n turns revolves fully within the gap of the electromagnet with angular velocity ω about a diameter perpendicular to the magnetic field which is assumed to be homogeneous. R is the resistance and L the self-inductance of the coil. Formulate an equation from which the current induced in the coil can be calculated.

Solution: We begin with Faraday's law,

$$\boldsymbol{\nabla} \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}.$$

The induced voltage V_i is

$$V_i = \int \mathbf{E} \cdot d\mathbf{l} = \int \boldsymbol{\nabla} \times \mathbf{E} \cdot d\mathbf{F} = -\frac{d}{dt}\int \mathbf{B} \cdot d\mathbf{F}.$$
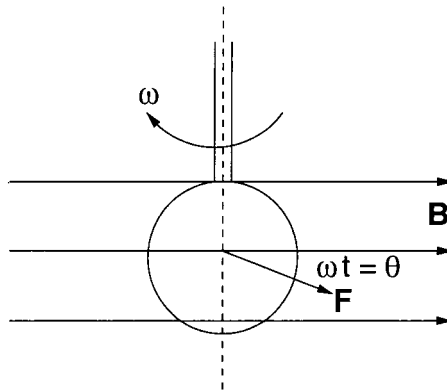
Fig. 7.7 The revolving coil.

The induced flux is $\phi = \int \mathbf{B} \cdot d\mathbf{F}$, i.e. $V_i = -d\phi/dt$. The inductances L_{kj} of the circuit are defined by $\phi_k = \sum_j L_{kj} I_j$ (cf. Chapter 8). Here we have two circuits – the one of the coil, as well as another one which is responsible for the varying flux through the area of the coil, i.e.

$$\phi = \underbrace{LI}_{\text{self-induction part}} + \underbrace{\tilde{\phi}}_{\text{the varying flux}} .$$

Here $\tilde{\phi} = nB \times$ projection of area F of one turn of the coil perpendicular to the magnetic field. Since

$$IR = \underbrace{V_{\text{generator}}}_{0} - \frac{d\phi}{dt},$$

we have

$$IR = -L\dot{I} - \frac{d\tilde{\phi}}{dt}.$$

In other words,

$$IR + L\dot{I} = -\frac{d\tilde{\phi}}{dt},$$

where $\tilde{\phi} = nBF \sin \omega t$, with B the field through area F and $F = \pi b^2$, as indicated in Fig. 7.7.

Example 7.6: The rectangular conductor loop

A magnetic field varying with time is given by the following expression: $\mathbf{B} = \mathbf{e}_y B_0 \cos \omega t$, with B_0 constant. Determine the voltage induced in a rectangular conductor loop in the (x, z)-plane with side lengths a in the direction of x, and b in the direction of z.

Solution: Let the unit vector $\mathbf{e}_y$ point into the paper. Then the magnetic flux into the paper is

$$\Phi = \int \mathbf{B} \cdot d\mathbf{F} = \int_{x=0}^{a} \int_{z=0}^{b} B_0 \cos \omega t \, \mathbf{e}_y \cdot dx dz \mathbf{e}_y,$$

i.e. $\Phi = abB_0 \cos \omega t$. The induced voltage is therefore

$$V = \oint_C \mathbf{E} \cdot d\mathbf{l} = -\frac{d}{dt} \Phi$$

$$= -\frac{d}{dt}[B_0 ab \cos \omega t]$$

$$= abB_o \omega \sin \omega t.$$
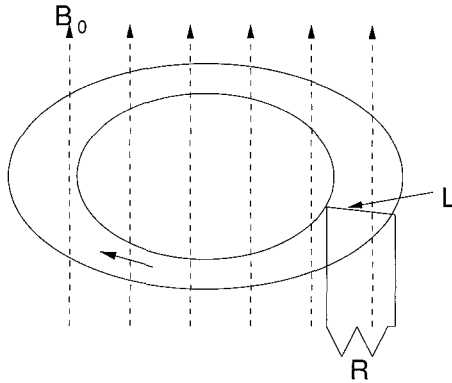
Fig. 7.8 The rails of the trolley with induced circuit.

Example 7.7: The metal trolley[¶]

A metal trolley (mass m) is placed on two parallel, circular, conducting rails (a distance of L meters apart), as indicated in Fig. 7.8. A uniform magnetic field B_0 points vertically out of the plane. With a rocket engine the trolley is given a constant thrust (i.e. force) F_0. The resistance between the rails is R. Determine the voltage V induced in the circuit linking the trolley with the resistance, and discuss its time dependence (Hint: It is advisable to start from Ampère's law of force).

Solution: Ampère's law of force is

$$\mathbf{F} = I \int d\mathbf{s} \times \mathbf{B}(\mathbf{r}).$$

In the present case we need to know the forces on the right side of Newton's equation which are responsible for the motion of the trolley, i.e. in

$$m\frac{dv}{dt} = F_0 - F_{\text{applied}}.$$

We obtain the force F_{applied} from Ampère's law (see discussion after Eq. (5.18)) as $B_0 I L$, by considering the current element with resistance R between the rails and identifying $d\mathbf{s}$ with L. We obtain the current I from Faraday's law or correspondingly from the formula for the induced potential difference,

$$V = \oint \left(d\mathbf{s} \times \frac{d\mathbf{l}}{dt} \right) \cdot \mathbf{B} = B_0 L v, \quad \text{i.e.} \quad I = \frac{V}{R} = \frac{B_0 v L}{R},$$

so that

$$m\frac{dv}{dt} = F_0 - \frac{B_0^2 L^2 v}{R} \quad \text{or} \quad m\frac{dv}{dt} + \frac{B_0^2 L^2 v}{R} = F_0.$$

This is an inhomogeneous differential equation of the first order which is readily solved (cf. Eq. (8.17)) with the initial condition $v = 0$ at time $t = 0$. One obtains after simple integrations the physically plausible result

$$v = \frac{RF_0}{L^2 B_0^2} \left[1 - e^{-t/(mR/L^2 B_0^2)} \right].$$

We see that in the course of time $(t \to \infty)$ the velocity becomes constant.

[¶]See also *The Electromagnetic Problem Solver* [21], Example 11–36, p. 621.

Chapter 8

Applications to Coils and Circuits

8.1   Introductory Remarks

In the following we consider circuits of currents, we define inductances L and consider coils and solenoids, and return to gauge transformations. Again examples are given. An important result will be that the product of self-inductance L and capacity C of a transmission line is equal to the product of dielectric constant and magnetic permeability of the surrounding medium. In Chapter 9 it will be shown that the latter product is, in fact, $1/c^2$, where c is the velocity of light. Thus the quantity $1/\sqrt{LC}$ represents the velocity of the electromagnetic wave travelling along the transmission line.

8.2   Inductances L_{ij}

We assume that several current-carrying circuits are given. For each, say the k-th circuit, we have

$$I_k R_k = (V_e)_k - \frac{\partial \Phi_k}{\partial t}. \tag{8.1}$$

Here $\Phi_k = \int_{F_k} \mathbf{B} \cdot d\mathbf{F}$ is the magnetic flux through the k-th conductor with enclosed area F_k, and external or generator voltage V_e. Equation (8.1) represents a combination of the usual Ohm's law with Faraday's law. If now the current I_j in circuit j is varied with time, then the magnetic flux through circuit k varies and hence according to the relation $(V_{\text{induced}})_k = -\partial \Phi_k / \partial t$ also the induced voltage $(V_{\text{induced}})_k$, and we can write (using the superposition

principle, i.e. the summability of contributions of all circuits)

$$(V_{\text{induced}})_k = -\sum_j L_{kj} \frac{\partial I_j}{\partial t},\tag{8.2}$$

where the coefficients of inductance, L_{kj}, are constants. Hence we can write

$$-\frac{\partial \Phi_k}{\partial t} = -\sum_j L_{kj} \frac{\partial I_j}{\partial t},$$

or

$$\Phi_k(t) = \sum_j L_{kj} I_j(t).\tag{8.3}$$

Thus we have, for instance, $\Phi_2(I_1) = L_{21} I_1$. If the current $I = I(t)$ varies with time t, also $\mathbf{j}$ and ρ must vary with t, i.e.

$$I(t) = \int \mathbf{j}(\mathbf{r}, t) \cdot d\mathbf{F}(\mathbf{r}) = \int \rho(\mathbf{r}, t) \frac{d\mathbf{s} \cdot d\mathbf{F}(\mathbf{r})}{dt}.$$

The equation of continuity must be preserved (in view of charge conservation), i.e.

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

We can obtain the relation (8.3) also as follows, thereby determining the coefficients L_{kj}. We have

$$\Phi_k(t) = \int_{F_k} \mathbf{B}(\mathbf{r}, t) \cdot d\mathbf{F} = \int_{F_k} \boldsymbol{\nabla} \times \mathbf{A}(\mathbf{r}, t) \cdot d\mathbf{F}(\mathbf{r}) \stackrel{\text{Stokes}}{=} \oint_k \mathbf{A}(\mathbf{r}, t) \cdot d\mathbf{s}_k.$$

But, as we saw previously, the vector potential at a point $\mathbf{r}$ resulting from current densities $\mathbf{j}(\mathbf{r}, t)$ is given by the following expression:

$$\mathbf{A}(\mathbf{r}, t) = \frac{\mu}{4\pi} \int \frac{\mathbf{j}(\mathbf{r}', t) d\mathbf{r}'}{|\mathbf{r} - \mathbf{r}'|} = \frac{\mu}{4\pi} \sum_i I_i(t) \int \frac{d\mathbf{s}_i}{|\mathbf{r} - \mathbf{r}_i|}$$

with $\mathbf{j}(\mathbf{r}', t) d\mathbf{r}' = I(t) d\mathbf{s}(\mathbf{r}') = \mathbf{I}(t) ds(\mathbf{r}')$, and ds_i is the length of the current element at $\mathbf{r}_i$. Inserting this into the preceding expression, we obtain

$$\Phi_k(t) = \sum_i L_{ki} I_i(t), \quad L_{ki} = \frac{\mu}{4\pi} \int_i \int_k \frac{d\mathbf{s}_i \cdot d\mathbf{s}_k}{|\mathbf{r}_k - \mathbf{r}_i|}.\tag{8.4}$$

Obviously $L_{ki} = L_{ik}$. The coefficient L_{kk} is called the *self-inductance* of conductor k. In its case, one has to take into account its cross section, since

otherwise the expression $1/|\mathbf{r}_k - \mathbf{r}_i|$ becomes problematic. Inserting Eq. (8.4) into Eq. (8.1), we obtain a relation of great importance for circuits:

$$I_k R_k + \sum_i L_{ki} \dot{I}_i = (Ve)_k. \tag{8.5}$$

The coefficients L_{ik} for $i \neq k$ are called *mutual inductances.* In the following examples we calculate inductances of both types, self-inductances and mutual inductances.

8.3 Examples
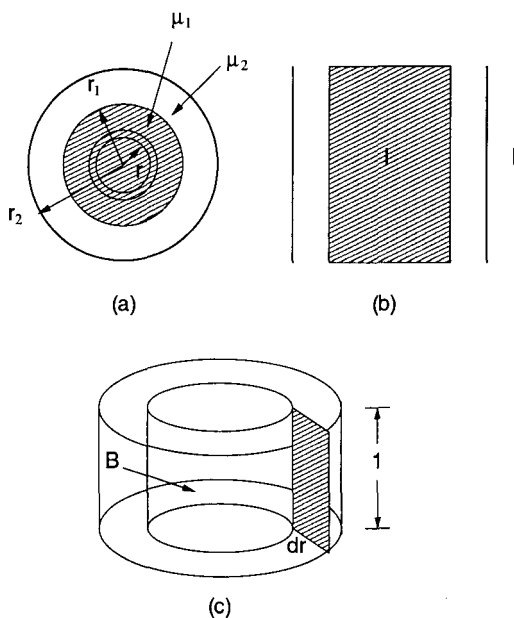
Fig. 8.1 The coaxial cable: (a) transverse cross section, (b) longitudinal cross section, (c) view to direction of flux.

Example 8.1: The solenoid
Determine the self-inductance of a long solenoid.

Solution: We let the length of the solenoid be l and assume it has n turns per unit length, and cross-sectional area S. The current in the solenoid is taken to be I. The magnetic flux through nl turns is nl times the flux through one turn and is $\Phi = \int \mathbf{B} \cdot d\mathbf{F} = \mu_0 H S n l$. The magnetic field strength H is according to Example 5.6 $H = nI$. Thus the self-inductance L_{11} is given by

$$L_{11} = \frac{\Phi}{I} = n^2 \mu_0 S l \quad \text{H,}$$

where H stands for *henry*, the unit in the MKSA-system of units, in which magnetic flux is given in *weber* (Wb). Evidently the condition for the existence of a nonvanishing inductance L is the existence of a nonvanishing flux Φ. If the current carrying conductors are "thin", i.e. of practically vanishing cross-sectional area, only the flux Φ resulting from the turns of the conductor is of practical significance. On the other hand, if the conductor is not "thin", i.e. "thick", the flux through its cross-sectional area also has to be taken into account. This, of course, complicates the considerations. We therefore consider an example of this case.

Example 8.2: Self-inductance of a coaxial cable
Figure 8.1 shows the cross section through the cable with total current zero whose self-inductance is to be calculated (see also Example 8.5).

Solution: (Important aspects of coaxial cables will be treated under the topic of Wave Guides and Resonators). We consider an element of the cable of unit length. The inner conductor carrying current I is taken to have radius r_1 and magnetic permeability μ_1. The outer conductor carrying current I in the opposite direction is taken to be shell-like, i.e. thin, and of radius r_2. The space in-between is filled with a substance of magnetic permeability μ_2. We have

$$I = \int_{r=0}^{r=r_1} \mathbf{j} \cdot d\mathbf{F}, \quad I = \pi r_1^2 j.$$

We consider a circular path of radius $r, r \leq r_1$, and apply the integral form of Ampère's law. Then

$$\oint_{\text{rad } r} \mathbf{H} \cdot dl = \int_{F=\pi r^2} \mathbf{j} \cdot d\mathbf{F}, \quad \text{i.e.} \quad 2\pi r H_r = \pi r^2 j = \frac{\pi r^2}{\pi r_1^2} I,$$

i.e. the field is determined by this inner fraction of the current I and is given by

$$H_r = \frac{Ir}{2\pi r_1^2}.$$

The lines of force of **B**, are circles around the axis of the cable, as indicated in Fig. 8.1 (recall that according to the Biot–Savart law $d\mathbf{B} \propto I d\mathbf{s} \times \mathbf{r}$). The magnetic flux through the area of breadth dr and of unit length is

$$\delta\Phi = \mathbf{B} \cdot d\mathbf{F} = \mu_1 H_r dr \times 1 = \frac{\mu_1 I r dr}{2\pi r_1^2}.$$

We recall that in the case of a solenoid with (say) n turns per unit length, we calculate the flux through one turn carrying current I and multiply by n to obtain the flux through n turns. In the present case we do not have multiples of I, but rather a fraction of I. Thus instead of multiplying $\delta\Phi$ by a number n representing a number of turns, we have to multiply $\delta\Phi$ by the fraction $\pi r^2/\pi r_1^2$, since the flux results only from this fraction of current. In order to understand this point better it may help to visualise the inner part of the cable as a section of a torus-like structure and to compare this with our treatment of the solenoid. The self-inductance of the inner part of the cable is therefore given by $L_1 = \Phi/I$, where

$$\Phi = \int_0^{r_1} I \frac{r^2}{r_1^2} \frac{\mu_1 r dr}{2\pi r_1^2} = I \int_0^{r_1} \frac{\mu_1 r^3 dr}{2\pi r_1^4} = I \left[\frac{\mu_1 r^4}{2\pi 4 r_1^4}\right]_0^{r_1} = I \frac{\mu_1}{8\pi}, \quad \text{i.e.} \quad L_1 = \frac{\mu_1}{8\pi}.$$

We now consider a circle of radius $r > r_1$:

$$2\pi r H_r = \oint \mathbf{H} \cdot dl = \int \mathbf{j} \cdot d\mathbf{F} = I,$$

i.e.

$$H_r = \frac{I}{2\pi r}.$$

The magnetic flux through an area of width dr and length 1 in the intermediate space (with no current) is therefore

$$\Phi = \int \mathbf{B} \cdot d\mathbf{F} = \int_{r_1}^{r_2} \mu_2 H_r dr \times 1 = \frac{\mu_2 I}{2\pi} \int_{r_1}^{r_2} \frac{dr}{r} = \frac{\mu_2 I}{2\pi} \ln \frac{r_2}{r_1}.$$

Hence

$$L_2 = \frac{\mu_2}{2\pi} \ln \frac{r_2}{r_1},$$

and the total self-inductance per unit length is

$$L = L_1 + L_2 = \frac{1}{4\pi} \left(\frac{1}{2}\mu_1 + 2\mu_2 \ln \frac{r_2}{r_1} \right).$$

This result is derived by a different method in Example 8.5.

Example 8.3: Self-inductance of a parallel wire cable of radius a

Calculate the self-inductance of a pair of parallel wires of radius a and a distance d apart. Figure 8.2 shows transverse and longitudinal cross sections through such a cable.

Solution: The space between the two parallel wires is taken to have magnetic permeability μ. The self-inductance is to be calculated.
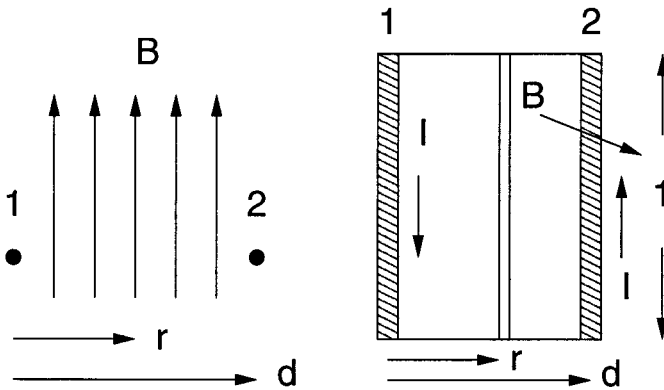
Fig. 8.2 The parallel wire cable.

We imagine the two wires connected at infinity to form a closed circuit. We first calculate the field H_r in a radial domain dr as indicated in Fig. 8.2. This field receives contributions from both wires. We saw above that the field H at a distance r from the centre of a wire is given by $I/2\pi r$. Thus the field pointing outward from the region between r and $r + dr$ is

$$H_r = \frac{I}{2\pi r} + \frac{I}{2\pi(d - r)}.$$

Hence we have

$$\Phi = \int_{d \times 1} \mathbf{B} \cdot d\mathbf{F} = \mu \int_a^{d-a} dr \times 1 \left(\frac{I}{2\pi r} + \frac{I}{2\pi(d - r)} \right)$$

$$= \frac{\mu I}{2\pi} \Big[\ln r - \ln(d - r) \Big]_a^{d-a} = 2\frac{\mu I}{2\pi} \ln \frac{d - a}{a}.$$

The self-inductance $L = \Phi/I$ is therefore given by

$$L = \frac{\mu}{\pi} \ln\left[\frac{d-a}{a}\right] \tag{8.6}$$

per unit length. This result neglects the flux through the cross sections of the wires (assumed to be small). In Example 8.10 the result (8.6) is obtained by a different method.

Example 8.4: Mutual inductance of one coil around another

A coil having n_2 turns per unit length and length l_2 is wound around a long solenoid with n_1 turns per unit length, cross-sectional area S_1 and current I_1. What is the mutual inductance L_{12}?

Solution: According to Eq. (8.4) the flux $\Phi_2(I_1)$ through circuit 2 as a consequence of the current I_1 in circuit 1 is given by

$$\Phi_2(I_1) = L_{21}I_1 = n_2 l_2 \int_{S_1} \mathbf{B} \cdot d\mathbf{F} = \mu_0(I_1 n_1)S_1 n_2 l_2$$
$$= n_1 n_2 \mu_0 S_1 l_2 I_1.$$

It follows that

$$L_{21} = n_1 n_2 \mu_0 S_1 l_2.$$

This mutual inductance can be calculated from its definition and as such is independent of whether an induced current actually flows in the other circuit. See also discussion below.

Consider, as in the preceding example, a coil of finite length wound around an infinitely long solenoid. In this case of the infinitely long inner solenoid, the field $\mathbf{B}$ is zero at all points of the coil wound around it. Can an observable current be induced in the outer coil? Since the field $\mathbf{B}$ is everywhere zero along the outer coil, there is no Lorentz force to produce the current. However, the vector potential $\mathbf{A}$ resulting from the current in the long solenoid does not vanish everywhere outside this solenoid (as will be shown below). The problem therefore requires a more detailed treatment and leads eventually to the so-called Aharonov–Bohm effect (see Chapter 19). We shall see that the effect of the vector potential around the long solenoid is a quantum mechanical phase effect which is observable only in appropriate interference experiments. We therefore investigate first the vector potential of a long solenoid.

8.4   The Vector Potential of a Long Solenoid

We consider a coil with coordinates as indicated in Fig. 8.3. Inside the solenoid the field is $\mathbf{B} = (0, 0, B_0)$, i.e. homogeneous along the z-axis. We obtain this expression with $\mathbf{B} = \nabla \times \mathbf{A}$ for the vector potential

$$\mathbf{A} = \frac{B_0}{2}(-y, x, 0) \equiv \frac{1}{2}\mathbf{B} \times \mathbf{r}.$$

The vector potential $\mathbf{A}$ is therefore (a) proportional to r, and (b) perpendicular to the z-axis and to $\mathbf{r}$. In the case of the infinitely long solenoid the field $\mathbf{B}$ outside is zero. Does this mean that $\mathbf{A}$ is also zero? No!
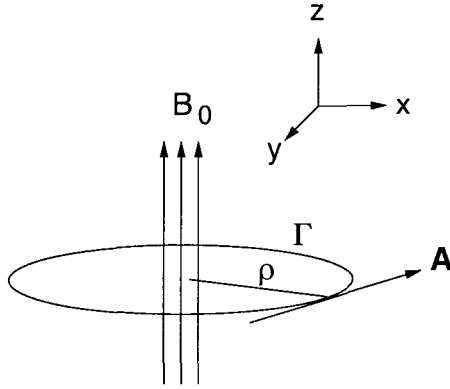
Fig. 8.3 The long solenoid with fields B_0 and $\mathbf{A}$.

We consider the path Γ of radius $\rho > a$ in Fig. 8.3. The flux through the area with the circle Γ as boundary is

$$\Phi = \int_{F(\Gamma)} \mathbf{B} \cdot d\mathbf{F} = B_0 \pi a^2 + B_{\text{outside}} \pi (\rho^2 - a^2) = B_0 \pi a^2,$$

so that

$$\Phi = \oint_\Gamma \mathbf{A} \cdot d\mathbf{l} = B_0 \pi a^2.$$

Thus the field $\mathbf{A}$ along the path Γ can be nonzero! We know the direction of $\mathbf{A}$ is that of the tangent to the path Γ. This suggests taking for $\mathbf{A}_{\text{outside}}$ the following ansatz:

$$\mathbf{A}_{\text{outside}} = k\left(-\frac{y}{x^2 + y^2}, \frac{x}{x^2 + y^2}, 0\right), \quad k = \text{const.},$$

because we know

$$\mathbf{A}_{\text{inside}} = \frac{1}{2} B_0(-y, x, 0).$$

First we have to verify that $\mathbf{B}_{\text{outside}} = \nabla \times \mathbf{A}_{\text{outside}} = 0$:

$$\nabla \times \mathbf{A}_{\text{outside}} = k \begin{vmatrix} \mathbf{e}_x & \mathbf{e}_y & \mathbf{e}_z \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ -\frac{y}{x^2+y^2} & \frac{x}{x^2+y^2} & 0 \end{vmatrix}.$$

For instance for the x-component:

$$\frac{\partial}{\partial y}(0) - \frac{\partial}{\partial z}\left(\frac{x}{x^2 + y^2}\right) = 0.$$

Similarly we obtain zero for the y-component. In the case of the z-component we require

$$\frac{\partial}{\partial x}\left(\frac{x}{x^2 + y^2}\right) + \frac{\partial}{\partial y}\left(\frac{y}{x^2 + y^2}\right)$$
$$= \frac{1}{x^2 + y^2} - \frac{2x^2}{(x^2 + y^2)^2} + \frac{1}{x^2 + y^2} - \frac{2y^2}{(x^2 + y^2)^2}$$
$$= \frac{2(x^2 + y^2) - 2x^2 - 2y^2}{(x^2 + y^2)^2} = 0,$$

hence $\mathbf{B}_{\text{outside}} = 0$. We determine the constant k such that $\mathbf{A}_{\text{outside}}$ yields the correct magnetic flux, i.e.

$$\int_\Gamma \mathbf{A}_{\text{outside}} \cdot d\mathbf{l} = B_0 \pi a^2.$$

But now

$$\rho = (x^2 + y^2)^{1/2} \quad \text{and} \quad |\mathbf{A}| = \frac{k}{\rho},$$

i.e.

$$\left(-\frac{y}{\sqrt{x^2 + y^2}}, \frac{x}{\sqrt{x^2 + y^2}}, 0\right)$$
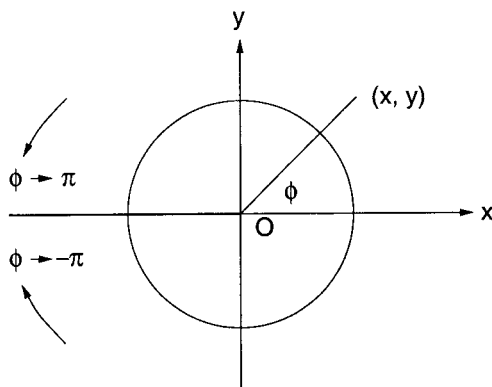
Fig. 8.4 The (x, y)-plane with discontinuity along the negative x-axis.

is a *unit vector parallel to* **A**. Hence

$$\int_\Gamma \mathbf{A}_{\text{outside}} \cdot d\mathbf{l} = \frac{k}{\rho} 2\pi\rho = 2\pi k = B_0 \pi a^2,$$

so that

$$k = \frac{1}{2} B_0 a^2$$

and

$$\mathbf{A}_{\text{outside}} = \frac{1}{2} B_0 \frac{a^2}{x^2 + y^2}(-y, x, 0).$$

It should be noted that at $r = a$ the field $\mathbf{A}_{\text{inside}}$ passes continuously over into $\mathbf{A}_{\text{outside}}$. We thus have a potential $\mathbf{A}_{\text{outside}} \neq 0$, where $\mathbf{B} = 0$. Is this strange? Why is it that $\mathbf{A}_{\text{outside}}$ does not vanish? Can we gauge it away, i.e. can we perform a gauge transformation so that the resulting vector potential vanishes? A suitable gauge transformation

$$\mathbf{A} \to \mathbf{A}' = \mathbf{A} + \boldsymbol{\nabla}\chi$$

seems to be one with the following function χ:

$$\chi(t, x, y, z) = -\frac{1}{2} B_0 a^2 \phi(x, y)$$

with

$$\phi(x, y) = \tan^{-1}\left(\frac{y}{x}\right), \quad \frac{y}{x} = \tan\phi.$$

The quantity ϕ is the polar angle shown in Fig. 8.4. At $x < 0, y > 0$ we have $\tan\phi < 0, \phi \to \pi$. At $x < 0, y < 0$ we have $\tan\phi > 0, \phi \to -\pi$. The values of ϕ on both sides of $y = 0, x < 0$ therefore differ by 2π. We put the cut as indicated in Fig. 8.4. The singularity at $x = 0$ with the cut along the negative semi-axis is described as a *Dirac string*.

We now demonstrate that $\mathbf{A}' = 0$, except along $y = 0, x < 0$. We have

$$\frac{\partial\phi}{\partial x} = \frac{\partial}{\partial x}\tan^{-1}\left(\frac{y}{x}\right) = \frac{-y/x^2}{1 + (y/x)^2} = -\frac{y}{x^2 + y^2},$$

$$\frac{\partial\phi}{\partial y} = \frac{\partial}{\partial y}\tan^{-1}\left(\frac{y}{x}\right) = \frac{1/x}{1 + (y/x)^2} = \frac{x}{x^2 + y^2}.$$

These expressions are very similar to A_x, A_y. The gauge transformation with

$$\chi(t, x, y, z) = -\frac{1}{2} B_0 a^2 \phi(x, y)$$

therefore leads to

$$\mathbf{A}' = \mathbf{A} + \boldsymbol{\nabla}\chi = \frac{B_0 a^2}{2}\begin{pmatrix} -\frac{y}{x^2+y^2} \\ \frac{x}{x^2+y^2} \\ 0 \end{pmatrix} - \frac{B_0 a^2}{2}\begin{pmatrix} -\frac{y}{x^2+y^2} \\ \frac{x}{x^2+y^2} \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

Thus $\mathbf{A}$ has been "*gauged away*", except where χ is singular: In the half-plane $y = 0, x < 0$. This means that at $y = 0$ the function χ possesses a discontinuity as indicated in Fig. 8.5. We recall that in the case of the step function θ: $d\theta(y)/dy = \delta(y)$, so that for $x < 0$:

$$\chi = f(x,y)\theta(-y) + g(x,y)\theta(y), \quad f(x,0) \neq g(x,0).$$
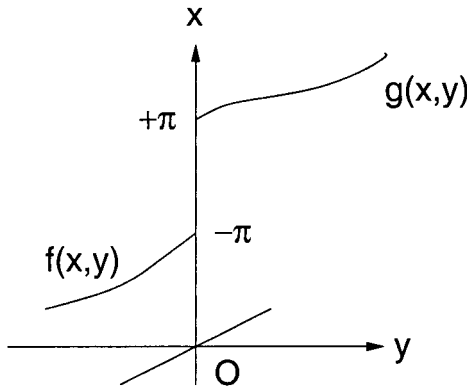
Fig. 8.5 The discontinuity of χ at $y = 0$.

Hence

$$\boldsymbol{\nabla}\chi = \left(\frac{\partial\chi}{\partial x}, \frac{\partial\chi}{\partial y}, \frac{\partial\chi}{\partial z}\right)$$
$$= \left\{ \frac{\partial f(x,y)}{\partial x}\theta(-y) + \frac{\partial g(x,y)}{\partial x}\theta(y),\right.$$
$$\left. \frac{\partial f(x,y)}{\partial y}\theta(-y) + \frac{\partial g(x.y)}{\partial x}\theta(y) + \delta(y)(g(x,y) - f(x,y)), 0 \right\}.$$

Thus the new gauge potential $\mathbf{A}' = \mathbf{A} + \boldsymbol{\nabla}\chi$ has a singularity of the type of a delta function in the y-component, and step function behaviour in the other components. It should be noted, however, that there is no necessity to resort to a singular gauge transformation!

The observation of this effect, called the *Aharonov–Bohm effect*, i.e. that an electron moving in a part of space where the field $\mathbf{B}$ is zero but the

vector potential is not, was achieved in quantum mechanical experiments. Older books do not cover this topic. For additional information we refer to Chapter 19 and to literature cited there.

8.5 Energy of a Self-Inductance

In many cases the self-inductance L can also be derived from the magnetic energy W^{mag}, in fact even more easily than otherwise. We have

$$
\begin{aligned}
-\frac{\partial W^{\text{mag}}}{\partial t} &= \text{loss of magnetic energy per unit time} \\
&= \text{energy in form of Joule heat per unit time} \\
&= I V_{\text{in}} \quad (V_{\text{in}} = \text{induced voltage}) \\
&= -I \frac{\partial \Phi}{\partial t} \\
&= -I \frac{\partial}{\partial t}(LI) \\
&= -\frac{1}{2} L \frac{\partial I^2}{\partial t}
\end{aligned}
\tag{8.7}
$$

provided the conductor was not deformed. The following relation therefore follows:

$$
W^{\text{mag}} = \frac{1}{2} L I^2.
\tag{8.8}
$$

In the case of several circuits the magnetic energy is correspondingly

$$
\begin{aligned}
W^{\text{mag}} &= \frac{1}{2} \sum_i L_i I_i^2 + \frac{1}{2} \sum_i \sum_j L_{ij} I_i I_j \\
&= \frac{1}{2} L_1 I_1^2 + \frac{1}{2} L_2 I_2^2 + L_{12} I_1 I_2,
\end{aligned}
\tag{8.9}
$$

if $L_{12} = L_{21}, L_{11} \equiv L_1, L_{22} \equiv L_2$, etc.

The *magnetic force* $\mathbf{F}$, which gives rise to a change of magnetic energy, is given by

$$
dW^{\text{mag}} = -\mathbf{F} \cdot d\mathbf{l} = \nabla W \cdot d\mathbf{l}.
$$

Here $d\mathbf{l}$ is the virtual displacement[*] of a circuit for a fixed current or fixed magnetic flux. Correspondingly large mechanical forces $\mathbf{F}^{\text{mech}}$ are needed, in order to prevent an actual displacement (so that for the virtual displacement in statics $-\delta W = \mathbf{F} \cdot \delta \mathbf{l} + \mathbf{F}^{\text{mech}} \cdot \delta \mathbf{l}^{\text{mech}} = 0$).

---

[*]Discussions of this can be found for example in Cheng [12], p. 252.

Example 8.5: Self-inductance of the coaxial cable

Derive the self-inductance of the coaxial cable defined in Example 8.2 from its magnetic energy.

Solution: The expression for the magnetic energy is

$$W = \frac{1}{2} \int d\mathbf{r}\, \mathbf{H} \cdot \mathbf{B},$$

where (in the present case) $d\mathbf{r} = r\,dr\,d\theta \times$ unit length. We calculate separately the magnetic energy of the inner and outer parts of the cable. Using the notation of Example 8.2 and the fields calculated there, we have for the inner part from $r = 0$ to $r = r_1$:

$$W_1 = \frac{1}{2} \int d\mathbf{r}\, \mu_1 \left(\frac{Ir}{2\pi r_1^2} \right)^2 = \frac{2\pi \mu_1 I^2}{2(2\pi r_1^2)^2} \int_0^{r_1} r^3\, dr = \frac{\mu_1 I^2}{2} \frac{1}{8\pi}.$$

Similarly we obtain for the part from r_1 to r_2:

$$W_2 = \frac{1}{2} \int d\mathbf{r}\, \mu_2 \left(\frac{I}{2\pi r} \right)^2 = \frac{2\pi I^2 \mu_2}{2(2\pi)^2} \int_{r_1}^{r_2} \frac{r\,dr}{r^2} = \frac{\mu_2 I^2}{2} \frac{\ln(r_2/r_1)}{2\pi}.$$

Comparing the sum of W_1 and W_2 with Eq. (8.9), i.e.

$$W = \frac{1}{2} L I^2,$$

we obtain the result of Example 8.2, i.e.

$$L = \frac{1}{4\pi} \left(\frac{1}{2}\mu_1 + 2\mu_2 \ln \frac{r_2}{r_1} \right).$$

Example 8.6: Coaxial coils

Two coaxial coils with respectively N_1 and N_2 turns and radii a_1 and a_2 carry currents I_1 and I_2. The distance separating their centres is $z \gg a_1, a_2$. Determine the magnetic force between the coils. Reexpress the result in terms of the magnetic moments of the coils.

Solution: We calculate first the vector potential $\mathbf{A}(\mathbf{r})$. We can calculate this in several ways. The quickest way is to use Eq.(6.4), i.e.

$$\mathbf{A}(\mathbf{r}) = k' \frac{\mathbf{m} \times \mathbf{r}}{r^3}, \quad k' = \frac{\mu_0}{4\pi},$$

(where the moment $\mathbf{m}$ is located at the origin, which has to be modified in our application here!) together with the expression for the magnetic moment as the product of current times area of the conducting loop.

It is instructive to derive the expression from first principles, i.e. from the Biot–Savart law or correspondingly from the expression for the vector potential derived therefrom, i.e.

$$\mathbf{A}(\mathbf{r}) = \frac{\mu_0}{4\pi} \int \frac{I d\mathbf{s}(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|}.$$

We distinguish between the planes of the two coils or rather first between the plane of the coil with current I_1 and a point P with position vector $\mathbf{r}$ on the circle with radius $a_2 = R \sin \theta$ in the plane of conductor "2", where R is the distance between the point P and the centre of the circular conductor "1" with radius a_1, and θ is the angle between $\mathbf{R}$ and the axis of the two circular conductors. The point $P'(\mathbf{r}')$ at element $d\mathbf{s}$ of the circular conductor "1" has coordinates in the latter's plane given by

$$\mathbf{r}'' = \mathbf{e}_{x''} a_1 \cos \phi' + \mathbf{e}_{y''} a_1 \sin \phi'.$$
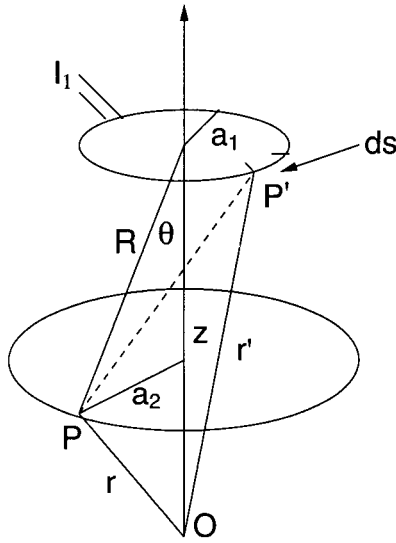
Fig. 8.6 Two coaxial coils.

Then

$$ds = \mathbf{r}''(\phi' + d\phi') - \mathbf{r}''(\phi') = d\phi' \frac{d\mathbf{r}''}{d\phi'} = d\phi' \left[-\mathbf{e}_{x''} a_1 \sin \phi' + \mathbf{e}_{y''} a_1 \cos \phi' \right].$$

Using the cosine-theorem we have

$$|\mathbf{r} - \mathbf{r}'|^2 = (PP')^2 = R^2 + a_1^2 - 2a_1 R \cos \psi$$

and (as may be deduced from the geometry)

$$R \cos \psi = R \sin \theta \cos(\phi - \phi').$$

Observe that when $\phi = \phi'$ the point P' has the same angular position on the upper circle as the point P on the lower, and angle ψ is in that case the complement of θ, i.e. $\theta = \pi/2 - \psi$. When $\phi \neq \phi'$, the factor $\cos(\phi - \phi')$ provides the appropriate projection. Hence

$$\mathbf{A}(\mathbf{r}) = \frac{\mu_0 I}{4\pi} \int \frac{d\phi'(-\mathbf{e}_{x''} a_1 \sin \phi' + \mathbf{e}_{y''} a_1 \cos \phi')}{\sqrt{R^2 + a_1^2 - 2a_1 R \sin \theta \cos(\phi - \phi')}}.$$

With polar coordinates ϕ, ρ in the plane of conductor "1" and

$$\mathbf{e}_{x''} = -\mathbf{e}_\phi \sin \phi + \mathbf{e}_\rho \cos \phi, \quad \mathbf{e}_{y''} = \mathbf{e}_\phi \cos \phi + \mathbf{e}_\rho \sin \phi,$$

and choosing the angle $\phi = 0$, this becomes

$$\mathbf{A}(\mathbf{r}) = \frac{\mu_0 I}{4\pi} \int_0^{2\pi} \frac{d\phi'(-a_1 \sin \phi' \mathbf{e}_\rho + a_1 \cos \phi' \mathbf{e}_\phi)}{\sqrt{R^2 + a_1^2 - 2a_1 R \sin \theta \cos \phi'}}.$$

Replacing ϕ' by $\phi' + \pi$ we obtain

$$\mathbf{A}(\mathbf{r}) = \frac{\mu_0 I}{4\pi} \int_{-\pi}^{\pi} \frac{d\phi'(a_1 \sin \phi' \mathbf{e}_\rho - a_1 \cos \phi' \mathbf{e}_\phi)}{\sqrt{R^2 + a_1^2 + 2a_1 R \sin \theta \cos \phi'}}$$

$$= -\frac{\mu_0 I}{4\pi} \int_0^{2\pi} \frac{d\phi' a_1 \cos \phi' \mathbf{e}_\phi}{\sqrt{R^2 + a_1^2 + 2a_1 R \sin \theta \cos \phi'}}.$$

For $R^2 \gg a_1^2$ this is

$$\mathbf{A}(\mathbf{r}) \simeq -\frac{\mu_0 I}{4\pi} \int_0^{2\pi} \frac{a_1 \cos \phi' \mathbf{e}_\phi d\phi'}{R} \left[1 - \frac{a_1^2 + 2a_1 R \sin \theta \cos \phi'}{2R^2}\right]$$

$$\simeq \frac{\mu_0 I}{4\pi} \frac{a_1^2}{R^2} \mathbf{e}_\phi \int_0^{2\pi} d\phi' \sin \theta \cos^2 \phi'.$$

Since $\int_0^{2\pi} \cos^2 \theta d\theta = \pi$, we have

$$\mathbf{A}(\mathbf{r}) \simeq \frac{\mu_0 I}{4\pi} \frac{\pi a_1^2}{R^2} \sin \theta \mathbf{e}_\phi + O(1/R^4). \tag{8.10}$$

With a different expansion of the denominator (for $R^2 + a_1^2 \gg a_1 R$) we can write

$$\mathbf{A}(\mathbf{r}) \simeq -\frac{\mu_0 I}{4\pi} \int_0^{2\pi} \frac{a_1 \cos \phi' \mathbf{e}_\phi d\phi'}{(R^2 + a_1^2)^{1/2}} \left[1 - \frac{a_1 R \sin \theta \cos \phi'}{(R^2 + a_1^2)}\right]$$

$$\simeq \frac{\mu_0 I}{4\pi} \frac{\pi a_1^2 R \sin \theta \mathbf{e}_\phi}{(R^2 + a_1^2)^{3/2}}$$

$$= \frac{\mu_0 I}{4\pi} \frac{\pi a_1^2 \rho \mathbf{e}_\phi}{(R^2 + a_1^2)^{3/2}} + O[1/(R^2 + a_1^2)^{5/2}] \tag{8.11}$$

with ρ now in the plane of conductor "2" as in Fig. 8.7.

We now return to our original problem. We assume the two coils to be arranged as depicted in Fig. 8.6. The electromagnetic vector potential $\mathbf{A}(\mathbf{r})$ at point $P(\mathbf{r})$ due to the current I_1 in the N_1 turns of the first coil is given by N_1 superpositions of the field of one turn, i.e. from Eq. (8.10)

$$\mathbf{A}(\mathbf{r}) = N_1 \frac{\mu_0 I_1}{4\pi} \frac{\pi a_1^2}{R^2} \sin \theta \mathbf{e}_\phi. \tag{8.12}$$

The magnetic flux Φ_2 through one turn of the second coil is given by

$$\Phi_2 = \int_{F_2} \mathbf{B} \cdot d\mathbf{F} = \int_{F_2} \boldsymbol{\nabla} \times \mathbf{A} \cdot d\mathbf{F}$$

$$= \oint_{C_2} \mathbf{A}(\mathbf{r}) \cdot d\mathbf{l}_2$$

$$= \oint_{C_2} \frac{N_1 \mu_0 I_1 \pi a_1^2 a_2}{4\pi (z^2 + a_2^2)^{3/2}} a_2 d\phi_2$$

$$= \frac{2\mu_0 I_1 N_1 \pi a_1^2 \pi a_2^2}{4\pi (z^2 + a_2^2)^{3/2}}, \tag{8.13}$$

where $dl_2 = a_2 d\phi_2$ and we used $\sin \theta = a_2/R$ and $R^2 = z^2 + a_2^2$.

The mutual inductance L_{12} is given by

$$N_2 \Phi_2 = L_{21} I_1,$$

i.e.

$$L_{21} = \frac{N_1 N_2 2\mu_0 \pi a_1^2 \pi a_2^2}{4\pi (z^2 + a_2^2)^{3/2}} \simeq \frac{2\mu_0 N_1 N_2 \pi a_1^2 \pi a_2^2}{4\pi z^3} \simeq L_{12} \quad \text{for} \quad z \gg a_1, a_2.$$

The interaction energy of the coils is given by

$$W = \frac{1}{2}\Big[L_{12}I_1 I_2 + L_{21}I_1 I_2\Big] \simeq L_{12}I_1 I_2.$$
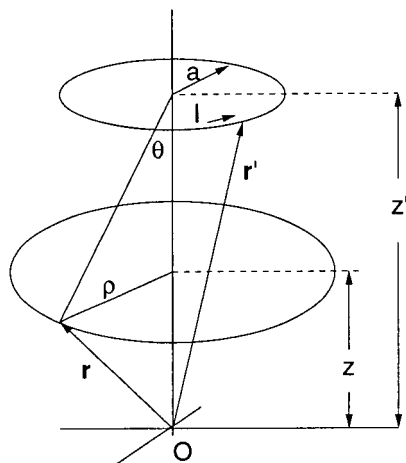
Fig. 8.7 Obtaining **A** at **r**.

Hence the magnetic force in the direction z is

$$\mathbf{F} = -\boldsymbol{\nabla}W = \mathbf{e}_z 6\mu_0 N_1 N_2 \frac{I_1 \pi a_1^2 I_2 \pi a_2^2}{4\pi z^4} = \mathbf{e}_z \frac{6\mu_0}{4\pi}\frac{m_1 m_2}{z^4}, \tag{8.14}$$

where

$$m_1 = N_1 I_1 \pi a_1^2, \quad m_2 = N_2 I_2 \pi a_2^2$$

are the magnetic moments of the coils.[†]

Example 8.7:[‡] **The solenoid of finite length**
Determine for a solenoid of finite length the radial component of the magnetic induction near the centre of the solenoid and near the axis.

Solution: In solving this problem we use the result of Eq. (8.11) obtained above for the vector potential $\mathbf{A}(\mathbf{r})$ at a point with position vector $\mathbf{r}$ due to current I in a circular conductor of radius a, as indicated in Fig. 8.7, i.e. (apart from a sign)

$$\mathbf{A}(\mathbf{r}) = \frac{\mu_0 I}{4\pi}\frac{\pi a^2 \rho \mathbf{e}_\phi}{[(z-z')^2 + a^2 + \rho^2]^{3/2}}. \tag{8.15}$$

This is an approximation for large $(z - z')^2 + a^2$.

[†]Similar considerations can be found in P. C. Clemmow [13], p. 154.
[‡]See also Jackson [3], problem 5.1, p. 205.

We consider a coil with N turns per unit length and of length L. We imagine the current I as circling around the cylindrical surface of the coil. Per unit length of the coil the current is then NI. The current in one slice of the solenoid of thickness dz' is $NIdz'$. This expression is to be identified with I in Eq. (8.15), if we compare the circular conductor with an infinitesimal ring of the solenoid. We then sum (i.e. integrate) over all ring-like elements of the solenoid, in order to obtain the expression for the total vector potential at the point $\mathbf{r}$. Then

$$\mathbf{A}(\mathbf{r}) = \int_{-L/2}^{L/2} \frac{\mu_0 NIdz'\pi a^2 \rho \mathbf{e}_\phi}{4\pi[(z-z')^2 + a^2 + \rho^2]^{3/2}} = \frac{\mu_0 NI\pi a^2 \rho}{4\pi} \int_{-L/2}^{L/2} \frac{dz'\mathbf{e}_\phi}{4\pi[(z-z')^2 + a^2 + \rho^2]^{3/2}}.$$

From Tables of Integrals we have

$$\int \frac{dz'}{[z'^2 + \rho^2]^{3/2}} = \frac{z'}{\rho^2\sqrt{z'^2 + \rho^2}},$$

i.e.

$$\begin{aligned}
\text{Int} :&= \int_{z'=-L/2}^{z'=L/2} \frac{d(z'-z)}{[(z-z')^2 + a^2 + \rho^2]^{3/2}}\\
&= \left[\frac{1}{(\rho^2 + a^2)}\frac{(z'-z)}{\sqrt{(z-z')^2 + a^2 + \rho^2}}\right]_{z'=-L/2}^{z'=L/2}\\
&= \frac{(\frac{L}{2}-z)}{(\rho^2 + a^2)\sqrt{(z-L/2)^2 + a^2 + \rho^2}} + \frac{(\frac{L}{2}+z)}{(\rho^2 + a^2)\sqrt{(z+L/2)^2 + a^2 + \rho^2}}\\
&= \frac{(\frac{L}{2}-z)}{(\rho^2 + a^2)\sqrt{z^2 + L^2/4 + a^2 + \rho^2}}\left[1 - \frac{Lz}{z^2 + L^2/4 + \rho^2 + a^2}\right]^{-1/2}\\
&\quad + \frac{(\frac{L}{2}+z)}{(\rho^2 + a^2)\sqrt{z^2 + L^2/4 + a^2 + \rho^2}}\left[1 + \frac{Lz}{z^2 + L^2/4 + \rho^2 + a^2}\right]^{-1/2}.
\end{aligned}$$

This can be expanded as

$$\frac{(\frac{L}{2}-z)}{(\rho^2 + a^2)\sqrt{z^2 + L^2/4 + a^2 + \rho^2}}\left[1 + \frac{Lz}{2(z^2 + L^2/4 + \rho^2 + a^2)} + \frac{3}{8}\frac{(Lz)^2}{(z^2 + L^2/4 + \rho^2 + a^2)^2} + \cdots\right]$$

$$+ \frac{(\frac{L}{2}+z)}{(\rho^2 + a^2)\sqrt{z^2 + L^2/4 + a^2 + \rho^2}}\left[1 - \frac{Lz}{2(z^2 + L^2/4 + \rho^2 + a^2)} + \frac{3}{8}\frac{(Lz)^2}{(z^2 + L^2/4 + \rho^2 + a^2)^2} + \cdots\right],$$

so that

$$\text{Int} = \frac{L}{(\rho^2 + a^2)\sqrt{z^2 + L^2/4 + a^2 + \rho^2}} - \frac{Lz^2}{(\rho^2 + a^2)(z^2 + L^2/4 + a^2 + \rho^2)^{3/2}}$$

$$+ \frac{3}{8}\frac{L(Lz)^2}{(\rho^2 + a^2)(z^2 + L^2/4 + a^2 + \rho^2)^{5/2}},$$

or

$$\text{Int} = \frac{L(z^2 + L^2/4 + a^2 + \rho^2) - Lz^2}{(\rho^2 + a^2)(z^2 + L^2/4 + a^2 + \rho^2)^{3/2}} + \frac{3}{8}\frac{L(Lz)^2}{(\rho^2 + a^2)(z^2 + L^2/4 + a^2 + \rho^2)^{5/2}}.$$

Thus

$$\mathbf{A}(\mathbf{r}) = \frac{\mu_0 NI\pi a^2 \rho}{4\pi(\rho^2 + a^2)}\left[\frac{L(L^2/4 + \rho^2 + a^2)}{(z^2 + L^2/4 + a^2 + \rho^2)^{3/2}} + \frac{3}{8}\frac{L(Lz)^2}{(z^2 + L^2/4 + a^2 + \rho^2)^{5/2}}\right]\mathbf{e}_\phi.$$

However,

$$\mathbf{B}(\mathbf{r}) = \nabla \times \mathbf{A}(\mathbf{r}), \quad B_\rho = \frac{1}{\rho}\frac{\partial A_z}{\partial \phi} - \frac{\partial A_\phi}{\partial z}.$$

Hence in our case

$$B_\rho = -\frac{\partial A_\phi}{\partial z} = -\frac{\partial}{\partial z}\frac{\mu_0 NI\pi a^2\rho}{4\pi(\rho^2 + a^2)}\left[\frac{L(L^2/4 + \rho^2 + a^2)}{(z^2 + L^2/4 + a^2 + \rho^2)^{3/2}} + \frac{3}{8}\frac{L(Lz)^2}{(z^2 + L^2/4 + a^2 + \rho^2)^{5/2}}\right]$$

$$= \frac{\mu_0}{4\pi}\frac{3}{2}\frac{2zNI\pi a^2\rho L(L^2/4 + \rho^2 + a^2)}{(\rho^2 + a^2)(z^2 + L^2/4 + a^2 + \rho^2)^{5/2}} - \frac{\mu_0}{4\pi}\frac{NI\pi a^2\rho\frac{3}{8}3zL^3}{(\rho^2 + a^2)(z^2 + L^2/4 + a^2 + \rho^2)^{5/2}}$$

$$+ O[1/(z^2 + L^2/4 + a^2 + \rho^2)^{7/2}].$$

For $\rho \ll a, |z| \ll L$ this becomes

$$B_\rho \simeq \frac{\mu_0}{4\pi}\frac{96zNI\pi\rho(L^2/4 + a^2)}{L^4} - \frac{\mu_0}{4\pi}\frac{24NI\pi\rho z}{L^2}$$

$$= \frac{\mu_0}{4\pi}\left[\frac{24zNI\pi\rho}{L^2} + \frac{96zNI\pi\rho a^2}{L^4} - \frac{24NI\pi\rho z}{L^2}\right]$$

$$= \frac{\mu_0}{4\pi}\left[\frac{96zNI\pi\rho a^2}{L^4}\right].$$

This result agrees with that stated by Jackson (see above).

8.6 Simple Current Circuits

8.6.1 Current Circuits with R and L

In the case of a single circuit with resistance R and self-inductance L we obtain from the general equation established earlier,

$$I_k R_k + \sum_j L_{kj}\dot{I}_j = V_k$$

the relation

$$RI + L\frac{dI}{dt} = V. \tag{8.16}$$

We consider two possibilities:

(1). $V = V_0 = $ const. In this case the solution of Eq. (8.16) is

$$I = \frac{V_0}{R}(1 - e^{-\frac{R}{L}t}),$$

so that after a longer period of time the current I assumes the value of Ohm's law.

(2). $V = V_0 \cos\omega t$. Here $\omega = 2\pi\nu$ is the alternating voltage frequency. The current is the real part of the solution $\tilde{I}$ of the equation

$$R\tilde{I} + L\frac{d\tilde{I}}{dt} = V_0 e^{i\omega t}.$$

We set

$$\tilde{I} = F(t)e^{i\omega t},$$

so that

$$RF + L(\dot{F} + i\omega t) = V_0,$$

i.e.

$$L\dot{F} + (R + Li\omega)F = V_0.$$

The general solution of the first order equation

$$\frac{dy}{dx} + Py = Q \qquad (8.17)$$

is

$$y = e^{-\int^x P dx} \int dx Q e^{\int^x P dx}.$$

This can be verified by simple differentiation. Hence

$$F = e^{-\int (i\omega + \frac{R}{L})dt} \int \frac{V_0}{L} e^{\int (i\omega + \frac{R}{L})dt} dt$$

$$= \frac{V_0}{L} \frac{1}{i\omega + \frac{R}{L}} = \frac{V_0(-i\omega + \frac{R}{L})}{L(\omega^2 + \frac{R^2}{L^2})}.$$

Consequently

$$\tilde{I} = \frac{V_0(R - i\omega L)}{(\omega^2 L^2 + R^2)} e^{i\omega t}$$

and

$$I = \Re(\tilde{I}) = \frac{V_0}{\omega^2 L^2 + R^2}[R\cos\omega t + \omega L\sin\omega t]$$

$$= \frac{V_0}{\sqrt{R^2 + \omega^2 L^2}} \cos(\omega t - \delta),$$

with

$$\cos\delta = \frac{R}{\sqrt{R^2 + \omega^2 L^2}}, \quad \sin\delta = \frac{\omega L}{\sqrt{R^2 + \omega^2 L^2}}.$$

The effective "resistance" $\sqrt{R^2 + \omega^2 L^2}$ is called *impedance*, and ωL is called *inductance* or *inductive resistance*.

8.6.2 Circuits with L, C and R

In this case the equation is

$$IR + L\dot{I} = V + (V_1 - V_2),$$

where $(V_1 - V_2)$ is the potential difference of the condenser in the circuit. (Note that the "current" from one condenser plate to the other is the *displacement current*). We saw earlier that

$$V_2 - V_1 = \frac{q}{C}, \quad I = \frac{dq}{dt}.$$

Hence

$$L\ddot{I} + R\dot{I} = \dot{V} - \frac{I}{C},$$

or

$$L\ddot{I} + R\dot{I} + \frac{1}{C}I = \dot{V}.$$

We set $V = V_0 e^{i\omega t}$. Then

$$L\ddot{I} + R\dot{I} + \frac{1}{C}I = i\omega V_0 e^{i\omega t}.$$

We also set $I = I_0 e^{i\omega t}$. Then

$$I_0\left(-\omega^2 L + i\omega R + \frac{1}{C}\right) = i\omega V_0,$$

i.e.

$$I_0 = \frac{V_0}{R + i(\omega L - \frac{1}{\omega C})}, \quad I = \frac{V_0 e^{i\omega t}}{R + i(\omega L - \frac{1}{\omega C})}.$$

For $V = V_0 \cos \omega t$ the real solution is similar to that in the previous case

$$I = \frac{V_0 \cos(\omega t - \delta)}{\sqrt{R^2 + (\omega L - \frac{1}{\omega C})^2}},$$

where now

$$\tan \delta = \frac{1}{R}(\omega L - \frac{1}{\omega C}).$$

The expression $1/\omega C$ is called *capacitive resistance*.

8.7   Self-Inductances: Conjugate Function Method

In Sec. 3.7 we saw how the capacities of essentially 2-dimensional systems can be calculated by the method of conjugate functions. Here we consider the corresponding steps to obtain self-inductances. We start from Eq. (5.28),

$$\mathbf{A}(\mathbf{r}) = k' \int \frac{\mathbf{j}(\mathbf{r}')dr'}{|\mathbf{r} - \mathbf{r}'|},$$

and choose $\mathbf{A}$ along the z-direction with $A_x = 0$, $A_y = 0$, $A_z = V$. Then

$$\mathbf{B} = \mathbf{\nabla} \times \mathbf{A} = \begin{vmatrix} \hat{x} & \hat{y} & \hat{z} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ 0 & 0 & V \end{vmatrix} = \left(\frac{\partial V}{\partial y}, -\frac{\partial V}{\partial x}, 0 \right),$$

and

$$B^2 = \left(\frac{\partial V}{\partial x} \right)^2 + \left(\frac{\partial V}{\partial y} \right)^2 = |\mathbf{\nabla} V|^2, \quad B = |\mathbf{\nabla} V|.$$

Then the magnetic energy in vacuum can be written

$$W = \frac{1}{2} \int d\mathbf{r}\, \mathbf{H} \cdot \mathbf{B} = \frac{1}{2\mu_0} \int d\mathbf{r} |\mathbf{\nabla} V|^2.$$

However

$$\mathbf{\nabla} \cdot (V\mathbf{\nabla} V) = (\mathbf{\nabla} V)^2 + V\nabla^2 V,$$

and so with Gauss' divergence theorem and neglecting the surface integral at infinity we obtain

$$W = -\frac{1}{2\mu_0} \int d\mathbf{r} V\nabla^2 V.$$

Using Eq. (5.35), i.e.

$$\triangle \mathbf{A}(\mathbf{r}) = -\mu_0 \mathbf{j}(\mathbf{r}), \quad \text{and so} \quad \triangle V = -\mu_0 j(\mathbf{r}),$$

we obtain

$$W = \frac{1}{2} \int d\mathbf{r}\, j(\mathbf{r}) V(\mathbf{r}).$$

Assuming the wires are thin, so that the flux through them is negligible and hence V constant across them, we can write

$$W = \frac{1}{2} V \int d\mathbf{r}\, j(\mathbf{r}) = \frac{1}{2} VI$$

per wire and per unit length in z-direction. In the case of two wires with antiparallel currents I_1, I_2, we have

$$W = \frac{1}{2}I[V_1 - V_2] \equiv \frac{1}{2}I[V].$$

But, by definition of the inductance L,

$$W = \frac{1}{2}LI^2.$$

Hence

$$L = \frac{[V]}{I}.$$

Our next aim is to express I in terms of the conjugate function U defined as in Sec. 3.7. We recall Eq. (6.22):

$$I = \oint_C \mathbf{H} \cdot ds = \frac{1}{\mu_0} \oint \mathbf{B} \cdot ds$$
$$= \frac{1}{\mu_0} \oint |\nabla V| ds = \frac{1}{\mu_0} \oint |\nabla U| ds$$
$$= \frac{[U]}{\mu_0}.$$

The direction of ∇U is that of constant V. Since for both wires $|I_1| = |I_2| \equiv I$, $[U]$ is also the same for both wires. Hence we obtain

$$L = \frac{[V]}{I} = \mu_0 \frac{[V]}{[U]} \tag{8.18}$$

per unit length in the z-direction. We now apply this result in some examples.

Example 8.8: The parallel plate condenser
Calculate the self-inductance by the method of conjugate functions.

Solution: We let the plates be a distance d apart, i.e. $d = [y]$. We choose (cf. Sec. 3.7)

$$W := U + iV = x + iy, \quad \text{i.e.} \quad U = x, \quad V = y.$$

Thus we obtain for the self-inductance

$$L = \mu_0 \frac{[V]}{[U]} = \frac{\mu_0 d}{[x]} = \mu_0 d$$

per unit area perpendicular to this paper. Using the result of Example 3.9 for the capacity C, where $k = 1/4\pi\epsilon_0$, we observe that

$$LC = \frac{\mu_0 d}{4\pi k d} = \epsilon_0 \mu_0.$$

Example 8.9: Coaxial cylinders

Calculate the self-inductance of coaxial cylinders with radii $a, b, a < b$.

Solution: As in Example 3.11 we choose

$$W := V + iU = \ln(re^{i\theta}), \quad U = \theta, \quad V = \ln r.$$

Then

$$L = \frac{\mu_0[\ln r]}{[\theta]} = \frac{\mu_0 \ln(b/a)}{2\pi}$$

per unit length. The result should be compared with that of Example 8.2. Multiplying L by C of Example 3.11 we again obtain $LC = \epsilon_0\mu_0$.

Example 8.10: Parallel wires, a distance D apart

Calculate the self-inductance of two parallel wires of radius a and a distance D apart.

Solution: We can use the result of Example 3.12,

$$C = \frac{1}{4k \cosh^{-1}(\frac{D}{2a})}$$

together with

$$L = \mu_0 \frac{[V]}{[U]}, \quad C = \frac{1}{4\pi k} \frac{[U]}{[V]}.$$

Thus

$$L = \frac{\mu_0}{4\pi kC} = \frac{\mu_0}{\pi} \cosh^{-1}\left(\frac{D}{2a}\right)$$

with $LC = \mu_0\epsilon_0$. Using $\cosh^{-1} x = \pm[\ln(2x) + O(1/x^2)]$ for $x > 1$, the self-inductance may be approximated as

$$L \simeq \frac{\mu_0}{\pi} \ln\left(\frac{D}{a}\right), \quad D \gg a.$$

This result should be compared with the result of Example 8.3.

Chapter 9

Electromagnetic Waves

9.1 Introductory Remarks

In the following we consider electromagnetic waves in vacuum and in media. This means, we first obtain the wave equations from Maxwell's equations and then obtain the very important vacuum relation already referred to in Chapter 1,

$$c^2 = \frac{1}{\epsilon_0 \mu_o}, \tag{9.1}$$

which determines the velocity of the electromagnetic wave in the vacuum (or more generally in a medium) as the reciprocal of the square root of the product of dielectric constant and magnetic permeability. One should note the difference to the case of Chapter 8, where the velocity was that of the wave in a transmission cable. We then investigate solutions in vacuum and in conducting media, and in the latter case we encounter the important *skin effect*.

9.2 Electromagnetic Waves in Vacuum

In this case we have $\mathbf{D} = \epsilon_0 \mathbf{E}, \mathbf{B} = \mu_0 \mathbf{H}, \mathbf{j} = 0, \rho = 0$, since we consider first the case of electromagnetic waves without charges and currents (simply as a matter of simplicity we leave out the source terms here, considered to be far away). With these provisions we write down Maxwell's equations:

$$\boldsymbol{\nabla} \times \mathbf{B} = \mu_0 \epsilon_0 \frac{\partial \mathbf{E}}{\partial t}, \;\; \boldsymbol{\nabla} \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}, \;\; \boldsymbol{\nabla} \cdot \mathbf{E} = 0, \;\; \boldsymbol{\nabla} \cdot \mathbf{B} = 0. \tag{9.2}$$

It is a general procedure to apply to the first two equations, which are *"curl"* equations, the curl operator once again in order to reduce them to Laplace

form. This is achieved with the help of the formula

$$\boldsymbol{\nabla} \times (\boldsymbol{\nabla} \times \mathbf{B}) = \boldsymbol{\nabla}(\boldsymbol{\nabla} \cdot \mathbf{B}) - \triangle \mathbf{B}. \tag{9.3}$$

Applying this to the first equation and using the second, we obtain

$$\boldsymbol{\nabla} \times (\boldsymbol{\nabla} \times \mathbf{B}) = \boldsymbol{\nabla}(\boldsymbol{\nabla} \cdot \mathbf{B}) - \triangle \mathbf{B}$$
$$= \mu_0 \epsilon_0 \frac{\partial}{\partial t}(\boldsymbol{\nabla} \times \mathbf{E})$$
$$= -\mu_0 \epsilon_0 \frac{\partial^2 \mathbf{B}}{\partial t^2},$$

i.e.

$$-\triangle \mathbf{B} = -\mu_0 \epsilon_o \frac{\partial^2 \mathbf{B}}{\partial t^2}$$

or

$$\Box \mathbf{B} = 0, \quad \Box = \triangle - \frac{1}{c^2}\frac{\partial^2}{\partial t^2}, \quad c^2 = \frac{1}{\mu_0 \epsilon_0}. \tag{9.4}$$

In the wave equation the parameter c obviously has the significance and dimension of the velocity of a wave of the field $\mathbf{B}$. In an analogous way we obtain from Eqs. (9.2) also $\Box \mathbf{E} = 0$. Thus the fields $\mathbf{E}, \mathbf{B}$ appear as waves propagating with the velocity c in vacuum. This velocity is independent of the source and depends only on ϵ_0, μ_o, whose product connects electric and magnetic properties to the *velocity of light* (here in vacuum). We write[*]

$$\Box = \sum_{\mu=0}^{3} \partial^\mu \partial_\mu = \sum_{\mu,\nu=0,1,2,3} \partial^\mu g_{\mu\nu}\partial^\nu$$

with

$$\partial^\mu \equiv \frac{\partial}{\partial x_\mu}, \quad x_0 = ct, \quad g_{\mu\nu} = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

One can also write

$$\Box = \sum_{\mu=1}^{4} \partial^\mu \partial^\mu, \quad x_4 = ict.$$

The differential operator $\Box$ is called *D'Alembert operator*; this operator is the spacetime generalisation of the Laplace operator $\triangle$.

---

[*]The justification will be seen later in the context of the Special Theory of Relativity.

Solutions Ψ of equations $\Box\Psi = 0$ are of the form $e^{\pm i(\mathbf{k}\cdot\mathbf{r}\mp\omega t)}$ with $k^2 = \omega^2/c^2$, since

$$\Box e^{\pm i(\mathbf{k}\cdot\mathbf{r}\mp\omega t)} = \left[-k^2 - \frac{1}{c^2}(-i\omega)^2\right]e^{\pm i(\mathbf{k}\cdot\mathbf{r}\mp\omega t)}$$

$$= \left[-k^2 + \frac{\omega^2}{c^2}\right]e^{\pm i(\mathbf{k}\cdot\mathbf{r}\mp\omega t)}.$$

A wave for which

$$f(\mathbf{r},t) \equiv \mathbf{k}\cdot\mathbf{r} - \omega t = \text{const.}$$

is called a *plane wave*. In its case

$$\boldsymbol{\nabla} f(\mathbf{r},t) = \mathbf{k},$$

i.e. *the wave vector* $\mathbf{k}$ is orthogonal to the plane defined by $f(\mathbf{r},t) = \text{const.}$ From $k^2 = \omega^2/c^2$ we deduce that $k = \omega/c = 2\pi/\lambda$, λ the wavelength. However, for an arbitrary curve $\mathbf{r} = \mathbf{r}(t)$ in the plane $f(\mathbf{r},t) = \text{const.}$ one has $\mathbf{k}\cdot\dot{\mathbf{r}} = \omega$, i.e. $\mathbf{k}\cdot\mathbf{v} = \omega, v = \omega/k$ for $\dot{\mathbf{r}} \parallel \mathbf{k}$, as we can choose. The vector $\mathbf{k}$ points in the direction of propagation of the wave. For electromagnetic waves in vacuum $v = c$.

We now consider a *general solution* of the equation $\Box\mathbf{E} = 0$. We can write this

$$\mathbf{E} = \mathbf{E}_{01}e^{i(\mathbf{k}\cdot\mathbf{r}-\omega t)} + \mathbf{E}_{02}e^{-i(\mathbf{k}\cdot\mathbf{r}+\omega t)} \tag{9.5}$$

("one time, several space directions"). We can restrict ourselves first to

$$\mathbf{E} = \mathbf{E}_0 e^{i(\mathbf{k}\cdot\mathbf{r}-\omega t)}$$

(or to $\mathbf{E} = \mathbf{E}_0\cos(\mathbf{k}\cdot\mathbf{r}-\omega t)$, since only the real part has physical significance). We let the vector $\mathbf{E}_0$ be constant in space and time. Then

$$\boldsymbol{\nabla}\cdot\mathbf{E} = \mathbf{E}_0(i\mathbf{k})e^{i(\mathbf{k}\cdot\mathbf{r}-\omega t)}, \quad \frac{\partial\mathbf{E}}{\partial t} = \mathbf{E}_0(-i\omega)e^{i(\mathbf{k}\cdot\mathbf{r}-\omega t)},$$

i.e. we have the replacements

$$\boldsymbol{\nabla} \to i\mathbf{k}, \quad \frac{\partial}{\partial t} \to -i\omega.$$

We consider $\mathbf{B}$ in a similar way. Maxwell's equations then become

$$\left.\begin{array}{l} \mathbf{k}\times\mathbf{B} = -\frac{\omega}{c^2}\mathbf{E}, \\ \mathbf{k}\times\mathbf{E} = \omega\mathbf{B}, \end{array}\right\} \therefore \mathbf{E}\perp\mathbf{B},\mathbf{k} \quad \text{and} \quad \mathbf{B}\perp\mathbf{E},\mathbf{k},$$

$$\left.\begin{array}{l} \mathbf{k}\cdot\mathbf{E} = 0, \\ \mathbf{k}\cdot\mathbf{B} = 0, \end{array}\right\} \therefore \mathbf{E},\mathbf{B}\perp\mathbf{k}. \tag{9.6}$$

It follows that the electromagnetic waves in vacuum are *transverse waves*, i.e. that the fields $\mathbf{E}$ and $\mathbf{B}$ are perpendicular to the propagation vector, and in addition the fields are orthogonal to each other, as illustrated in Fig. 9.1. We infer from the first two of the four relations (9.6) by scalar multiplication by $\mathbf{E}$ and $\mathbf{B}$ the important (but not universally valid) consequence that

$$\frac{1}{c^2}|\mathbf{E}|^2 = |\mathbf{B}|^2.$$
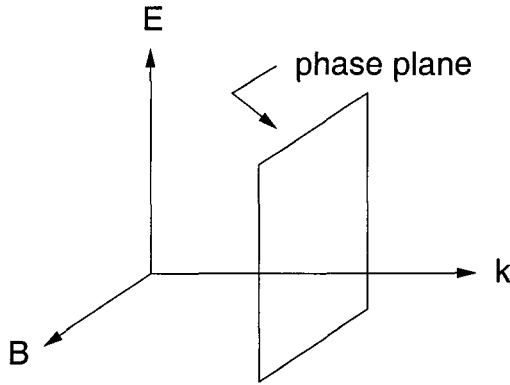
Fig. 9.1 The phase plane $\perp$ vector of propagation.

The vector $\mathbf{E}_0$ in the expression $\mathbf{E} = \mathbf{E}_0 e^{i(\mathbf{k}\cdot\mathbf{r}-\omega t)}$ defines the direction of oscillation. If $\mathbf{E}_0$ is constant in time one says, the wave is *linearly polarised*. The transport of energy per unit area (cf. the Poynting vector) takes place in the direction of the propagation vector $\mathbf{k} \propto \mathbf{E} \times \mathbf{B}$.

We let ϵ_1 and ϵ_2 be unit vectors, called *polarisation vectors*, which span the plane orthogonal to $\mathbf{k}$. Then, in a general case,

$$\mathbf{E} = \mathbf{E}_1 + \mathbf{E}_2, \quad |\mathbf{E}| = |\mathbf{E}_0|, \quad \mathbf{B} = \mathbf{B}_1 + \mathbf{B}_2$$

with $(i = 1, 2)$

$$\mathbf{E}_i(\mathbf{r}, t) = \epsilon_i E_i e^{i(\mathbf{k}\cdot\mathbf{r}-\omega t)}, \quad \mathbf{B}_i(\mathbf{r}, t) = \epsilon_i B_i e^{i(\mathbf{k}\cdot\mathbf{r}-\omega t)}.$$

Since $\mathbf{E} \perp \mathbf{B}$, it suffices and is convenient, to consider the vector $\mathbf{E}$ as a representative of the entire wave. $\mathbf{E}_1, \mathbf{E}_2$ have different phases, if we permit E_1, E_2 to be complex, i.e.

$$E_1 = |E_1|e^{i\theta_1}, \quad E_2 = |E_2|e^{i\theta_2}.$$

Since only the relative phase is of interest, we can choose $\theta_1 = 0$. If $\theta_2 = 0$, the wave is said to be *linearly polarised*. In this case we have

$$\mathbf{E} = (|E_1|\boldsymbol{\epsilon}_1 + |E_2|\boldsymbol{\epsilon}_2)e^{i(\mathbf{k}\cdot\mathbf{r}-\omega t)}, \text{ or } \Re\mathbf{E} = (|E_1|\boldsymbol{\epsilon}_1 + |E_2|\boldsymbol{\epsilon}_2)\cos(\mathbf{k}\cdot\mathbf{r}-\omega t).$$

If $\theta_2 = \varphi \neq 0$, one says the wave is *elliptically polarised*,

$$\mathbf{E}(\mathbf{r},t) = \boldsymbol{\epsilon}_1 E_1 e^{i(\mathbf{k}\cdot\mathbf{r}-\omega t)} + \boldsymbol{\epsilon}_2 E_2 e^{i(\mathbf{k}\cdot\mathbf{r}-\omega t+\varphi)}.$$

A *circularly polarised* wave is obtained, if $|E_1| = |E_2|$ *and* $\varphi = \pm\pi/2$, i.e.

$$\mathbf{E}(\mathbf{r},t) = |E_1|(\boldsymbol{\epsilon}_1 \pm i\boldsymbol{\epsilon}_2)e^{i(\mathbf{k}\cdot\mathbf{r}-\omega t)}$$
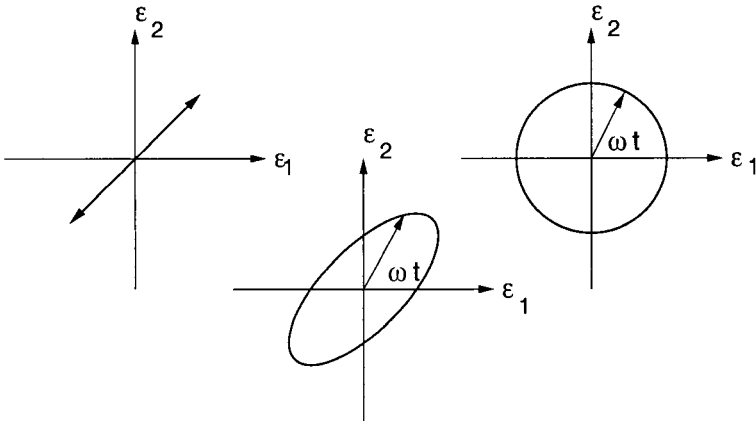
Fig. 9.2 Linear, circular and elliptic polarisation.

(left and right polarised) or

$$\Re\mathbf{E}(\mathbf{r},t) = |E_1|\boldsymbol{\epsilon}_1 \cos(\mathbf{k}\cdot\mathbf{r}-\omega t) \pm |E_1|\boldsymbol{\epsilon}_2 \sin(\mathbf{k}\cdot\mathbf{r}-\omega t).$$

For example, when $\mathbf{k}$ is parallel to the z-axis, $\boldsymbol{\epsilon}_1$ parallel to $\mathbf{x}$, and $\boldsymbol{\epsilon}_2$ parallel to $\mathbf{y}$, we have:

$$\Re E_x(\mathbf{r},t) = |E_1|\cos(kz-\omega t), \ \Re E_y(\mathbf{r},t) = \pm|E_1|\sin(\omega t - kz).$$

In the case of the elliptically polarised wave $|E_1| \neq |E_2|$. These cases are illustrated in Fig. 9.2.

9.3   Electromagnetic Waves in Media

We now consider the case of a *homogeneous, infinitely extended, uncharged but conducting medium* with medium parameters ϵ, μ and a current density $\mathbf{j}$ as source. In this case we have

$$\mathbf{D} = \epsilon\mathbf{E}, \ \mathbf{B} = \mu\mathbf{H}, \ \mathbf{j} = \sigma\mathbf{E}, \ \rho = 0.$$

The Maxwell equations in this case are

$$\nabla \cdot \mathbf{E} = 0, \quad \nabla \cdot \mathbf{H} = 0, \quad \nabla \times \mathbf{E} = -\mu\frac{\partial \mathbf{H}}{\partial t}, \quad \nabla \times \mathbf{H} = \epsilon\frac{\partial \mathbf{E}}{\partial t} + \sigma\mathbf{E}.$$

It follows that:

$$\underbrace{\nabla \times (\nabla \times \mathbf{E})}_{\underbrace{\nabla(\nabla \cdot \mathbf{E})}_{0} - \nabla^2\mathbf{E}} = -\mu\frac{\partial}{\partial t}\underbrace{\nabla \times \mathbf{H}}_{\epsilon\frac{\partial \mathbf{E}}{\partial t} + \sigma\mathbf{E}}.$$

We thus obtain the following equation for $\mathbf{E}$ and analogously the corresponding equation for $\mathbf{H}$, i.e.

$$\triangle\mathbf{E} = \epsilon\mu\frac{\partial^2\mathbf{E}}{\partial t^2} + \mu\sigma\frac{\partial\mathbf{E}}{\partial t},$$

$$\triangle\mathbf{H} = \epsilon\mu\frac{\partial^2\mathbf{H}}{\partial t^2} + \mu\sigma\frac{\partial\mathbf{H}}{\partial t}. \tag{9.7}$$

These equations are also known as *telegraph equations*. In the case of dielectrics ($\sigma = 0$) they become:

$$\left(\triangle - \frac{1}{c'^2}\frac{\partial^2}{\partial t^2}\right)\mathbf{E} = 0, \quad \left(\triangle - \frac{1}{c'^2}\frac{\partial^2}{\partial t^2}\right)\mathbf{H} = 0, \tag{9.8}$$

where

$$c' = \frac{1}{\sqrt{\epsilon\mu}} = \sqrt{\frac{\epsilon_0\mu_0}{\epsilon\mu}}c.$$

We can see that c' is the *phase velocity*, i.e. $\omega/\sqrt{\mathbf{k}^2}$, by setting $\mathbf{E}, \mathbf{H} \propto e^{i(\mathbf{k}\cdot\mathbf{r}-\omega t)}$, so that $(\triangle + \mathbf{k}^2)\mathbf{E}, \mathbf{H} = 0$ with $\sqrt{\mathbf{k}^2} = \omega/c' = \omega\sqrt{\epsilon\mu}$, and the phase velocity is $\omega/k = 1/\sqrt{\epsilon\mu}$.

The *refractive index* n known from optics, is defined by

$$n = \frac{c}{c'} = \frac{\sqrt{\epsilon\mu}}{\sqrt{\epsilon_0\mu_0}} = c\sqrt{\epsilon\mu}. \tag{9.9}$$

9.4 Frequency Dependence of ϵ and σ

At high frequencies of electromagnetic radiation the parameters ϵ, μ and σ become frequency dependent. When we introduced these parameters we were concerned with static fields. We therefore face the problem of determining their frequency dependence. We consider here only the two more important cases, i.e. those of ϵ and σ, and take in most cases simply $\mu \sim$ constant.

9.4.1 The Generalised Dielectric Constant

We set
$$\mathbf{E} = \mathbf{E}_0 e^{i(\mathbf{k}\cdot\mathbf{r}-\omega t)}, \quad \mathbf{H} = \mathbf{H}_0 e^{i(\mathbf{k}\cdot\mathbf{r}-\omega t)}. \tag{9.10}$$

Then the Maxwell equation

$$\boldsymbol{\nabla} \times \mathbf{H} = \epsilon\frac{\partial \mathbf{E}}{\partial t} + \sigma\mathbf{E}$$

implies
$$\boldsymbol{\nabla} \times \mathbf{H} = -i\epsilon\omega\mathbf{E} + \sigma\mathbf{E} \equiv -i\omega\eta\mathbf{E}.$$

The quantity
$$\eta = \epsilon + i\frac{\sigma}{\omega} \tag{9.11}$$

is called the *generalised (complex) dielectric constant.* Substituting now the Eqs. (9.10), (9.11) into the Maxwell equations, we obtain

$$\boldsymbol{\nabla} \times \mathbf{H} = \epsilon\frac{\partial \mathbf{E}}{\partial t} + \sigma\mathbf{E} \rightarrow \mathbf{k} \times \mathbf{H} = -\omega\eta\mathbf{E},$$
$$\boldsymbol{\nabla} \times \mathbf{E} = -\mu\frac{\partial \mathbf{H}}{\partial t} \rightarrow \mathbf{k} \times \mathbf{E} = \omega\mu\mathbf{H},$$
$$\boldsymbol{\nabla} \cdot \mathbf{H} = 0 \rightarrow \mathbf{k} \cdot \mathbf{H} = 0,$$
$$\boldsymbol{\nabla} \cdot \mathbf{E} = 0 \rightarrow \mathbf{k} \cdot \mathbf{E} = 0.$$

Thus again we have $\mathbf{H}, \mathbf{E} \perp \mathbf{k}$, and *no longer* $|\mathbf{E}| = |\mathbf{H}|$. But

$$\underbrace{\mathbf{k} \times (\mathbf{k} \times \mathbf{H})}_{\underbrace{(\mathbf{k} \cdot \mathbf{H})}_{0}\mathbf{k}-k^2\mathbf{H}} = -\omega\eta\underbrace{\mathbf{k} \times \mathbf{E}}_{\omega\mu\mathbf{H}}.$$

It follows that
$$k^2\mathbf{H} = \omega^2\eta\mu\mathbf{H} = \mu\epsilon\omega^2\left(1 + i\frac{\sigma}{\epsilon\omega}\right)\mathbf{H}.$$

Provided $\mathbf{H} \neq 0$ everywhere (hence the reference to infinitely extended media at the beginning) we have therefore

$$k^2 = \mu\epsilon\omega^2\left(1 + i\frac{\sigma}{\epsilon\omega}\right). \tag{9.12}$$

This relation is called the *dispersion relation of the conducting medium*. We can solve it for $k = \alpha + i\beta$, where for conductivity $\sigma = 0$:

$$k = \omega\sqrt{\epsilon\mu} = \frac{\omega}{c}n,$$

where n is the refractive index introduced earlier. As a *generalised refractive index p* one defines (instead of $n = \sqrt{\epsilon\mu}$) the quantity

$$p = \sqrt{\eta\mu}c = p_1 + ip_2.$$

A wave moving in the direction of x contains the factor

$$e^{i(kx-\omega t)} = e^{i(\alpha x-\omega t)}e^{-\beta x}.$$
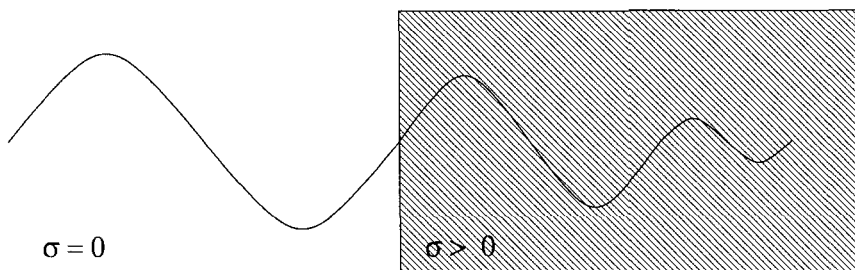
Fig. 9.3 Damping of a wave in a conducting medium.

Here β is of the order of σ for small σ. We see therefore that the wave in the conductor is damped, as indicated in Fig. 9.3. The distance $|1/\beta|$ is termed the *penetration depth* or *skin depth*.

In the next subsection we determine the frequency dependence of the conductivity σ. The result will be inserted into Eq. (9.11) for the generalised dielectric constant.

9.4.2 Frequency Dependence of σ

The conductivity of metals has its explanation in the existence of free electrons. We assume we have n such electrons in a volume $V, N = n/V$. Then

Newton's equation of motion for one such electron, say electron "k" with velocity $\mathbf{v}_k$, is ($e =$ charge)

$$m\frac{d}{dt}(e\mathbf{v}_k) = e^2\mathbf{E}_k - \xi e\mathbf{v}_k,$$

where $\mathbf{E}_k$ is the field of the other electrons and ξ is the frictional resistance of the lattice atoms. Hence

$$\frac{1}{V}\sum_k m\frac{d}{dt}(e\mathbf{v}_k) = \frac{1}{V}\sum_k e^2\mathbf{E}_k - \frac{1}{V}\xi\sum_k e\mathbf{v}_k. \tag{9.13}$$

But

$$\mathbf{j}(\mathbf{r},t) = \sum_k e\dot{\mathbf{r}}_k(t)\delta(\mathbf{r} - \mathbf{r}_k(t)),$$

so that

$$\int_V \mathbf{j}(\mathbf{r},t)d\mathbf{r} = \sum_k e\dot{\mathbf{r}}_k(t).$$

We now define a mean current density by the expression (note the averaging over space, and the time dependence of the mean value!)

$$\langle\mathbf{j}(t)\rangle V := \int_V \mathbf{j}(\mathbf{r},t)dV. \tag{9.14}$$

We also define a mean field strength $\mathbf{E}$ by

$$N\mathbf{E} := \frac{1}{V}\sum_k \mathbf{E}_k,$$

and obtain with Eq. (9.14),

$$m\frac{d}{dt}\langle\mathbf{j}(t)\rangle = N\mathbf{E}e^2 - \xi\langle\mathbf{j}(t)\rangle.$$

In the following we write

$$\langle\mathbf{j}(\mathbf{r},t)\rangle \rightarrow \mathbf{j}(t).$$

Then the equation is

$$m\frac{d\mathbf{j}}{dt} = N\mathbf{E}e^2 - \xi\mathbf{j}.$$

For $\mathbf{j} =$ const. we have

$$\mathbf{j} = \sigma_0\mathbf{E}, \quad \text{i.e.} \quad \sigma_0 = \frac{Ne^2}{\xi}.$$

The quantity σ_0 is called *direct current conductivity*. It follows that $\xi = Ne^2/\sigma_0$. Setting

$$\frac{1}{\tau} := \frac{\xi}{m},$$

we see that τ has the dimension of time and is interpreted as a mean time interval between successive collisions of the electron with the lattice, called the *relaxation time*. We thus obtain the equation

$$\frac{d\mathbf{j}}{dt} + \frac{1}{\tau}\mathbf{j} = \frac{Ne^2}{m}\mathbf{E}. \qquad (9.15)$$

Assuming $\mathbf{E} \propto e^{-i\omega t}, \mathbf{j} \propto e^{-i\omega t}$, we have

$$-i\omega\mathbf{j} + \frac{1}{\tau}\mathbf{j} = \frac{Ne^2}{m}\mathbf{E},$$

i.e.

$$\mathbf{j} = \sigma(\omega)\mathbf{E}$$

with

$$\sigma(\omega) = \frac{Ne^2/m}{-i\omega + \frac{1}{\tau}} = \frac{\sigma(0)}{1 - i\omega\tau},$$

where $\sigma(0) = \sigma_0$. For small frequencies ω, we have $\sigma(0) = $ constant. $\sigma(\omega)$ is real also for $\omega\tau \ll 1$, and almost purely imaginary for $\omega\tau \gg 1$ (this implies a phase difference of $\pi/2$ between $\mathbf{j}$ and $d\mathbf{j}/dt$, or between $\mathbf{j}$ and $\mathbf{E}$).

We can now insert the expression for $\sigma(\omega)$ into Eq. (9.11) in order to obtain the complete frequency dependence of the generalised dielectric constant, i.e.

$$\eta = \epsilon + i\frac{\sigma(\omega)}{\omega}. \qquad (9.16)$$

Fields which are called "high frequency fields" in practice are those for which

$$\omega\tau \ll 1, \quad \omega \ll 1/\tau,$$

where $1/\tau$ is large in the case of strong damping. Thus the term "high frequency" used in practice does not correspond to the limiting case of $\omega \rightarrow \infty$. In the case $\omega\tau \ll 1$ we have σ real, so that damping occurs. This is the reason why these so-called "high frequency fields" hardly penetrate into a conductor. In general, however, ϵ and σ are frequency dependent.

We consider now the ratio of *displacement current to conduction current*, i.e. that of their densities

$$\frac{j_d}{j} = \frac{\partial D/\partial t}{\sigma E} = \frac{\epsilon \partial E/\partial t}{\sigma E}.$$

For $E = E_0 e^{i\omega t}$ one has $\sigma \sim 5.8 \times 10^7 \ \Omega^{-1}$ per meter (or older: mhos per meter). This is such a high value that for all frequencies in ordinary circuits $|j_d/j| < 1\%$. We look at this relationship more closely. We define the *plasma frequency* ω_P by

$$\omega_P^2 = \frac{Ne^2}{\epsilon m}.$$

Then

$$\eta(\omega) = \epsilon + i\frac{\sigma(\omega)}{\omega} = \epsilon + i\frac{Ne^2}{m\omega}\frac{1}{-i\omega + \frac{1}{\tau}}$$

$$= \epsilon\left[1 + \frac{i\omega_P^2}{\omega(-i\omega + \frac{1}{\tau})}\right] = \epsilon\left[1 - \frac{\omega_P^2}{\omega^2(1 - \frac{1}{i\omega\tau})}\right].$$

We consider the case $\omega\tau \gg 1$.
In this case

$$\eta(\omega) = \epsilon\left[1 - \frac{\omega_P^2}{\omega^2}\right].$$

We distinguish between two subsidiary cases:

(a) $\omega \ll \omega_P$:

$$\eta(\omega) \simeq -\epsilon\frac{\omega_P^2}{\omega^2}.$$

The index of refraction

$$p = c\sqrt{\eta\mu} \simeq ic\sqrt{\epsilon\mu}\frac{\omega_P}{\omega}$$

is in this case pure imaginary. In order to obtain the penetration depth we have to calculate

$$k^2 = \omega^2\eta\mu \simeq \omega^2\left(-\epsilon\frac{\omega_P^2}{\omega^2}\right)\mu = -\epsilon\mu\omega_P^2,$$

so that

$$e^{ikz} = e^{i\sqrt{-\epsilon\mu\omega_P^2}z} = e^{\pm\sqrt{\epsilon\mu}\omega_P z} \equiv e^{\pm z/z_0},$$

i.e. the penetration depth

$$z_0 = \frac{1}{\sqrt{\epsilon\mu}\omega_P} = \sqrt{\frac{m}{Ne^2\mu}}$$

is *independent of the frequency.*
     The conductivity is

$$\sigma(\omega) = \frac{Ne^2/m}{-i\omega + \frac{1}{\tau}} = \frac{Ne^2\tau/m}{-i\omega\tau + 1}.$$

In the case under consideration ($\omega\tau \gg 1$), we have

$$\sigma(\omega) \simeq \frac{Ne^2/m}{-i\omega\tau} = i\frac{Ne^2}{m\omega}.$$

The factor "i" implies, that $\mathbf{j}$ and $\mathbf{E}$ of

$$\mathbf{j} = \sigma(\omega)\mathbf{E}$$

have a phase difference of $\pi/2$. Since τ does not appear here, we would have obtained this case also without the friction term in the equation for $\mathbf{j}$, i.e. from

$$\frac{d\mathbf{j}}{dt} = \frac{Ne^2}{m}\mathbf{E}.$$

Thus this case is normally not applicable to rigid bodies or metals (these possess an atomic lattice which gives rise to the friction term). The case finds application, however, for instance in the case of ions (charge e) in the stratosphere.

(b) $\omega \gg \omega_P$:
In this case

$$\eta(\omega) = \epsilon\left(1 - \frac{\omega_P^2}{\omega^2}\right)$$

and

$$p = c\sqrt{\eta\mu} \simeq c\sqrt{\epsilon\mu}\sqrt{1 - \frac{\omega_P^2}{\omega^2}}$$

is real. Hence

$$k^2 = \omega^2\eta\mu = \omega^2\epsilon\mu\left(1 - \frac{\omega_P^2}{\omega^2}\right)$$

is positive, so that

$$k = \omega\sqrt{\epsilon\mu}\sqrt{1 - \frac{\omega_P^2}{\omega^2}} \qquad (9.17)$$

is real. With $\sqrt{\epsilon\mu} = 1/c'^2$ this relation can be written

$$k^2c'^2 = \omega^2 - \omega_P{}^2.$$

This means, there is no damping. On the other hand, again

$$\sigma(\omega) = i\frac{Ne^2}{m\omega} = i\frac{\epsilon\omega_P^2}{\omega} \propto \frac{1}{\omega},$$

so that

$$\mathbf{j} = \sigma(\omega)\mathbf{E} \sim 0 \quad \text{for} \ \omega \gg \omega_P.$$

Then there is practically no conduction current. Since the resistance of the lattice is practically zero (no damping) ($1/\tau \ll \omega$, hence $1/\tau$ negligible), the explanation has to be found in the conduction electrons — one can say, at high frequencies ω they acquire *"inertia"*, so that $\mathbf{j} \to 0$.[†] Thus fields called "high frequency fields" in practice are not those with $\omega \gg \omega_P$ and $\omega\tau \gg 1$, but those for which $\omega\tau \ll 1$. In the latter case $\sigma \sim$ is real, and damping occurs. We see therefore: For $\omega\tau \ll 1$ there is little dissipation of energy in the metal (the wave is almost completely reflected), and only for $\omega > \omega_P$ one has transmission and little reflection.

9.5   The (Normal) Skin Effect

Let electromagnetic radiation with frequency $\omega/2\pi$ fall in the direction of x on an uncharged metal of conductivity σ and magnetic permeability μ. We assume the time dependence of the fields to be given by $\mathbf{E}, \mathbf{H} \propto e^{i\omega t}$. The relevant Maxwell equations together with

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

are then (neglecting the displacement current),

$$(1) \ \ \boldsymbol{\nabla} \times \mathbf{H} = \mathbf{j}, \quad (2) \ \ \boldsymbol{\nabla} \times \mathbf{E} = -\mu\frac{\partial \mathbf{H}}{\partial t},$$
$$(3) \ \ \boldsymbol{\nabla} \cdot \mathbf{H} = 0, \quad (4) \ \ \boldsymbol{\nabla} \cdot \mathbf{E} = 0.$$

Taking the curl of Eq. (2) we obtain

$$\boldsymbol{\nabla} \times (\boldsymbol{\nabla} \times \mathbf{E}) = \boldsymbol{\nabla}(\boldsymbol{\nabla} \cdot \mathbf{E}) - \triangle\mathbf{E} = -\mu\frac{\partial}{\partial t}\underbrace{(\boldsymbol{\nabla} \times \mathbf{H})}_{\mathbf{j}},$$

i.e.

$$\triangle\mathbf{j} = \mu\sigma\frac{\partial \mathbf{j}}{\partial t}.$$

Since $\mathbf{j} \propto e^{i\omega t}$, it follows that

$$\triangle\mathbf{j} = \gamma^2 \mathbf{j}, \ \ \gamma^2 = i\omega\mu\sigma$$

with

$$\gamma = (1+i)\sqrt{\frac{\mu\sigma\omega}{2}} \equiv (1+i)\lambda, \quad \lambda = \sqrt{\frac{\mu\sigma\omega}{2}} > 0. \tag{9.18}$$

[†]See the discussion in Greiner [4], p. 300: For $\omega \to \infty$ metals become "transparent", i.e. the metal behaves like a dielectric permitting the transmission of radiation.

For

$$\mathbf{j}(x,t) = e^{i\omega t}\mathbf{j}(x)$$

we therefore have $(j_0 = j(0))$:

$$\frac{d^2}{dx^2}j(x) = j_0(x)e^{-\gamma x} = j_0 e^{-\lambda x}e^{-i\lambda x}.$$

For obvious physical reasons we have to choose the damped solution. The expression $1/\lambda$ is called penetration or *skin depth*.

Chapter 10

Moving Charges in Vacuum

10.1   Introductory Remarks

In the following we study moving charges in a vacuum, i.e. we consider the Maxwell equations with nonvanishing charge density (hence also with nonvanishing current density). We investigate the solutions of the wave equations and thus obtain potentials and fields. Finally we consider the radiation of an oscillating dipole (consisting of two charges). In particular we shall see that accelerations of charges produce the dominant field contributions at large distances (called *"radiation fields"*) and that these are mutually orthogonal. These considerations are of eminent importance for an understanding of electrodynamical phenomena, also because the dipole is often referred to as a kind of idealised classical model of an atom and thus offers the explanation why this picture is wrong — the dipole radiates off energy and exhausts itself therewith.

10.2   Maxwell's Equations for Moving Charges

Here $\mathbf{D} = \epsilon_0 \mathbf{E}, \mathbf{B} = \mu_0 \mathbf{H}$, so that Maxwell's equations become

$$\boldsymbol{\nabla} \cdot \mathbf{E} = \frac{1}{\epsilon_0} \rho(\mathbf{r}, t), \quad \boldsymbol{\nabla} \cdot \mathbf{H} = 0$$

and

$$\boldsymbol{\nabla} \times \mathbf{E} = -\mu_0 \frac{\partial \mathbf{H}}{\partial t} = -\frac{\partial}{\partial t}(\mu_0 \mathbf{H}) \tag{10.1}$$

and

$$\boldsymbol{\nabla} \times \mathbf{H} = \mathbf{j}(\mathbf{r}, t) + \epsilon_0 \frac{\partial \mathbf{E}}{\partial t}. \tag{10.2}$$

Expressing $\mathbf{B}$ in terms of the vector potential $\mathbf{A}$, we have

$$\mu_0 \mathbf{H} = \boldsymbol{\nabla} \times \mathbf{A}$$

and from Eq. (10.1) we have

$$\boldsymbol{\nabla} \times \mathbf{E} = -\boldsymbol{\nabla} \times \frac{\partial \mathbf{A}}{\partial t},$$

i.e.

$$\mathbf{E} = -\frac{\partial \mathbf{A}}{\partial t} - \boldsymbol{\nabla}\phi,$$

as already familiar. With $\boldsymbol{\nabla} \cdot \mathbf{E} = \rho/\epsilon_0$ we obtain

$$\frac{1}{\epsilon_0}\rho = -\frac{\partial}{\partial t}\boldsymbol{\nabla} \cdot \mathbf{A} - \boldsymbol{\nabla} \cdot \boldsymbol{\nabla}\phi \tag{10.3}$$

and with Eq. (10.2) we obtain

$$\boldsymbol{\nabla} \times \mathbf{H} = \frac{1}{\mu_0}\underbrace{\boldsymbol{\nabla} \times (\boldsymbol{\nabla} \times \mathbf{A})}_{\boldsymbol{\nabla}(\boldsymbol{\nabla}\cdot\mathbf{A})-\triangle\mathbf{A}} = \mathbf{j} + \epsilon_0\frac{\partial}{\partial t}\left(-\frac{\partial \mathbf{A}}{\partial t} - \boldsymbol{\nabla}\phi \right). \tag{10.4}$$

Equation (10.3) reduces to

$$-\triangle\phi - \frac{\partial}{\partial t}\boldsymbol{\nabla} \cdot \mathbf{A} = \frac{1}{\epsilon_0}\rho \tag{10.5}$$

and Eq. (10.4) to

$$-\triangle\mathbf{A} + \mu_0\epsilon_0\frac{\partial^2 \mathbf{A}}{\partial t^2} + \boldsymbol{\nabla}\left(\boldsymbol{\nabla} \cdot \mathbf{A} + \mu_0\epsilon_0\frac{\partial\phi}{\partial t} \right) = \mu_0\mathbf{j}, \tag{10.6}$$

where $\mu_0\epsilon_0 = 1/c^2$.

We have seen earlier that the observable fields $\mathbf{E}$ and $\mathbf{H}$ are independent of the particular choice or gauge of the vector potential. Previously, however, we encountered the transformation from one gauge potential to another only in the context of electrostatics and magnetostatics (see, however, Sec. 7.5). If the fields depend also on time, the gauge transformations are given by

$$\mathbf{A} = \mathbf{A}' + \boldsymbol{\nabla}\chi, \quad \phi = \phi' - \frac{\partial\chi}{\partial t} \tag{10.7}$$

for arbitrary functions χ. We apply these transformations to Eqs. (10.5) and (10.6) and obtain

$$-\triangle(\phi' - \frac{\partial\chi}{\partial t}) - \frac{\partial}{\partial t}\boldsymbol{\nabla} \cdot (\mathbf{A}' + \boldsymbol{\nabla}\chi) = \frac{1}{\epsilon_0}\rho,$$

$$-\triangle(\mathbf{A}' + \boldsymbol{\nabla}\chi) + \frac{1}{c^2}\frac{\partial^2}{\partial t^2}(\mathbf{A}' + \boldsymbol{\nabla}\chi)$$

$$+\boldsymbol{\nabla}\left(\boldsymbol{\nabla} \cdot \mathbf{A}' + \boldsymbol{\nabla} \cdot \boldsymbol{\nabla}\chi + \frac{1}{c^2}\frac{\partial\phi'}{\partial t} - \frac{1}{c^2}\frac{\partial^2\chi}{\partial t^2} \right) = \mu_0\mathbf{j},$$

i.e.

$$-\triangle\phi' - \frac{\partial}{\partial t}\boldsymbol{\nabla}\cdot\mathbf{A}' = \frac{1}{\epsilon_0}\rho \qquad (10.8)$$

and

$$-\triangle\mathbf{A}' + \frac{1}{c^2}\frac{\partial^2}{\partial t^2}\mathbf{A}' + \boldsymbol{\nabla}\left(\boldsymbol{\nabla}\cdot\mathbf{A}' + \frac{1}{c^2}\frac{\partial\phi'}{\partial t}\right) = \mu_0\mathbf{j}. \qquad (10.9)$$

We observe: The terms in χ cancel out, i.e. the equations are invariant under the gauge transformation. We also observe: For a special gauge, which means if we consider potentials $(\mathbf{A}', \phi')$ which satisfy the *additionally chosen condition*

$$\boldsymbol{\nabla}\cdot\mathbf{A}' + \frac{1}{c^2}\frac{\partial\phi'}{\partial t} = 0, \qquad (10.10)$$

called *Lorentz gauge*, the equations simplify to

$$\Box\phi' = -\frac{1}{\epsilon_0}\rho, \quad \Box\mathbf{A}' = -\mu_0\mathbf{j}. \qquad (10.11)$$

In fact, that it is even necessary to demand a condition like the Lorentz gauge* can be seen by a more detailed investigation of the solutions of Eqs. (10.8), (10.9) and (10.11). In order to see that the gauge condition implies the *transversality* of the electromagnetic field, and instead of the four components of the potentials $(\mathbf{A}', \phi')$ the independence of only two mutually orthogonal field components $\mathbf{E}$ and $\mathbf{H}$, also requires further detailed consideration.[†] In the following we write $(\mathbf{A}', \phi')$ again as $(\mathbf{A}, \phi)$. Thus we have

$$\boldsymbol{\nabla}\cdot\mathbf{A} + \frac{1}{c^2}\frac{\partial\phi}{\partial t} = 0, \qquad (10.12)$$

called Lorentz condition, and we have with $\Box \equiv \triangle - \frac{1}{c^2}\frac{\partial^2}{\partial t^2}$ the equations

$$\Box\phi = -\frac{1}{\epsilon_0}\rho, \quad \Box\mathbf{A} = -\mu_0\mathbf{j}. \qquad (10.13)$$

Example 10.1: Coulomb gauge

Starting from Maxwell's equations for moving charges in vacuum, derive the wave equations of the potentials $\phi, \mathbf{A}$ in Coulomb gauge ($\boldsymbol{\nabla}\cdot\mathbf{A} = 0$) and show that a transverse current is the source of the potential $\mathbf{A}$.

Solution: From the above we have the Eqs. (10.8) and (10.9), i.e.

$$\triangle\phi + \frac{\partial}{\partial t}\boldsymbol{\nabla}\cdot\mathbf{A} = -\frac{1}{\epsilon_0}\rho,$$

---

[*]A different gauge, called Coulomb gauge, $\boldsymbol{\nabla}\cdot\mathbf{A} = 0$, is treated, for example, in Schwabl [16], p. 127.
[†]See Chapter 18.

and

$$\triangle \mathbf{A} - \frac{1}{c^2}\frac{\partial^2}{\partial t^2}\mathbf{A} - \boldsymbol{\nabla}\left(\boldsymbol{\nabla}\cdot\mathbf{A} + \frac{1}{c^2}\frac{\partial\phi}{\partial t}\right) = -\mu_0\mathbf{j}.$$

With Coulomb gauge $\boldsymbol{\nabla}\cdot\mathbf{A} = 0$ and $\square = \triangle - \frac{1}{c^2}(\partial^2/\partial t^2)$, it follows that

$$\triangle\phi = -\frac{1}{\epsilon_0}\rho, \quad \square\mathbf{A} = -\mu_0\mathbf{j}_t,$$

where (with $\mathbf{D} = \epsilon_0\mathbf{E}$)

$$\begin{aligned}
\mathbf{j}_t &= \mathbf{j} - \frac{1}{\mu_0 c^2}\boldsymbol{\nabla}\frac{\partial\phi}{\partial t} = \mathbf{j} - \frac{1}{\mu_0 c^2}\frac{\partial}{\partial t}(\boldsymbol{\nabla}\phi) \\
&= \mathbf{j} + \frac{\partial\mathbf{D}}{\partial t} + \epsilon_0\frac{\partial^2\mathbf{A}}{\partial t^2} \quad (\text{with} \quad \mathbf{E} = -\frac{\partial\mathbf{A}}{\partial t} - \boldsymbol{\nabla}\phi) \\
&= \boldsymbol{\nabla}\times\mathbf{H} + \epsilon_0\frac{\partial^2\mathbf{A}}{\partial t^2}.
\end{aligned}$$

Together with Ampère's law $\boldsymbol{\nabla}\times\mathbf{B} + \triangle\mathbf{A} = 0$, we also have

$$\triangle\mathbf{A} = -\mu_0\boldsymbol{\nabla}\times\mathbf{H} \equiv -\mu_0\mathbf{j}_{tt}$$

(this defines $\mathbf{j}_{tt}$). Hence

$$\mathbf{j}_t = \mathbf{j}_{tt} + \epsilon_0\frac{\partial^2\mathbf{A}}{\partial t^2}.$$

We construct the divergence of this with Coulomb gauge. The operator $-i\boldsymbol{\nabla}$ applied to $\mathbf{A} \propto e^{i\mathbf{k}\cdot\mathbf{r}}$ yields $\mathbf{k}$, the propagation vector. Since

$$\boldsymbol{\nabla}\cdot\mathbf{j}_{tt} = \boldsymbol{\nabla}\cdot(\boldsymbol{\nabla}\times\mathbf{H}) = 0,$$

we have $\mathbf{j}_{tt} \perp \mathbf{k}$, i.e. $\boldsymbol{\nabla}\cdot\mathbf{j}_t = 0$, which means $\mathbf{j}_t$ is transverse.

In order to be able to solve Eqs. (10.13) we need the appropriate *Green's function* $G(\mathbf{r},t;\mathbf{r}',t')$, which is an inhomogeneous solution of

$$\square G(\mathbf{r},t;\mathbf{r}',t') = \delta(\mathbf{r}-\mathbf{r}')\delta(t-t'). \tag{10.14}$$

In the following Example 10.2 we calculate that particular Green's function, which represents the propagation of a light signal or disturbance of some kind as a spherical wave spreading out with velocity c from $\mathbf{r}'$ at time t' to $\mathbf{r}$ at time $t > t'$. This Green's function which takes into account causality is called the *retarded Green's function*, and the corresponding potentials are called the *retarded potentials*. According to the following Example 10.2 this *retarded Green's function* is given by

$$G(\mathbf{r},t;\mathbf{r}',t') = -\frac{\delta(t-t'-\frac{|\mathbf{r}-\mathbf{r}'|}{c})}{4\pi|\mathbf{r}-\mathbf{r}'|}. \tag{10.15}$$

Hence

$$\phi(\mathbf{r},t) = -\frac{1}{\epsilon_0}\int G(\mathbf{r},t;\mathbf{r}',t')\rho(\mathbf{r}',t')d\mathbf{r}'dt', \tag{10.16}$$

implies

$$\phi(\mathbf{r}, t) = \frac{1}{4\pi\epsilon_0} \int d\mathbf{r}' \frac{\rho(\mathbf{r}', t - \frac{|\mathbf{r}-\mathbf{r}'|}{c})}{|\mathbf{r} - \mathbf{r}'|}. \tag{10.17}$$

Similarly one obtains

$$\mathbf{A}(\mathbf{r}, t) = \frac{\mu_0}{4\pi} \int d\mathbf{r}' \frac{\mathbf{j}(\mathbf{r}', t - \frac{|\mathbf{r}-\mathbf{r}'|}{c})}{|\mathbf{r} - \mathbf{r}'|}. \tag{10.18}$$

Example 10.2: The retarded Green's function

Calculate the Green's function G which takes into account causality, i.e. that a cause must precede its effect.

Solution: We have

$$\Box = \triangle - \frac{1}{c^2}\frac{\partial^2}{\partial t^2}$$

and the Fourier representation of the 4-dimensional delta function, i.e.

$$\delta(\mathbf{r} - \mathbf{r}')\delta(t - t') = \frac{1}{(2\pi)^4} \int_{-\infty}^{\infty} d\mathbf{k} \int_{-\infty}^{\infty} d\omega e^{i\mathbf{k}\cdot(\mathbf{r}-\mathbf{r}')} e^{-i\omega(t-t')}. \tag{10.19}$$

The Fourier representation of the time-dependent Green's function can be written

$$G(\mathbf{r}, t; \mathbf{r}', t') = \int_{-\infty}^{\infty} d\mathbf{k} \int_{-\infty}^{\infty} d\omega g(\mathbf{k}, \omega) e^{i\mathbf{k}\cdot(\mathbf{r}-\mathbf{r}')} e^{-i\omega(t-t')},$$

where the spectral function $g(\mathbf{k}, \omega)$ is still to be determined. First we have

$$\Box G(\mathbf{r}, t; \mathbf{r}', t') = \int_{-\infty}^{\infty} d\mathbf{k} \int_{-\infty}^{\infty} d\omega g(\mathbf{k}, \omega) \left(\frac{\omega^2}{c^2} - k^2\right) e^{i\mathbf{k}\cdot(\mathbf{r}-\mathbf{r}')} e^{-i\omega(t-t')}. \tag{10.20}$$

Inserting (10.19) and (10.20) into the equation for G, Eq. (10.14), and comparing coefficients, we obtain

$$g(\mathbf{k}, \omega) = -\frac{1}{(2\pi)^4} \frac{1}{k^2 - \frac{\omega^2}{c^2}} \tag{10.21}$$

and therefore

$$G(\mathbf{r}, t; \mathbf{r}', t') = -\frac{1}{(2\pi)^4} \int_{-\infty}^{\infty} d\mathbf{k} \int_{-\infty}^{\infty} d\omega \frac{c^2}{c^2 k^2 - \omega^2} e^{i[\mathbf{k}\cdot(\mathbf{r}-\mathbf{r}')-\omega(t-t')]}. \tag{10.22}$$

The integrand has simple poles at $\omega = \pm ck$. The integration is performed with the help of Cauchy's residue theorem, i.e.

$$\oint f(z)dz = 2\pi i \sum_{k=1}^{n} \text{Res}_k f(z).$$

The Green's function represents an instantaneous and local disturbance or event or something similar which occurs at radius $\mathbf{r}'$ at time $t = t'$ and propagates as a spherical wave with velocity c. A causality-preserving wave propagating forward in time must be such that

$$G(\mathbf{r}, t; \mathbf{r}', t') = 0 \quad \text{for} \quad t < t'.$$

Thus for $t > t'$ the function G represents a wave which spreads into the future. The singularities of $g(\mathbf{k}, \omega)$ are at $\omega = \pm ck$. We now consider the plane of complex ω with $\omega = \omega_1 + i\omega_2$, and try to

define the causal waves by a suitable choice of the integration contour in this plane. To ensure that $G = 0$ for $t < t'$, we have to circumvent the poles at $\omega = \pm ck$. This can be achieved by displacing them by an infinitesimal amount $(-i\epsilon)$, which means replacing in the above equation ω by $\omega + i\epsilon$, so that we write

$$G(\mathbf{r}, t; \mathbf{r}', t') = -\frac{1}{(2\pi)^4} \int_{-\infty}^{\infty} d\mathbf{k} \int_{-\infty}^{\infty} d\omega \frac{e^{i[\mathbf{k}\cdot(\mathbf{r}-\mathbf{r}')-\omega(t-t')]}}{k^2 - \frac{(\omega+i\epsilon)^2}{c^2}}. \tag{10.23}$$

We consider first the integration with respect to ω, i.e. the integral

$$I(k) = \int_{-\infty}^{\infty} d\omega \frac{e^{-i\omega(t-t')}}{k^2 - \frac{(\omega+i\epsilon)^2}{c^2}}.$$

The idea is, to add to the integral from $-\infty$ to $+\infty$ the integral along the infinite semi-circle either in the upper or in the lower half-plane so that as a consequence of the displacement of the poles by $-i\epsilon, \epsilon > 0$, the integral along this semi-circle falls off exponentially towards infinity and hence does not contribute. This means, for $t - t' > 0$ we have to choose the semi-circle in the lower half-plane, and for $t - t' > 0$ in the upper half-plane. Since we have two poles of the first order, we obtain two residues. We let the poles be $\omega_1 = ck - i\epsilon$ and $\omega_2 = -ck - i\epsilon$, both in the lower half-plane. Their residues are the coefficients of the factors $1/(\omega - \omega_1)$ and $1/(\omega - \omega_2)$ in the integrand. These are easily obtained as

$$\text{residue at } \omega_1 = -\frac{c}{2k}e^{-ick(t-t')}, \quad \text{residue at } \omega_2 = \frac{c}{2k}e^{ick(t-t')}.$$

We thus obtain with Cauchy's theorem

$$I(k) = -2\pi i\left(-\frac{c}{2k}e^{-ick(t-t')} + \frac{c}{2k}e^{ick(t-t')}\right) = \frac{2\pi c}{k}\sin(ck(t-t')).$$

Hence we have thus far

$$G(\mathbf{r}, t; \mathbf{r}', t') = \begin{cases} 0 & : \quad t < t', \\ -\frac{c}{(2\pi)^3}\int_{-\infty}^{\infty} d\mathbf{k}\, e^{i\mathbf{k}\cdot(\mathbf{r}-\mathbf{r}')}k^{-1}\sin(ck(t-t')) & : \quad t > t'. \end{cases} \tag{10.24}$$

We still have to perform the integration with respect to $\mathbf{k}$. To this end we introduce spherical coordinates. Let θ be the polar angle to the vector $\mathbf{r} - \mathbf{r}'$ as z-axis. We then have

$$d\mathbf{k} = k^2 \sin\theta\, dk\, d\theta\, d\phi, \quad \mathbf{k}\cdot(\mathbf{r} - \mathbf{r}') = k|\mathbf{r} - \mathbf{r}'|\cos\theta.$$

Inserting this into the expression for G we obtain

$$G(\mathbf{r}, t; \mathbf{r}', t') = -\frac{c}{(2\pi)^3}\int d\phi d\theta dk k \sin\theta \sin(ck(t-t'))e^{ik\cos\theta|\mathbf{r}-\mathbf{r}'|}.$$

Integration with respect to ϕ yields the factor 2π, and we have

$$G(\mathbf{r}, t; \mathbf{r}', t') = \frac{c}{(2\pi)^2}\int_{k=0}^{\infty} k \sin(ck(t-t'))dk \int_{\theta=0}^{\pi} e^{ik\cos\theta|\mathbf{r}-\mathbf{r}'|}d\cos\theta.$$

The integral over θ yields $-2\sin(k|\mathbf{r} - \mathbf{r}'|)/k|\mathbf{r} - \mathbf{r}'|$, so that

$$G(\mathbf{r}, t; \mathbf{r}', t') = \frac{-c}{2\pi^2|\mathbf{r} - \mathbf{r}'|}\int_0^{\infty} \sin(k|\mathbf{r} - \mathbf{r}'|)\sin(ck(t-t'))dk.$$

Now we use the relation

$$c \int_0^\infty dk \sin Akc \sin Bkc = -\frac{\pi}{2}[\delta(A+B) - \delta(A-B)],$$

which can be verified by replacing the sines by exponentials and using the integral representation of the one-dimensional delta function. Thus G can be expressed as the sum of two delta function contributions, i.e. we obtain

$$G(\mathbf{r}, t; \mathbf{r}', t') = \frac{1}{4\pi|\mathbf{r}-\mathbf{r}'|} \left[\delta\left(t - t' + \frac{|\mathbf{r}-\mathbf{r}'|}{c}\right) - \delta\left(t - t' - \frac{|\mathbf{r}-\mathbf{r}'|}{c}\right)\right].$$

Since the argument of the second delta function is nowhere zero for $t > t'$ this cannot contribute to G. We thus obtain

$$G(\mathbf{r}, t; \mathbf{r}', t') = -\frac{\delta(t - t' - \frac{|\mathbf{r}-\mathbf{r}'|}{c})}{4\pi|\mathbf{r}-\mathbf{r}'|}. \tag{10.25}$$

This Green's function describes the effect at time $t > t'$ and at radius $r > r'$ of a delta function-like cause at time $t = t'$ and radius $r = r'$ propagating from the source with the velocity of light.

10.3 The Liénard–Wiechert Potentials of a Moving Point Charge

Our aim is now to calculate the potentials $\phi(\mathbf{r}, t), \mathbf{A}(\mathbf{r}, t)$ for the case of a point charge moving with velocity $\mathbf{v}(\mathbf{t})$. The very important results we obtain are the so-called *Liénard–Wiechert potentials*. We expect, of course, to regain in the limit of vanishing velocity the well known static expressions we had before.

We let $\mathbf{r}_0(t)$ be the vector representing the path or trajectory of the point charge e and set

$$\mathbf{v}(t) = \dot{\mathbf{r}}_0(t)$$

for its velocity. The charge density can then be written

$$\rho(\mathbf{r}, t) = e\delta(\mathbf{r} - \mathbf{r}_0(t)) \tag{10.26}$$

and the current density

$$\mathbf{j}(\mathbf{r}, t) = \rho(\mathbf{r}, t)\mathbf{v}(t) = e\mathbf{v}(t)\delta(\mathbf{r} - \mathbf{r}_0(t)). \tag{10.27}$$

We consider first the scalar potential ϕ. Inserting into Eq. (10.17) the

expression for ρ, we obtain

$$
\phi(\mathbf{r}, t) = \frac{e}{4\pi\epsilon_0} \int d\mathbf{r}' \frac{\delta\left(\mathbf{r}' - \mathbf{r}_0(t - \frac{|\mathbf{r} - \mathbf{r}'|}{c})\right)}{|\mathbf{r} - \mathbf{r}'|}
$$

$$
= \frac{e}{4\pi\epsilon_0} \int dt' \int d\mathbf{r}' \frac{\delta(\mathbf{r}' - \mathbf{r}_0(t'))\delta\left(t' - t + \frac{|\mathbf{r} - \mathbf{r}'|}{c}\right)}{|\mathbf{r} - \mathbf{r}'|}
$$

$$
= \frac{e}{4\pi\epsilon_0} \int dt' \frac{\delta\left(t' - t + \frac{|\mathbf{r} - \mathbf{r}_0(t')|}{c}\right)}{|\mathbf{r} - \mathbf{r}_0(t')|}.
$$

We define a time difference u by

$$
u = t' - t + \frac{|\mathbf{r} - \mathbf{r}_0(t')|}{c} = t' - t + \frac{\sqrt{(\mathbf{r} - \mathbf{r}_0(t'))^2}}{c}, \tag{10.28}
$$

so that

$$
\frac{du}{dt'} = 1 - \frac{\mathbf{n}(t') \cdot \mathbf{v}(t')}{c}, \quad \mathbf{n} = \frac{\mathbf{r} - \mathbf{r}_0(t')}{|\mathbf{r} - \mathbf{r}_0(t')|}
$$

and

$$
dt' = \frac{du}{1 - \frac{\mathbf{n}(t') \cdot \mathbf{v}(t')}{c}}.
$$

It follows that

$$
\phi(\mathbf{r}, t) = \frac{e}{4\pi\epsilon_0} \int \frac{du}{1 - \frac{\mathbf{n}(t') \cdot \mathbf{v}(t')}{c}} \frac{\delta(u)}{|\mathbf{r} - \mathbf{r}_0(t')|},
$$

i.e.

$$
\phi(\mathbf{r}, t) = \frac{e/4\pi\epsilon_0}{|\mathbf{r} - \mathbf{r}_0(t')|\left(1 - \mathbf{n}(t') \cdot \frac{\mathbf{v}(t')}{c}\right)}, \tag{10.29}
$$

where

$$
t' = t - \frac{1}{c}|\mathbf{r} - \mathbf{r}_0(t')|, \quad \mathbf{v}(t') \equiv \dot{\mathbf{r}}_0(t').
$$

In a similar way we obtain

$$
\mathbf{A}(\mathbf{r}, t) = \frac{e\mathbf{v}(t')\mu_0/4\pi}{|\mathbf{r} - \mathbf{r}_0(t')|\left(1 - \mathbf{n}(t') \cdot \frac{\mathbf{v}(t')}{c}\right)}. \tag{10.30}
$$

These expressions, called *Liénard–Wiechert potentials*, are *exact solutions* of the above wave equations *for a pointlike, charged particle* moving along the trajectory $\mathbf{r}_0(t')$. For

$$\mathbf{r} - \mathbf{r}_0(t') \equiv \mathbf{r}(t') \tag{10.31}$$

the potential ϕ is

$$\phi(\mathbf{r}, t) = \left(\frac{e/4\pi\epsilon_0}{r(t') - \frac{\mathbf{r}(t') \cdot \mathbf{v}(t')}{c}} \right)_{t' = t - \frac{r(t')}{c}}, \tag{10.32}$$

i.e. $|\mathbf{r}(t')| = c(t - t')$. One should note, that the vector $\mathbf{r}$ is here to be considered as fixed and $\mathbf{r} = \mathbf{r}_0(t') + \mathbf{r}(t')$. The expression

$$t' = t - \frac{r(t')}{c}$$

is called *retarded time* and is defined implicitly by this equation (e.g. when $r(t') = vt' + a$ we have $t' = t - (vt' + a)/c, t - t' = (a/c + vt/c)/(1 + v/c)$). If c were infinite (which is *not* possible according to the theory of relativity), we would have $t' = t$ for $r(t') = $ finite. This means the radiation emitted by the charge would reach any other point in space instantaneously (which is nonsense). However, for $c = $ finite, the radiation emitted by the charge reaches other points only with delay, or, as one says, with *retarded time*. This means the radiation (i.e. the fields) observed at the point $\mathbf{r}$ at time t was emitted by the charge at its earlier position $t - t'$ seconds before.

Analogously to Eq. (10.32) we obtain

$$\mathbf{A}(\mathbf{r}, t) = \frac{e\mu_0}{4\pi} \left(\frac{\mathbf{v}(t')}{r(t') - \frac{\mathbf{r}(t') \cdot \mathbf{v}(t')}{c}} \right)_{t' = t - \frac{r(t')}{c}}. \tag{10.33}$$

These are the potentials at position $\mathbf{r}$ (or $\mathbf{r}(t')$ away from $\mathbf{r}_0(t')$) and at time t in the case of a point charge e. One should observe our notation, in order not to confuse $\mathbf{r}$ and $\mathbf{r}(t')$. In the following we apply these potentials to the oscillating electric dipole, called *Hertz dipole*, with dipole moment $\mathbf{p}(t)$.

Example 10.3: Lorentz 4-form of the Liénard–Wiechert potentials
(This example assumes familiarity with Chapter 17). Let the components of $\mathbf{r}_0(t')$ be the coordinates of the charge e and the components of $\mathbf{r}$ those of an observation point with light vector (i.e. one with $c^2 = (\text{distance})^2/(\text{time})^2$) $U^\mu = (c(t - t'), \mathbf{r} - \mathbf{r}_0(t'))$ given by $U^\mu U_\mu = 0$. Verify that the Liénard–Wiechert potentials can be combined into the 4-form

$$A^\mu = \frac{e}{4\pi\epsilon_0 c} \frac{u^\mu}{U_\nu u^\nu}$$

where u^μ is the 4-velocity defined (later) by Eq. (17.64), i.e. with components $u_0 = c/\sqrt{1 - \beta^2}, u_i = -v_i/\sqrt{1 - \beta^2}, \beta^2 = v^2/c^2$.

Solution: The light vector is given by $c^2(t-t')^2 - (\mathbf{r} - \mathbf{r}_0(t'))^2 = 0$, i.e. $c(t-t')/\sqrt{(\mathbf{r} - \mathbf{r}_0(t'))^2} = 1$. Hence Eqs. (10.29) and (10.30) with (17.31) and $\epsilon_0 \mu_0 c^2 = 1$ yield the expression

$$
\begin{aligned}
A^\mu &= \left(\frac{1}{c} \phi, \mathbf{A} \right) \\
&= \frac{e}{4\pi\epsilon_0 c^2} \left(\frac{c}{|\mathbf{r} - \mathbf{r}_0|(1 - \frac{\mathbf{n} \cdot \mathbf{v}}{c})}, \frac{\mathbf{v}}{|\mathbf{r} - \mathbf{r}_0|(1 - \frac{\mathbf{n} \cdot \mathbf{v}}{c})} \right) \\
&= \frac{e}{4\pi\epsilon_0 c^2} \frac{(c, \mathbf{v})}{|\mathbf{r} - \mathbf{r}_0| \left(\frac{c(t-t')}{|\mathbf{r}-\mathbf{r}_0|} - \frac{\mathbf{n} \cdot \mathbf{v}}{c} \right)} \\
&= \frac{e}{4\pi\epsilon_0 c} \frac{(c, \mathbf{v})/\sqrt{1-\beta^2}}{(c^2(t-t') - \mathbf{v} \cdot (\mathbf{r} - \mathbf{r}_0))/\sqrt{1-\beta^2}} \\
&= \frac{e}{4\pi\epsilon_0 c} \frac{u^\mu}{U_\nu u^\nu}.
\end{aligned}
$$

10.4   The Fields $\mathbf{E}, \mathbf{B}$ of a Moving Point Charge

We obtain the fields $\mathbf{E}$ and $\mathbf{B}$ from the equations

$$
\mathbf{E} = -\boldsymbol{\nabla}\phi - \frac{\partial \mathbf{A}}{\partial t}, \qquad \mathbf{B} = \boldsymbol{\nabla} \times \mathbf{A}. \tag{10.34}
$$

Thus we have to perform complicated differentiations. We start with $\mathbf{E}$:

$$
\mathbf{E} = -\boldsymbol{\nabla} \frac{e/4\pi\epsilon_0}{|\mathbf{r} - \mathbf{r}_0(t')| \left(1 - \frac{\mathbf{n}(t') \cdot \mathbf{v}(t')}{c} \right)}
$$
$$
- \frac{\partial}{\partial t} \left[\frac{e\mathbf{v}(t')\mu_0/4\pi}{|\mathbf{r} - \mathbf{r}_0(t')| \left(1 - \frac{\mathbf{n}(t') \cdot \mathbf{v}(t')}{c} \right)} \right]. \tag{10.35}
$$

Here $t' = t - |\mathbf{r} - \mathbf{r}_0(t')|/c$, so that

$$
\frac{\partial t'}{\partial t} = 1 - \frac{1}{c} \left(\frac{\partial}{\partial t'} |\mathbf{r} - \mathbf{r}_0(t')| \right) \frac{\partial t'}{\partial t},
$$
$$
\frac{\partial}{\partial t'} |\mathbf{r} - \mathbf{r}_0(t')| = \frac{\partial}{\partial t'} \sqrt{(\mathbf{r} - \mathbf{r}_0(t'))^2}
$$
$$
= -\mathbf{v}(t') \cdot \mathbf{n}(t'), \tag{10.36}
$$

and hence

$$
\frac{\partial t'}{\partial t} = \frac{1}{\left(1 - \frac{\mathbf{n}(t') \cdot \mathbf{v}(t')}{c} \right)}. \tag{10.37}
$$

Similarly (since $\nabla |\mathbf{r}| = \mathbf{r}/r$)

$$\nabla t' = -\frac{1}{c}(\nabla)_{t'=\text{const.}}|\mathbf{r} - \mathbf{r}_0(t')| - \frac{1}{c}\left(\frac{\partial}{\partial t'}|\mathbf{r} - \mathbf{r}_0(t')|\right)\nabla t'$$

$$= -\frac{\mathbf{n}(t')}{c} - \frac{1}{c}\underbrace{\left(\frac{\partial}{\partial t'}|\mathbf{r} - \mathbf{r}_0(t')|\right)}_{-\mathbf{v}(t')\cdot\mathbf{n}(t')}\nabla t'$$

and hence

$$\nabla t' = \frac{-\mathbf{n}(t')}{c\left(1 - \frac{\mathbf{n}(t')\cdot\mathbf{v}(t')}{c}\right)}. \tag{10.38}$$

Moreover, with Eq. (10.36),

$$\frac{\partial \mathbf{n}}{\partial t'} = \frac{\partial}{\partial t'}\left(\frac{\mathbf{r} - \mathbf{r}_0(t')}{|\mathbf{r} - \mathbf{r}_0(t')|}\right)$$

$$= \frac{(\mathbf{v}(t')\cdot\mathbf{n}(t'))(\mathbf{r} - \mathbf{r}_0(t'))}{|\mathbf{r} - \mathbf{r}_0(t')|^2} - \frac{\mathbf{v}(t')}{|\mathbf{r} - \mathbf{r}_0(t')|}$$

$$= \frac{\mathbf{n}(t') \times (\mathbf{n}(t') \times \mathbf{v}(t'))}{|\mathbf{r} - \mathbf{r}_0(t')|}. \tag{10.39}$$

Hence, with $\mathbf{a}(t') = d\mathbf{v}(t')/dt'$ and Eqs. (10.36) and (10.39), we obtain

$$\frac{\partial}{\partial t'}\left[|\mathbf{r} - \mathbf{r}_0(t')|\left(1 - \frac{\mathbf{v}(t')\cdot\mathbf{n}(t')}{c}\right)\right]$$

$$= -\mathbf{v}(t')\cdot\mathbf{n}(t')\left(1 - \frac{\mathbf{v}(t')\cdot\mathbf{n}(t')}{c}\right) - |\mathbf{r} - \mathbf{r}_0(t')|\frac{\mathbf{a}(t')\cdot\mathbf{n}(t')}{c}$$

$$- \frac{\mathbf{v}(t')}{c}\cdot[\mathbf{n}(t') \times (\mathbf{n}(t') \times \mathbf{v}(t'))]$$

$$= -\mathbf{v}(t')\cdot\left(\mathbf{n}(t') - \frac{\mathbf{v}(t')}{c}\right) - \frac{|\mathbf{r} - \mathbf{r}_0(t')|}{c}\mathbf{a}(t')\cdot\mathbf{n}(t'). \tag{10.40}$$

Now, $\phi = \phi(\mathbf{r}, t'(\mathbf{r}))$, so that

$$\nabla\phi = (\nabla\phi)_{t'=\text{const.}} + \frac{\partial\phi}{\partial t'}\nabla t'.$$

Thus we obtain with (10.29) and (10.38) (as well as $\nabla(1/|\mathbf{r}|) = -\mathbf{n}/r^2$) (skipping here two steps with rearrangements of terms):

$$-\nabla\phi = -\nabla\left[\frac{e/4\pi\epsilon_0}{|\mathbf{r}-\mathbf{r}_0(t')|\left(1-\frac{\mathbf{v}(t')\cdot\mathbf{n}(t')}{c}\right)}\right]_{t'=\text{const.}}$$

$$-\frac{\partial}{\partial t'}\left[\frac{e/4\pi\epsilon_0}{|\mathbf{r}-\mathbf{r}_0(t')|\left(1-\frac{\mathbf{v}(t')\cdot\mathbf{n}(t')}{c}\right)}\right]\left[\frac{-\mathbf{n}(t')}{c\left(1-\frac{\mathbf{v}(t')\cdot\mathbf{n}(t')}{c}\right)}\right],$$

and with Eq. (10.40)

$$-\nabla\phi = \frac{(e/4\pi\epsilon_0)[\mathbf{n}\left(1-\frac{v^2}{c^2}\right)-\mathbf{n}(\frac{\mathbf{n}\cdot\mathbf{v}}{c})(1-\frac{\mathbf{n}\cdot\mathbf{v}}{c})]}{|\mathbf{r}-\mathbf{r}_0|^2\left(1-\frac{\mathbf{n}\cdot\mathbf{v}}{c}\right)^3}$$

$$+\frac{(e/4\pi\epsilon_0)\mathbf{n}(t')[\mathbf{a}(t')\cdot\mathbf{n}(t')]}{c^2\left(1-\frac{\mathbf{n}\cdot\mathbf{v}}{c}\right)^3|\mathbf{r}-\mathbf{r}_0|}. \tag{10.41}$$

Next we have to calculate $\partial\mathbf{A}/\partial t$. We have first with Eq. (10.30):

$$\frac{\partial\mathbf{A}}{\partial t'} = \frac{\partial}{\partial t'}\left[\frac{(e\mu_0/4\pi)\mathbf{v}(t')}{|\mathbf{r}-\mathbf{r}_0(t')|\left(1-\frac{\mathbf{v}(t')\cdot\mathbf{n}(t')}{c}\right)}\right]$$

$$= \frac{(e\mu_0/4\pi)\mathbf{a}(t')}{|\mathbf{r}-\mathbf{r}_0(t')|\left(1-\frac{\mathbf{v}(t')\cdot\mathbf{n}(t')}{c}\right)}$$

$$-\frac{(e\mu_0/4\pi)\mathbf{v}(t')\frac{\partial}{\partial t'}\left\{|\mathbf{r}-\mathbf{r}_0(t')|\left(1-\frac{\mathbf{v}(t')\cdot\mathbf{n}(t')}{c}\right)\right\}}{|\mathbf{r}-\mathbf{r}_0(t')|^2\left(1-\frac{\mathbf{v}(t')\cdot\mathbf{n}(t')}{c}\right)^2}, \tag{10.42}$$

and then with Eq. (10.40):

$$\frac{\partial\mathbf{A}}{\partial t'} = -\frac{(e\mu_0/4\pi)\mathbf{v}(t')\left\{-\mathbf{v}(t')\cdot\left(\mathbf{n}(t')-\frac{\mathbf{v}(t')}{c}\right)-\frac{|\mathbf{r}-\mathbf{r}_0(t')|}{c}\mathbf{a}(t')\cdot\mathbf{n}(t')\right\}}{|\mathbf{r}-\mathbf{r}_0(t')|^2\left(1-\frac{\mathbf{v}(t')\cdot\mathbf{n}(t')}{c}\right)^2}$$

$$+\frac{(e\mu_0/4\pi)\mathbf{a}(t')}{|\mathbf{r}-\mathbf{r}_0(t')|\left(1-\frac{\mathbf{v}(t')\cdot\mathbf{n}(t')}{c}\right)}. \tag{10.43}$$

Using Eq. (10.37), we have

$$\frac{\partial \mathbf{A}}{\partial t} = \frac{\partial \mathbf{A}}{\partial t'} \frac{\partial t'}{\partial t}$$

and hence

$$\frac{\partial \mathbf{A}}{\partial t} = \frac{(e\mu_0/4\pi)\mathbf{v}(t')\left\{\mathbf{v}(t') \cdot \left(\mathbf{n}(t') - \frac{\mathbf{v}(t')}{c}\right)\right\}}{|\mathbf{r} - \mathbf{r}_0(t')|^2 \left(1 - \frac{\mathbf{v}(t')\cdot\mathbf{n}(t')}{c}\right)^3}$$

$$+ \frac{(e\mu_0/4\pi)\left\{\mathbf{a}(t') - \mathbf{n}(t') \times \left(\mathbf{a}(t') \times \frac{\mathbf{v}(t')}{c}\right)\right\}}{|\mathbf{r} - \mathbf{r}_0(t')|\left(1 - \frac{\mathbf{v}(t')\cdot\mathbf{n}(t')}{c}\right)^3}. \qquad (10.44)$$

With this we can calculate the electric field strength **E**, using also the relation $c^2\epsilon_0\mu_0 = 1$:

$$\mathbf{E} = -\boldsymbol{\nabla}\phi - \frac{\partial \mathbf{A}}{\partial t}$$

$$= \frac{e}{4\pi\epsilon_0}\left[\frac{\mathbf{n}\left(1 - \frac{v^2}{c^2}\right) - \mathbf{n}(\frac{\mathbf{n}\cdot\mathbf{v}}{c})\left(1 - \frac{\mathbf{n}\cdot\mathbf{v}}{c}\right) - \frac{\mathbf{v}(t')}{c}\left\{\frac{\mathbf{v}(t')}{c} \cdot \left(\mathbf{n}(t') - \frac{\mathbf{v}(t')}{c}\right)\right\}}{|\mathbf{r} - \mathbf{r}_0|^2\left(1 - \frac{\mathbf{n}\cdot\mathbf{v}}{c}\right)^3}\right.$$

$$\left.+ \frac{\mathbf{n}(t')\left\{\mathbf{a}(t') \cdot \mathbf{n}(t')\right\} - \left\{\mathbf{a}(t') - \mathbf{n}(t') \times \left(\mathbf{a}(t') \times \frac{\mathbf{v}(t')}{c}\right)\right\}}{c^2\left(1 - \frac{\mathbf{n}\cdot\mathbf{v}}{c}\right)^3|\mathbf{r} - \mathbf{r}_0|}\right]. \qquad (10.45)$$

For later considerations we observe here, that $\mathbf{n} \cdot \mathbf{E} = O(1/|\mathbf{r} - \mathbf{r}_0|^2)$ and that because $\mathbf{n} \times \mathbf{n} = 0$

$$\mathbf{n} \times \mathbf{E} = -\frac{e}{4\pi\epsilon_0}\frac{\mathbf{n}(t') \times \left\{\mathbf{a}(t') - \mathbf{n}(t') \times \left(\mathbf{a}(t') \times \frac{\mathbf{v}(t')}{c}\right)\right\}}{c^2\left(1 - \frac{\mathbf{n}\cdot\mathbf{v}}{c}\right)^3|\mathbf{r} - \mathbf{r}_0|} + O\left(\frac{1}{|\mathbf{r} - \mathbf{r}_0|^2}\right).$$

$$(10.46)$$

Our next step is to evaluate $\boldsymbol{\nabla} \times \mathbf{A}$ for the calculation of **B**. First we observe that

$$\boldsymbol{\nabla} \times \mathbf{A} = \boldsymbol{\nabla}_{t'=\text{const.}} \times \mathbf{A} + (\boldsymbol{\nabla}t')\frac{\partial}{\partial t'} \times \mathbf{A}$$

$$= (\boldsymbol{\nabla} \times \mathbf{A})_{t'=\text{const.}} - \frac{\partial \mathbf{A}}{\partial t'} \times \boldsymbol{\nabla}t'. \qquad (10.47)$$

Here

$$(\boldsymbol{\nabla} \times \mathbf{A})_{t'=\text{const.}} \overset{(10.30)}{=} \left[\boldsymbol{\nabla} \times \frac{e\mathbf{v}(t')\mu_0/4\pi}{|\mathbf{r} - \mathbf{r}_0(t')|\left(1 - \mathbf{n}(t') \cdot \frac{\mathbf{v}(t')}{c}\right)} \right]_{t'=\text{const.}}.$$

In the following vector product the gradient acts on the expression in the denominator; writing out the components and performing the simple differentiation one obtains

$$\boldsymbol{\nabla} \times \frac{\mathbf{v}(t')}{|\mathbf{r}|} = -\frac{1}{r^3}(\mathbf{r} \times \mathbf{v}(t')) = -\frac{1}{r^2}(\mathbf{n} \times \mathbf{v}(t')).$$

Hence we obtain

$$(\boldsymbol{\nabla} \times \mathbf{A})_{t'=\text{const.}} = -\frac{(e\mu_0/4\pi)(\mathbf{n}(t') \times \mathbf{v}(t'))}{|\mathbf{r} - \mathbf{r}_0(t')|^2\left(1 - \mathbf{n}(t') \cdot \frac{\mathbf{v}(t')}{c}\right)}. \tag{10.48}$$

We also require in Eq. (10.47) the differentiation of $\mathbf{A}$ with respect to t', i.e.

$$\frac{\partial \mathbf{A}}{\partial t'} = \frac{\partial}{\partial t'} \left[\frac{(e\mu_0/4\pi)\mathbf{v}(t')}{|\mathbf{r} - \mathbf{r}_0(t')|\left(1 - \mathbf{n}(t') \cdot \frac{\mathbf{v}(t')}{c}\right)} \right], \tag{10.49}$$

which we already have with Eq. (10.43), i.e.

$$\frac{\partial \mathbf{A}}{\partial t'} = \frac{(e\mu_0/4\pi)\mathbf{v}(t')\left\{\mathbf{v}(t') \cdot \left(\mathbf{n}(t') - \frac{\mathbf{v}(t')}{c}\right) + \frac{|\mathbf{r}-\mathbf{r}_0(t')|}{c}\mathbf{a}(t') \cdot \mathbf{n}(t')\right\}}{|\mathbf{r} - \mathbf{r}_0(t')|^2\left(1 - \mathbf{n}(t') \cdot \frac{\mathbf{v}(t')}{c}\right)^2}$$

$$+ \frac{(e\mu_0/4\pi)\mathbf{a}(t')}{|\mathbf{r} - \mathbf{r}_0(t')|\left(1 - \mathbf{n}(t') \cdot \frac{\mathbf{v}(t')}{c}\right)} \tag{10.50}$$

and finally (again skipping one intermediate rearrangement step)

$$\frac{\partial \mathbf{A}}{\partial t'} = \frac{(e\mu_0/4\pi)\mathbf{v}(t')\left\{\mathbf{v}(t') \cdot \left(\mathbf{n}(t') - \frac{\mathbf{v}(t')}{c}\right)\right\}}{|\mathbf{r} - \mathbf{r}_0(t')|^2\left(1 - \mathbf{n}(t') \cdot \frac{\mathbf{v}(t')}{c}\right)^2}$$

$$+ \frac{(e\mu_0/4\pi)\left\{\mathbf{a}(t') - \mathbf{n}(t') \times \left(\mathbf{a}(t') \times \frac{\mathbf{v}(t')}{c}\right)\right\}}{|\mathbf{r} - \mathbf{r}_0(t')|\left(1 - \mathbf{n}(t') \cdot \frac{\mathbf{v}(t')}{c}\right)^2}. \tag{10.51}$$

Using Eq. (10.47) we can now obtain the field **B**. With a rearrangement of terms we obtain

$$
\mathbf{B} = \frac{e\mu_0}{4\pi} \frac{(\mathbf{v}(t') \times \mathbf{n}(t')) \left[\left(1 - \mathbf{n}(t') \cdot \frac{\mathbf{v}(t')}{c} \right)^2 + \frac{\mathbf{v}(t')}{c} \cdot \left(\mathbf{n}(t') - \frac{\mathbf{v}(t')}{c} \right) \right]}{|\mathbf{r} - \mathbf{r}_0(t')|^2 \left(1 - \mathbf{n}(t') \cdot \frac{\mathbf{v}(t')}{c} \right)^3}
$$

$$
+ \frac{e\mu_0}{4\pi} \frac{\left\{ \mathbf{a}(t') - \mathbf{n}(t') \times \left(\mathbf{a}(t') \times \frac{\mathbf{v}(t')}{c} \right) \right\} \times \mathbf{n}(t')}{c|\mathbf{r} - \mathbf{r}_0(t')| \left(1 - \mathbf{n}(t') \cdot \frac{\mathbf{v}(t')}{c} \right)^3}. \tag{10.52}
$$

We now have at our disposal the desired expressions for the field strengths **E** and **B** at a point **r** for the case when the charge e moves with velocity **v** along its trajectory $\mathbf{r}_0$. We see that for velocities considerably smaller than c the above expressions yield the nonrelativistic approximations

$$
\mathbf{E} \rightarrow \frac{e}{4\pi\epsilon_0} \frac{\mathbf{n}}{|\mathbf{r} - \mathbf{r}_0(t)|^2}, \quad \mathbf{B} \rightarrow \frac{e\mu_0}{4\pi} \frac{(\mathbf{v} \times \mathbf{n})}{|\mathbf{r} - \mathbf{r}_0(t)|^2}, \tag{10.53}
$$

which we recognise as the Coulomb and Biot–Savart laws.

We also see, that both field strengths consist of two contributions – one decreasing like $1/|\mathbf{r}|$ with $r \rightarrow \infty$ and another decreasing like $1/|\mathbf{r}|^2$ with $r \rightarrow \infty$. Thus the dominant contributions are those of the first type. These contributions contain the acceleration **a** and are therefore referred to as *radiation fields*. Hence *accelerated charges emit electromagnetic radiation*. Furthermore, one can convince oneself, that **E** and **B** are mutually orthogonal and that their moduli are equal. In the case of motion with vanishing acceleration, the velocity is constant. This is the case of a charge in an *inertial frame*; in this case the fields transform in accordance with the transformation from one inertial frame to another, these transformations being those of Lorentz transformations (see Chapter 17).

Obviously the next quantity to evaluate is the Poynting vector **S**. Taking into account Eq. (10.46) and the preceding remarks, we obtain

$$
\mathbf{S} = \mathbf{E} \times \mathbf{H} \overset{(10.46)}{=} \mathbf{E} \times \left(\frac{\mathbf{n} \times \mathbf{E}}{\mu_0 c} \right)
$$

$$
= \frac{E^2}{c\mu_0} \mathbf{n} - \frac{(\mathbf{E} \cdot \mathbf{n})\mathbf{E}}{c\mu_0}
$$

$$
= \frac{E^2}{c\mu_0} \mathbf{n} + O(1/|\mathbf{r} - \mathbf{r}_0|^3). \tag{10.54}
$$

We deduce from the equation for $\mathbf{E}$, i.e. Eq. (10.45), that for $\mathbf{r}$ large and velocities small compared to c

$$\mathbf{E} = \frac{e}{4\pi\epsilon_0} \frac{\mathbf{n}(t')\left\{\mathbf{a}(t') \cdot \mathbf{n}(t')\right\} - \mathbf{a}(t')}{c^2|\mathbf{r} - \mathbf{r}_0(t)|}. \tag{10.55}$$

We see that

$$\mathbf{a}^* := \mathbf{n}(t')\left\{\mathbf{a}(t') \cdot \mathbf{n}(t')\right\} - \mathbf{a}(t') \equiv \mathbf{n} \times (\mathbf{n} \times \mathbf{a}),$$

since $\mathbf{a}^* \cdot \mathbf{n} = 0$ is the acceleration perpendicular to $\mathbf{n}$, i.e. to the vector of the "direction of observation", i.e. the direction to the point P, at which we want to know the value of the fields. These directions are indicated in Fig. 10.1. The modulus of the electric field strength at the point P is therefore given by the projection of $\mathbf{a}$ in the direction of observation, i.e.

$$|\mathbf{E}| = \frac{e}{4\pi\epsilon_0} \frac{a\sin\theta}{c^2|\mathbf{r} - \mathbf{r}_0(t)|}, \tag{10.56}$$

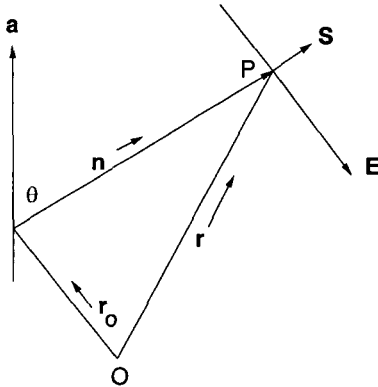which is zero along $\mathbf{n}$ (i.e. for $\theta = 0$).

Fig. 10.1 The directions of $\mathbf{a}$ and $\mathbf{S}$.

For the modulus of the Poynting vector we obtain (using $\epsilon_0\mu_0 c^2 = 1$)

$$|\mathbf{S}| = \frac{e^2 a^2 \sin^2\theta}{4\pi\epsilon_0 4\pi c^3 |\mathbf{r} - \mathbf{r}_0(t)|^2}. \tag{10.57}$$

The total energy W radiated off per unit time at time t', $\partial W/\partial t$, is obtained

by integrating over all angles at radius $|\mathbf{r} - \mathbf{r}_0(t)|$, i.e. the rate is

$$
\begin{aligned}
P = \frac{dW}{dt} &= \int_0^\pi 2\pi \sin\theta d\theta |\mathbf{r} - \mathbf{r}_0(t)|^2 |\mathbf{S}| \\
&= \frac{e^2 a^2}{8\pi\epsilon_0 c^3} \int_0^\pi \sin^3\theta d\theta \\
&= \frac{2}{3} \frac{e^2 a^2}{4\pi\epsilon_0 c^3} \text{ watt,}
\end{aligned}
\tag{10.58}
$$

where we put $\cos\theta = z$ and evaluated the integral as

$$
\int_0^\pi \sin^3\theta d\theta = \int_{-1}^1 dz(1 - z^2) = 4/3.
$$

The final result, Eq. (10.58), is called *Larmor formula* in the case of radiation of a single charged particle.[‡] Since the main quantity appearing in this Larmor formula is the acceleration a, exercises or applications based on this formula generally require some acceleration.

Example 10.4: Radiation emitted by a proton

A proton (charge q) is given a constant acceleration in a van de Graaff accelerator by a potential difference of 700 kilovolt. The acceleration region has a length of 3 meters. Calculate the ratio of the energy emitted by the proton to its final kinetic energy, and derive with the help of the Larmor formula a relation between the emitted energy per unit time and the kinetic energy.

Solution: The energy emitted by a charge q in time t is according to Eq. (10.58) the energy

$$
U_{\text{radiation}} = Pt = \frac{dU}{dt} t = \frac{2q^2 a^2 t}{4\pi\epsilon_0 3c^3}.
$$

Since $v = at$, the length of the proton's runway to achieve its constant acceleration a is $s = at^2/2 = vt/2$, so that $t = 2s/v$, $a = v/t = v^2/2s$, and $a^2 = (v^2/2s)(v/t)$, $a^2 t = v^3/2s$, and hence

$$
U_{\text{radiation}} = \frac{q^2 v^3}{4\pi\epsilon_0 3c^3 s}.
$$

The kinetic energy at the end is (with $M=$ mass of the proton, V the potential):

$$
U_k = \frac{1}{2} M v^2 = qV, \qquad v = \left(\frac{2qV}{M} \right)^{1/2}.
$$

Hence

$$
\frac{U_{\text{radiation}}}{U_k} = \frac{2q^2 v}{4\pi\epsilon_0 3Mc^3 s} = \frac{2q^2}{4\pi\epsilon_0 3Mc^3 s} \left(\frac{2qV}{M} \right)^{1/2}.
$$

In MKSA-units:

$$
q = 1.6 \times 10^{-19} \text{ coulomb}, \qquad V = 7 \times 10^5 \text{ volt},
$$

$$
c = 3 \times 10^8 \text{ meter per second}, \qquad M = 1.67 \times 10^{-27} \text{C/(V m)}.
$$

[‡]Below another Larmor formula is obtained, but for dipole radiation.

It follows that

$$\frac{U_{\text{radiation}}}{U_k} = 1.31 \times 10^{-20}.$$

Thus this ratio is extremely small!
Second part: $U_k = Mv^2/2$. Hence

$$\frac{dU_k}{dx} = Mv\frac{dv}{dx} = Mv\frac{dv}{dt}\frac{dt}{dx} = Ma.$$

$$\therefore a = \frac{1}{M}\frac{dU_k}{dx}.$$

We insert this into the Larmor formula and find:

$$\frac{dU_{\text{radiation}}}{dt} = \frac{2q^2}{4\pi\epsilon_0 3M^2c^3}\left(\frac{dU_k}{dx}\right)^2.$$

This is trivial for a constant acceleration; otherwise of general significance.

Example 10.5:[§] Model of bremsstrahlung

As a simplified model of bremsstrahlung we consider the following scattering process. Electrons having velocity $\mathbf{v}$ are scattered off ions with charge Q. With the help of the Coulomb law (which supplies the acceleration!) and the Larmor formula derive an expression for the energy W per unit time emitted as bremsstrahlung (let N_e be the density of the electrons).

Solution: Let m be the mass of the electrons. The acceleration a then follows from the Coulomb law as

$$a = \frac{1}{4\pi\epsilon_0}\frac{qQ}{mr^2},$$

where r is the distance between the electron and the ion.
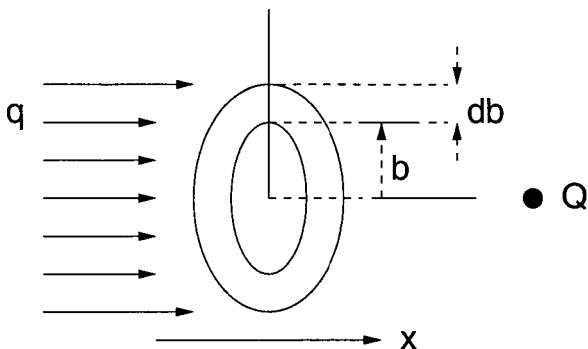

Fig. 10.2 The impact parameter b.

The Larmor formula for the radiation power P of a single charged particle is

$$P = \frac{2q^2a^2}{4\pi\epsilon_0 3c^3} \equiv \frac{dW}{dt}.$$

[§]See also G. Bekefi and A. H. Barrett, *Electromagnetic Vibrations* (Cambridge Mass., MIT Press, 1977), pp. 278–283.

With $r^2 = x^2 + b^2, x = vt$ the formula yields the expression $P = \frac{1}{(4\pi\epsilon_0)^3} \frac{2q^4Q^2}{3m^2c^3(v^2t^2+b^2)^2}$. Here b is, as indicated in Fig.10.2, the so-called *impact parameter* of the scattering process. In order to obtain the total energy emitted by one electron, we integrate over the time t of this electron and obtain

$$
\begin{aligned}
W &= \frac{1}{(4\pi\epsilon_0)^3} \frac{2q^4Q^2}{3m^2c^3} \int_{-\infty}^{\infty} \frac{dt}{(v^2t^2+b^2)^2} \\
&= \frac{1}{(4\pi\epsilon_0)^3} \frac{2q^4Q^2}{3m^2c^3} \frac{\pi}{2vb^3}.
\end{aligned}
\tag{10.59}
$$

If N_e is the number of electrons per unit volume as given, their current is $N_e v$. The entire energy emitted per unit time in the form of bremsstrahlung is therefore

$$
P_{\text{brems}} = \int_{b_0}^{\infty} W N_e v \underbrace{2\pi b\, db}_{\text{cylindrical element of area}},
$$

where b_0 is the smallest impact parameter. This result is incorrect only in the region of small impact parameters b, since for these the path of the electrons is not linear (in view of strong deflection close to the ion). With this approximation the result is

$$
P_{\text{brems}} = \frac{1}{(4\pi\epsilon_0)^3} \frac{2N_e q^4 Q^2 \pi^2}{3m^2c^3} \int_{b_0}^{\infty} \frac{db}{b^2} = \frac{1}{(4\pi\epsilon_0)^3} \frac{2N_e q^4 Q^2 \pi^2}{3m^2c^3} \frac{1}{b_0}.
$$

Example 10.6: Cyclotron radiation
Derive an expression for the energy per unit time emitted by an electron which is circling in a homogeneous magnetic field (cyclotron radiation).

Solution: An electron (charge q) injected horizontally into a vertical magnetic field B_0 performs circular orbits around the lines of force of this field. Even with constant circular velocity the electron experiences a centripetal acceleration and hence emits electromagnetic radiation. In the case of low energy electrons, one refers to this as cyclotron radiation, in the case of relativistic electrons as synchrotron radiation. Let v be the velocity of the electron on its circular orbit. The magnetic field subjects the electron to the Lorentz force $q(\mathbf{v} \times \mathbf{B}_0)$ resulting in an acceleration v^2/R, where R is the radius of its circular orbit. Thus $qvB_0 = mv^2/R$ and hence $R = mv/qB_0$. The acceleration is therefore

$$
a = v\frac{qB_0}{m}.
$$

The cyclotron frequency ω_c is defined by $v = \omega_c R$, so that

$$
\omega_c = \frac{qB_0}{m} \text{ rad/s}
$$

The acceleration is therefore $a = v\omega_c$. Inserting this expression into the single particle Larmor formula (10.58), we obtain for the power (i.e. energy per unit time) emitted by the electron the result

$$
P_c = \frac{q^2\omega_c^2 v^2}{6\pi\epsilon_0 c^3} \text{ watt per electron.}
$$

If the emitting substance contains N_e electrons per meter^3, the power per meter^3 is $P = N_e P_c$, i.e. $P = 6.21 \times 10^{-20} N_e B_0^2 u$ W/m^3, where u is the energy of the electron expressed in electron volts.

10.5   The Hertz Dipole

We assume an arrangement of charges as depicted in Fig. 10.3. Our notation corresponds to that used earlier in this Chapter.
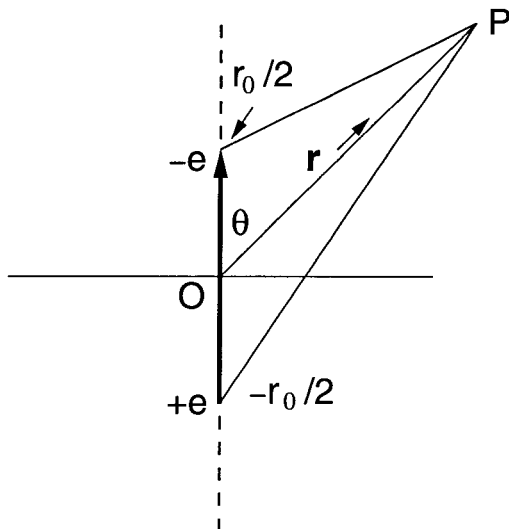
Fig. 10.3 The Hertz dipole.

We assume that

$$\frac{v}{c} \ll 1, \quad \mathbf{r}(t') = \mathbf{r} - \mathbf{r}_0(t') \simeq \mathbf{r};$$

for $r \sim$ large the latter seems a plausible approximation. We consider first the potentials and subsequently the field strengths.

10.5.1   The Potentials

By superposition of the potentials of the two charges $+e, -e$ of the dipole at the points $\mathbf{r}_0/2, -\mathbf{r}_0/2$ we obtain from Eq. (10.33) for the vector potential

$$\mathbf{A}(\mathbf{r}, t) = -\frac{e\mu_0}{4\pi}\left[+\frac{\frac{1}{2}\dot{\mathbf{r}}_0}{r - \frac{\mathbf{r}\cdot\dot{\mathbf{r}}_0}{2c}}\right]_{t'=t-\frac{r(t')}{c}\simeq t-\frac{r}{c}}$$

$$+\frac{e\mu_0}{4\pi}\left[-\frac{\frac{1}{2}\dot{\mathbf{r}}_0}{r + \frac{\mathbf{r}\cdot\dot{\mathbf{r}}_0}{2c}}\right]_{t'=t-\frac{r(t')}{c}\simeq t-\frac{r}{c}}.$$

For large values of r one would naturally approximate this by

$$\mathbf{A}(\mathbf{r}, t) \simeq -\frac{\mu_0}{4\pi} \left(\frac{e\dot{\mathbf{r}}_0(t')}{r} \right)_{t'=t-\frac{r}{c}}$$

$$= \frac{\mu_0}{4\pi} \frac{\dot{\mathbf{p}}(t - \frac{r}{c})}{r} + O(|\dot{\mathbf{r}}_0/c|). \qquad (10.60)$$

Here we content ourselves first with *only this contribution*, because — as we shall see — we then obtain $\mathbf{H}$ of the same order as $\mathbf{E}$. In a similar manner we calculate with the help of Eq. (10.32) the scalar potential ϕ of the dipole:[¶]

$$\phi(\mathbf{r}, t) = -\frac{e}{4\pi\epsilon_0} \left(\frac{1}{r - \frac{\mathbf{r}\cdot\mathbf{r}_0}{2c}} \right)_{t'=t-\frac{r}{c}} + \frac{e}{4\pi\epsilon_0} \left(\frac{1}{r + \frac{\mathbf{r}\cdot\mathbf{r}_0}{2c}} \right)_{t'=t-\frac{r}{c}}$$

$$\overset{r \text{ large}}{\simeq} -\frac{e}{4\pi\epsilon_0 r} \left(1 + \frac{\mathbf{r} \cdot \mathbf{r}_0}{2rc} + \cdots \right)_{t'=t-\frac{r}{c}}$$

$$+ \frac{e}{4\pi\epsilon_0 r} \left(1 - \frac{\mathbf{r} \cdot \mathbf{r}_0}{2rc} + \cdots \right)_{t'=t-\frac{r}{c}}$$

$$\simeq \frac{1}{4\pi\epsilon_0} \left(\frac{\dot{\mathbf{p}} \cdot \mathbf{r}}{cr^2} \right)_{t'=t-\frac{r}{c}} = O(1/r). \qquad (10.61)$$

Thus in this case the dominant contributions cancel out — which is different from the case of $\mathbf{A}(\mathbf{r}, t)$. It is therefore necessary to examine the approximations made in the denominator more carefully.

On the other hand, *for smaller values of r (i.e. of $O(1/r^2)$) we expect the familiar expression for the potential of the static dipole*, i.e.

$$\phi = k\frac{\mathbf{d} \cdot \mathbf{r}}{r^3}, \quad k = \frac{1}{4\pi\epsilon_0}.$$

We observe, however, that ϕ does not supply this expression! This is due to the fact, that in deriving this expression, we replaced the factor $1/|\mathbf{r} - \mathbf{r}_0(t')|$ by $1/|\mathbf{r}| = 1/r$, i.e. $|\mathbf{r} - \mathbf{r}_0(t')| \simeq r$. We therefore return to Eq. (10.32), i.e. to

$$\phi = \left(\frac{e/4\pi\epsilon_0}{r(t') - \frac{\mathbf{r}(t')\cdot\mathbf{v}(t')}{c}} \right)_{t'=t-\frac{r(t')}{c}}, \qquad (10.62)$$

or

$$\phi(\mathbf{r}, t) = \frac{e}{4\pi\epsilon_0} \frac{1}{\left[|\mathbf{r} - \mathbf{r}_0(t')| - (\mathbf{r} - \mathbf{r}_0(t')) \cdot \frac{\mathbf{v}(t')}{c} \right]} \Bigg|_{t'=t-\frac{|\mathbf{r}-\mathbf{r}_0(t')|}{c}}. \qquad (10.63)$$

[¶]In the second step we also neglect contributions of $O(|\dot{\mathbf{r}}_0/c|^n)$ for $n \geq 2$.

Now

$$|\mathbf{r} - \mathbf{r}_0(t')| = \sqrt{r^2 - 2\mathbf{r} \cdot \mathbf{r}_0(t') + \mathbf{r}_0(t')^2},$$

so that for $r^2 \gg r_0^2$

$$|\mathbf{r} - \mathbf{r}_0(t')| \simeq \sqrt{r^2 - 2\mathbf{r} \cdot \mathbf{r}_o(t')} \simeq r\left(1 - \frac{\mathbf{r} \cdot \mathbf{r}_0(t')}{r^2}\right)$$

$$= r - \mathbf{r}_0(t') \cdot \frac{\mathbf{r}}{r}.$$

But from the definition

$$\mathbf{n} = \frac{\mathbf{r} - \mathbf{r}_0(t')}{|\mathbf{r} - \mathbf{r}_0(t')|},$$

we obtain

$$\mathbf{r} = \mathbf{n}|\mathbf{r} - \mathbf{r}_0(t')| + \mathbf{r}_0(t') \simeq \mathbf{n}\left(r - \mathbf{r}_0(t') \cdot \frac{\mathbf{r}}{r}\right) + \mathbf{r}_0(t')$$

and hence

$$\mathbf{r}_0(t') \cdot \frac{\mathbf{r}}{r} \simeq \mathbf{r}_0(t') \cdot \mathbf{n}\left(1 - \mathbf{r}_0(t') \cdot \frac{\mathbf{r}}{r^2}\right) + \frac{\mathbf{r}_0(t')^2}{r} \simeq \mathbf{r}_0(t') \cdot \mathbf{n} + O(1/r).$$

Thus

$$|\mathbf{r} - \mathbf{r}_0(t')| \simeq r - \mathbf{r}_0(t') \cdot \mathbf{n} + O(1/r).$$

Here the second term on the right (with the scalar product) was missing in the first approximation. We insert this expression now in ϕ, keeping in mind, that $\mathbf{v}(t') = \dot{\mathbf{r}}_0(t')$. Then

$$\phi(\mathbf{r}, t) \simeq \frac{e}{4\pi\epsilon_0} \frac{1}{\left[r - \mathbf{r}_0(t') \cdot \mathbf{n} - (\mathbf{r} - \mathbf{r}_0(t')) \cdot \frac{\dot{\mathbf{r}}_0(t')}{c}\right]}\Bigg|_{t'=t-\frac{r(t')}{c}\simeq t-\frac{r}{c}}$$

$$\simeq \frac{e}{4\pi\epsilon_0} \frac{1}{\left[r - \underbrace{\mathbf{r}_0(t') \cdot \mathbf{n}}_{\text{not in first approx.}} - \frac{\mathbf{r} \cdot \dot{\mathbf{r}}_0(t')}{c}\right]}\Bigg|_{t'=t-\frac{r(t')}{c}\simeq t-\frac{r}{c}}. \qquad (10.64)$$

Applying this general formula to the special charges of the dipole we have to take into account that the coordinates of the charges, i.e. $\mathbf{r}_0(t')$, have to be

replaced by $\pm\mathbf{r}_0(t')/2$. With these changes we obtain instead of Eq. (10.61):

$$
\begin{aligned}
\phi(\mathbf{r},t) &= \left[-\frac{e}{4\pi\epsilon_0}\left(\frac{1}{r-\frac{\mathbf{r}_0\cdot\mathbf{n}}{2}-\frac{\mathbf{r}\cdot\dot{\mathbf{r}}_0}{2c}}\right)+\frac{e}{4\pi\epsilon_0}\left(\frac{1}{r+\frac{\mathbf{r}_0\cdot\mathbf{n}}{2}+\frac{\mathbf{r}\cdot\dot{\mathbf{r}}_0}{2c}}\right)\right]_{t'=t-\frac{r}{c}}\\
&\stackrel{r\ \text{large}}{\simeq} -\frac{e}{4\pi\epsilon_0 r}\left(1+\frac{\mathbf{r}_0\cdot\mathbf{n}}{2r}+\frac{\mathbf{r}\cdot\dot{\mathbf{r}}_0}{2rc}+\cdots\right)_{t'=t-\frac{r}{c}}\\
&\quad +\frac{e}{4\pi\epsilon_0 r}\left(1-\frac{\mathbf{r}_0\cdot\mathbf{n}}{2r}-\frac{\mathbf{r}\cdot\dot{\mathbf{r}}_0}{2rc}+\cdots\right)_{t'=t-\frac{r}{c}}\\
&\simeq \frac{1}{4\pi\epsilon_0}\left(\frac{\mathbf{p}\cdot\mathbf{n}}{r^2}+\frac{\dot{\mathbf{p}}\cdot\mathbf{r}}{cr^2}\right)_{t'=t-\frac{r}{c}}\\
&= \frac{1}{4\pi\epsilon_0}\left(\frac{\mathbf{p}\cdot\mathbf{r}}{r^3}+\frac{\dot{\mathbf{p}}\cdot\mathbf{r}}{cr^2}\right)_{t'=t-\frac{r}{c}}. \tag{10.65}
\end{aligned}
$$

We therefore have the situation, that for $r\to\infty$ the second term ($\propto 1/r$) dominates, but for smaller r the first term ($\propto 1/r^2$). One should note that on the whole the expansion corresponds to the multi-pole expansion, i.e. Coulomb contribution + dipole contribution + $\cdots$.

10.5.2 The Field Strengths

We obtain the field strengths $\mathbf{E}$ and $\mathbf{H}$ from the equations

$$\mathbf{E}=-\frac{\partial\mathbf{A}}{\partial t}-\boldsymbol{\nabla}\phi,\quad \mu_0\mathbf{H}=\boldsymbol{\nabla}\times\mathbf{A}. \tag{10.66}$$

We evaluate first $-\boldsymbol{\nabla}\phi$. For this we need

$$\frac{\partial}{\partial\mathbf{r}}(r)=\left(\frac{\partial}{\partial x},\frac{\partial}{\partial y},\frac{\partial}{\partial z}\right)\sqrt{x^2+y^2+z^2}=\frac{\mathbf{r}}{r},$$

$$\frac{\partial}{\partial\mathbf{r}}\left(\frac{1}{r^2}\right)=-\frac{2}{r^3}\frac{\partial r}{\partial\mathbf{r}}=-\frac{2\mathbf{r}}{r^4},\quad \frac{\partial}{\partial\mathbf{r}}\left(\frac{1}{r^3}\right)=-\frac{3}{r^4}\frac{\partial r}{\partial\mathbf{r}}=-\frac{3\mathbf{r}}{r^5}.$$

Then

$$
\begin{aligned}
-\boldsymbol{\nabla}\phi &= -\frac{1}{4\pi\epsilon_0}\frac{\partial}{\partial\mathbf{r}}\left\{\left[\frac{\mathbf{p}\cdot\mathbf{r}}{r^3}+\frac{\dot{\mathbf{p}}\cdot\mathbf{r}}{cr^2}\right]_{t'=t-\frac{r(t')}{c}}\right\}\\
&= -\frac{1}{4\pi\epsilon_0}\frac{\partial}{\partial\mathbf{r}}\left\{\left[\frac{\mathbf{p}(t-\frac{r(t')}{c})\cdot\mathbf{r}}{r^3}+\frac{\dot{\mathbf{p}}(t-\frac{r(t')}{c})\cdot\mathbf{r}}{cr^2}\right]_{t'=t-\frac{r(t')}{c}}\right\}\\
&= \frac{1}{4\pi\epsilon_0}\left[-\frac{\dot{\mathbf{p}}\cdot\mathbf{r}}{r^3}\frac{\partial}{\partial\mathbf{r}}\left(-\frac{r(t')}{c}\right)-\frac{\mathbf{p}}{r^3}-\mathbf{p}\cdot\mathbf{r}\frac{\partial}{\partial\mathbf{r}}\left(\frac{1}{r^3}\right)\right.\\
&\quad \left.-\frac{\ddot{\mathbf{p}}\cdot\mathbf{r}}{cr^2}\frac{\partial}{\partial\mathbf{r}}\left(-\frac{r(t')}{c}\right)-\frac{\dot{\mathbf{p}}}{cr^2}-\frac{\dot{\mathbf{p}}\cdot\mathbf{r}}{c}\frac{\partial}{\partial\mathbf{r}}\left(\frac{1}{r^2}\right)\right],
\end{aligned}
$$

where

$$r(t') = |\mathbf{r} - \mathbf{r}_0(t')| = \sqrt{(\mathbf{r} - \mathbf{r}_0(t'))^2}$$

and

$$\frac{\partial r(t')}{\partial \mathbf{r}} = \frac{\mathbf{r} - \mathbf{r}_0(t')}{|\mathbf{r} - \mathbf{r}_0(t')|} \simeq \frac{\mathbf{r}}{r}.$$

Hence

$$-\boldsymbol{\nabla}\phi = \frac{1}{4\pi\epsilon_0}\left[\frac{\dot{\mathbf{p}}\cdot\mathbf{r}}{cr^3}\cdot\frac{\mathbf{r}}{r} - \frac{\mathbf{p}}{r^3} - \mathbf{p}\cdot\mathbf{r}\left(-\frac{3\mathbf{r}}{r^5}\right) + \frac{\ddot{\mathbf{p}}\cdot\mathbf{r}}{c^2r^2}\cdot\frac{\mathbf{r}}{r}\right.$$
$$\left. - \frac{\dot{\mathbf{p}}}{cr^2} - \frac{\dot{\mathbf{p}}\cdot\mathbf{r}}{c}\left(-\frac{2\mathbf{r}}{r^4}\right)\right]$$
$$= \frac{1}{4\pi\epsilon_0}\left[3\frac{(\dot{\mathbf{p}}\cdot\mathbf{r})\mathbf{r}}{cr^4} - \frac{\mathbf{p}}{r^3} + \frac{3(\mathbf{p}\cdot\mathbf{r})\mathbf{r}}{r^5} + \frac{(\ddot{\mathbf{p}}\cdot\mathbf{r})\mathbf{r}}{c^2r^3} - \frac{\dot{\mathbf{p}}}{cr^2}\right].$$

This together with Eq. (10.60) and Eq. (10.66) yields

$$\mathbf{E} = -\frac{\mu_0}{4\pi}\frac{\ddot{\mathbf{p}}}{r} + \frac{1}{4\pi\epsilon_0}\left[-\frac{\dot{\mathbf{p}}}{cr^2} + \frac{3(\dot{\mathbf{p}}\cdot\mathbf{r})\mathbf{r}}{cr^4} + \frac{(\ddot{\mathbf{p}}\cdot\mathbf{r})\mathbf{r}}{c^2r^3} + \frac{3(\mathbf{r}\cdot\mathbf{p})\mathbf{r}}{r^5} - \frac{\mathbf{p}}{r^3}\right] \quad (10.67)$$

and

$$\mu_0\mathbf{H} = \boldsymbol{\nabla}\times\mathbf{A}$$
$$= \frac{\mu_0}{4\pi}\boldsymbol{\nabla}\times\frac{\dot{\mathbf{p}}(t-\frac{r}{c})}{r} = \frac{\mu_0}{4\pi}\frac{\mathbf{r}}{r}\frac{\partial}{\partial r}\times\frac{\dot{\mathbf{p}}(t-\frac{r}{c})}{r}$$
$$= -\frac{\mu_0}{4\pi c}\frac{\mathbf{r}}{r}\times\frac{\ddot{\mathbf{p}}}{r} - \frac{\mu_0}{4\pi}\frac{\mathbf{r}}{r}\times\frac{\dot{\mathbf{p}}}{r^2}$$
$$= \frac{\mu_0}{4\pi}\frac{\ddot{\mathbf{p}}\times\mathbf{r}}{cr^2} + \frac{\mu_0}{4\pi}\frac{\dot{\mathbf{p}}\times\mathbf{r}}{r^3}, \quad (10.68)$$

so that

$$\mathbf{H} = \frac{1}{4\pi}\left[\frac{\ddot{\mathbf{p}}\times\mathbf{r}}{cr^2} + \frac{\dot{\mathbf{p}}\times\mathbf{r}}{r^3}\right]. \quad (10.69)$$

In the so-called *near field region*, (i.e. for $r \sim$ small), we have

$$\mathbf{E} = \frac{1}{4\pi\epsilon_0}\left[\frac{3(\mathbf{p}\cdot\mathbf{r})\mathbf{r}}{r^5} - \frac{\mathbf{p}}{r^3}\right], \quad \mathbf{H} = \frac{1}{4\pi}\frac{\dot{\mathbf{p}}\times\mathbf{r}}{cr^3}. \quad (10.70)$$

However, in the so-called *far field region*, (i.e. for $r \sim$ large), we obtain

$$\mathbf{E} \simeq -\frac{\mu_0}{4\pi}\frac{\ddot{\mathbf{p}}}{r} + \frac{1}{4\pi\epsilon_0}\frac{(\ddot{\mathbf{p}}\cdot\mathbf{r})\mathbf{r}}{c^2r^3} = \frac{1}{4\pi\epsilon_0}\frac{(\ddot{\mathbf{p}}\times\mathbf{r})\times\mathbf{r}}{c^2r^3} \quad (10.71)$$

and

$$\mathbf{H} \simeq \frac{1}{4\pi}\frac{\ddot{\mathbf{p}}\times\mathbf{r}}{cr^2}, \quad (10.72)$$

where

$$p = p(t')|_{t'=t-\frac{r(t')}{c}} \simeq t - \frac{r}{c}.$$

Since $\epsilon_0 = 1/\mu_0 c^2$, we deduce from Eqs. (10.71) and (10.72), that in the far field region

$$\mathbf{E} \perp \mathbf{r}, \mathbf{H}.$$

We see that in this region

$$\mathbf{r} \cdot \mathbf{E} \propto [(\ddot{\mathbf{p}} \times \mathbf{r}) \times \mathbf{r}] \cdot \mathbf{r} = (\ddot{\mathbf{p}} \times \mathbf{r}) \cdot (\mathbf{r} \times \mathbf{r}) = 0$$

and

$$\mathbf{r} \cdot \mathbf{H} \propto (\ddot{\mathbf{p}} \times \mathbf{r}) \cdot \mathbf{r} = \ddot{\mathbf{p}} \cdot (\mathbf{r} \times \mathbf{r}) = 0.$$

It follows that $\mathbf{r} \perp \mathbf{E}, \mathbf{H}$. Since also $\mathbf{E} \perp \mathbf{r}, \mathbf{H}$, the vectors $\mathbf{r}, \mathbf{E}, \mathbf{H}$ define at *large distances* an orthogonal system. At large distances

$$\mathbf{E} = \frac{1}{c\epsilon_0} \mathbf{H} \times \frac{\mathbf{r}}{r}. \tag{10.73}$$

Thus

$$
\begin{aligned}
-\mathbf{E} \times \frac{\mathbf{r}}{r} &= -\frac{1}{c\epsilon_0} \left(\mathbf{H} \times \frac{\mathbf{r}}{r} \right) \times \frac{\mathbf{r}}{r} \\
&= -\frac{1}{c\epsilon_0} \left(\underbrace{\left\{ \mathbf{H} \cdot \frac{\mathbf{r}}{r} \right\}}_{=0} \frac{\mathbf{r}}{r} - \frac{r^2}{r^2} \mathbf{H} \right) = \frac{1}{c\epsilon_0} \mathbf{H}.
\end{aligned}
$$

10.6 Current Element *Ids* and Dipole Radiation

We now consider the following. An element of current Ids of a conducting wire with alternating current I can be looked at as an electric dipole of length ds with charges $\pm q = \pm q_0 e^{i\omega t}$ at its endpoints. Since

$$\mathbf{p} = q d\mathbf{s}, \quad \mathbf{p}_0 = q_0 d\mathbf{s},$$

we can put formally (here $d\mathbf{p} = dq d\mathbf{s}$)

$$I d\mathbf{s} = \frac{dq}{dt} d\mathbf{s} \equiv \dot{\mathbf{p}}.$$

Then

$$\dot{\mathbf{p}} = \dot{q} d\mathbf{s} = i\omega \mathbf{p} \equiv i\omega \mathbf{p}_0 e^{i\omega t}, \quad \ddot{\mathbf{p}} = i\omega \dot{\mathbf{p}} = i\omega I d\mathbf{s}.$$

Only the real part is physically meaningful, so that

$$\dot{\mathbf{p}} \to -\omega \mathbf{p}_0 \sin \omega t.$$

Thus in the far-away region[||] (cf. Eq. (10.71))

$$\mathbf{E} = \frac{(\ddot{\mathbf{p}} \times \mathbf{r}) \times \mathbf{r}}{4\pi\epsilon_0 c^2 r^3} = \frac{i\omega(Ids \times \mathbf{r}) \times \mathbf{r}}{4\pi\epsilon_0 c^2 r^3}.$$

Using the relations $c = \nu\lambda, \omega = 2\pi\nu, \omega = 2\pi c/\lambda$, we obtain

$$\mathbf{E} = i\frac{2\pi(Ids \times \mathbf{r}) \times \mathbf{r}}{4\pi\epsilon_0 \lambda c r^3} \tag{10.74}$$

and similarly (cf. Eq. (10.72))

$$\mathbf{H} = \frac{\ddot{\mathbf{p}} \times \mathbf{r}}{4\pi c r^2} = \frac{i\omega Ids \times \mathbf{r}}{4\pi c r^2} = i\frac{2\pi Ids \times \mathbf{r}}{\lambda 4\pi r^2}. \tag{10.75}$$

10.6.1   The Power of an Oscillating Dipole

The radiation emitted by Ids or $\mathbf{p}$ is therefore called *dipole radiation*. In propagating radially away from the dipole, the wave transports energy away. The *power*, i.e. the loss of energy per unit time, can be calculated with the help of the Poynting vector $\mathbf{S}$. With Eq. (10.73) we obtain

$$\begin{aligned}
\mathbf{S} &= \mathbf{E} \times \mathbf{H} \quad \text{watt/meter}^2 \\
&\overset{\text{far region}}{=} \frac{1}{c\epsilon_0}\left(\mathbf{H} \times \frac{\mathbf{r}}{r}\right) \times \mathbf{H} \\
&= \frac{1}{c\epsilon_0}(\mathbf{H} \cdot \mathbf{H})\frac{\mathbf{r}}{r} - \frac{1}{c\epsilon_0}\underbrace{\left(\mathbf{H} \cdot \frac{\mathbf{r}}{r}\right)}_{0}\mathbf{H},
\end{aligned}$$

and hence

$$\begin{aligned}
\mathbf{S} &= \frac{1}{c\epsilon_0}H^2\frac{\mathbf{r}}{r} = \frac{1}{c\epsilon_0}\left(\frac{\ddot{\mathbf{p}} \times \mathbf{r}}{4\pi c r^2}\right)^2\frac{\mathbf{r}}{r} \\
&= \frac{(\ddot{p})^2 \sin^2\theta}{(4\pi)^2\epsilon_0 c^3 r^2}\frac{\mathbf{r}}{r}. \tag{10.76}
\end{aligned}$$

We see that $\mathbf{S}$ points in the direction of $\mathbf{r}$. *The oscillating dipole does not emit radiation in the direction of its axis, but perpendicular to this maximally, i.e. in the equatorial plane*, as indicated in Fig. 10.4. We now put, along the lines given above,

$$\dot{\mathbf{p}} = -\omega\mathbf{p}_0\sin\omega t, \quad \ddot{\mathbf{p}} = -\omega^2\mathbf{p}_0\cos\omega t$$

[||]Like previously $\mathbf{p} = \mathbf{p}(t')$, where $t' = t - r/c$, but for reasons of simplicity we write simply $p(t)$, where $t \to t - r/c$.

(actually here $t \to t - r/c$, retarded time, but this is irrelevant here) and insert this in **S**. Then

$$|\mathbf{S}| = \frac{p_0^2 \omega^4}{\epsilon_0 (4\pi)^2 c^3 r^2} \cos^2 \omega t \sin^2 \theta. \tag{10.77}$$
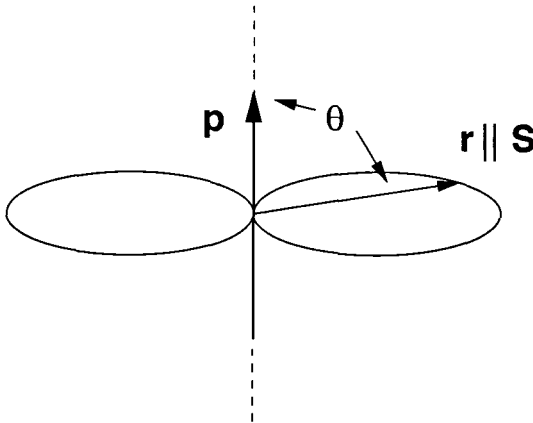
Fig. 10.4 The radiating dipole.

The *average energy emitted per unit time and unit area* is obtained by calculating the average of $\cos^2 \omega t$ over one complete period of oscillation $T = 2\pi/\omega$, i.e. by considering on the one hand

$$\frac{1}{T} \int_0^T \cos^2 \omega t\, dt = \frac{1}{T} \int_0^T \cos^2 \frac{2\pi t}{T}\, dt = \frac{1}{2\pi} \int_0^{2\pi} \cos^2 x\, dx = \frac{\pi}{2\pi} = \frac{1}{2},$$

but also (for a little later)

$$\frac{1}{T} \int_0^T \left\{ \cos^2 \omega t + \sin^2 \omega t \right\} dt = \frac{1}{T} T = 1 \quad \text{and} \quad \frac{1}{T} \int_0^T \sin^2 \omega t\, dt = \frac{1}{2}.$$

On the other hand, we have to integrate $|\mathbf{S}|$ over all directions of a sphere with radius r. Hence the average energy emitted per unit time is ($d\Omega = d\varphi \sin\theta d\theta$)

$$
\begin{aligned}
L &= \frac{1}{2} \int |\mathbf{S}| r^2 d\Omega \\
&= \frac{1}{2} \int_0^{2\pi} d\varphi \int_0^{\pi} |\mathbf{S}| r^2 \sin\theta d\theta \\
&= \frac{1}{2} \frac{2\pi p_0^2 \omega^4}{\epsilon_0 (4\pi)^2 c^3} \int_0^{\pi} \sin^3 \theta d\theta \\
&= \frac{p_0^2 \omega^4}{4 c^3 4\pi \epsilon_0} \int_{-1}^{1} (1 - z^2) dz.
\end{aligned} \tag{10.78}
$$

Here we put $z = \cos\theta$, so that $dz = -\sin\theta d\theta, (1 - z^2)dz = -\sin^3\theta d\theta$. Moreover, $\int_{-1}^{1}(1 - z^2)dz = [z - z^3/3]_{-1}^{1} = 4/3$, so that

$$L = \frac{p_0^2\omega^4}{4\pi\epsilon_0 3c^3} \quad \text{watt.} \tag{10.79}$$

This result is also known as a *Larmor formula*; we can call it the *Larmor dipole formula*, in order to distinguish it from the Larmor formula (10.58) which we obtained for a *single* charged particle. It implies that $L \propto p_0^2$, i.e. proportional to the square of the amplitude p_0, but also that $L \propto \omega^4$, the fourth power of the frequency. Expression (10.79) should be compared with the Rayleigh–Jeans law in statistical mechanics, which gives the classical limit of black body radiation.** The radiation emitted by atoms is of dipole-type; however, the classical picture of an atom as a dipole like the one discussed here, is — as we know — wrong (and in any case would not explain the stability of atoms).

10.6.2 Radiation Resistance

We saw that

$$I d\mathbf{s} = \dot{\mathbf{p}} = -\omega\mathbf{p}_0\sin\omega t.$$

We now put

$$|\mathbf{p}_0\omega| \equiv |I_0 d\mathbf{s}|, \quad \text{or} \quad p_0^2\omega^2 = (I_0 d\mathbf{s})^2.$$

We can say

$$\frac{1}{T}\int_0^T (I d\mathbf{s})^2 dt = \frac{1}{T}\int_0^T \omega^2 p_0^2 \sin^2\omega t\, dt = \frac{1}{2}\omega^2 p_0^2 = \frac{1}{2}I_0^2(d\mathbf{s})^2,$$

$$I_0^2 = \frac{2}{T}\int_0^T I^2(t)dt \equiv 2\bar{I}^2,$$

i.e. $I_0 \sim \sqrt{2}\times$ mean current $\bar{I}$. Then

$$L = \frac{(I_0 d\mathbf{s})^2\omega^2}{3c^3 4\pi\epsilon_0} = \frac{(\sqrt{2}\bar{I}d\mathbf{s})^2\omega^2}{3c^3 4\pi\epsilon_0}. \tag{10.80}$$

**This is a body with absorption coefficient $a(\omega) = 1$ in the case of thermal equilibrium with its surroundings so that because of the densely lying levels, at every frequency per unit area and per unit time and per direction of polarisation the amount of energy absorbed is equal to the amount of energy emitted $(1 - a(\omega)$ is the fraction of reflected energy). Note: The number n of photonic states in the wave number interval $(\kappa, \kappa + d\kappa)$ is (since there are 2 polarisation directions) $n = 2V4\pi\kappa^2 d\kappa$. Since $\kappa \propto 1/\lambda$, it follows that $n \propto d\lambda/\lambda^4$. In the case of monochromatic radiation $\int \delta(\lambda - \lambda_0)d\lambda/\lambda^4 = 1/\lambda_0^4 \propto \omega_0^4$.

This is to be compared with (see below) the *Ohmic power of a resistance R*, i.e. with

$$L = I^2 R. \tag{10.81}$$

This expression is obtained as follows. We saw earlier that (see Sec. 7.7):

Loss of energy per unit volume and per unit time

$$= -\frac{\partial}{\partial t}\left\{\frac{1}{2}(\mathbf{E} \cdot \mathbf{D} + \mathbf{H} \cdot \mathbf{B})\right\} = \nabla \cdot (\mathbf{E} \times \mathbf{H}) + \mathbf{j} \cdot \mathbf{E}.$$

Integrating this over the volume and using Gauss' divergence theorem, we obtain

Loss of energy per unit time
$= Transport\ of\ energy\ through\ the\ surface + Ohmic\ power.$

Hence the Ohmic power is (with $\mathbf{j} = \sigma\mathbf{E}$, $\mathbf{j}dV = I d\mathbf{s}$) the volume integral of $\mathbf{j} \cdot \mathbf{E}$, i.e.

$$\int \mathbf{j} \cdot \mathbf{E} dV = \frac{1}{\sigma}\int \mathbf{j}^2 dV = \frac{1}{\sigma}\int \mathbf{j} \cdot I d\mathbf{s}$$

$$= \frac{1}{\sigma}\int I^2 \frac{d\mathbf{s} \cdot d\mathbf{s}}{dV} = \frac{I^2}{\sigma}\int \frac{dsds}{dV}$$

$$= \frac{I^2}{\sigma}\int \frac{ds}{F} = I^2 \rho \int \frac{ds}{F}$$

$$= I^2 R,$$

where $R = \rho s/F$. We thus have the above expression for L. The quantity $L/\bar{I}^2$ in Eq. (10.80) is therefore (by comparison with $L = I^2 R$) referred to as *radiation resistance*. The radiation resistance is a quantity which depends on the spatial dimensions of the radiator.

10.7 Further Examples

Example 10.7: Comparison of powers of rod and loop antennas
With a very approximate calculation compare the power of a rod antenna with that of a loop antenna.

Solution: We return to the far-field result for $\mathbf{E}$ given by Eq. (10.74), i.e.

$$\mathbf{E} = i\frac{2\pi(I_1 d\mathbf{s}_1 \times \mathbf{r}) \times \mathbf{r}}{4\pi\epsilon_0 \lambda c r^3}. \tag{10.82}$$

This expression gives the electric field strength at a distance $\mathbf{r}$ far from the current element $I_1 d\mathbf{s}_1$. The electromotive force and hence *voltage* induced in a conductor element $d\mathbf{s}_2$ at the point $\mathbf{r}$ is

$$\mathbf{E} \cdot d\mathbf{s}_2 = V_2,$$

i.e.

$$V_2 = i\frac{2\pi(I_1 ds_1 \times \mathbf{r}) \times \mathbf{r} \cdot ds_2}{4\pi\epsilon_0 \lambda c r^3}$$

$$= i\frac{2\pi I_1 (ds_1 \times \mathbf{r}) \cdot (\mathbf{r} \times ds_2)}{4\pi\epsilon_0 \lambda c r^3}. \tag{10.83}$$

This expression is completely symmetric in ds_1 and ds_2; thus if ds_2 were carrying current I_1, then V_2 would be the voltage induced in ds_1.

The simplest way to receive radio waves is with the help of a loop antenna, i.e. a closed antenna (in general with several turns). The *induced electromotive force* or voltage is

$$V = \oint \underbrace{\mathbf{E}}_{\mathbf{E}\ \text{of wave}} \cdot ds = \int \boldsymbol{\nabla} \times \mathbf{E} \cdot d\mathbf{F} = \int \left(-\frac{\partial \mathbf{B}}{\partial t}\right) \cdot d\mathbf{F}.$$

Let us assume $\mathbf{B} \propto e^{i\omega t}$ with angular frequency ω, i.e.

$$\frac{\partial \mathbf{B}}{\partial t} = i\omega \mathbf{B}.$$

Then, if the area A of the loop is small, $A \sim l^2$, so that $\mathbf{B}$ is approximately uniform over A (we assume: $l \ll \lambda \ll r$) and

$$V = -i\omega \mathbf{B} \cdot \mathbf{A}.$$

We attempt a very rough comparison of the power of the loop antenna with that of a rod antenna of length l. In the case of the loop

$$V_{\text{loop}} = -i\omega BA,$$

if $\mathbf{B} \parallel \mathbf{A}$ (then maximal — we assume an optimal orientation of the antenna). In the case of the rod-antenna (see Eq. (10.73)) we have

$$V_{\text{rod}} = \mathbf{E} \cdot \mathbf{l} = \left(\mathbf{H} \times \frac{\mathbf{r}}{r}\right) \cdot \frac{1}{c\epsilon_0} = \frac{Hl}{c\epsilon_0} \text{ or } \frac{Bl}{c\mu_0\epsilon_0}.$$

Thus the ratio is

$$\left|\frac{V_{\text{loop}}}{V_{\text{rod}}}\right| = \mu_0\epsilon_0 c\frac{\omega BA}{Bl}.$$

Since $c = \nu\lambda, \nu = \omega/2\pi$, and so $c = \omega\lambda/2\pi, \omega = 2\pi c/\lambda$, it follows that

$$\left|\frac{V_{\text{loop}}}{V_{\text{rod}}}\right| = \mu_0\epsilon_0 c\frac{BA2\pi c}{\lambda Bl} = \mu_0\epsilon_0 c^2\frac{2\pi A}{l\lambda} = \frac{2\pi A}{l\lambda}$$

for optimal reception. Now if $A \sim O(l^2)$, the ratio is $\sim l/\lambda$. Since we assumed that the dimensions of the loop are very small compared with λ, i.e. $l \ll \lambda$, the rod is the better receiver in these considerations. In view of our earlier treatment of ring-shaped conductors we can say, that the loop with oscillating current can also be looked at as a *magnetic dipole*.

Example 10.8: The Thomson scattering cross section[††]

A plane, monochromatic wave with electric vector

$$\mathbf{E} = \mathbf{e}_x E_0 \cos(\omega t - kz)$$

[††]A cross section is a quantity with dimension of (length)^2. The obvious possibility to define this here can be seen from the definition of the Poynting vector as a quantity with dimension of (energy)(time)$^{-1}$(length)$^{-2}$.

(where $\mathbf{e}_x$ is a unit vector in the direction of x) falls on a free electron (charge q, mass m) at the origin of coordinates. Calculate:
(a) the radiation power P of a single electron,
(b) the mean radiation power $\langle P \rangle$ of such an electron (averaged over one period of oscillation),
(c) the rate of incoming radiation $\mathbf{S}$,
(d) the mean rate of incoming radiation $\langle S_z \rangle$, and
(e) show that in vacuum $\langle P \rangle = \sigma \langle S_z \rangle$, where

$$\sigma = \frac{2}{3} \left(\frac{q^2}{mc^2} \right)^2 \sqrt{\frac{\mu}{\epsilon}} \frac{c}{4\pi\epsilon_0},$$

is the so-called Thomson scattering cross section.

Solution: We have $E_x = E_0 \cos(\omega t - kz)$. An electron with charge q at the origin $z = 0$ is therefore given an acceleration (in Newton's nonrelativistic treatment) given by the applied force divided by m,

$$a_x = \frac{q}{m} E_0 \cos \omega t$$

with solution $x \propto \cos \omega t$, so that

$$a_x = -\omega^2 x, \ qa_x = -q\omega^2 x \equiv -\omega^2 p, \ p := qx.$$

The power radiated by a single electron is given by the single particle Larmor formula of Eq. (10.58) (see above, here still without the averaging)

$$P(t) = \frac{2q^2 a_x^2(t')}{4\pi\epsilon_0 3c^3} \bigg|_{t'=t-\frac{r}{c}} = \frac{2q^4}{3m^2 c^3 4\pi\epsilon_0} E_0^2 \left\{ \cos\left(\omega t - \frac{r\omega}{c} \right) \right\}^2.$$

Averaging over one period of oscillation $T = 2\pi/\omega$ we obtain the mean value

$$\langle P \rangle = \frac{2q^4 E_0^2}{4\pi\epsilon_0 3m^2 c^3} \left\langle \cos^2 \left(\omega t - \frac{r\omega}{c} \right) \right\rangle,$$

where

$$\left\langle \cos^2 \left(\omega t - \frac{r\omega}{c} \right) \right\rangle = \frac{1}{T} \int_0^T \cos^2 \left(\omega t - \frac{r\omega}{c} \right) dt = \frac{1}{2},$$

i.e.

$$\langle P \rangle = \frac{q^4 E_0^2}{3m^2 c^3 4\pi\epsilon_0} \quad \text{watt}.$$

The rate at which radiation (this is the given electric field) falls on the electron is obtained with the help of the Poynting vector, $\mathbf{S} = \mathbf{E} \times \mathbf{H}$. Since $\mathbf{E} \rightarrow E_x \mathbf{e}_x = E_0 \cos(\omega t - kz) \mathbf{e}_x$, the Maxwell equations imply

$$\mathbf{B} = \sqrt{\epsilon_0 \mu} \frac{\mathbf{k} \times \mathbf{E}}{k}, \ \ \mathbf{k} = k\mathbf{e}_z,$$

so that with $B = \mu H$

$$\mathbf{B} \rightarrow B_y = \sqrt{\epsilon_0 \mu} E_0 \cos(\omega t - kz), \qquad H_y = \sqrt{\frac{\epsilon_0}{\mu}} E_0 \cos(\omega t - kz).$$

It follows that

$$\mathbf{S} \rightarrow S_z = \sqrt{\frac{\epsilon_0}{\mu}} E_0^2 \cos^2(\omega t - kz).$$

The time-averaged rate of incoming radiation is (with $T = 2\pi/\omega$)

$$\langle S_z \rangle = \frac{1}{T} \int_0^T dt S_z = \sqrt{\frac{\epsilon_0}{\mu}} E_0^2 \frac{\omega}{2\pi} \int_0^{2\pi/\omega} dt \cos^2(\omega t - kz)$$

$$= \sqrt{\frac{\epsilon_0}{\mu}} E_0^2 \frac{1}{2} = \frac{1}{2}\sqrt{\frac{\epsilon_0}{\mu}} E_0^2 \quad \text{W/m}^2.$$

Hence

$$E_0^2 = 2\sqrt{\frac{\mu}{\epsilon_0}} \langle S_z \rangle,$$

so that

$$\langle P \rangle = \frac{2q^4}{3m^2 c^3} \sqrt{\frac{\mu}{\epsilon_0}} \frac{\langle S_z \rangle}{4\pi\epsilon_0} \equiv \sigma \langle S_z \rangle,$$

where

$$\sigma = \frac{2}{3}\left(\frac{q^2}{mc^2}\right)^2 \sqrt{\frac{\mu}{\epsilon_0}} \frac{c}{4\pi\epsilon_0} \quad \text{meter}^2.$$

This cross section is seen to be independent of the frequency ω of the incident radiation. Generalisations are considered in Examples 10.9 and 10.10.

Example 10.9: The electron bound in an atom

In Example 10.8 we considered the free electron. We now consider an electron bound in an atom subject to a binding force proportional to x but neglecting damping effects. The problem is to derive the new expression for the Thomson cross section.

Solution: The equation of motion of the electron of Example 10.8 now has an additional term $\omega_0^2 x$ representing the binding of the electron to an atom (harmonic oscillator models are always the easiest to construct!),

$$\frac{d^2 x}{dt^2} + \omega_0^2 x = \frac{q}{m} E_0 \cos(\omega t). \quad \therefore x = \frac{q/m}{\omega_0^2 - \omega^2} E_0 \cos \omega t.$$

$$\therefore a_x = \ddot{x} = -\frac{(q/m)\omega^2}{\omega_0^2 - \omega^2} E_0 \cos \omega t.$$

Thus we obtain $\langle P \rangle = \sigma \langle S_z \rangle$ with

$$\sigma = \frac{2}{3}\left(\frac{q^2}{mc^2}\right)^2 \left(\frac{\omega^2}{\omega_0^2 - \omega^2}\right)^2 \sqrt{\frac{\mu}{\epsilon_0}} \frac{c}{4\pi\epsilon_0}.$$

We see that σ is now a function of ω. In the case of scattering of light off atoms and molecules in air $\omega_0 \gg \omega$, and hence

$$\langle P \rangle \simeq \frac{2}{3}\left(\frac{q^2}{mc^2}\right)^2 \left(\frac{\omega}{\omega_0}\right)^4 \sqrt{\frac{\mu}{\epsilon_0}} \frac{c}{4\pi\epsilon_0}.$$

Blue light with $\lambda_{\text{blue}} = 3900\text{Å} \ll \lambda_{\text{red}} = 7600\text{Å}$ is therefore scattered much more than red light, because:

$$c = \nu\lambda, \ \omega = 2\pi\nu, \quad \therefore \omega \propto \frac{1}{\lambda}.$$

The scattering off *free* electrons is (see Example 10.8) independent of the frequency. As a consequence the light scattered by the corona of the sun is white.

Example 10.10: Line broadening

(a) Calculate $\langle W_0 \rangle$, the work done by the force E_x of Example 10.8 applied to an electron averaged over one period of oscillation T. What is the meaning of the result?

(b) Add $\alpha \cos \omega t$ to the velocity of the electron, calculate the power $\langle W_0 \rangle / T$, which is the power of radiated energy. What is the corrected equation of motion of the electron?
(c) How does the loss of radiated energy become evident?

Solution:
(a) We have

$$m\ddot{x} = qE_0 \cos \omega t = F_{\text{applied}}, \quad \dot{x} = \frac{qE_0}{m\omega} \sin \omega t. \tag{10.84}$$

Hence the work done is W_0:

$$W_0 = \int F_{\text{applied}} dx = \int F_{\text{applied}} \dot{x} dt$$

and

$$\frac{\langle W_0 \rangle}{T} = \frac{1}{T} \int_0^T F_{\text{applied}} \frac{qE_0}{m\omega} \sin(\omega t) dt = \frac{q^2 E_0^2}{m\omega T} \int_0^T \cos(\omega t) \sin(\omega t) dt = 0.$$

Thus the work done by the applied force is not equal to the radiated energy. Where does the radiated energy come from? Equation (10.84) must contain a term, i.e. a force, whose work does not average out to zero.
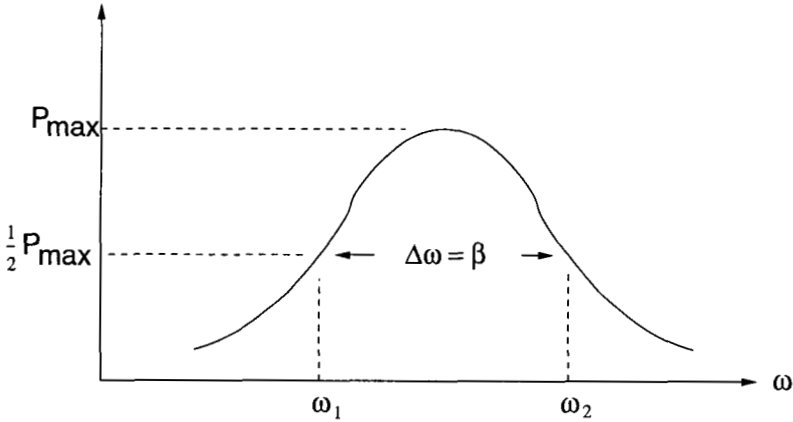
Fig. 10.5 Line broadening by radiation damping.

(b) With addition of the extra term we have

$$\dot{x} = \frac{qE_0}{m\omega} \sin \omega t + \alpha \cos \omega t, \tag{10.85}$$

so that

$$\frac{\langle W_0 \rangle}{T} = \frac{1}{T} \int_0^T F_{\text{applied}} \alpha \cos(\omega t) dt \equiv \frac{1}{2} qE_0 \alpha.$$

In Example 10.8 we calculated the mean power of the radiated energy:

$$\langle P \rangle = \frac{q^4 E_0^2}{3m^2 c^3 4\pi\epsilon_0} \equiv \frac{1}{2} qE_0 \alpha, \quad \therefore \alpha = \frac{2q^3 E_0}{3m^2 c^3 4\pi\epsilon_0}.$$

Hence

$$\dot{x} = \frac{qE_0}{m\omega} \sin(\omega t) + \frac{2q^3 E_0}{3m^2 c^3 4\pi\epsilon_0} \cos(\omega t), \tag{10.86}$$

so that

$$m\ddot{x} = \underbrace{qE_0\cos(\omega t)}_{F_{\text{applied}}} - \underbrace{\frac{2q^3E_0\omega}{3mc^3 4\pi\epsilon_0}\sin(\omega t)}_{F_{\text{radiation}}},$$

where $F_{\text{radiation}}$ is the radiative reaction. This equation is known as the *Abraham–Lorentz equation of motion*. The ratio

$$\frac{F_{\text{radiation}}}{F_{\text{applied}}} = \frac{2q^2\omega}{3mc^3 4\pi\epsilon_0}\tan(\omega t)$$

is in general very small.

(c) The loss of radiated energy is due to a damping or *"line broadening"* (speaking spectroscopically). Let us write the radiation reaction as a "friction force" proportional to velocity, i.e. we put the friction force $= F_{\text{radiation}}$, i.e.

$$m\beta\frac{dx}{dt} \equiv \frac{2q^3E_0\omega}{3mc^3 4\pi\epsilon_0}\sin(\omega t). \tag{10.87}$$

From Eq. (10.85) we obtain

$$\frac{dx}{dt} \sim \frac{qE_0}{m\omega}\sin(\omega t) + (\text{something negligible}).$$

Inserting the latter into Eq. (10.87) we obtain

$$\beta = \frac{2q^2\omega^2}{3mc^3 4\pi\epsilon_0}.$$

Solving the equation

$$m\ddot{x} + m\beta\dot{x} + m\omega_0^2 = qE_0\cos(\omega t),$$

we find that the term $m\beta\dot{x}$ produces a damping factor $\exp(-\beta t)$ (*"radiation damping"*) with the so-called line width $\beta \equiv \triangle\omega$ as indicated in Fig. 10.5, where

$$\omega_{1,2} = \sqrt{\omega_0^2 + (\beta/2)^2} \mp \beta/2.$$

Chapter 11

The Laws of Optics

11.1 Introductory Remarks

With the realisation that light is nothing but electromagnetic radiation which reaches our eyes, it is clear that basic optical phenomena must also find their explanation in electrodynamics and hence in Maxwell's equations. The subject of this chapter is therefore the derivation of the well known laws of reflection and refraction of optics from Maxwell's equations. In subsequent chapters these will then be applied to metals, radio waves and wave guides.[*]

11.2 Continuity Conditions and Definition of the Surface Current

Our procedure will be to arrive at the laws of optics from a consideration of the behaviour of an electromagnetic wave at the interface between two media. First, however, we recapitulate and summarise the essential continuity conditions and those of their validity for normal and tangential components of the fields, with indices (1) and (2) referring to two different media.

(a) $\nabla \cdot \mathbf{D} = \rho$. Here ρ is the density of the *"true" charges*. With this equation we have:

$$\int_V \nabla \cdot \mathbf{D} \, dV = \oint_{F_V} \mathbf{D} \cdot d\mathbf{F} = \int_V \rho \, dV = Q. \tag{11.1}$$

In the case of the infinitesimal volume element of thickness $d \to 0$ as shown in Fig. 11.1 we obtain from $\oint_{F_V} \mathbf{D} \cdot d\mathbf{F}$ (observe that in the limit $d \to 0$ the

---

243

quantity $\triangle F$ is an element of area *in* the boundary surface)

$$(\mathbf{D}^{(1)} - \mathbf{D}^{(2)}) \cdot \triangle \mathbf{F} = (D_n^{(1)} - D_n^{(2)})\triangle F = \triangle Q, \tag{11.2}$$

so that (see Fig. 11.1),

$$D_n^{(1)} - D_n^{(2)} = \sigma, \;\; \sigma = \frac{\triangle Q}{\triangle F}. \tag{11.3}$$

In the case of the electric field strength $\mathbf{E}$ we now obtain (Q_P being the polarisation charge with density ρ_P) with the equations (see Chapter 4)

$$\mathbf{D} = \mathbf{E}\epsilon_0 + \mathbf{P}, \;\; \boldsymbol{\nabla} \cdot \mathbf{P} = -\rho_P, \tag{11.4}$$

the integrated expressions

$$\oint_{F_V} \mathbf{E} \cdot d\mathbf{F} = \frac{1}{\epsilon_0} \oint_{F_V} (\mathbf{D} - \mathbf{P}) \cdot d\mathbf{F}$$
$$= \frac{1}{\epsilon_0} \int_V (\boldsymbol{\nabla} \cdot \mathbf{D} - \boldsymbol{\nabla} \cdot \mathbf{P}) dV$$
$$= \frac{1}{\epsilon_0} \int_V (\rho + \rho_P) dV = \frac{1}{\epsilon_0}(Q + Q_P). \tag{11.5}$$

Hence at the interface

$$E_n^{(1)} - E_n^{(2)} = \frac{1}{\epsilon_0}(\sigma + \sigma_P). \tag{11.6}$$

Here σ_P is the induced surface charge density.

(b) $\boldsymbol{\nabla} \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}$ (Faraday). Here we have in the case of a cross-sectional area F' with $\triangle F' = l \times d$ as in Fig. 11.1:
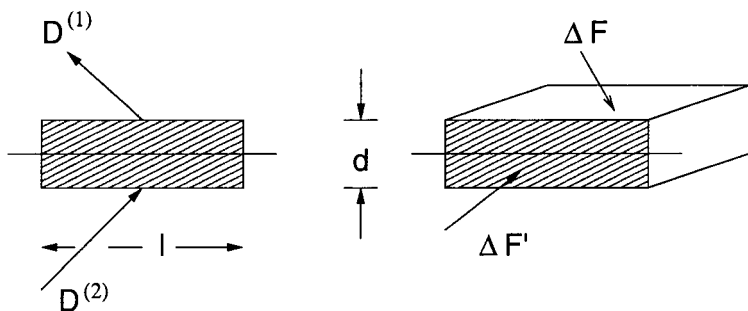
Fig. 11.1 Elements of area and volume.

$$\int_{F'} (\boldsymbol{\nabla} \times \mathbf{E}) \cdot d\mathbf{F} = -\frac{\partial}{\partial t} \int_{F'} \mathbf{B} \cdot d\mathbf{F} = -\frac{\partial}{\partial t} |\mathbf{B}_n| \triangle F' \to 0,$$

since $\triangle F' = ld \to 0$ with $d \to 0$. Using the Stokes theorem we obtain

$$\int_{F'} (\boldsymbol{\nabla} \times \mathbf{E}) \cdot d\mathbf{F} = \oint \mathbf{E} \cdot d\mathbf{s}$$
$$= (\mathbf{E} \cdot \mathbf{s})^{(1)} + (\mathbf{E} \cdot \mathbf{s})^{(2)}$$
$$= (\mathbf{E}^{(1)} \cdot \mathbf{l}) - (\mathbf{E}^{(2)} \cdot \mathbf{l})$$
$$= (E_t^{(1)} - E_t^{(2)})l, \tag{11.7}$$

and with the preceding result

$$E_t^{(1)} - E_t^{(2)} = 0. \tag{11.8}$$
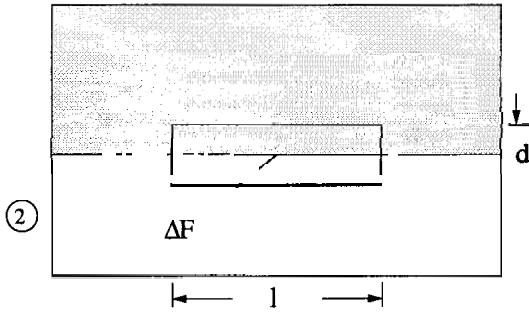
Fig. 11.2 The element of area $\perp$ to the interface.

(c) $\boldsymbol{\nabla} \cdot \mathbf{B} = 0$. From this we obtain immediately

$$\int \boldsymbol{\nabla} \cdot \mathbf{B} dV = \oint \mathbf{B} \cdot d\mathbf{F} = 0, \tag{11.9}$$

so that

$$B_n^{(1)} - B_n^{(2)} = 0. \tag{11.10}$$

(d) $\boldsymbol{\nabla} \times \mathbf{H} = \frac{\partial \mathbf{D}}{\partial t} + \mathbf{j}$. If there are no surface currents K (see below), i.e. if $\int \mathbf{j} \cdot d\mathbf{F} = 0$, then as under (b):

$$H_t^{(1)} - H_t^{(2)} = 0, \quad (K = 0). \tag{11.11}$$

The *definition of the above surface current* is given as follows. We consider the element of area dF perpendicular to the interface between the two media as illustrated in Fig. 11.2. Then dF is a vector in the plane of the interface. We now consider equation (d) integrated over this area, i.e.

$$\int_{\triangle F} (\boldsymbol{\nabla} \times \mathbf{H}) \cdot d\mathbf{F} = \int_{\triangle F} d\mathbf{F} \cdot \left[\frac{\partial \mathbf{D}}{\partial t} + \mathbf{j} \right].$$

The first contribution on the right is

$$\frac{\partial}{\partial t} |\mathbf{D}| \triangle F = \frac{\partial}{\partial t} |\mathbf{D}| l d \to 0 \quad \text{with} \quad d \to 0. \tag{11.12}$$

We rewrite the second contribution as

$$\int_{\triangle F} d\mathbf{F} \cdot \mathbf{j} = \mathbf{j} \cdot \triangle \mathbf{F} = j dl. \tag{11.13}$$

For conductors we have Ohm's law

$$\mathbf{j} = \sigma \mathbf{E}, \quad \text{i.e.} \quad j = \sigma E. \tag{11.14}$$

For perfect conductors $\sigma = \infty$ and $\mathbf{E} = -\boldsymbol{\nabla}\phi = 0$. Then it is possible that

$$j dl = \sigma E dl \neq 0 \quad \text{and finite,} \tag{11.15}$$

and one defines

$$\lim_{\sigma \to \infty,\, d \to 0} \sigma E d := K, \tag{11.16}$$

i.e. $jd \sim K$, where K is a *surface current density* (see Jackson [3], p. 336); K is also called *"linear current density"* (see Lim [2], p. 42).

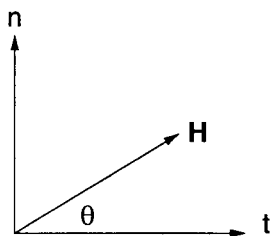Fig. 11.3 Normal and tangential directions.

If the conductivity σ is finite, $K = 0$, i.e. in this case there is really no surface current (see Jackson [3], p. 336). The continuity condition for H is therefore, treating $\int (\boldsymbol{\nabla} \times \mathbf{H}) \cdot d\mathbf{F}$ as for $\mathbf{E}$, i.e. from (d),

$$\int_{\triangle F} \mathbf{H} \cdot d\mathbf{l} = l j d = K l,$$

and it follows that

$$H_t^{(1)} - H_t^{(2)} = K \quad (= 0 \quad \text{for} \quad \sigma \quad \text{finite}). \tag{11.17}$$

For later purposes we can write this also

$$\mathbf{n} \times (\mathbf{H}^{(1)} - \mathbf{H}^{(2)}) = \mathbf{K}, \tag{11.18}$$

where $\mathbf{n}$ is a unit vector perpendicular to the interface, and $\mathbf{K}$ is a vector in the plane of the interface. One should note that as indicated in Fig. 11.3:

$$H_t = \mathbf{H} \cdot \mathbf{e}_t = |\mathbf{H}| \cos \theta$$
$$= |\mathbf{H}| \sin(\frac{\pi}{2} - \theta) = |\mathbf{H} \times \mathbf{n}|. \tag{11.19}$$

The limits $d \to 0, \sigma \to \infty$ apply to an idealised case, and therefore the question arises whether a surface current does actually exist or not. We shall see later, in Chapter 14 on wave guides, that a current density $\mathbf{j}$ which is restricted to a very thin region corresponds to an *effective surface current* $\mathbf{K}_{\text{eff}}$ (see also Jackson [3], p. 339).

Example 11.1: Change of direction of E at an interface
The electric field $\mathbf{E}$ in medium (1) with dielectric constant $\epsilon_1 = 7$ falls at an angle of 60° to the surface normal on the interface to a medium (2) with dielectric constant $\epsilon_2 = 2$. What is the angle between $\mathbf{E}$ and the surface normal in medium (2)?

Solution: We have

$$E_{t1} = E_{t2}, \quad D_{n1} - D_{n2} = 0 \text{ (charge)}, \quad \therefore \epsilon_1 E_{n1} = \epsilon_2 E_{n2}.$$

We are also given : $E_{t1} = \tan 60° E_{n1}$, $E_{t2} = \tan \theta_2 E_{n2}$. The quantity requested is θ_2:

$$\tan \theta_2 = \frac{E_{t2}}{E_{n2}} = \frac{E_{t1}}{\epsilon_1 E_{n1}/\epsilon_2} = \frac{\tan 60° E_{n1}}{\epsilon_1 E_{n1}/\epsilon_2} = \frac{\epsilon_2}{\epsilon_1} \tan 60° = \frac{\tan 60°}{3.5} = 0.495,$$

so that $\theta_2 = 26.4°$.

11.3   Electromagnetic Waves in Media

In the following we consider a *plane, electromagnetic* wave, which falls at an angle from medium (1) on the interface to medium (2), as indicated in Fig. 11.4.

We recapitulate first some of our earlier considerations for the case of *electromagnetic waves in a medium* with dielectric constant ϵ, magnetic permeability μ and conductivity σ. We had the equations

$$\mathbf{D} = \epsilon \mathbf{E}, \quad \mathbf{B} = \mu \mathbf{H}, \quad \mathbf{j} = \sigma \mathbf{E}. \tag{11.20}$$

In the case of charge density $\rho = 0$, the equations are

$$\nabla \cdot \mathbf{H} = 0, \qquad \nabla \times \mathbf{H} = \epsilon \frac{\partial \mathbf{E}}{\partial t} + \sigma \mathbf{E},$$

$$\nabla \cdot \mathbf{E} = 0, \qquad \nabla \times \mathbf{E} = -\mu \frac{\partial \mathbf{H}}{\partial t}. \tag{11.21}$$

From the first two equations we obtain after forming

$$\nabla \times (\nabla \times \mathbf{H}) = \nabla(\nabla \cdot \mathbf{H}) - (\nabla \cdot \nabla)\mathbf{H}$$

with $\nabla \cdot \mathbf{H} = 0$:

$$\triangle \mathbf{H} = \epsilon\mu \frac{\partial^2 \mathbf{H}}{\partial t^2} + \mu\sigma \frac{\partial \mathbf{H}}{\partial t}. \tag{11.22a}$$

With $\nabla \times (\nabla \times \mathbf{E})$, one obtains analogously

$$\triangle \mathbf{E} = \epsilon\mu \frac{\partial^2 \mathbf{E}}{\partial t^2} + \mu\sigma \frac{\partial \mathbf{E}}{\partial t}. \tag{11.22b}$$
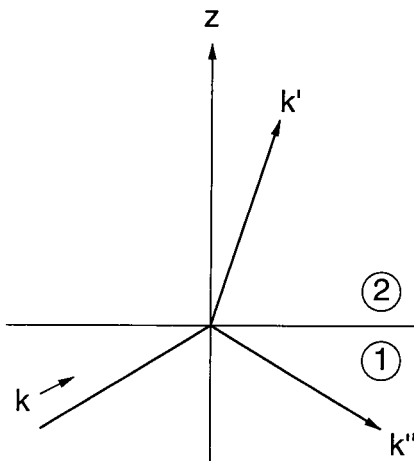

Fig. 11.4 An electromagnetic wave falling at an inclination
on the interface between two media.

For $\sigma = 0$ and propagation velocity

$$c' = \frac{1}{\sqrt{\epsilon\mu}} = \sqrt{\frac{\epsilon_0 \mu_0}{\epsilon\mu}} c \tag{11.23}$$

this is

$$\triangle \mathbf{E} - \frac{1}{c'^2} \frac{\partial^2 \mathbf{E}}{\partial t^2} = 0. \tag{11.24}$$

In the case of *high frequency radiation*, the parameters ϵ, μ, σ become frequency-dependent. Considering *plane waves*, we write for $\mathbf{E}$ and $\mathbf{H}$:

$$\mathbf{E} = \mathbf{E}_0 e^{i(\mathbf{k} \cdot \mathbf{r} - \omega t)}, \qquad \mathbf{H} = \mathbf{H}_0 e^{i(\mathbf{k} \cdot \mathbf{r} - \omega t)}. \tag{11.25}$$

Inserting these into Eq. (11.21) we obtain

$$\nabla \times \mathbf{H} = i\mathbf{k} \times \mathbf{H} = -\epsilon i\omega \mathbf{E} + \sigma \mathbf{E} \equiv -i\omega \eta \mathbf{E},$$

where

$$\eta = \epsilon + i\frac{\sigma}{\omega}, \tag{11.26}$$

which is the *generalised dielectric constant*, in the case $\sigma \neq 0$ for a conducting medium. It follows that we can write Maxwell's equations:

$$\mathbf{k} \times \mathbf{H} = -\omega \eta \mathbf{E}, \tag{11.27}$$
$$\mathbf{k} \times \mathbf{E} = \omega \mu \mathbf{H}, \tag{11.28}$$
$$\mathbf{k} \cdot \mathbf{E} = 0, \tag{11.29}$$
$$\mathbf{k} \cdot \mathbf{H} = 0. \tag{11.30}$$

The last two equations show that here also we have transversality of the electromagnetic field (but not $|\mathbf{E}| = |\mathbf{H}|$). From these equations we determine $|\mathbf{k}|$:

$$\underbrace{\mathbf{k} \times (\mathbf{k} \times \mathbf{H})}_{(\mathbf{k} \cdot \mathbf{H})\mathbf{k} - k^2 \mathbf{H}} = -\omega \eta \underbrace{\mathbf{k} \times \mathbf{E}}_{\omega \mu \mathbf{H}},$$

where $\mathbf{k} \cdot \mathbf{H} = 0$. Hence for $\mathbf{H} \neq 0$:

$$k^2 = \omega^2 \eta \mu = \omega^2 \mu \left(\epsilon + i\frac{\sigma}{\omega} \right),$$

i.e.

$$k^2 = \mu \epsilon \omega^2 \left(1 + i\frac{\sigma}{\omega \epsilon} \right). \tag{11.31}$$

This relation is called the *dispersion relation* of a conducting medium. Setting $k = \alpha + i\beta$, we obtain

$$k = \sqrt{\mu \epsilon} \omega \left[\left\{ \frac{1}{2} \left(\sqrt{1 + \left(\frac{\sigma}{\omega \epsilon} \right)^2} + 1 \right) \right\}^{1/2} \right.$$
$$\left. + i \left\{ \frac{1}{2} \left(\sqrt{1 + \left(\frac{\sigma}{\omega \epsilon} \right)^2} - 1 \right) \right\}^{1/2} \right]. \tag{11.32}$$

For $\sigma = 0$ we have k real; i.e. the wave is not damped, and the medium is said to be *"transparent"* (allows radiation to pass through).

We now put
$$k = |\mathbf{k}|e^{i\varphi},$$

and find
$$|k| = \sqrt{\alpha^2 + \beta^2} = \sqrt{\epsilon\mu}\,\omega \left[1 + \left(\frac{\sigma}{\omega\epsilon}\right)^2\right]^{1/4},$$
$$\varphi = \tan^{-1}\frac{\beta}{\alpha} = \frac{1}{2}\tan^{-1}\left(\frac{\sigma}{\omega\epsilon}\right). \tag{11.33}$$

The *(generalised) refractive index* is defined as the quantity
$$p = c\sqrt{\mu\eta} = p_1 + ip_2. \tag{11.34}$$

Hence for the *vacuum* we have $(\sigma = 0)$:
$$p = \sqrt{\mu_0\epsilon_0}\,c, \quad \text{i.e.} \quad p = 1 \quad \text{and} \quad k = \frac{\omega}{c}. \tag{11.35}$$

11.4 Kinematical Aspects of Reflection and Refraction: Snell's Law

11.4.1 Preliminary Remarks

We now consider reflection and refraction of a wave at a boundary surface between two media (1) and (2) as indicated in Fig. 11.5 which are not metals.

For the our kinematical considerations here we make the following *assumptions*:

(a) there are no surface charges or surface currents,

(b) both media have conductivity $\sigma = 0,^\dagger$ (i.e. are transparent),

(c) in order that $B = \mu H$ be valid, the media are assumed not to be ferromagnetic ($B = \mu H$ is *not* valid for these!),

(d) the incident wave is a *plane wave* with
$$\mathbf{E} = \mathbf{E}_0 e^{i(\mathbf{k}\cdot\mathbf{r} - \omega t)} \tag{11.36}$$

(in the case of a plane wave $f(\mathbf{r}, t) := \mathbf{k}\cdot\mathbf{r} - \omega t = \text{const.}, \mathbf{k} = \boldsymbol{\nabla} f(\mathbf{r}, t)$).

---

†Metals will be treated in Chapter 12. Most of the considerations here for $\sigma = 0$ are also valid for $\sigma \neq 0$ and finite. Then also the surface current density $K = 0$.

The equations
$$\mathbf{k} \times \mathbf{E} = \omega\mu\mathbf{H} = \omega\mathbf{B}, \quad k = \omega\sqrt{\epsilon\mu}$$
then imply for the magnetic induction
$$\mathbf{B} = \sqrt{\epsilon\mu}\frac{\mathbf{k} \times \mathbf{E}}{k}. \tag{11.37}$$

We assume that there are waves in both media, *reflected waves* in one
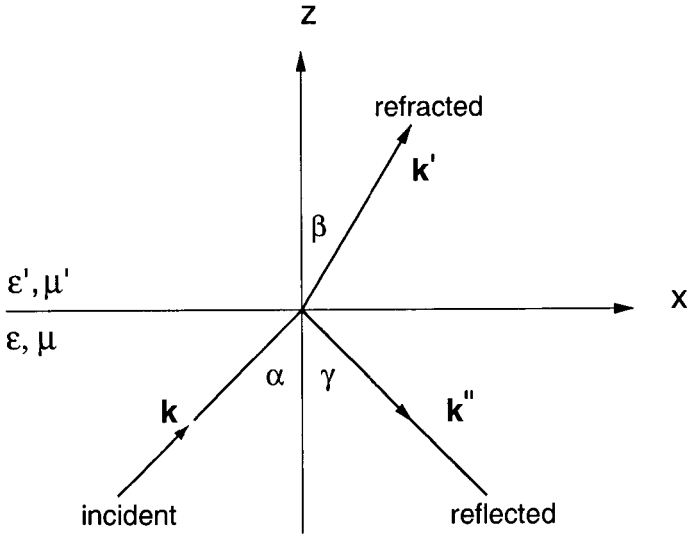
Fig. 11.5 Directions of incident, refracted and reflected waves
with angular inclinations to the normal.

medium and *refracted waves* in the other. For these we take the ansatze

(1) *in the case of reflection*:
$$\mathbf{E}'' = \mathbf{E}_0'' e^{i(\mathbf{k}''\cdot\mathbf{r}-\omega''t)}, \quad \mathbf{B}'' = \sqrt{\epsilon\mu}\frac{\mathbf{k}'' \times \mathbf{E}''}{k''}, \tag{11.38}$$

(2) *in the case of refraction*:
$$\mathbf{E}' = \mathbf{E}_0' e^{i(\mathbf{k}'\cdot\mathbf{r}-\omega't)}, \quad \mathbf{B}' = \sqrt{\epsilon'\mu'}\frac{\mathbf{k}' \times \mathbf{E}'}{k'}. \tag{11.39}$$

In the *case of normal incidence* of the wave and in view of the transversality (see equations preceding Eq. (11.30)) we can write in self-evident notation using Eqs. (11.37), (11.38) and (11.39):

$$\sqrt{\frac{\mu}{\epsilon}}H_{\text{in}} = E_{\text{in}}, \quad \sqrt{\frac{\mu}{\epsilon}}H_{\text{refl}} = E_{\text{refl}}, \quad \sqrt{\frac{\mu'}{\epsilon'}}H_{\text{refr}} = E_{\text{refr}}. \tag{11.40}$$

11.4.2   Kinematical Aspects of Reflection and Refraction

The kinematical aspects of reflection and refraction which we consider now are consequences of the *general boundary condition*, that the conditions of continuity summarised above (whatever they express in detail) must be valid *at every time t and at every point of the interface.* Consider the conditions

$$H_t^{(1)} - H_t^{(2)} = 0, \quad E_t^{(1)} - E_t^{(2)} = 0 \quad \text{at} \quad z = 0. \tag{11.41}$$

If these conditions hold at every time t and at all points of the interface plane (x, y), then we must have for (e.g.) $\mathbf{E}$:

$$[(\mathbf{E}_0)_t e^{i(\mathbf{k}\cdot\mathbf{r}-\omega t)}]_{z=0} + [(\mathbf{E}_0'')_t e^{i(\mathbf{k}''\cdot\mathbf{r}-\omega'' t)}]_{z=0} = [(\mathbf{E}_0')_t e^{i(\mathbf{k}'\cdot\mathbf{r}-\omega' t)}]_{z=0},$$

i.e.

$$E_{0t} e^{i(k_x x + k_y y - \omega t)} + E_{0t}'' e^{i(k_x'' x + k_y'' y - \omega'' t)} = E_{0t}' e^{i(k_x' x + k_y' y - \omega' t)}. \tag{11.42}$$

If this is to hold at all times t, we must have

$$e^{i(\omega'-\omega)t} = 1 = e^{i(\omega'-\omega'')t},$$

i.e.

$$\omega = \omega' = \omega''$$

(possibly for $\omega t = \omega'' t + 2m\pi, m = 0, \pm 1, \ldots$, but demanding validity for all t, hence also for $t = 0$, implies $m = 0$). Similarly we must have for projections onto the (x, y) plane

$$\mathbf{k} \cdot \mathbf{r} = \mathbf{k}' \cdot \mathbf{r} = \mathbf{k}'' \cdot \mathbf{r} \quad \text{at} \quad z = 0. \tag{11.43}$$

Thus the projections of $\mathbf{k}, \mathbf{k}', \mathbf{k}''$ on any arbitrarily chosen vector $\mathbf{r}$ in the interface plane must be equal. However, this is possible only if (see below) $\mathbf{k}, \mathbf{k}', \mathbf{k}''$ lie in one and the same plane, the so-called *plane of incidence*. This can be seen as follows. On the one hand, if we replace $\mathbf{r}$ by a unit vector along the x-axis, the x-components have to be equal, and analogously the y-components have to be equal. On the other hand, if two of the three vectors $\mathbf{k}, \mathbf{k}', \mathbf{k}''$ lie in the (x, z)-plane, and so have zero y-components, (also at $z = 0$), also the y-component of the third vector is zero, implying that this vector must also lie in the (x, z)-plane, i.e.

$$\{\mathbf{k} \cdot \mathbf{r} = \mathbf{k}' \cdot \mathbf{r} = \mathbf{k}'' \cdot \mathbf{r}\}_{z=0} = 0.$$

If $\mathbf{k}, \mathbf{k}''$ span the (x, z)-plane, i.e. if $k_y = 0 = k_y''$, then

$$k_x x = k_x' x + k_y' y = k_x'' x;$$

but because $k_x = k'_x = k''_x$, it follows that $k'_y = 0$. Furthermore we deduce from Eq. (11.43) for the x-components (as we can see from Fig. 11.5)

$$k \sin \alpha = k' \sin \beta = k'' \sin \gamma. \tag{11.44}$$

Since $\omega = \omega''$ and in general for $\sigma = 0 : k = \sqrt{\epsilon\mu}\omega$, it follows that[‡] $k = k''$ in the same medium, and hence

$$\alpha = \gamma,$$

which implies that *the incident angle is equal to the reflection angle.* In addition *Snell's law* must hold, i.e.

$$\frac{\sin \alpha}{\sin \beta} = \frac{k'}{k} = \sqrt{\frac{\epsilon'\mu'}{\epsilon\mu}} = \frac{n'}{n} = \frac{c_1}{c_2} = \frac{\lambda_1}{\lambda_2}, \tag{11.45}$$

where c_i, λ_i are velocity of light and wavelength in medium 'i', and the refractive index is

$$n = \frac{c}{c_1} = \sqrt{\frac{\epsilon\mu}{\epsilon_0\mu_0}} \quad \text{and} \quad n' = \frac{c}{c_2} = \sqrt{\frac{\epsilon'\mu'}{\epsilon_0\mu_0}}. \tag{11.46}$$

11.5 Dynamical Aspects of Reflection and Refraction: The Fresnel Formulae

11.5.1 The Conditions

The dynamical aspects of reflection and refraction follow from the *specific boundary conditions* for $\mathbf{D}, \mathbf{E}, \mathbf{B}, \mathbf{H}$. We assume that the media are isotropic (i.e. $\mathbf{D} = \epsilon\mathbf{E}$) and that there are neither surface charges nor surface currents. We use the relations found above

$$\frac{\sqrt{\epsilon\mu}}{k} = \frac{1}{\omega} = \frac{1}{\omega'} = \frac{\sqrt{\epsilon'\mu'}}{k'}. \tag{11.47}$$

We begin with the relation $E_t^{(1)} - E_t^{(2)} = 0$. We let $\mathbf{n}$ be the unit vector parallel to $\mathbf{e}_z$ in the direction of the z-axis. Then

$$A_x = \mathbf{A} \cdot \mathbf{e}_x = |\mathbf{A}| \cos \theta = |\mathbf{A}| \sin \left(\frac{\pi}{2} - \theta \right) = |\mathbf{A} \times \mathbf{n}|.$$

With this we can rewrite the above tangential condition as

(a) $$\qquad\qquad (\mathbf{E}_0 + \mathbf{E}''_0 - \mathbf{E}'_0) \times \mathbf{n} = 0. \tag{11.48}$$

[‡]In the case of different media the ω's are equal, but the k's, $\sqrt{\epsilon\mu}\omega$, are different.

The next condition we consider is $B_n^{(1)} - B_n^{(2)} = 0$. With the relations (11.37), (11.38) and (11.39) for $\mathbf{B}, \mathbf{B}'$ and $\mathbf{B}''$, as well as

$$\mathbf{E} = \mathbf{E}_0 e^{i(\mathbf{k}\cdot\mathbf{r}-\omega t)}, \quad \mathbf{E}' = \mathbf{E}'_0 e^{i(\mathbf{k}'\cdot\mathbf{r}-\omega' t)}, \quad \mathbf{E}'' = \mathbf{E}''_0 e^{i(\mathbf{k}''\cdot\mathbf{r}-\omega'' t)},$$

we have

$$\left[\sqrt{\epsilon\mu}\left(\frac{\mathbf{k} \times \mathbf{E}_0}{k} + \frac{\mathbf{k}'' \times \mathbf{E}''_0}{k''} \right) - \underbrace{\sqrt{\epsilon'\mu'}}_{\frac{k'}{k}\sqrt{\epsilon\mu}} \frac{\mathbf{k}' \times \mathbf{E}'_0}{k'} \right] \cdot \mathbf{n} = 0.$$

With (11.47) and $k = k''$ (reflection) we obtain from this

(b) $\qquad\qquad (\mathbf{k} \times \mathbf{E}_0 + \mathbf{k}'' \times \mathbf{E}''_0 - \mathbf{k}' \times \mathbf{E}'_0) \cdot \mathbf{n} = 0.$ $\qquad\qquad$ (11.49)

The scalar product with $\mathbf{n}$ gives us the normal component of the vector in the bracket.

Next we consider the condition $H_t^{(1)} = H_t^{(2)}$. From this we obtain as in case (b) with $\mathbf{B} = \mu\mathbf{H}$

(c) $\qquad \left[\frac{1}{\mu}(\mathbf{k} \times \mathbf{E}_0 + \mathbf{k}'' \times \mathbf{E}''_0) - \frac{1}{\mu'}(\mathbf{k}' \times \mathbf{E}'_0) \right] \times \mathbf{n} = 0.$ $\qquad$ (11.50)

The vector product with $\mathbf{n}$ gives us the tangential component of the vector in the bracket.

Finally we obtain from the condition $D_n^{(1)} = D_n^{(2)}$ the relation

(d) $\qquad\qquad [\epsilon(\mathbf{E}_0 + \mathbf{E}''_0) - \epsilon'\mathbf{E}'_0] \cdot \mathbf{n} = 0.$ $\qquad\qquad$ (11.51)

In the next section we consider two cases of linear polarisation (linearly polarised implies: $\mathbf{E}, \mathbf{B}$ constant in direction and modulus).

11.5.2   Two Cases of Linear Polarisation

Case 1: $\mathbf{E}_0$ perpendicular to the plane of incidence

The plane of incidence is the plane spanned by the propagation vectors $\mathbf{k}, \mathbf{k}', \mathbf{k}''$. The field vectors are, as we saw, orthogonal to these (cf. transversality). According to the given condition the incident field vector $\mathbf{E}_0$ is perpendicular to this plane and therefore lies in the tangential plane. Since $E_t^{(1)} - E_t^{(2)} = 0$, there is no change of direction of the field $\mathbf{E}$ at the interface of the media for either reflection or refraction, i.e. with $E_{0x} = 0 = E_{0z}$ we also have $E'_{0x} = 0 = E'_{0z}$ and $E''_{0x} = 0 = E''_{0z}$. The relation (a) implies

$$(\mathbf{E}_0 + \mathbf{E}''_0)_t = (\mathbf{E}')_t.$$

The tangential plane is the plane of the interface, i.e. the (x, y)-plane. A general vector in this plane would be $\mathbf{r}_t = x\mathbf{e}_x + y\mathbf{e}_y$. Since in the case of the E-fields here, only the y-components are not zero, we can rewrite the tangential relation also as

$$E_{0y} + E_{0y}'' = E_{0y}',$$

or
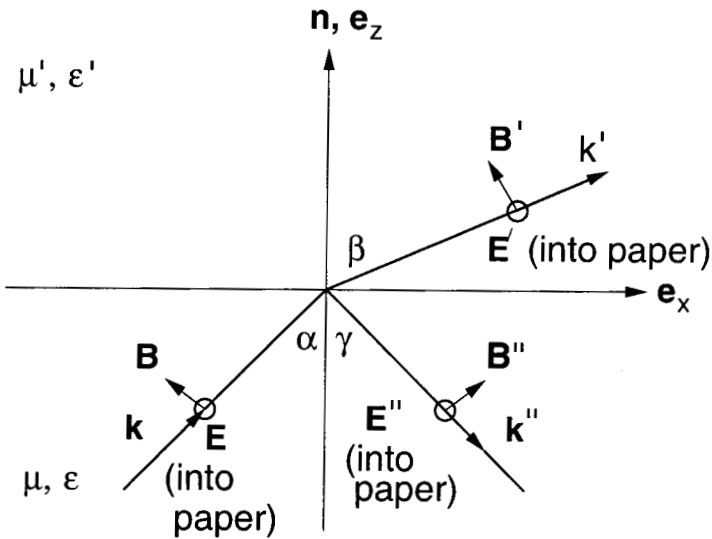
Fig. 11.6 Case 1: The directions of the field components at incidence, refraction and reflection.

(A) $$E_0 + E_0'' = E_0'. \tag{11.52}$$

One should note that E_0, E_0'', E_0' are *amplitudes*, which can be complex. We thus have the situation depicted in Fig. 11.6. The orientation of the **B**-vectors has been chosen in Fig. 11.6 such that the transport of energy $\mathbf{E} \times \mathbf{H}$ is positive in the direction of the wave vectors.

We can rewrite the relation (b), i.e. Eq. (11.49), using the property of scalar triple products

$$(\mathbf{A} \times \mathbf{B}) \cdot \mathbf{C} = (\mathbf{C} \times \mathbf{A}) \cdot \mathbf{B}.$$

Thus Eq. (11.49) can be written as

$$(\mathbf{n} \times \mathbf{k}) \cdot \mathbf{E}_0'' + (\mathbf{n} \times \mathbf{k}'') \cdot \mathbf{E}_0'' - (\mathbf{n} \times \mathbf{k}') \cdot \mathbf{E}_0' = 0,$$

i.e.

$$k \sin \alpha \mathbf{e}_y \cdot \mathbf{E}_0 + k'' \sin \gamma \mathbf{e}_y \cdot \mathbf{E}_0'' - k' \sin \beta \mathbf{e}_y \cdot \mathbf{E}_0' = 0.$$

However, according to Snell's law

$$k \sin \alpha = k' \sin \beta, \quad \alpha = \gamma, \quad k = k'',$$

so that

$$k \sin \alpha \mathbf{e}_y \cdot (\mathbf{E}_0 + \mathbf{E}_0'' - \mathbf{E}_0') = 0.$$

Since the sum of the vectors $\mathbf{E}$ (as above) lies in the tangential plane, we conclude as we concluded from (a) that $E_0 + E_0'' - E_0' = 0$. Hence this case yields no new relation.

Consider now relation (c). We handle Eq. (11.50) as follows. We have

$$\frac{1}{\mu}(\mathbf{k} \times \mathbf{E}_0) \times \mathbf{n} = \frac{1}{\mu}[(\mathbf{n} \cdot \mathbf{k})\mathbf{E}_0 - \mathbf{k}\underbrace{(\mathbf{n} \cdot \mathbf{E}_0)}_{0}].$$

Hence for the incident wave

$$\frac{1}{\mu}(\mathbf{k} \times \mathbf{E}_0) \times \mathbf{n} = \frac{1}{\mu}\underbrace{k}_{\omega\sqrt{\epsilon\mu}}\mathbf{E}_0 \cos \alpha = \omega\sqrt{\frac{\epsilon}{\mu}}\mathbf{E}_0 \cos \alpha.$$

Similarly we obtain for the reflected and refracted waves

$$\frac{1}{\mu}(\mathbf{k}'' \times \mathbf{E}_0'') \times \mathbf{n} = -\omega''\sqrt{\frac{\epsilon}{\mu}}\mathbf{E}_0'' \cos \gamma.$$

The minus sign on the right follows from the fact that the projection of $\mathbf{k}'' \times \mathbf{E}_0''$ is antiparallel to $\mathbf{n}$. Hence also

$$\frac{1}{\mu'}(\mathbf{k}' \times \mathbf{E}_0') \times \mathbf{n} = \omega'\sqrt{\frac{\epsilon'}{\mu'}}\mathbf{E}_0' \cos \beta.$$

But $\omega = \omega' = \omega''$ and $\alpha = \gamma$. Thus Eq. (11.50) yields the relation

$$\omega\sqrt{\frac{\epsilon}{\mu}}\mathbf{E}_0 \cos \alpha - \omega\sqrt{\frac{\epsilon}{\mu}}\mathbf{E}_0'' \cos \alpha - \omega\sqrt{\frac{\epsilon'}{\mu'}}\mathbf{E}_0' \cos \beta = 0,$$

i.e. (since — see above — the direction of the field $\mathbf{E}$ does not change),

$$(B) \qquad (E_0 - E_0'')\sqrt{\frac{\epsilon}{\mu}}\cos \alpha - E_0'\sqrt{\frac{\epsilon'}{\mu'}}\cos \beta = 0. \qquad (11.53)$$

This is a new relation.

Finally we consider relation (d). The relation (11.51) yields nothing, since we have

$$(\mathbf{E}_0, \mathbf{E}_0'', \mathbf{E}_0') \cdot \mathbf{n} = 0.$$

We have thus obtained the results (A) and (B).

Eliminating E_0' from (A) and (B), we obtain

$$(E_0 - E_0'')\sqrt{\frac{\epsilon}{\mu}} \cos\alpha - (E_0 + E_0'')\sqrt{\frac{\epsilon'}{\mu'}} \cos\beta = 0,$$

or

$$\frac{E_0''}{E_0} = \left[1 - \sqrt{\frac{\epsilon'\mu}{\epsilon\mu'}} \frac{\cos\beta}{\cos\alpha}\right]\left[1 + \sqrt{\frac{\epsilon'\mu}{\epsilon\mu'}} \frac{\cos\beta}{\cos\alpha}\right]^{-1}.$$

Using the following relation obtained earlier

$$\frac{\sin\alpha}{\sin\beta} = \sqrt{\frac{\epsilon'\mu'}{\epsilon\mu}},$$

we can rewrite the ratio as

$$\frac{E_0''}{E_0} = \left[1 - \frac{\mu}{\mu'} \frac{\sin\alpha}{\sin\beta} \frac{\cos\beta}{\cos\alpha}\right]\left[1 + \frac{\mu}{\mu'} \frac{\sin\alpha}{\sin\beta} \frac{\cos\beta}{\cos\alpha}\right]^{-1}$$

$$= \left[1 - \frac{\mu}{\mu'} \frac{\tan\alpha}{\tan\beta}\right]\left[1 + \frac{\mu}{\mu'} \frac{\tan\alpha}{\tan\beta}\right]^{-1}.$$

Eliminating E_0'' from (A) and (B), we obtain

$$\frac{E_0'}{E_0} = 2\left[1 + \sqrt{\frac{\epsilon'\mu}{\epsilon\mu'}} \frac{\cos\beta}{\cos\alpha}\right]^{-1}$$

$$= 2\left[1 + \frac{\mu}{\mu'} \frac{\tan\alpha}{\tan\beta}\right]^{-1}. \tag{11.54}$$

For the so-called optical frequencies (i.e. those in the visible range of frequencies) we have $\mu(\omega) \simeq \mu'(\omega)$. In this case we obtain the *Fresnel formulae* for light polarised vertically to the plane of incidence:

$$\frac{E_0''}{E_0} = \frac{\sin(\beta - \alpha)}{\sin(\alpha + \beta)}, \qquad \frac{E_0'}{E_0} = \frac{2\sin\beta\cos\alpha}{\sin(\alpha + \beta)}. \tag{11.55}$$

Case 2: E_0 in the plane of incidence

In this case the directions of the fields are as indicated in Fig. 11.7. The relation (a) implies at $z = 0$

$$E_0 \cos \alpha - E_0'' \cos \gamma - E_0' \cos \beta = 0,$$

or, since $\alpha = \gamma$,

(A') $\qquad\qquad (E_0 - E_0'') \cos \alpha - E_0' \cos \beta = 0.$ $\qquad\qquad$ (11.56)

The relation (b) gives nothing, since $\mathbf{k} \times \mathbf{E} \perp \mathbf{n}$. Relation (d) yields

$$\epsilon(E_0 + E_o'') \sin \alpha - \epsilon' E_0' \sin \beta = 0,$$

or since

$$\frac{\sin \alpha}{\sin \beta} = \sqrt{\frac{\epsilon' \mu'}{\epsilon \mu}} = \frac{n'}{n},$$

we have

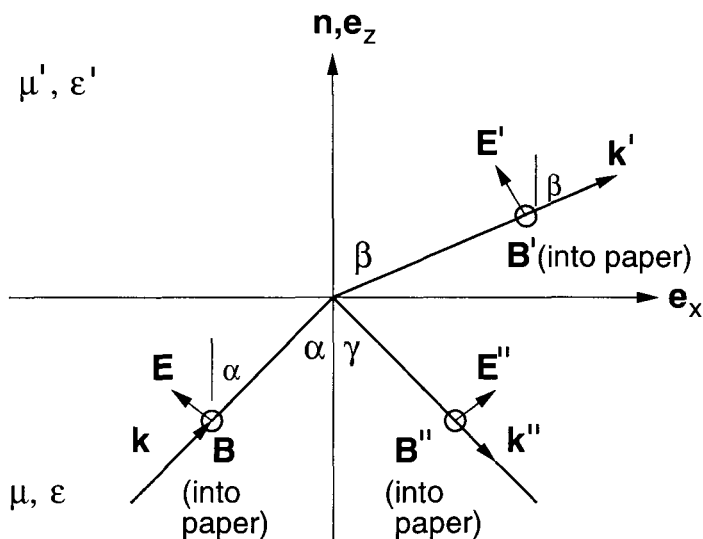$$(E_0 + E_0'') \frac{\epsilon}{\epsilon'} \sqrt{\frac{\epsilon' \mu'}{\epsilon \mu}} - E_0' = 0,$$

Fig. 11.7 Case 2: The directions of the field components at incidence, refraction and reflection.

so that

$$(B') \qquad\qquad (E_0 + E_0'')\sqrt{\frac{\epsilon}{\mu}} - E_0'\sqrt{\frac{\epsilon'}{\mu'}} = 0. \qquad\qquad (11.57)$$

The relation (c) also yields (B').

From both of the relations found we obtain the next pair of *Fresnel formulae*, using again the relation

$$\sqrt{\frac{\mu'}{\mu}} = \sqrt{\frac{\epsilon}{\epsilon'}}\frac{\sin \alpha}{\sin \beta}.$$

Eliminating E_0' we obtain

$$\frac{E_0''}{E_0} = \left[1 - \sqrt{\frac{\epsilon\mu'}{\epsilon'\mu}\frac{\cos \beta}{\cos \alpha}}\right]\left[1 + \sqrt{\frac{\epsilon\mu'}{\epsilon'\mu}\frac{\cos \beta}{\cos \alpha}}\right]^{-1}$$

$$= \left[1 - \frac{\epsilon}{\epsilon'}\frac{\tan \alpha}{\tan \beta}\right]\left[1 + \frac{\epsilon}{\epsilon'}\frac{\tan \alpha}{\tan \beta}\right]^{-1},$$

and eliminating E_0'' we obtain

$$\frac{E_0'}{E_0} = 2\sqrt{\frac{\epsilon\mu'}{\epsilon'\mu}}\left[1 + \sqrt{\frac{\epsilon\mu'}{\epsilon'\mu}\frac{\cos \beta}{\cos \alpha}}\right]^{-1}$$

$$= 2\frac{n'}{n}\frac{\epsilon}{\epsilon'}\left[1 + \frac{\epsilon}{\epsilon'}\frac{\tan \alpha}{\tan \beta}\right]^{-1}.$$

For $\mu = \mu'$ we obtain (as can be verified by replacing the tangent by sine and cosine)

$$\frac{E_0''}{E_0} = \frac{\tan(\alpha - \beta)}{\tan(\alpha + \beta)}, \qquad \frac{E_0'}{E_0} = \frac{2 \sin \beta \cos \alpha}{\sin(\alpha + \beta) \cos(\alpha - \beta)}. \qquad (11.58)$$

For the discussion of *"total reflection"* below we observe here that if we set $\beta = \pi/2$ in Eqs. (11.55) and (11.58) (or alternatively in a preceding formula) we obtain

$$\frac{E_0''}{E_0} = 1, \qquad \frac{E_0'}{E_0} = 2\sqrt{\frac{\epsilon}{\epsilon'}} \neq 0.$$

11.5.3   The Brewster Angle and the Case without Reflection

We observe that when $\alpha + \beta = \pi/2$ in the case of the vector $\mathbf{E}_0$ lying in the plane of incidence, we have (cf. (11.58)) $\tan(\alpha + \beta) = \infty$ and $E_0'' = 0$, but also $\sin \beta = \cos \alpha$. Snell's law then implies

$$\frac{n'}{n} = \frac{\sin \alpha}{\sin \beta} = \tan \alpha_B. \tag{11.59}$$
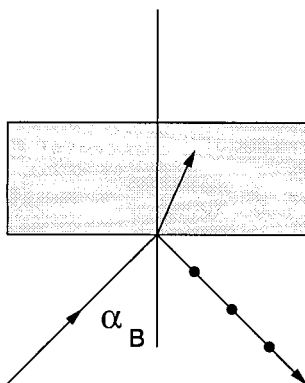
Fig. 11.8 Polarisation at angle of incidence $=\alpha_B$.

The angle α_B is known as *Brewster's angle*. In particular when $\alpha = \pi/4$, also $\beta = \pi/4$, i.e. one has straight passage. On the other hand, if the refractive indices n, n' are equal, we have $\alpha = \beta$ and hence

$$E_0'' = 0, \quad E_0' = E_0,$$

and there is no reflection — as expected. Thus at Brewster's angle with $\mathbf{E}_0$ in the plane of incidence there is no reflected ray whose electric vector oscillates in the plane of incidence. If arbitrary (i.e. unpolarised) light is incident at Brewster's angle, the reflected ray is thus linearly polarised perpendicular to the plane of incidence, as indicated in Fig. 11.8.

11.5.4   The General Case

The general case of an elliptically polarised wave can be dealt with by superposition of the cases 1 and 2, i.e. with

$$\mathbf{E}_1 = \boldsymbol{\epsilon}_1 E_1 e^{i(\mathbf{k} \cdot \mathbf{r} - \omega t)}, \quad \mathbf{E}_2 = \boldsymbol{\epsilon}_2 E_2 e^{i(\mathbf{k} \cdot \mathbf{r} - \omega t)}, \quad \boldsymbol{\epsilon}_1 \cdot \boldsymbol{\epsilon}_2 = 0, \tag{11.60}$$

where $\epsilon_i, i = 1, 2$ are *polarisation vectors* (i.e. unit vectors along the respective direction of polarisation) and

$$\mathbf{B}_j = \sqrt{\mu\epsilon}\frac{\mathbf{k} \times \mathbf{E}_j}{k}, \quad j = 1, 2, \tag{11.61}$$

we have for the most general homogeneous wave in the direction of $\mathbf{k} = k\mathbf{n}_k$:

$$\mathbf{E}(\mathbf{r}, t) = (\epsilon_1 E_1 + \epsilon_2 E_2)e^{i(\mathbf{k}\cdot\mathbf{r} - \omega t)}. \tag{11.62}$$

Here E_1, E_2 are complex numbers, which also permit a phase difference between waves of different polarisation. If E_1 and E_2 have the same phase, E describes a linearly polarised wave with a polarisation vector which makes with ϵ_1 an angle $\theta = \tan^{-1}(E_2/E_1)$ and has modulus $E = \sqrt{E_1^2 + E_2^2}$. If E_1, E_2 have different phases, $\mathbf{E}$ is elliptically polarised. If $|E_1| = |E_2|$ and if their phases differ by $\pi/2$, i.e. if $\mathbf{E}$ can be written

$$\mathbf{E} = E_0(\epsilon_1 \pm i\epsilon_2)e^{i(\mathbf{k}\cdot\mathbf{r} - \omega t)},$$

then $\mathbf{E}$ is circularly polarised.

11.5.5 Total Reflection

It remains to consider the case of total reflection. Let n, n' be the refractive indices of the media on either side of an interface and $n' < n$. Then, see Fig. 11.9, there is a *critical angle* α_0, at which $\beta \to \beta_0 = \pi/2$ (reflected and refracted fields in the same direction x).
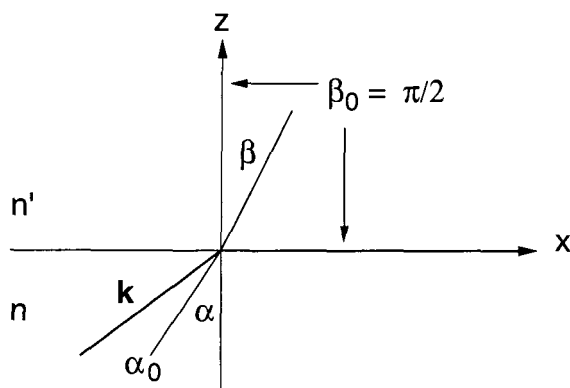

Fig. 11.9 The case of total reflection.

Also (as we deduce from the above formulae) independent of the direction of polarisation of the electric field (cf. Eq. (11.56))

$$E_0 = E_0'', \quad \text{but} \quad E_0' \neq 0.$$

This case is described as *total reflection.* Since

$$\frac{\sin \alpha'}{\sin \beta} = \frac{n'}{n},$$

it follows that $\sin \alpha_0 = n'/n$.

We can consider this case also from another point of view, by demonstrating that there is no transport of energy into the other medium. For real angles β the maximal value of $\sin \beta$ is, of course, 1 (for $\beta = \pi/2$). The maximal value of α, which satisfies the condition

$$\frac{\sin \alpha}{\sin \beta} = \frac{n'}{n} < 1$$

for real values of β, is therefore given by

$$\sin \alpha_0 = \frac{n'}{n}$$

($\sin \beta_0 = 1$). For $\alpha > \alpha_0$, hence for $n'/n = \sin \alpha_0$, we write the "angle of refraction" β which is then complex as

$$\beta = \frac{\pi}{2} + i\delta.$$

Then $\sin \beta = \cosh \delta$ and

$$\delta = \cosh^{-1}(\sin \beta) = \cosh^{-1}\left(\frac{n}{n'}\sin \alpha\right).$$

Moreover

$$\cos \beta = \sqrt{1 - \sin^2 \beta} = +i\sqrt{\cosh^2 \delta - 1} = i \sinh \delta$$

and

$$e^{i\mathbf{k}'\cdot\mathbf{r}} = e^{ik'z\cos\beta} = e^{-k'z\sinh\delta} \to 0 \text{ for } z \to \infty.$$

It follows that the field of the refracted radiation, i.e. $\mathbf{E}'$, goes to 0 for $z \to \infty$, but for z finite, as we can show, $\mathbf{n} \cdot (\mathbf{E}' \times \mathbf{H}') = 0$. This means, there is no transport of energy into the other medium. We see this as follows. Recalling the transversality condition and Eq. (11.61), we have

$$\mathbf{S} \cdot \mathbf{n} = \Re[\mathbf{n} \cdot \underbrace{(\mathbf{E}' \times \mathbf{H}')}_{\sqrt{\frac{\epsilon}{\mu}}\frac{\mathbf{E}'^2}{k}\mathbf{k}}]$$

$$= \Re\left\{\sqrt{\frac{\epsilon}{\mu}}|E_0'|^2\frac{\mathbf{n} \cdot \mathbf{k}'}{k'}\right\} = 0,$$

the latter because of "i" in

$$\mathbf{n} \cdot \mathbf{k}' = k' \cos \beta = ik' \sinh \delta.$$

We have total reflection. This total reflection is also evident from the fact that $E_0 = E_0''$, i.e. $\mathbf{E}_0^2 = \mathbf{E}_0''^2$.

We return once again to the general case:

$$\cos \beta = \pm(1 - \sin^2 \beta)^{1/2}$$

$$= \pm\sqrt{1 - \left(\frac{n}{n'}\right)^2 \sin^2 \alpha}$$

$$= \pm\sqrt{1 - \left(\frac{\sin \alpha}{\sin \alpha_0}\right)^2}.$$

For $\alpha < \alpha_0$ we have $\cos \beta$ real and hence β. For $\alpha > \alpha_0$, however, $\cos \beta$ is purely imaginary, i.e. β complex:

$$\cos \beta = \pm i \left[\left(\frac{\sin \alpha}{\sin \alpha_0} \right)^2 - 1 \right]^{\frac{1}{2}}.$$

The phase of the refracted wave is therefore ($\mathbf{k}'$ in the (x, z)-plane)

$$
\begin{aligned}
e^{i\mathbf{k}' \cdot \mathbf{r}} &= \exp\{i(k_x' x + k_z' z)\} \\
&= \exp\{i(k' x \sin \beta + k' z \cos \beta)\} \\
&= \exp\left\{ ik' x \frac{n}{n'} \sin \alpha \right\} \exp\left\{ \mp k' z \sqrt{\left(\frac{\sin \alpha}{\sin \alpha_0}\right)^2 - 1} \right\} \\
&\propto \exp(-z/z_0).
\end{aligned}
\tag{11.63}
$$

Thus the wave is damped in the direction of z if we make the physically plausible choice of selecting the exponentially decreasing function. In general

$$z_0 \equiv \frac{1}{k'\sqrt{\left(\frac{\sin \alpha}{\sin \alpha_0}\right)^2 - 1}}$$

is of the order of several wavelengths, i.e. the refractive wave is damped so strongly, that in general there is no refractive wave, unless the thickness of the weaker medium is less than one wavelength. In such a case light can tunnel through the film, and one has to take a combination of both exponential functions. We see that when $\alpha = \alpha_0$, there is no damping. One has the scattering of light at the interface, which can also be observed, e.g. in an experiment with a sphere immersed in a medium, the sphere containing a

lamp and permitting light to escape through some windows, as indicated in Fig. 11.10.
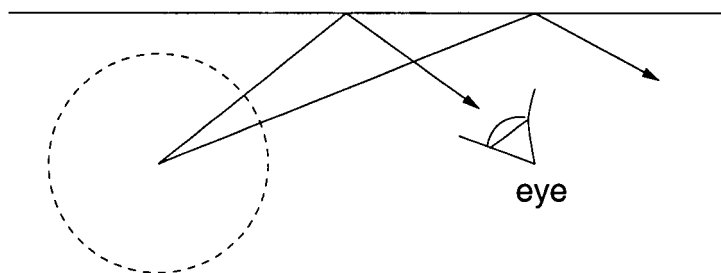
Fig. 11.10　Scattering of light at an interface.

11.6　Useful Formulation of the Fresnel Formulae

We set

$$N = \sqrt{\frac{\epsilon'\mu}{\epsilon\mu'}}, \quad R = \frac{E_0''}{E_0}, \quad T = \frac{E_0'}{E_0}. \tag{11.64}$$

Then in both cases of above we can rewrite the results as follows:

Case (1): $E_0 \perp$ plane of incidence

$$(A) \quad 1 + R - T = 0, \qquad (B) \quad (1 - R)\cos\alpha - NT\cos\beta = 0,$$

so that $(1 - R)\cos\alpha = N(1 + R)\cos\beta$. Hence

$$R = \frac{\cos\alpha - N\cos\beta}{\cos\alpha + N\cos\beta}, \quad T = \frac{2\cos\alpha}{\cos\alpha + N\cos\beta}. \tag{11.65}$$

The expressions R and T are respectively called *reflection coefficient* and *transmission coefficient*.

Case (2): E_0 in the plane of incidence

$$(A') \quad (1 - R)\cos\alpha = T\cos\beta, \qquad (B') \quad (1 + R) = NT,$$

so that $(1 - R)\cos\alpha = \frac{(1+R)}{N}\cos\beta$, and

$$R = \frac{N\cos\alpha - \cos\beta}{N\cos\alpha + \cos\beta}, \quad T = \frac{2\cos\alpha}{N\cos\alpha + \cos\beta}. \tag{11.66}$$

Here we have defined R and T as ratios of amplitudes; this is not the most common formulation. Some authors such as Lim [2] define R and T as ratios of the energy of the reflected ray, and that of the transmitted ray, to the energy of the incident ray. In the book of Jackson [3] (Chapter 7) the coefficients are not introduced. Since, as we emphasised, $\mathbf{E}_0, \mathbf{E}_0', \mathbf{E}_0''$ are amplitudes, which can still be complex, also R and T can be complex. We introduce the coefficients R and T here in analogy to their counterparts in quantum mechanics or quantum mechanical scattering problems, which seems to suggest itself.

Example 11.2 Light falling on glass
(a) The refractive index of glass is $n' = 1.5$. Calculate what fraction of light intensity vertically incident on a surface of clean glass is reflected.
(b) What are the phases of the vectors $\mathbf{E}, \mathbf{H}$ at the interface between air (n) and glass $(n', n' > n)$ in the case of vertical incidence?

Solution:
(a) We have (with $\mu = \mu', N = n'/n$) for vertical incidence and irrespective of whether $\mathbf{E}_0$ lies in the plane of incidence or is perpendicular to this

$$\left|\frac{E_0''}{E_0}\right|^2 = |R|^2 = \left|\frac{n-n'}{n+n'}\right|^2 = \left(\frac{\frac{1}{2}}{\frac{5}{2}}\right)^2 = \left(\frac{1}{5}\right)^2 = \frac{1}{25} = 4\%.$$

(b) For $\mathbf{E}_0 \perp$ plane of incidence: $\alpha = \beta = 0$. Hence

$$R = \frac{1-N}{1+N} = \frac{n-n'}{n+n'} < 0, \quad R = \frac{E_0''}{E_0}.$$

It follows that

$$E_0'' = e^{i\pi}|R|E_0,$$

i.e. $\mathbf{E}_0''$ points in a direction opposite to that of $\mathbf{E}_0$. For the direction of transport of energy (given by $(\mathbf{E} \times \mathbf{H})$) of the reflected wave to be opposite to that of the incident wave, the phase of $\mathbf{H}$ must remain unchanged.

Example 11.3:[§] Calculation of the angle of refraction
Polysterene has a relative permittivity of 2.7. An electromagnetic wave in air is incident at an angle of 30° to the normal on a surface of polysterene. Calculate the angle of refraction and repeat the calculation for the reverse case.

Solution: For the passage from air to polysterene we set:

$$\epsilon_1 := \epsilon_0, \quad \epsilon_2 := 2.7\epsilon_0, \quad \mu_1 := \mu_0, \quad \mu_2 := \mu_0.$$

From $\sin\theta_2 = \frac{n_1}{n_2}\sin\theta_1$ it follows that

$$\sin\theta_2 = \sqrt{\frac{1}{2.7}}(0.5) = 0.304, \quad \theta_2 = 17.7°.$$

From polysterene to air:

$$\epsilon_1 := 2.7\epsilon_0, \quad \epsilon_2 := \epsilon_0, \quad \mu_1 = \mu_0, \quad \mu_2 = \mu_0,$$

and

$$\sin\theta_2 = \sqrt{2.7}(0.5) = 0.822, \quad \theta_2 = 55,2°.$$

[§]See also *The Electromagnetic Problem Solver* [21], p. 649.

This page intentionally left blank

Chapter 12

Metals

12.1 Introductory Remarks

Our considerations in the preceding chapter were restricted to media with conductivity $\sigma = 0$. We now proceed to corresponding considerations for metals.[*]

12.2 Reflection and Absorption of Plane Waves by Metals

In the case now under consideration we have

$$\nabla \times \mathbf{H} = \frac{\partial \mathbf{D}}{\partial t} + \mathbf{j}. \tag{12.1}$$

With

$$\mathbf{D} = \epsilon \mathbf{E}, \quad \mathbf{j} = \sigma \mathbf{E}, \quad \mathbf{E} \propto e^{i\omega t} \tag{12.2}$$

we have

$$\nabla \times \mathbf{H} = i\omega \epsilon \mathbf{E} + \sigma \mathbf{E} = i\omega \mathbf{E}\left(\epsilon + \frac{\sigma}{i\omega}\right), \tag{12.3}$$

i.e. in the case of conducting media we have the generalised dielectric constant

$$\eta = \epsilon + \frac{\sigma}{i\omega}. \tag{12.4}$$

Thus, although now $\mathbf{j} \neq 0, \int \mathbf{j} \cdot d\mathbf{F} \neq 0$, we can again proceed as before except that ϵ has to be replaced by η, because we saw earlier, that for finite

---

[*]Although now old, the most important monograph on this subject was for a long time that of A. H. Wilson, *The Theory of Metals*, 2nd ed. (Cambridge University Press, 1953).

conductivity σ the "linear current density" $K = 0$. Hence we have the case with no surface currents.

We consider now *reflection in the case of vertical incidence* of the electromagnetic wave, as indicated in Fig. 12.1.
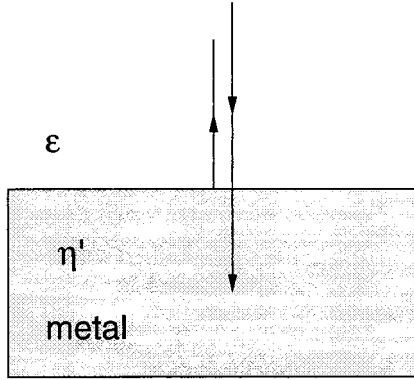
Fig. 12.1 Vertical incidence on a metal.

In this case we have incident angle $\alpha = 0$ and (from $\sin \alpha = (\epsilon'\mu'/\epsilon\mu)^{1/2} \sin\beta$)

$$\beta = 0.$$

Hence the reflected intensity is in both cases (1) and (2) considered in Sec. 11.5.2:

$$|R|^2 = \left|\frac{N-1}{N+1}\right|^2, \tag{12.5}$$

where with $\epsilon' \to \eta'$

$$N = \sqrt{\frac{\eta'\mu}{\epsilon\mu'}}. \tag{12.6}$$

For *typical metals* $|N| \gg 1$, so that we can approximate $|R|^2$ as follows:

$$
\begin{aligned}
|R|^2 &= \left|\frac{1-1/N}{1+1/N}\right|^2 \simeq \left|1 - \frac{2}{N}\right|^2 \\
&= \left|1 - 2\left\{\Re\left(\frac{1}{N}\right) + i\Im\left(\frac{1}{N}\right)\right\}\right|^2 \\
&= \left\{1 - 2\Re\left(\frac{1}{N}\right)\right\}^2 + \left\{2\Im\left(\frac{1}{N}\right)\right\}^2 \\
&\simeq 1 - 4\Re\left(\frac{1}{N}\right) + O\left(\frac{1}{N^2}\right).
\end{aligned}
\tag{12.7}
$$

The *absorptivity* A is a quantity defined by

$$A \equiv 1 - |R|^2, \tag{12.8}$$

so that for $|N| \gg 1$

$$A \simeq 4\Re\left(\frac{1}{N}\right). \tag{12.9}$$

But now — note that the source of ϵ' is the displacement current, whereas the source of $\sigma'/i\omega$ is the electron current —

$$N = \sqrt{\frac{\mu}{\epsilon\mu'}}\sqrt{\epsilon' + \frac{\sigma'}{i\omega}} \tag{12.10}$$

$$= \sqrt{\frac{\mu}{\epsilon\mu'}}\sqrt{\frac{\sigma'}{i\omega}}\sqrt{1 + \frac{i\omega\epsilon'}{\sigma'}}. \tag{12.11}$$

Hence with

$$i^{1/2} = e^{i\pi/4} = \frac{1}{\sqrt{2}}(1 + i)$$

we have

$$4\frac{1}{N} = 4\sqrt{\frac{\omega\epsilon\mu'}{\sigma'\mu}}\frac{i^{1/2}}{\sqrt{1 + i\omega\epsilon'/\sigma'}}$$

$$= 4\sqrt{\frac{\omega}{2\sigma'}\frac{\epsilon\mu'}{\mu}}\frac{(1 + i)}{\sqrt{1 + i\omega\epsilon'/\sigma'}}. \tag{12.12}$$

For *low frequencies* we have

$$\sigma = \text{const.} = \text{real}, \quad \text{i.e.} \quad \sigma(\omega) \simeq \sigma(0), \tag{12.13}$$

so that

$$A \simeq \sqrt{\frac{8\omega}{\sigma'}\frac{\epsilon\mu'}{\mu}} \propto \sqrt{\omega}\sqrt{\rho'}, \tag{12.14}$$

where $\rho' = 1/\sigma'$ is the resistivity. This is the so-called *Hagen–Rubens approximation* of the absorptivity at low frequencies (experimentally investigated by Hagen and Rubens).[†]

[†]E. Hagen and H. Rubens, *Ann. Physik* **11** (1903) 873. A discussion of the Hagen–Rubens formula can be found e.g. in J. A. Stratton [15], p. 508.

12.3   The Theory of Drude

At *low frequencies* ($\omega \to 0$) the equation of motion of a free electron subjected to an applied electric field E can be taken to be approximately given by

$$m\ddot{x} + b\dot{x} = eE. \qquad (12.15)$$

Here m is the mass of the electron, and $b = m/\tau$, where τ has the dimension of time and is called *relaxation time*. Here τ has the physical meaning of an average scattering or collision time, i.e. the average time interval between successive collisions of an electron with atoms of the metal lattice. In the case of a bound electron, one would add a term $\omega_0^2 x$ to $b\dot{x}$ representing a binding of the electron to an atom like that of an oscillator potential (this is a purely classical model at this stage!).

At *high frequencies* the field $\mathbf{E}$ alternates its direction so rapidly, that the electron has only a small chance to get sufficiently far away in order to be able to collide with an ion, as indicated in Fig. 12.2. This means, in this case the equation of motion can be approximated by

$$m\ddot{x} = eE_0 e^{i\omega t} = eE. \qquad (12.16)$$

In the *intermediate domain* the equation of motion is therefore approximately

$$m\ddot{x} + \frac{m}{\tau}\dot{x} = eE_0 e^{i\omega t} = eE \qquad (12.17)$$

with solution[‡]

$$\dot{x} = \frac{eE}{m(i\omega + \frac{1}{\tau})} \qquad (12.18)$$
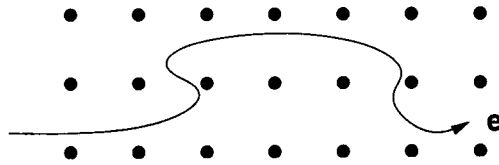
Fig. 12.2 A conduction electron in a metal lattice.

But now[§] *current density* $\mathbf{j}$ = charge of one electron × number of electrons per unit volume × velocity = $ne\dot{x}$, i.e.

$$\mathbf{j} = \frac{ne^2 \mathbf{E}}{m(i\omega + \frac{1}{\tau})} = \sigma \mathbf{E}, \qquad (12.19)$$

[‡]$\ddot{x} = i\omega\dot{x}, \;\; \therefore m\ddot{x} + \frac{m}{\tau}\dot{x} = m(i\omega + \frac{1}{\tau})\dot{x} = eE.$

[§]Initially the free electrons wander around arbitrarily and in any direction. After application of the field $\mathbf{E}$, however, they are forced to move antiparallel to the direction of $\mathbf{E}$.

so that

$$\sigma = \frac{ne^2\tau}{m(1 + i\omega\tau)}. \tag{12.20}$$

We see therefore: The expression for σ is in general complex and hence the Hagen–Rubens formula an unsatisfactory approximation.

We can write σ also

$$\sigma = \frac{\sigma_{dc}}{1 + i\omega\tau}, \quad \sigma_{dc} \equiv \frac{ne^2\tau}{m} \text{ (ohm meter)}^{-1}. \tag{12.21}$$

Here σ_{dc} is called the *direct current conductivity* and τ could be taken as $\bar{l}/\bar{v}$, the ratio of an average length of the mean free path, $\bar{l}$, divided by an average velocity, $\bar{v}$.

We return now to the calculation of the absorptivity and replace σ' in the first of expressions (12.12) by

$$\sigma' = \frac{\sigma'_{dc}}{1 + i\omega\tau},$$

so that from the middle expression in Eq. (12.12),

$$A = \left(\frac{16\omega\epsilon\mu'}{\sigma'_{dc}\mu}\right)^{1/2} \Re\left\{ \sqrt{i(1 + i\omega\tau)} \frac{1}{\sqrt{1 + \frac{i\epsilon'\omega(1+i\omega\tau)}{\sigma'_{dc}}}} \right\}.$$

But

$$\sqrt{i(1 + i\omega\tau)} = \sqrt{\frac{\sqrt{1 + \omega^2\tau^2} - \omega\tau}{2}} + i\sqrt{\frac{\sqrt{1 + \omega^2\tau^2} + \omega\tau}{2}},$$

so that for $\epsilon'\omega/\sigma'_{dc} \ll 1$,

$$A \simeq \sqrt{\frac{8\omega\epsilon\mu'}{\sigma'_{dc}\mu}} \sqrt{\sqrt{1 + \omega^2\tau^2} - \omega\tau}$$

$$= \text{Hagen–Rubens expression} \times \sqrt{\sqrt{1 + \omega^2\tau^2} - \omega\tau}, \tag{12.22}$$

and hence for $\omega\tau \gg 1$

$$A = \sqrt{\frac{8\omega\epsilon\mu'}{\sigma'_{dc}\mu}} \sqrt{\omega\tau\left(1 + \frac{1}{2\omega^2\tau^2} + \cdots\right) - \omega\tau}$$

$$\simeq \sqrt{\frac{8\omega\epsilon\mu'}{\sigma'_{dc}\mu} \frac{\omega\tau}{2\omega^2\tau^2}}$$

$$= \sqrt{\frac{4\epsilon\mu'}{\mu} \frac{1}{\tau\sigma'_{dc}}}, \quad (\omega\tau \gg 1). \tag{12.23}$$

The result (12.22) is known as *Drude–Kronig formula*. We infer from these results, that at low frequencies, i.e. $\omega\tau \ll 1$,

$$A \propto \sqrt{\omega} \qquad (\text{ in the extreme infrared region}),$$

whereas at high frequencies (infrared region as compared to the extreme infrared region) A behaves like (12.23), i.e. is approximately constant. This behaviour of the absorptivity is depicted in Fig. 12.3.
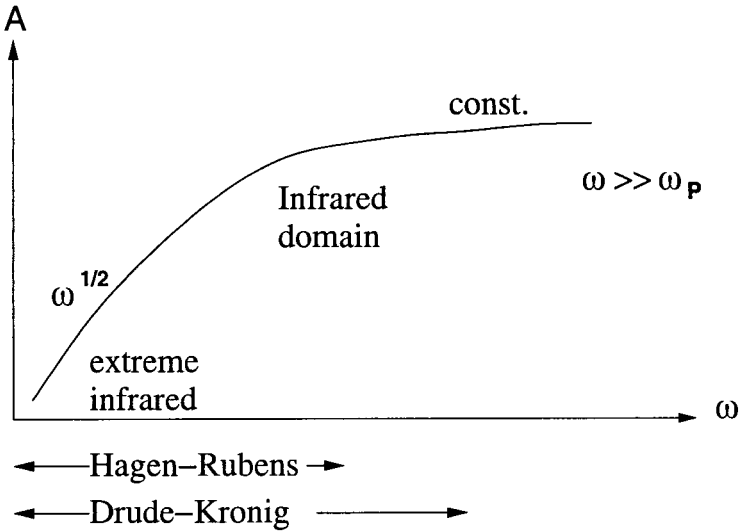
Fig. 12.3 Absorptivity A as a function of frequency ω.

The *ultraviolet region* is characterised by high frequencies. In this case the displacement current dominates over the electron current as we can see from Eq. (12.11). This means, in this case the metal behaves like a dielectric with reduced dielectric constant (see below). We saw that the effective dielectric constant of a metal is given by Eq. (12.4), and that at high frequencies ω

$$\sigma = \frac{ne^2\tau}{m(1 + i\omega\tau)} \simeq \frac{ne^2}{im\omega}, \qquad (12.24)$$

so that

$$\eta = \underbrace{\epsilon - \frac{ne^2}{m\omega^2}}_{\text{reduced dielectric const.}} \qquad (12.25)$$

In general the above *Drude–Kronig formula* cannot be correct. We know that due to the skin effect the electric field **E** varies considerably with penetration depth into the metal. This implies that the relation

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

must be modified, if the length of the mean free path of the electrons is $\sim v/\omega$ and larger or of the order of the skin depth z_0, or if the relaxation time is very large. The above current density formula will then have to be replaced by a relation of the form

$$\mathbf{j}(\mathbf{r}) = \int d\mathbf{r}' g(\mathbf{r}, \mathbf{r}') E(\mathbf{r}')$$

in order to take into account the rapid change of the field over the region of the mean path of an electron. In the course of motion of the electrons towards or away from the surface, electromagnetic energy is directly transferred to the conduction electrons. If one averages over all possible directions of these electrons, and if one takes into account, that the electrons are fermions — their maximum velocity v_0 at absolute temperature $T = 0$ being that of the Fermi boundary — one obtains a correction contribution. The velocity v_0 is determined as follows. Number of states = number of particles (according to the Pauli principle) $= N = 2V\kappa^3/(2\pi)^3, \mathbf{p} = \hbar\kappa, \hbar = h/2\pi$ (the factor '2' takes into account the two possible spin directions of the electron as a fermion). This means

$$N = 2\frac{V}{(2\pi)^3}\left(\frac{4}{3}\pi\frac{p^3 8\pi^3}{h^3}\right), \tag{12.26}$$

or, if $p = mv_0, n = N/V$,

$$v_0 = \frac{h}{m}\left(\frac{3n}{8\pi}\right)^{1/3}$$

(as for the Fermi distribution at the absolute zero point of temperature). Taking these effects into account a further modification of the formula for the absorptivity of electromagnetic radiation in a metal results and is known as *Dingle–Holstein formula*.[¶] This formula can be written as

$$A_{\text{DH}} = \text{Drude–Kronig expression} + \frac{3}{4}\frac{v_0}{c}, \tag{12.27}$$

where c is the velocity of light in vacuum. The additional term dominates at low temperatures. At higher temperatures and for good conductors like Cu and Ag both contributions of the formula are of comparable magnitude.

[¶] R. B. Dingle [18], R. B. Dingle, *Physica* **19** (1953) 311 (see in particular p. 342) and T. Holstein, *Phys. Rev.* **82** (1951) 1427.

Example 12.1: Absorptivity of silver

Calculate the absorptivity of silver with direct current conductivity of 3×10^7 mhos/meter at a radiation frequency of $\nu = 10^{10}$ second^{-1} (microwave region, $\lambda \simeq 3$ cm).

Solution: We use the Hagen–Rubens relation

$$A \simeq \sqrt{\frac{8\omega\epsilon\mu'}{\sigma'\mu}},$$

where we insert

$$\mu = \mu', \quad \omega = 2\pi \times 10^{10} \quad \text{per second}, \quad \epsilon = \epsilon_0 = 8.854 \times 10^{-12}, \quad \sigma' = 3 \times 10^7.$$

The result is $A = 3.9 \times 10^{-4}$.

Example 12.2: Skin effect and Drude theory

Electromagnetic radiation with

$$\mathbf{E} = E(z)e^{i\omega t}\mathbf{e}_x, \quad \mathbf{H} = H(z)e^{i\omega t}\mathbf{e}_y$$

($\mathbf{e}_x, \mathbf{e}_y$ being unit vectors) falls vertically on a plate of metal. Show that the electric field in the plate at a distance z from the surface and at low frequencies ω is given by the following expression:

$$E(z) = E(0)\exp\left\{-\frac{(1+i)z}{\delta\sqrt{1+i\omega\tau}}\right\},$$

where δ, the penetration depth for small ω, is given by

$$\delta = \sqrt{\frac{2}{\omega\mu\sigma_{\mathrm{dc}}}}.$$

Here σ_{dc} is the direct current conductivity and τ the relaxation time. Finally evaluate $E(0)$ with the help of Fresnel's formulae.

Solution: We start from the equation

$$\boldsymbol{\nabla} \times \mathbf{H} = \mathbf{j} + \frac{\partial \mathbf{D}}{\partial t},$$

where in the present case

$$\boldsymbol{\nabla} \times \mathbf{H} = \mathbf{e}_x\left(-\frac{\partial H_y}{\partial z}\right) + \mathbf{e}_z\left(\underbrace{\frac{\partial H_y}{\partial x}}_{0}\right),$$

since $H_y = H(z)$. Therefore,

$$-\frac{\partial H_y}{\partial z} = j_x + \epsilon\frac{\partial E_x}{\partial t},$$

or

$$-\frac{dH(z)}{dz} = j_x + i\omega\epsilon E(z). \tag{12.28}$$

Similarly with

$$\boldsymbol{\nabla} \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} = -i\omega\mu\mathbf{H} = \mathbf{e}_y\frac{\partial E_x}{\partial z} + \mathbf{e}_z\left(\underbrace{-\frac{\partial E_x}{\partial y}}_{0}\right)$$

we obtain

$$\frac{dE(z)}{dz} = -i\omega\mu H. \tag{12.29}$$

From (12.28) and (12.29) we eliminate H and obtain:

$$\frac{d^2 E(z)}{dz^2} = -i\omega\mu\frac{dH}{dz} = i\omega\mu(j_x + i\omega\epsilon E),$$

i.e.

$$\frac{d^2 E(z)}{dz^2} + \omega^2\epsilon\mu E(z) = i\omega\mu j_x. \tag{12.30}$$

Thus we have *one* equation in $E(z)$ and j_x. We obtain a second relation from $\mathbf{j} = \sigma\mathbf{E}$, where $\mathbf{j} = ne\dot{\mathbf{r}}$ and with the simple Drude theory

$$m\ddot{\mathbf{r}} + \frac{m}{\tau}\dot{\mathbf{r}} = e\mathbf{E} = e\mathbf{E}_0 e^{i\omega t},$$

so that

$$\dot{\mathbf{r}} = \frac{e\mathbf{E}\tau}{m(1 + i\omega\tau)}, \quad \mathbf{j} = \frac{ne^2\tau\mathbf{E}}{m(1 + i\omega\tau)}$$

and hence

$$\mathbf{j} = \frac{ne^2\bar{l}\tau\mathbf{E}}{m\bar{l}(1 + i\omega\tau)},$$

where $\bar{l}$ is the length of a mean free path and $\bar{v} = \bar{l}/\tau$ a mean velocity. It follows that

$$\sigma = \frac{ne^2\bar{l}}{m\bar{v}(1 + i\omega\tau)}.$$

With $\sigma_{\text{dc}} \equiv ne^2\bar{l}/m\bar{v}$ we can write $\mathbf{j} = \sigma_{\text{dc}}\mathbf{E}/(1+i\omega\tau)$. With this we obtain as our *second* equation

$$j_x = \frac{\sigma_{\text{dc}}E(z)}{1 + i\omega\tau}. \tag{12.31}$$

Equations (12.30) and (12.31) now imply:

$$\frac{d^2 E(z)}{dz^2} + \left(\epsilon\mu\omega^2 - \frac{i\mu\omega\sigma_{\text{dc}}}{1 + i\omega\tau}\right)E(z) = 0. \tag{12.32}$$

We neglect the displacement current contribution $\epsilon\mu\omega^2$ and consider *low frequencies*. Moreover, $i = (\sqrt{i})^2 = [(1 + i)/\sqrt{2}]^2$. Hence the physically sensible, i.e. decreasing, solution is

$$\mathbf{E} = E(z)e^{i\omega t}\mathbf{e}_x, \quad \text{with} \quad E(z) = E(0)\exp\left\{-\frac{(1+i)z}{\delta\sqrt{1 + i\omega\tau}}\right\},$$

where

$$\delta = \sqrt{\frac{2}{\mu\omega\sigma_{\text{dc}}}}$$

(this is the penetration depth of the refracted wave at low frequencies in agreement with Eq. (9.17)). We have here the case of the vector $\mathbf{E}$ lying in the plane of incidence. The appropriate Fresnel formulae of Sec. 11.6 are

$$(1 - R)\cos\alpha = T\cos\beta, \quad (1 + R) = NT, \quad \text{where} \quad R = \frac{E_0''}{E_0}, T = \frac{E_0'}{E_0}, \quad N = \sqrt{\frac{\epsilon'\mu}{\epsilon\mu'}},$$

where E_0, E_0'', E_0' are the amplitudes of the incident, reflected and refracted waves. We are interested in T with $E_0 = 1$, i.e. $T = E_0' = E(0)$. Vertical incidence ($\alpha = 0 \to \beta = 0$) implies $T = 1 - R$. R follows from the above relations as $R = (N-1)/(N+1)$, so that $T = 1 - R = 2/(N+1) \simeq 2/N$ for $|N| \gg 1$. Neglecting the contribution of the displacement current, we have

$$N \simeq \sqrt{\frac{\mu}{\epsilon \mu'}} \sqrt{\frac{\sigma'}{i\omega}},$$

or (with $\epsilon = \epsilon_0$, $\mu' = \mu$, $\sigma' \simeq \sigma_{\mathrm{dc}}$)

$$\frac{1}{N} = \sqrt{\frac{\omega \epsilon_0}{\sigma'}} i^{1/2},$$

so that

$$E(0) = T \simeq \frac{2}{N} \simeq 2\sqrt{\frac{\omega \epsilon_0}{\sigma_{\mathrm{dc}}}} i^{1/2}.$$

Chapter 13

Propagation of Radio Waves in the Ionosphere

13.1 Introductory Remarks

The phenomenon we consider here,[*] takes place several hundred kilometers above the surface of the Earth. The gaseous atoms in the upper atmosphere are ionised by extraterrestrial radiation (e.g. by cosmic rays, ultraviolet radiation from the sun etc.). The ionosphere consists of several layers of electrons and ions with first increasing and then decreasing densities. Consider first only one layer. We let the dielectric constant near the surface of the Earth be ϵ_0 and in the ionised region ($\omega\tau \gg 1$, see Chapter 9)

$$\eta = \epsilon_0 - \frac{ne^2}{m\omega^2} \equiv \epsilon_0 - \frac{\omega_P^2}{\omega^2}\epsilon_0. \tag{13.1}$$

Here

n is the number of electrons per unit volume,

m the mass of an electron,

ω_P the plasma frequency, $\omega_P^2 = ne^2/m\epsilon_0$.

The effect of the positively charged ions (i.e. their contribution to the conductivity and hence to the effective dielectric constant) can be neglected, since the mass of the ions is approximately 2000 times larger than that of an electron. We also ignore at the beginning (see later) the effect of the terrestrial magnetic field which in general cannot be neglected (see Jackson [3], pp. 292–294).

---

[*]A brief description — different from that here — can be found in Stratton [17]. A more complete account is given in H.R. Mimno, *Rev. Mod. Phys.* **9** (1937) 1–43 and in *Ergebnisse der exakten Wissenschaften*, Vol. **17** (Springer, 1938).

For the ray of a radio wave to return from the ionosphere back to Earth, the density of electrons cannot increase indefinitely, since otherwise the ray would be bent more and more towards the vertical and hence would eventually escape. Thus beyond a certain maximum value, the refractive index must decrease again, and this means the density of electrons, enabling the ray to be bent away from the vertical. In any case the wave has a reflected part and a refracted part with coefficients R and T in the case of $\mathbf{E} \perp$ to the plane of incidence given by (cf. Chapter 11)

$$R = \frac{\cos\alpha - N\cos\beta}{\cos\alpha + N\cos\beta}, \quad T = \frac{2\cos\alpha}{\cos\alpha + N\cos\beta}, \qquad (13.2)$$

where α and β are incident and refractive angles as before, and N is (cf. Eq. (11.64)) now with generalised dielectric constants

$$N = \sqrt{\frac{\eta'\mu}{\eta\mu'}}.$$
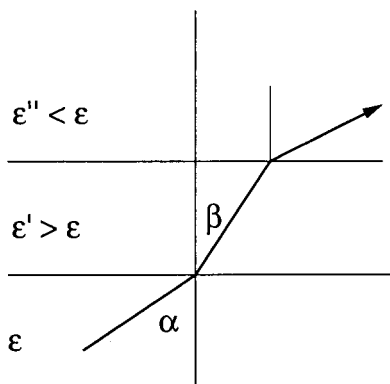
Fig. 13.1 Refraction of a ray.   Fig. 13.2 Reflection after refraction.
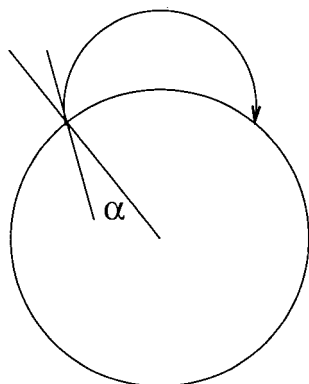
13.2 Condition for Return of Waves

Along its ascent into the ionosphere the ray is refracted, as also on its way back. However, the highest point of the ray's path must be a point of reflection. Naturally this situation is somewhat idealised. For example some waves wander even directly from the emitter to the receiver. There are also waves which circle around the Earth (with reflection from the ionosphere), and thus give rise to radio echos, which have been observed. Since the electron density

increases with increasing height above the Earth only very slowly, one has initially almost exclusively refraction. Since

$$N = \sqrt{\frac{\eta'\mu}{\eta\mu'}} \sim 1,$$

this means that for a ray starting vertically upwards $R = 0$ (since with Snell's law also $\beta = 0$). After attaining its maximum value, the electron density decreases rapidly with further increase in height. In the plasma the dielectric constant is given by Eq. (13.1). We recall that we had defined the generalised refractive index as

$$p = c\sqrt{\eta\mu} \equiv p_1 + ip_2, \qquad (13.3)$$

and that

$$k = \frac{\omega}{c}p, \qquad e^{i(kx-\omega t)} = e^{i\frac{\omega}{c}(px-ct)}. \qquad (13.4)$$

We observe that for $\eta < 0$, i.e. p complex, the wave is strongly damped. For large masses m (of the ions) $\eta \simeq \epsilon_0$. For $\eta < 0$ the wave transmitted beyond the reflection point is strongly damped (i.e. practically no transmitted wave), and one has effectively only reflection. For this to occur the density must at least be large enough so that $\eta = 0$, i.e.

$$\frac{n}{\epsilon_0} \gtrsim \frac{m\omega^2}{e^2}. \qquad (13.5)$$

We now apply Snell's law to different layers of the atmosphere/ionosphere. For the case of three layers as illustrated in Fig. 13.3 we have (beginning with an angle of incidence whose angle of refraction is equal to the angle of incidence at the next layer and so on)

$$\sqrt{\epsilon\mu}\sin\alpha = \sqrt{\epsilon'\mu'}\sin\beta = \sqrt{\epsilon''\mu''}\sin\gamma = \cdots . \qquad (13.6)$$

In the present case the electron density depends on the height y above the ground, i.e. we have to write

$$n = n(y).$$

Let θ_b be the angle of incidence at ground level and θ the angle of refraction at height y. Then we can write

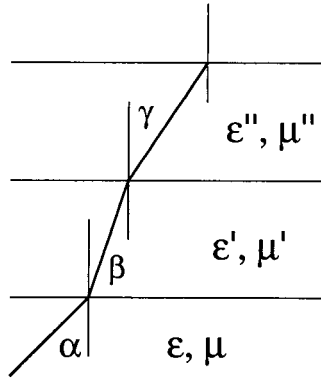$$\left[\epsilon - \frac{n(y)e^2}{m\omega^2}\right]^{1/2}\sin\theta = \sqrt{\epsilon}\sin\theta_b, \qquad (13.7)$$

Fig. 13.3 Refraction of the ray on its way upwards.

and for $\epsilon = \epsilon_0$

$$\left[\epsilon_0 - \frac{n(y)e^2}{m\omega^2}\right]^{1/2} \sin\theta = \sqrt{\epsilon_0}\sin\theta_b, \tag{13.8}$$

so that

$$\left(1 - \frac{n(y)e^2}{\epsilon_0 m\omega^2}\right)\sin^2\theta = \sin^2\theta_b = 1 - \cos^2\theta_b,$$

or

$$\cos^2\theta_b - \frac{n(y)e^2}{\epsilon_0 m\omega^2}(1 - \cos^2\theta) = 1 - \sin^2\theta = \cos^2\theta$$

or

$$\cos^2\theta_b - \frac{n(y)e^2}{\epsilon_0 m\omega^2} = \cos^2\theta\left(1 - \frac{n(y)e^2}{\epsilon_0 m\omega^2}\right),$$

or

$$\cos\theta\left[1 - \frac{n(y)e^2}{\epsilon_0 m\omega^2}\right]^{1/2} = \cos\theta_b\left[1 - \frac{n(y)e^2}{\epsilon_0 m\omega^2 \cos^2\theta_b}\right]^{1/2}. \tag{13.9}$$

From Eqs. (13.8) and (13.9), we obtain

$$\tan\theta = \frac{\tan\theta_b}{\sqrt{1 - \dfrac{n(y)e^2}{\epsilon_0 m\omega^2 \cos^2\theta_b}}}. \tag{13.10}$$

We let x and y be horizontal and vertical coordinates. The highest point of the trajectory of the ray is a point of reflection; reflection of the ray therefore requires, as shown, a maximal height y_{max}. Near the ground practically no reflection takes place, but only around y_{max}. We have

$$\tan\theta = \frac{dx}{dy},$$

so that with (13.10)

$$x = \tan \theta_b \int_0^y \frac{dy}{\sqrt{1 - n(y)e^2/\epsilon_0 m\omega^2 \cos^2 \theta_b}}. \tag{13.11}$$
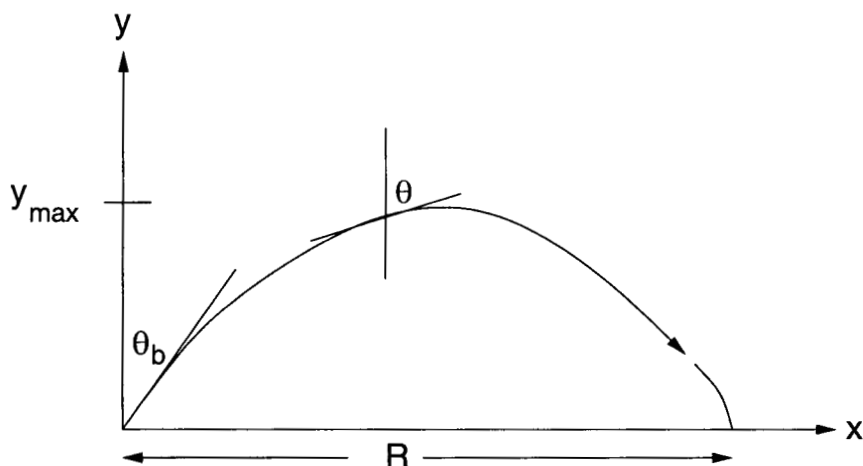
Fig. 13.4 The trajectory of the ray.

With this we obtain for the range R indicated in Fig. 13.4

$$R = 2 \tan \theta_b \int_0^{y_{max}} \frac{dy}{\sqrt{1 - n(y)e^2/\epsilon_0 m\omega^2 \cos^2 \theta_b}}. \tag{13.12}$$

At y_{max} we have $\theta = \pi/2$, i.e. $\tan \theta = \infty$, i.e. at this point

$$\cos \theta_b = \sqrt{\frac{n(y_{max})e^2}{\epsilon_0 m\omega^2}} = \frac{\omega_P}{\omega}. \tag{13.13a}$$

This equation may not always have solutions for any value of θ_b. If for all y

$$1 - \frac{n(y)e^2}{\epsilon_0 m\omega^2 \cos^2 \theta_b} > 0,$$

there is no y_{max}. In this case $n(y)$ cannot grow indefinitely, and there must be a value n_{max} with

$$\frac{n_{max}e^2}{\epsilon_0 m\omega^2} \equiv \cos^2 \theta_0 \tag{13.13b}$$

(this expression defines θ_0), at which

$$1 - \frac{\cos^2 \theta_0}{\cos^2 \theta_b} > 0,$$

i.e.

$$\cos^2\theta_b > \cos^2\theta_0, \quad \theta_b < \theta_0.$$

This means, with this condition the rays *cannot* return to Earth. *For return to be possible*, we must have

$$\theta_b > \theta_0 \left(\text{also } 1 - \frac{n_{\max}e^2}{\epsilon m \omega^2 \cos^2\theta_b} < 0\right). \tag{13.14}$$

This means, rays satisfying this condition, can be reflected back to Earth. One can express this in another way by saying, that

$$\omega^2 < \frac{\omega_P^2}{\cos^2\theta_b}$$

for a fixed angle θ_b. Thus for a fixed angle θ_b the frequency

$$\omega_{\max} = \frac{\omega_P}{\cos\theta_b}$$

is the *maximal frequency* possible for a return to Earth of the ray of radio waves.

13.3   Effect of Terrestrial Magnetic Field

We now want to take into account the effect of the magnetic field of the Earth on the propagation of radio waves. In our treatment here we again consider only the electric component of the electromagnetic field of the waves. This suffices since the magnetic field component is much weaker, as we can see from the fact that $\mathbf{B} = O(E/c)$. We also simplify the terrestrial magnetic field to one along the z-direction parallel to the normal to the surface of the Earth (the latter being assumed more or less flat for our purposes here). Hence we write for the terrestrial field

$$\mathbf{B}_e = (0, 0, B_e).$$

We can consider a more general case, i.e. that of the terrestrial field at some angle, by replacing in the following the z-component $\mathbf{B}_e$ by $B_e \cos\theta_e$. The equation of motion of a free particle with charge q (i.e. an electron in the ionosphere) is then with $\mathbf{r}' = (x', y', z')$ given by

$$m\left(\ddot{\mathbf{r}}' + \frac{1}{\tau}\dot{\mathbf{r}}'\right) = q[\underbrace{\mathbf{E}}_{\text{ray}} + \dot{\mathbf{r}}' \times \underbrace{\mathbf{B}_e}_{\text{Earth}}], \tag{13.15}$$

where τ is the relaxation time, i.e. the mean time interval between successive collisions of the particle with other particles. We put

$$\xi = x' + iy', \quad \text{so that} \quad (\dot{\mathbf{r}}' \times \mathbf{B}_e)_x = \dot{y}' B_e \tag{13.16}$$

and

$$E \equiv E_x + iE_y \tag{13.17}$$

and define the *Larmor frequency* ω_L by

$$qB_e = m\omega_L \quad (\text{or} \quad m\omega_L \cos\theta_e). \tag{13.18}$$

The equation of motion then leads to

$$m\left(\ddot{\xi} + \frac{1}{\tau}\dot{\xi}\right) = qE - im\omega_L\dot{\xi},$$

or

$$\ddot{\xi} + \left(\frac{1}{\tau} + i\omega_L\right)\dot{\xi} = \frac{q}{m}E. \tag{13.19}$$

Setting

$$E = \frac{1}{2}E_0 e^{i(\omega t - kz)},$$

$$E_x = \frac{1}{2}E_0\cos(\omega t - kz), \quad E_y = \pm\frac{1}{2}E_0\sin(\omega t - kz), \tag{13.20}$$

where the latter are the x,y-components of the real part of ($\mathbf{e}_x, \mathbf{e}_y$ again unit vectors)

$$\mathbf{E}_{\substack{R\\L}}(z,t) = \frac{1}{2}E_0(\mathbf{e}_x \mp i\mathbf{e}_y)e^{i(\omega t - kz)}, \tag{13.21}$$

i.e.

$$\Re\mathbf{E}_{\substack{R\\L}}(z,t) = \frac{1}{2}E_0\cos(\omega t - kz)\mathbf{e}_x \pm \frac{1}{2}E_0\sin(\omega t - kz)\mathbf{e}_y$$

$$= E_x\mathbf{e}_x \pm E_y\mathbf{e}_y.$$

These waves are described as right polarised and left polarised respectively. The sum

$$\mathbf{E}_R + \mathbf{E}_L = E_0 e^{i(\omega t - kz)}\mathbf{e}_x,$$

or its real part

$$\Re(\mathbf{E}_R + \mathbf{E}_L) = E_0\cos(\omega t - kz)\mathbf{e}_x$$

represents a linearly polarised wave in the direction of x. With

$$E = \frac{1}{2}E_0 e^{\pm i(\omega t - kz)}$$

we make the following ansatz for the solution of Eq. (13.19):

$$\xi = \xi_0 e^{\pm i(\omega t - kz)} \tag{13.22}$$

and obtain

$$\xi = -\frac{q}{m\omega} \frac{E(z,t)}{\omega \pm \omega_L \mp \frac{i}{\tau}}. \tag{13.23}$$

Recalling the current density

$$\mathbf{j} = (j_x, j_y, j_z),$$

we define the complex quantity

$$j = j_x + ij_y = nq\dot{\xi} = -\frac{nq^2}{m\omega} \frac{\dot{E}}{\omega \pm \omega_L \mp \frac{i}{\tau}},$$

$$\dot{j} = \frac{nq^2}{m\omega} \frac{\ddot{E}}{\omega \pm \omega_L \mp \frac{i}{\tau}}. \tag{13.24}$$

Now, we know from Maxwell's equations (cf. e.g. Eq. (9.7)) that every component of the electric vector satisfies the equation

$$\nabla^2 E_i - \epsilon\mu\ddot{E}_i - \mu\dot{j}_i = 0,$$

so that for the complex quantities

$$E = E_x + iE_y, \qquad \dot{j} = \dot{j}_x + i\dot{j}_y$$

the following equation results:

$$\nabla^2 E - \epsilon_0\mu_0\ddot{E} - \mu_0\dot{j} = 0,$$

or

$$\nabla^2 E - \epsilon_0\mu_0 \left[1 - \frac{\omega_P^2}{\omega(\omega \pm \omega_L \mp \frac{i}{\tau})}\right]\ddot{E} = 0, \tag{13.25}$$

where

$$\omega_P = \sqrt{\frac{nq^2}{m\epsilon_0}}, \quad p = c\sqrt{\eta\mu}, \quad \eta\mu = \frac{1}{c'^2}.$$

Here ω_P is the plasma frequency of the ionosphere. The definition of the generalised refractive index implies therefore for the refractive index of the ionosphere

$$p = \left[1 - \frac{\omega_P^2}{\omega(\omega \pm \omega_L \mp \frac{i}{\tau})}\right]^{1/2}. \tag{13.26}$$

For

$$p = p_1 + ip_2 \quad \text{and} \quad \omega \gg \frac{1}{\tau}$$

it follows that

$$p_1 \simeq \left[1 - \frac{\omega_P^2}{\omega(\omega \pm \omega_L)} \right]^{1/2}, \quad p_2 \simeq \frac{1}{2\omega\tau} \frac{\omega_P^2}{(\omega \pm \omega_L)^2}.$$

For right circularly polarised electromagnetic waves and left circularly polarised electromagnetic waves, i.e.

$$E = \frac{1}{2} E_0 e^{\pm i(\omega t - kz)}, \quad k = \frac{\omega}{c} p = \frac{\omega}{c}(p_1 + ip_2)$$

the refractive index p therefore differs. A linearly polarised wave with

$$\mathbf{E} = (E_x, 0, 0), \quad E_x = E_0 \cos(\omega t - kz)$$

can be considered as a superposition of two oppositely circularly polarised waves. We see therefore, that in the case of an originally linearly polarised wave, one of its circularly polarised components is absorbed more strongly (i.e. damped) than the other, so that the original wave becomes partially circularly polarised.

Here the origin of p_2 is different from that in the case of metals with $\sigma \neq 0$. In the present case $p_2 \propto 1/\tau$, and $p_2 \to 0$ for $\tau \to \infty$. τ is the relaxation time and thus originates from a term (like friction) which implies a loss of energy in the form of heat.

Example 13.1: The pulsar

A pulsar emits a pulse with frequencies ω_1, ω_2 (i.e. emitted simultaneously), which are considerably larger than the plasma frequency ω_P of the interstellar medium. The times of arrival of the pulses with frequencies ω_1, ω_2 are measured. Show that this permits the determination of the electron density of the medium integrated over a distance L, i.e. $\int_0^L n(l)dl$.

Solution: Since $\omega \gg \omega_P$, we have (cf. Eq. (9.17)) the dispersion relation

$$\frac{k^2}{\sqrt{\epsilon\mu}} = \omega^2 - \omega_P^2, \quad \text{or} \quad k^2 c^2 = \omega^2 - \omega_P^2.$$

$(\epsilon, \mu \sim \epsilon_0, \mu_0)$. Moreover

$$\omega_P^2 = \frac{ne^2}{m\epsilon_0}, \quad n = n(l).$$

The group velocity v_g is given by

$$v_g = \frac{d\omega}{dk}, \quad \text{so that} \quad v_g = c\sqrt{1 - \left(\frac{\omega_P}{\omega}\right)^2}.$$

The time of flight T of a pulse (or wave packet which travels with the group velocity, see Chapter 15) from its source to the observer is

$$T = \int_0^T \frac{dl}{v_g} = \frac{1}{c} \int_0^T \frac{dl}{\sqrt{1 - (\frac{\omega_P}{\omega})^2}} \simeq \frac{1}{c} \int_0^T \left(1 + \frac{\omega_P^2}{2\omega^2}\right) dl.$$

Let $T \to T_i$ for $\omega \to \omega_i$, $i = 1, 2$. Then

$$\triangle T = T_2 - T_1 = \frac{1}{2c} \int_0^T \left(\frac{\omega_P^2}{\omega_2^2} - \frac{\omega_P^2}{\omega_1^2}\right) dl = \frac{e^2}{2mc\epsilon_0} \left(\frac{1}{\omega_2^2} - \frac{1}{\omega_1^2}\right) \int_0^L n(l)dl.$$

This means

$$\int_0^L n(l)dl = \frac{2mc\epsilon_0}{e^2} \left(\frac{\omega_1^2 \omega_2^2}{\omega_1^2 - \omega_2^2}\right) \triangle T.$$

Example 13.2: Scattering of radio waves

Radio waves with frequency $\omega = 3 \times 10^7$ per second are emitted from a point on the surface of the Earth (radius 6.37×10^6 m) and are reflected back from the ionosphere approximately 3×10^5 m above the surface of the Earth. Calculate the largest angle of reflection θ, at which the wave is reflected back to the surface of the Earth. Also calculate the maximal density of the electrons determined by this angle.

Solution: Since the angle of reflection is equal to the angle of incidence, it is also the largest angle of incidence at layers in the stratosphere which is asked for, i.e. the extreme to vertical incidence. Hence we draw a tangent to the surface of the Earth and extend it to the point of reflection.
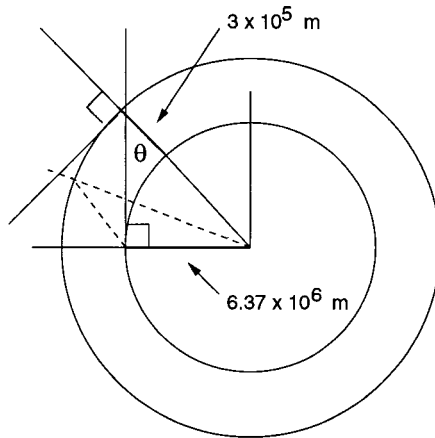
Fig. 13.5 The requested angle θ.

We therefore have the geometry shown in Fig. 13.5. Hence

$$\sin \theta = \frac{6.37 \times 10^6}{6.37 \times 10^6 + 3 \times 10^5} = \frac{6.37}{6.37 + 0.3} = 0.9550, \quad \theta \simeq 73°.$$

We use the following relation obtained above in the text (cf. Eq. (13.13b))

$$\theta = \theta_0 : \cos \theta = \sqrt{\frac{ne^2}{m\epsilon_0 \omega^2}}.$$

With
$$\epsilon_0 = 8.854 \times 10^{-12} \text{ A s/(V m)},$$

this yields for $\omega = 3 \times 10^7$ per second the density

$$n \simeq 8.5 \times 10^{11}$$

of free electrons per meter^3.

This page intentionally left blank

Chapter 14

Wave Guides and Resonators

14.1 What are Wave Guides?

Wave guides are cylindrically shaped objects of metal (almost ideal conductors) which are open at both ends. Resonators, also called resonant cavities, differ from these in being closed at both ends. The interior of such objects is filled with some homogeneous material with electromagnetic constants ϵ, μ; alternatively the interior can be composed of layers of different materials. We have seen before, that the larger the conductivity σ of a medium or the frequency of the incident electromagnetic wave, the smaller the skin depth, i.e. the depth of penetration of the radiation into the conducting medium. This follows from the dispersion relation that we encountered several times earlier, i.e.

$$\mathbf{k}^2 = \mu\epsilon\frac{\omega^2}{c^2}\left(1 + i\frac{\sigma}{\omega\epsilon}\right), \tag{14.1}$$

together with the plane wave ansatz

$$\mathbf{E} = \mathbf{E}_0 e^{i(\mathbf{k}\cdot\mathbf{r}-\omega t)}. \tag{14.2}$$

The square root of $\mathbf{k}^2$ is

$$k = \pm\sqrt{\mu\epsilon}\frac{\omega}{c}\sqrt{1 + i\frac{\sigma}{\omega\epsilon}}. \tag{14.3}$$

Here we have to choose the physically relevant sign. For instance for the wave $\exp(ikz)$ and a medium extended indefinitely in the direction of z, the sign has to be chosen such that the wave is damped, so that the current density $\mathbf{j} = \sigma\mathbf{E}$ cannot grow arbitrarily (which would be nonsensical). It is because of this skin effect that it is possible to confine radiation in hollow bodies or

to use them for the propagation of waves (ideal conductors with $\sigma = \infty$ do not permit any penetration of radiation into the walls). The radiation is fed into the wave guide or resonator with a sender or emitter. We assume here that sender and receiver of electromagnetic waves or signals are effectively at infinity, so that there are no charge or current sources in any finite part of space which would have to be taken into account in our considerations. A signal is a wave packet propagating with group velocity, and will be discussed in more detail later. Such a signal given at (say) time $t = 0$ and at a particular point of the conductor characterises the external charges and currents which generate the corresponding fields $\mathbf{E}$ and $\mathbf{B}$ (charges for $\mathbf{E}$ and currents for $\mathbf{B}$). Correspondingly one can investigate the properties of wave guides and resonators in terms of current I along the wave guide and the potential difference or voltage V across it instead of the fields E and B. We will derive later the corresponding equations for the important case of the coaxial transmission cable, for which we calculated earlier the capacity and the self-inductance per unit length. The consideration of wave guides is therefore very instructive because this involves the full interplay between radiation and boundaries.

We now consider electromagnetic waves in the hollow space enclosed by a wave guide (we consider resonators separately at the end). For $\mathbf{E}, \mathbf{H} \propto e^{-i\omega t}$, and (as explained above) $\mathbf{j} = 0 = \rho, \mathbf{D} = \epsilon\mathbf{E}$ we have

$$\boldsymbol{\nabla} \times \mathbf{E} = i\omega\mathbf{B},$$
$$\boldsymbol{\nabla} \times \mathbf{B} = -i\mu\epsilon\omega\mathbf{E},$$
$$\boldsymbol{\nabla} \cdot \mathbf{E} = 0,$$
$$\boldsymbol{\nabla} \cdot \mathbf{B} = 0. \tag{14.4}$$

We observe here the symmetry of the equations in exchanges $\mathbf{E} \leftrightarrow \mathbf{B}$ with $\omega \leftrightarrow -\omega\mu\epsilon$, which we will exploit later. We have seen in earlier chapters that by performing the "curl–curl" operation, i.e. $\boldsymbol{\nabla} \times \boldsymbol{\nabla}$, on Maxwell's equations, the following equations result

$$\nabla^2\mathbf{E} + \mu\epsilon\omega^2\mathbf{E} = 0,$$
$$\nabla^2\mathbf{B} + \mu\epsilon\omega^2\mathbf{B} = 0. \tag{14.5}$$

In view of the cylindrical symmetry of the problem (cylinder axis along the z-axis) we expect waves, which travel along the z-direction, i.e. we set

$$\mathbf{E}(x, y, z, t) = \mathbf{E}(x, y)e^{\pm ikz - i\omega t},$$
$$\mathbf{B}(x, y, z, t) = \mathbf{B}(x, y)e^{\pm ikz - i\omega t}. \tag{14.6}$$

One should note that the directions of $\mathbf{E}$ and $\mathbf{B}$ are still arbitrary. In general it is also possible that $\mathbf{E}(x,y)$ and $\mathbf{B}(x,y)$ can depend on z. We now set for convenience

$$\nabla^2 = \nabla_\perp^2 + \nabla_z^2, \quad \text{where} \quad \nabla_\perp^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}, \tag{14.7}$$

and $\perp$ means "transverse", so that

$$(\nabla^2 + \mu\epsilon\omega^2)\mathbf{E} = (\nabla_\perp^2 + \mu\epsilon\omega^2 - k^2)\mathbf{E}(x,y)e^{\pm i(kz-\omega t)} = 0,$$

and hence we obtain

$$\begin{aligned} (\nabla_\perp^2 + \mu\epsilon\omega^2 - k^2)\mathbf{E}(x,y) &= 0, \\ (\nabla_\perp^2 + \mu\epsilon\omega^2 - k^2)\mathbf{B}(x,y) &= 0, \end{aligned} \tag{14.8}$$

with the second equation for $\mathbf{B}$ resulting from the symmetry pointed out above (see Eq. (14.5)). We set in addition

$$\mathbf{E} = \mathbf{E}_z + \mathbf{E}_\perp, \quad \mathbf{B} = \mathbf{B}_z + \mathbf{B}_\perp. \tag{14.9}$$

14.2 Transverse Fields Derived from Longitudinal Fields

We now show that it suffices to know the z-components $\mathbf{E}_z, \mathbf{B}_z$; the transverse fields $\mathbf{E}_\perp, \mathbf{B}_\perp$ can then be obtained from these. We consider the two curl-equations of Eq. (14.4):

Case (1). We have

$$\nabla \times \mathbf{E} = i\omega\mathbf{B} \tag{14.10}$$

or with (14.9) and

$$\nabla_\perp = \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}\right), \quad \nabla_z = \mathbf{e}_z\frac{\partial}{\partial z},$$

the following equation

$$(\nabla_\perp + \nabla_z) \times (\mathbf{E}_\perp + \mathbf{E}_z) = i\omega(\mathbf{B}_\perp + \mathbf{B}_z).$$

In the direction of z

$$\nabla_\perp \times \mathbf{E}_\perp = i\omega\mathbf{B}_z \tag{14.11}$$

and in the (x, y)-plane

$$\boldsymbol{\nabla} \times \mathbf{E}_\perp + \boldsymbol{\nabla}_\perp \times \mathbf{E}_z = i\omega \mathbf{B}_\perp. \tag{14.12}$$

We multiply Eq. (14.12) from the left by $\boldsymbol{\nabla}_z \times$ and obtain (see comments at the end)

$$i\omega \boldsymbol{\nabla}_z \times \mathbf{B}_\perp = \boldsymbol{\nabla}_z \times (\boldsymbol{\nabla}_z \times \mathbf{E}_\perp) + \boldsymbol{\nabla}_z \times (\boldsymbol{\nabla}_\perp \times \mathbf{E}_z) \tag{14.13}$$
$$= \boldsymbol{\nabla}_z(\boldsymbol{\nabla}_z \cdot \mathbf{E}_\perp) - (\boldsymbol{\nabla}_z \cdot \boldsymbol{\nabla}_z)\mathbf{E}_\perp$$
$$+ \boldsymbol{\nabla}_\perp(\boldsymbol{\nabla}_z \cdot \mathbf{E}_z) - (\boldsymbol{\nabla}_z \cdot \boldsymbol{\nabla}_\perp)\mathbf{E}_z. \tag{14.14}$$

In the case of the first two expressions on the right of Eq.(14.14) we used the relation *"curl curl = grad div − div grad"*, in the case of the remaining terms the relation

$$\mathbf{A} \times (\mathbf{B} \times \mathbf{C}) = (\mathbf{A} \cdot \mathbf{C})\mathbf{B} - (\mathbf{A} \cdot \mathbf{B})\mathbf{C}, \tag{14.15}$$

taking care of the ordering. The first and the last contributions on the right of Eq. (14.14) are scalar products of orthogonal vectors and therefore vanish. Hence we are left with

$$-\boldsymbol{\nabla}_z^2 \mathbf{E}_\perp + \boldsymbol{\nabla}_\perp \left(\frac{\partial E_z}{\partial z} \right) = i\omega \boldsymbol{\nabla}_z \times \mathbf{B}_\perp.$$

But

$$-\boldsymbol{\nabla}_z^2 \mathbf{E}_\perp = -\frac{\partial^2}{\partial z^2}(\mathbf{E}_\perp(x, y)e^{\pm ikz - i\omega t}) = k^2 \mathbf{E}_\perp,$$

so that

$$k^2 \mathbf{E}_\perp + \boldsymbol{\nabla}_\perp \left(\frac{\partial E_z}{\partial z} \right) = i\omega \boldsymbol{\nabla}_z \times \mathbf{B}_\perp. \tag{14.16}$$

Case (2). In a corresponding way we deal with

$$\boldsymbol{\nabla} \times \mathbf{B} = -i\omega\mu\epsilon\mathbf{E}, \tag{14.17}$$

i.e.

$$(\boldsymbol{\nabla}_\perp + \boldsymbol{\nabla}_z) \times (\mathbf{B}_\perp + \mathbf{B}_z) = -i\omega\mu\epsilon(\mathbf{E}_\perp + \mathbf{E}_z),$$

where the left hand side is the following sum

$$(\boldsymbol{\nabla}_\perp \times \mathbf{B}_\perp) + (\boldsymbol{\nabla}_\perp \times \mathbf{B}_z) + (\boldsymbol{\nabla}_z \times \mathbf{B}_\perp) + \underbrace{(\boldsymbol{\nabla}_z \times \mathbf{B}_z)}_{0},$$

so that in the direction of z:

$$\boldsymbol{\nabla}_\perp \times \mathbf{B}_\perp = -i\omega\mu\epsilon\mathbf{E}_z \tag{14.18}$$

and transversally

$$\nabla_\perp \times \mathbf{B}_z + \nabla_z \times \mathbf{B}_\perp = -i\omega\mu\epsilon\mathbf{E}_\perp,$$

i.e.

$$\nabla_z \times \mathbf{B}_\perp = -\nabla_\perp \times \mathbf{B}_z - i\omega\mu\epsilon\mathbf{E}_\perp. \tag{14.19}$$

We insert this in Eq. (14.16) and obtain

$$k^2\mathbf{E}_\perp + \nabla_\perp\left(\frac{\partial E_z}{\partial z}\right) = i\omega[-\nabla_\perp \times \mathbf{B}_z - i\omega\mu\epsilon\mathbf{E}_\perp],$$

i.e. arranging the contributions in a different way,

$$\nabla_\perp\left(\frac{\partial E_z}{\partial z}\right) + i\omega(\nabla_\perp \times \mathbf{B}_z) = (\omega^2\mu\epsilon - k^2)\mathbf{E}_\perp.$$

Hence

$$(\mu\epsilon\omega^2 - k^2)\mathbf{E}_\perp = \left(\nabla_\perp\frac{\partial E_z}{\partial z} - i\omega(\mathbf{e}_z \times \nabla_\perp)B_z\right). \tag{14.20}$$

Owing to the symmetry of the Maxwell equations (14.4) in $\mathbf{E}$ and $\mathbf{B}$ we have also

$$(\mu\epsilon\omega^2 - k^2)\mathbf{B}_\perp = \left(\nabla_\perp\frac{\partial B_z}{\partial z} + i\omega\mu\epsilon(\mathbf{e}_z \times \nabla_\perp)E_z\right). \tag{14.21}$$

We see therefore: The transverse components of the fields can be obtained from the longitudinal components, in fact in each case from both longitudinal components. The relations (14.20) and (14.21) can also be derived in a different way, i.e. by immediate substitution from

$$\nabla \cdot \mathbf{E} = 0, \quad \nabla \cdot \mathbf{B} = 0$$

and

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}, \quad \nabla \times \mathbf{H} = \frac{\partial \mathbf{D}}{\partial t}.$$

in the interior of the wave guide (i.e. not in the wall).*

---

*In these equations the derivatives with respect to t and z are replaced by $-i\omega$ and ik, and the equations are then solved for the transverse components of the fields.

14.3   Boundary Conditions

We now come to the important aspect of boundary conditions. We recall for this reason the continuity conditions at a boundary surface. We had in particular

$$B_n^{(1)} = B_n^{(2)}, \quad E_t^{(1)} = E_t^{(2)}, \quad \text{with} \quad \mathbf{k} \times \mathbf{E} = \omega\mu\mathbf{H} = \omega\mathbf{B}, \qquad (14.22)$$

where "n" stands for "normal component" and "t" for "tangential component". The two conditions in Eq. (14.22) are exact. The other two conditions which we obtained and used previously are not needed in the present case since surface charge and surface current are here zero (for $\sigma = \infty$). The two conditions above can also be written ($\mathbf{e}_n \cdot \mathbf{e}_z = 0$)

$$\mathbf{n} \cdot \mathbf{B} = 0, \quad \mathbf{n} \times \mathbf{E} = 0 \qquad (14.23)$$

with $B^{(1)} = B, E^{(1)} = E$ in the interior of the wave guide and $B^{(2)} = 0 = E^{(2)}$ outside (again for $\sigma = \infty$, i.e. ideal conductors). We assume here that the wave guides are infinitely long in order to avoid finite end effects.

We can write the two boundary conditions:

$$\mathbf{n} \cdot (\mathbf{B}_\perp + \mathbf{B}_z) = 0 \quad \text{and} \quad \mathbf{n} \times (\mathbf{E}_\perp + \mathbf{E}_z) = 0,$$

i.e. (since $\mathbf{n} \perp \mathbf{B}_z$):

$$\mathbf{n} \cdot \mathbf{B}_\perp = 0 \qquad (14.24)$$

and

$$\underbrace{\mathbf{n} \times \mathbf{E}_\perp}_{\text{vector}\|\mathbf{e}_z} = 0, \qquad (14.25)$$

$$\underbrace{\mathbf{n} \times \mathbf{E}_z}_{\text{vector}\perp\mathbf{e}_z} = 0, \quad \text{i.e.} \quad E_z = 0. \qquad (14.26)$$

These conditions hold at the boundary surface (i.e. not anywhere else). Since we saw that the fields $\mathbf{E}_\perp, \mathbf{B}_\perp$ follow from E_z, B_z, we are interested in the boundary conditions and the differential equations of E_z, B_z. Thus we have to find the boundary condition of B_z. This is our next step.

Vector multiplication of the Maxwell equation (14.19) by $\mathbf{e}_z$ gives

$$\mathbf{e}_z \times (\mathbf{\nabla}_z \times \mathbf{B}_\perp) = -\mathbf{e}_z \times (\mathbf{\nabla}_\perp \times \mathbf{B}_z) - i\omega\mu\epsilon(\mathbf{e}_z \times \mathbf{E}_\perp).$$

Using Eq. (14.15) it follows that

$$0 - (\mathbf{e}_z \cdot \mathbf{\nabla}_z)\mathbf{B}_\perp = -\mathbf{\nabla}_\perp B_z + 0 - i\omega\mu\epsilon(\mathbf{e}_z \times \mathbf{E}_\perp),$$

i.e.

$$\frac{\partial \mathbf{B}_\perp}{\partial z} = \boldsymbol{\nabla}_\perp B_z + i\omega\mu\epsilon(\mathbf{e}_z \times \mathbf{E}_\perp). \tag{14.27}$$

Scalar multiplication of this equation by **n** yields

$$\frac{\partial}{\partial z}\mathbf{n} \cdot \mathbf{B}_\perp = \mathbf{n} \cdot \boldsymbol{\nabla}_\perp B_z + i\omega\mu\epsilon\mathbf{n} \cdot (\mathbf{e}_z \times \mathbf{E}_\perp).$$

Since

$$\mathbf{B} = \mathbf{B}(x, y)e^{\pm ikz - i\omega t},$$

i.e.

$$\frac{\partial \mathbf{B}_\perp}{\partial z} = \pm ik\mathbf{B}_\perp,$$

we obtain

$$\pm ik(\mathbf{n} \cdot \mathbf{B}_\perp) = \frac{\partial B_z}{\partial n} - i\omega\mu\epsilon(\mathbf{n} \times \mathbf{E}_\perp) \cdot \mathbf{e}_z. \tag{14.28}$$

Thus at the surface of the wave guide we have with Eqs. (14.24) and (14.28)

$$\left.\frac{\partial B_z}{\partial n}\right|_{\text{surface}} = 0, \quad \left(\frac{\partial}{\partial n} \equiv \mathbf{n} \cdot \boldsymbol{\nabla}_\perp\right). \tag{14.29}$$

Above we reduced the Maxwell equations to the following two equations which apply to each of the three components of the fields, i.e.

$$(\boldsymbol{\nabla}_\perp^2 + \mu\epsilon\omega^2 - k^2)\mathbf{E}(x, y) = 0, \quad (\boldsymbol{\nabla}_\perp^2 + \mu\epsilon\omega^2 - k^2)\mathbf{B}(x, y) = 0. \tag{14.30}$$

Moreover, we saw that the transverse components $\mathbf{E}_\perp, \mathbf{B}_\perp$ can be derived from a knowledge of the longitudinal components E_z, B_z. Thus there remains the investigation of the equations

$$(\boldsymbol{\nabla}_\perp^2 + \mu\epsilon\omega^2 - k^2)E_z(x, y) = 0, \quad (\boldsymbol{\nabla}_\perp^2 + \mu\epsilon\omega^2 - k^2)B_z(x, y) = 0 \tag{14.31}$$

with the boundary conditions

$$\left. E_z\right|_{\text{surface}} = 0, \quad \left.\frac{\partial B_z}{\partial n}\right|_{\text{surface}} = 0, \tag{14.32}$$

where the surfaces here are two-dimensional. Equations (14.30) with these conditions define an *eigenvalue problem*. In general the boundary conditions cannot be satisfied simultaneously.

14.4 Wave Guides and their TEM Fields

We consider wave guides. The frequency ω is given; the eigenvalue problem then determines the permissible (axial) wave numbers k. Depending on the boundary conditions one distinguishes between fields of different types:

TM, *transverse magnetic*: $B_z = 0$ everywhere (hence the terminology) and $E_z|_{\text{surface}} = 0$,

TE, *transverse electric*: $E_z = 0$ everywhere (hence the terminology) and $\partial B_z/\partial n|_{\text{surface}} = 0$,

TEM, *transverse electric-magnetic*: $B_z = 0 = E_z$ everywhere.

We consider first the case of TEM fields.

14.4.1 TEM Fields

In these cases we obtain from Eqs. (14.20) and (14.21) the equations

$$(\mu\epsilon\omega^2 - k^2)\mathbf{E}_\perp = 0, \quad (\mu\epsilon\omega^2 - k^2)\mathbf{B}_\perp = 0. \tag{14.33}$$

These equations have the trivial solution

$$\mathbf{E}_\perp = 0 = \mathbf{B}_\perp \quad \text{for} \quad \mu\epsilon\omega^2 - k^2 \neq 0. \tag{14.34}$$

Hence for a nontrivial solution we have in general

$$\mu\epsilon\omega^2 - k^2 = 0, \quad k = \sqrt{\mu\epsilon}\omega. \tag{14.35}$$

Equations (14.20) and (14.21) imply $0 \times \mathbf{E}_\perp, \mathbf{B}_\perp = 0$, i.e. $\mathbf{E}_\perp, \mathbf{B}_\perp$ remain undetermined. For this reason we now write the TEM fields

$$\mathbf{E}_{TEM} = (\mathbf{E}_\perp, E_z = 0), \quad \mathbf{B}_{TEM} = (\mathbf{B}_\perp, B_z = 0). \tag{14.36}$$

According to Eqs. (14.31) we now have (with $\nabla_\perp^2 \equiv \triangle_\perp$):

$$\triangle_\perp\mathbf{E}_\perp = 0, \quad \triangle_\perp\mathbf{B}_\perp = 0, \tag{14.37}$$

i.e. $\mathbf{E}_{TEM}, \mathbf{B}_{TEM}$ are solutions of the two-dimensional Laplace equation. Before we investigate these equations we show that

$$\mathbf{E}_{TEM} \perp \mathbf{B}_{TEM}. \tag{14.38}$$

This is *not* trivial, because so far we established the transversality of electromagnetic radiation only for the case of an unlimited medium.

From

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial \mathbf{t}} \quad \text{with } \mathbf{B} \propto e^{-i\omega t}$$

we get

$$i\omega \mathbf{B} = \nabla \times \mathbf{E} = (\nabla_\perp + \nabla_z) \times \mathbf{E}$$
$$= \nabla_\perp \times \mathbf{E} + \frac{\partial}{\partial z}(\mathbf{e}_z \times \mathbf{E}).$$

For $\mathbf{E}, \mathbf{B} \to \mathbf{E}_{TEM}, \mathbf{B}_{TEM}$, i.e. $B_z = 0 = E_z$ the fields $\mathbf{E}, \mathbf{B}$ are vectors in the (x, y)-plane so that

$$\nabla_\perp \times \mathbf{E}_{TEM} \perp (x, y)-\text{plane},$$

i.e. parallel to the z-axis, i.e.

$$i\omega \mathbf{B}_{TEM} = \frac{\partial}{\partial z}(\mathbf{e}_z \times \mathbf{E}_{TEM}) \tag{14.39}$$

and

$$\nabla_\perp \times \mathbf{E}_{TEM} = 0 \quad (\to \mathbf{E}_{TEM} = \nabla_\perp \phi). \tag{14.40}$$

From Eq. (14.39) we deduce that

$$\mathbf{B}_{TEM} = \frac{1}{i\omega}\frac{\partial}{\partial z}(\mathbf{e}_z \times \mathbf{E}_{TEM}).$$

With

$$\mathbf{E}_{TEM} = \mathbf{E}_{0TEM}(x, y)e^{i(kz-\omega t)}$$

we obtain

$$\mathbf{B}_{TEM} = \frac{k}{\omega}(\mathbf{e}_z \times \mathbf{E}_{TEM}),$$

i.e.

$$\mathbf{B}_{TEM} = \sqrt{\mu\epsilon}(\mathbf{e}_z \times \mathbf{E}_{TEM}). \tag{14.41}$$

Thus $\mathbf{E}_{TEM}$ and $\mathbf{B}_{TEM}$ obey the same transversality connection as in the case of the unlimited medium.

However, *TEM waves do not propagate in hollow wave guides, which are made of ideal conductor material* (i.e. have $\sigma = \infty$ and hence represent equipotential surfaces). This can be seen as follows. Inside the wave guide

$$\nabla \cdot \mathbf{E} = 0, \quad \nabla \cdot \mathbf{B} = 0,$$

i.e. for $E_z = 0 = B_z$:

$$\frac{\partial E_x}{\partial x} + \frac{\partial E_y}{\partial y} = 0, \quad \frac{\partial B_x}{\partial x} + \frac{\partial B_y}{\partial y} = 0. \tag{14.42}$$

On the other hand, from Eq. (14.40) or from the z-component of

$$\boldsymbol{\nabla} \times \mathbf{E} = i\omega\mathbf{B}, \quad \boldsymbol{\nabla} \times \mathbf{B} = -i\mu\epsilon\omega\mathbf{E} \quad \text{for} \quad E_z = 0 = B_z,$$

we obtain

$$\frac{\partial E_y}{\partial x} - \frac{\partial E_x}{\partial y} = i\omega B_z = 0,$$

$$\frac{\partial B_y}{\partial x} - \frac{\partial B_x}{\partial y} = -i\mu\epsilon\omega E_z = 0. \tag{14.43}$$

We write the solutions of these equations with potentials $\phi = \phi(x,y), \psi = \psi(x,y)$

$$\left.\begin{array}{ll} E_x = \partial\phi/\partial x, & E_y = \partial\phi/\partial y \\ B_x = \partial\psi/\partial x, & B_y = \partial\psi/\partial y \end{array}\right\}, \quad \mathbf{E}_\perp = \boldsymbol{\nabla}\phi, \quad \mathbf{B}_\perp = \boldsymbol{\nabla}\psi. \tag{14.44}$$

Then Eqs. (14.42) and (14.43) become

$$\frac{\partial^2\phi}{\partial x^2} + \frac{\partial^2\phi}{\partial y^2} = 0, \quad \frac{\partial^2\psi}{\partial x^2} + \frac{\partial^2\psi}{\partial y^2} = 0,$$

$$\frac{\partial^2\phi}{\partial x\partial y} = \frac{\partial^2\phi}{\partial y\partial x}, \quad \frac{\partial^2\psi}{\partial x\partial y} = \frac{\partial^2\psi}{\partial y\partial x},$$

i.e. ϕ and ψ are solutions of the 2-dimensional Laplace equations

$$\triangle_\perp\phi = 0, \quad \triangle_\perp\psi = 0. \tag{14.45}$$

Since, however, the conducting surface of the wave guide is an equipotential surface, on which the potential of $\mathbf{E}$ is constant, we deduce from Gauss' theorem and the fact that the interior of the wave guide does not contain any charges, that inside the wave guide

$$\mathbf{E}_\perp = \boldsymbol{\nabla}_\perp\phi = 0$$

(from $\int_{F(V)} \mathbf{E} \cdot d\mathbf{F} = \int_V \boldsymbol{\nabla} \cdot \mathbf{E} dV = \int_V \rho dV/\epsilon = 0$). The argument here is analogous to that which we used in considerations of spherical condensers and cylindrical condensers in electrostatics. Since the components of $\mathbf{B}$ are linked with components of $\mathbf{E}$ via the transversality condition (14.41), it follows that also $B_x = 0 = B_y$. Thus *TEM waves do not propagate in hollow and empty pipes.*

However, this is different if the wave guide contains inside another symmetric surface, like a concentric metal cylinder of circular cross section; a construction of this type which is no longer simply connected is known as a *coaxial cable*. In this case according to Fig. 14.1 $\phi_2 \neq \phi_1$. In this case the Laplace equation has a solution for the fields $\mathbf{E}, \mathbf{B}$ which does not vanish.

14.4.2 The Coaxial Transmission Cable

We have for the TEM wave of the coaxial cable

$$k = \sqrt{\mu\epsilon}\omega, \quad \triangle_\perp \mathbf{E}_{TEM} = 0, \quad \mathbf{B}_{TEM} = \sqrt{\mu\epsilon}(\mathbf{e}_z \times \mathbf{E}_{TEM}). \qquad (14.46)$$

Also

$$\mathbf{E}_{TEM} = (\boldsymbol{\nabla}_\perp \phi)e^{-i(\omega t - kz)},$$

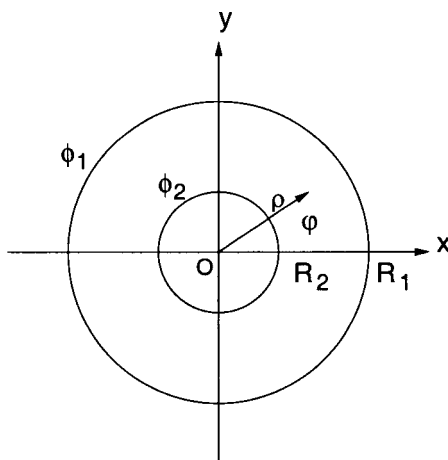and we have to solve

$$\triangle_\perp \phi = 0.$$

Fig. 14.1 Cross section of a coaxial cable.

We choose a cable with circular cross section and radii R_1, R_2 (with $R_2 < R_1$). The cylindrical symmetry of the cable suggests separation of the equation in cylindrical coordinates, here in polar coordinates, so that we have

$$\frac{1}{\rho}\frac{\partial}{\partial\rho}\left(\rho\frac{\partial\phi}{\partial\rho}\right) + \frac{1}{\rho^2}\frac{\partial^2\phi}{\partial\varphi^2} = 0. \qquad (14.47)$$

This equation has to be solved with the boundary conditions

$$\phi(\rho = R_1, \varphi) = \phi_1 = \text{const.},$$
$$\phi(\rho = R_2, \varphi) = \phi_2 = \text{const.} \quad (R_2 < R_1). \qquad (14.48)$$

Because of the symmetry of the cylinder

$$\phi(\rho, \varphi_1) = \phi(\rho, \varphi_2) = \cdots = \phi(\rho),$$

i.e. $\partial\phi/\partial\varphi = 0$, so that

$$\frac{1}{\rho}\frac{\partial}{\partial\rho}\left(\rho\frac{\partial\phi}{\partial\rho}\right) = 0, \tag{14.49}$$

and hence

$$\phi(\rho) = A\ln\rho + B, \tag{14.50}$$

and so

$$\phi_1 = \phi(R_1) = A\ln R_1 + B, \quad \phi_2 = \phi(R_2) = A\ln R_2 + B,$$

i.e.

$$A = \frac{\phi_1 - \phi_2}{\ln(R_1/R_2)}, \quad B = \frac{\phi_2\ln R_1 - \phi_1\ln R_2}{\ln R_1 - \ln R_2}. \tag{14.51}$$

Thus we have

$$\nabla_\perp\phi = \frac{\partial\phi}{\partial\rho}\mathbf{e}_\rho,$$

i.e.

$$\mathbf{E}_{TEM}(\rho,\varphi,z,t) = \frac{A}{\rho}\mathbf{e}_\rho e^{-i(\omega t - kz)},$$

$$\mathbf{B}_{TEM}(\rho,\varphi,z,t) = \sqrt{\mu\epsilon}\frac{A}{\rho}\mathbf{e}_\varphi e^{-i(\omega t - kz)}. \tag{14.52}$$

Expressed in terms of Cartesian coordinates the relations are

$$\mathbf{E}_{TEM}(x,y,z,t) = \frac{A}{x^2 + y^2}(x\mathbf{e}_x + y\mathbf{e}_y)e^{-i(\omega t - kz)},$$

$$\mathbf{B}_{TEM}(x,y,z,t) = \frac{\sqrt{\mu\epsilon}A}{x^2 + y^2}(-y\mathbf{e}_x + x\mathbf{e}_y)e^{-i(\omega t - kz)}. \tag{14.53}$$

One should note that, for $R_2 \to 0$ (removal of the inner cylinder) $A \to 0$ and $B \to \phi_1$, i.e. in this limiting case $\phi = $ const., and there is no TEM wave in the wave guide, as we observed earlier.

14.5 Fundamental Equations for the Coaxial Cable

We mentioned already previously that the investigation of properties of wave guides can also be carried out by focusing on the induced charges and currents, i.e. we infer from the fields $\mathbf{E}$ and $\mathbf{H}$ the presence of charge q and current I, described as induced, which are the same as those which give rise to the fields $\mathbf{E}$ and $\mathbf{H}$. We now derive the corresponding equations. These

describe the propagation of the potential difference or voltage $V = V(t, z)$ and the current $I = I(t, z)$ along the cable. The equations that we obtain, are more general than our derivation here, and retain their validity for many other types of wave guides. Thus we consider a coaxial cable with radii a, b (with $a < b$). We let $V(t, z)$ be the voltage between the outer and the inner cylinders at z-coordinate z, and we let $I(t, z)$ be the current induced in the inner cylinder (or the current corresponding to an appropriate surface current). Then we know that at distance r with $r > a$, from the axis of this cylinder, as indicated in Fig. 14.2, the magnetic field strength $\mathbf{H}$ is given by

$$\mathbf{H} = \frac{I(t, z)}{2\pi r} \mathbf{e}_\varphi, \tag{14.54}$$

as may be verified, for instance, by referring back to Example 8.2. We now apply Faraday's law

$$\oint_C \mathbf{E} \cdot d\mathbf{l} = -\frac{\partial}{\partial t} \int_{F(C)} \mathbf{B} \cdot d\mathbf{F} \tag{14.55}$$

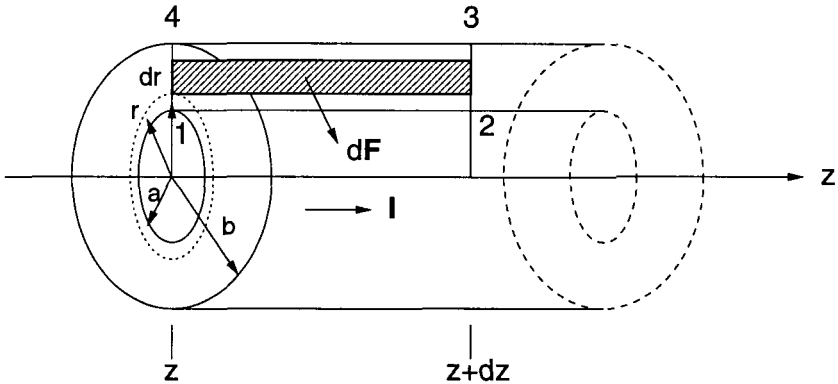to the area (1234) in the interior of the cable as shown in Fig. 14.2.

Fig. 14.2 The coaxial transmission cable.

We obtain

$$\int_1^2 \mathbf{E} \cdot d\mathbf{l} + \int_2^3 \mathbf{E} \cdot d\mathbf{l} + \int_3^4 \mathbf{E} \cdot d\mathbf{l} + \int_4^1 \mathbf{E} \cdot d\mathbf{l} = -\frac{\partial}{\partial t} \int_a^b \mu \frac{I(t, z)}{2\pi r} dr dz.$$

We set

$$\int_4^1 \mathbf{E} \cdot d\mathbf{l} \equiv -V(t, z). \tag{14.56}$$

Then, since the field $\mathbf{E}$ is radial and inside the conductor zero,

$$V(t, z + dz) - V(t, z) = -\frac{\partial}{\partial t} \int_a^b \mu \frac{I(t, z)}{2\pi r} dr dz$$

implying

$$\frac{\partial V(t, z)}{\partial z} = -\frac{\mu}{2\pi} \ln(b/a) \frac{\partial I(t, z)}{\partial t},$$

or

$$\frac{\partial V(t, z)}{\partial z} = -L_0 \frac{\partial I(t, z)}{\partial t} \tag{14.57}$$

with *self-inductance* (see Example 8.2):

$$L_0 = \frac{\mu}{2\pi} \ln(b/a) \quad \text{per unit length.} \tag{14.58}$$

This result is independent of our choice of directions in Fig. 14.3.

A second equation is obtained from Ampère's law to which Maxwell added his displacement current, i.e.

$$\oint \mathbf{B} \cdot dl = \mu \int \left(\mathbf{j} + \frac{\partial \mathbf{D}}{\partial t} \right) \cdot d\mathbf{F}. \tag{14.59}$$

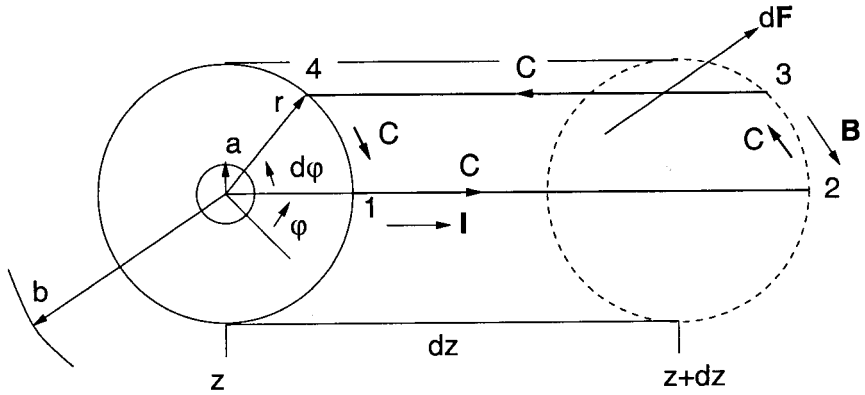
Fig. 14.3 Surface piece of the coaxial transmission cable.

We apply this to a section of the cylindrical surface of the cable at radius $r, a < r < b$, as indicated in Fig. 14.3. The magnetic field $\mathbf{B}$ is entirely due to the current on the internal cylinder (as we observed earlier), and is given by

$$\mathbf{B}(t, z) = \mu \frac{I(t, z)}{2\pi r} \mathbf{e}_\varphi.$$

We thus have along the path C indicated in Fig. 14.4, where $d\mathbf{F} = -rd\varphi\mathbf{e}_\varphi \times dz\mathbf{e}_z$,

$$
\oint_C \mathbf{B} \cdot d\mathbf{l} = \left(\int_1^2 + \int_2^3 + \int_3^4 + \int_4^1 \right) \mathbf{B} \cdot d\mathbf{l}
$$

$$
= 0 + rd\varphi \frac{\mu}{2\pi r} I(t, z + dz) + 0 - \frac{\mu}{2\pi r} I(t,z)(rd\varphi)
$$

$$
= \frac{\mu}{2\pi r} dz r d\varphi \frac{\partial I(t,z)}{\partial z}. \tag{14.60a}
$$

Evaluating the other side of the Ampère–Maxwell equation (14.59) we obtain, since $\mathbf{j} = 0$ in the intermediate space and in the case of the cylinder $E = q/2\pi\epsilon r$, q the induced charge per unit length,

$$
\mu \int \frac{\partial \mathbf{D}}{\partial t} \cdot d\mathbf{F} = \epsilon\mu \frac{\partial}{\partial t} \int_{\triangle F} \mathbf{E} \cdot d\mathbf{F} = -\epsilon\mu \frac{\partial}{\partial t} \left(\frac{q}{2\pi\epsilon r} dz r d\varphi \right)
$$

$$
= -\epsilon\mu dz d\varphi \frac{\partial}{\partial t} \left(\frac{q}{2\pi\epsilon} \right)
$$

$$
= -\epsilon\mu dz d\varphi \frac{\partial}{\partial t} \left(\frac{V(t,z)}{\ln(b/a)} \right), \tag{14.60b}
$$

since $V = q\ln(b/a)/2\pi\epsilon$ per unit length, as may be checked by consulting Example 2.12. From the results (14.60a) and (14.60b) we obtain

$$
\frac{\mu}{2\pi r} dz r d\varphi \frac{\partial I(t,z)}{\partial z} = -\epsilon\mu dz d\varphi \frac{\partial}{\partial t} \left(\frac{V(t,z)}{\ln(b/a)} \right),
$$

and so

$$
\frac{\partial I(t,z)}{\partial z} = -\frac{2\pi\epsilon}{\ln(b/a)} \frac{\partial V(t,z)}{\partial t},
$$

or

$$
\frac{\partial I(t,z)}{\partial z} = -C_0 \frac{\partial V(t,z)}{\partial t}, \tag{14.61}
$$

where

$$
C_0 = \frac{2\pi\epsilon}{\ln(b/a)} \tag{14.62}
$$

is the capacity of the cable per unit length obtained in Example 2.12. One should note again the relation

$$
L_0 C_0 = \frac{\mu}{2\pi} \ln(b/a) \frac{2\pi\epsilon}{\ln(b/a)} = \mu\epsilon = \left(\frac{1}{c'} \right)^2 \tag{14.63}
$$

(in vacuum $c' = c$). Differentiation of the *fundamental equations* (14.57) and (14.61) with respect to t and z yields

$$\frac{\partial^2 V(t,z)}{\partial t \partial z} = -L_0 \frac{\partial^2 I(t,z)}{\partial t^2}, \quad \frac{\partial^2 V(t,z)}{\partial z^2} = -L_0 \frac{\partial^2 I(t,z)}{\partial t \partial z},$$

$$\frac{\partial^2 I(t,z)}{\partial z \partial t} = -C_0 \frac{\partial^2 V(t,z)}{\partial t^2}, \quad \frac{\partial^2 I(t,z)}{\partial z^2} = -C_0 \frac{\partial^2 V(t,z)}{\partial z \partial t},$$

and hence the equations

$$\frac{\partial^2 V(t,z)}{\partial z^2} - L_0 C_0 \frac{\partial^2 V(t,z)}{\partial t^2} = 0,$$

$$\frac{\partial^2 I(t,z)}{\partial z^2} - L_0 C_0 \frac{\partial^2 I(t,z)}{\partial t^2} = 0. \tag{14.64}$$

These equations show that current and voltage propagate wavelike along the cable with the phase velocity $c' = 1/\sqrt{L_0 C_0}$ because

$$I(t,z), V(t,z) \propto e^{i(kz - \omega t)}, \quad k^2 = L_0 C_0 \omega^2 = \frac{\omega^2}{c'^2}. \tag{14.65}$$

14.6 TM and TE Waves in Wave Guides

14.6.1 General Considerations

We first demonstrate that TM and TE waves are not transversal, i.e. they *do* possess longitudinal components. In the cases of TE and TM waves the expressions for $\mathbf{E}_\perp$ and $\mathbf{B}_\perp$ in Eqs. (14.20) and (14.21) also simplify. Setting

$$\gamma^2 = \mu \epsilon \omega^2 - k^2, \quad \mathbf{E}, \mathbf{B} \propto e^{ikz}, \quad k^2 > 0, \tag{14.66}$$

we obtain from Eqs. (14.20) and (14.21) for $\gamma^2 \neq 0$:

TM: $B_z = 0$ everywhere,

$$\mathbf{E}_\perp = \frac{1}{\gamma^2} ik \boldsymbol{\nabla}_\perp E_z, \quad \mathbf{B}_\perp = \frac{1}{\gamma^2} i\mu\epsilon\omega (\mathbf{e}_z \times \boldsymbol{\nabla}_\perp) E_z; \tag{14.67}$$

TE: $E_z = 0$ everywhere,

$$\mathbf{E}_\perp = -\frac{i\omega}{\gamma^2} (\mathbf{e}_z \times \boldsymbol{\nabla}_\perp) B_z, \quad \mathbf{B}_\perp = \frac{ik}{\gamma^2} \boldsymbol{\nabla}_\perp B_z. \tag{14.68}$$

TM: Inserting in the second of Eqs. (14.67) the expression for $\boldsymbol{\nabla}_\perp E_z$ from the first of Eqs. (14.67), we obtain

$$\mathbf{B}_\perp = \frac{\mu\epsilon\omega}{k} \mathbf{e}_z \times \mathbf{E}_\perp. \tag{14.69}$$

TE: Inserting similarly in Eqs. (14.68) the expression of the equation for $\nabla_\perp B_z$ into the other, we obtain

$$\mathbf{E}_\perp = -\frac{\omega}{k}\mathbf{e}_z \times \mathbf{B}_\perp. \tag{14.70}$$

We conclude: $E_z, B_z \neq 0$, because from Eqs. (14.67) for TM

$$\nabla_\perp E_z = \frac{\gamma^2}{ik}\mathbf{E}_\perp$$

and from Eqs. (14.68) for TE

$$\nabla_\perp B_z = \frac{\gamma^2}{ik}\mathbf{B}_\perp.$$

Thus the components E_z (TM), B_z (TE) are not both zero; i.e. *the TM and TE waves are not transversal waves*, they possess longitudinal components E_z and B_z.

We saw previously (see Eqs. (14.30)), that $\mathbf{E}(x,y)$, $\mathbf{B}(x,y)$ satisfy the equations

$$(\nabla_\perp^2 + \gamma^2)\mathbf{E}(x,y) = 0, \quad (\nabla_\perp^2 + \gamma^2)\mathbf{B}(x,y) = 0 \tag{14.71}$$

(i.e. for each of the three components of $\mathbf{E}$ and $\mathbf{B}$). We have therefore in particular for the z-components

TM:

$$(\nabla_\perp^2 + \gamma^2)\mathbf{E}_z(x,y) = 0, \quad E_z|_{\text{surface}} = 0, \quad (B_z = 0 \text{ everywhere}), \tag{14.72}$$

TE:

$$(\nabla_\perp^2 + \gamma^2)\mathbf{B}_z(x,y) = 0, \quad \frac{\partial B_z}{\partial n}\bigg|_{\text{surface}} = 0, \quad (E_z = 0 \text{ everywhere}). \tag{14.73}$$

Equations (14.72) and (14.73) together with their boundary conditions define an eigenvalue problem for the determination of E_z and B_z; we observed earlier that once both of these are known, one can derive from them the transverse field components. For the solution of Eqs. (14.72) and (14.73) one requires precise information about the geometry of the wave guide.

14.6.2 Wave Guides with Rectangular Cross Section

As an example we consider a wave guide with rectangular cross section as illustrated in Fig. 14.4. We show in particular, how the waves propagate

through the wave guide by continuous reflection from the walls. We begin with **TM modes**.

The hollow inside of the wave guide is bounded by the planes

$$x = 0, \quad x = a, \quad y = 0, \quad y = b.$$

These are the surfaces on which the *boundary* condition $E_z|_{\text{surface}} = 0$ has to be imposed. In Eq. (14.72) for E_z the variables x and y can be separated. We therefore make for $E_z(x, y)$ the ansatz

$$E_z(x, y) = f(x)g(y), \tag{14.74}$$

so that

$$\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \gamma^2 \right) f(x)g(y) = 0,$$

or for $x \neq 0, a$ and $y \neq 0, b$:

$$\frac{1}{f(x)} \frac{\partial^2 f(x)}{\partial x^2} + \frac{1}{g(y)} \frac{\partial^2 g(y)}{\partial y^2} + \gamma^2 = 0,$$

or

$$\frac{\partial^2 f(x)}{\partial x^2} + \rho^2 f(x) = 0 \tag{14.75}$$

and

$$\frac{\partial^2 g(y)}{\partial y^2} + (\gamma^2 - \rho^2)g(y) = 0, \tag{14.76}$$

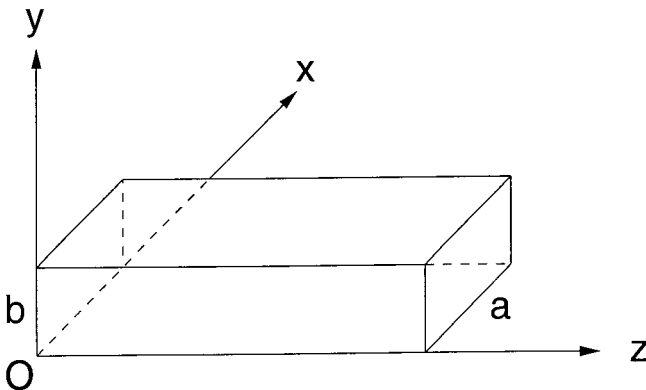where $\rho^2 = \text{const.}$

Fig. 14.4 Wave guide with rectangular cross section.

In order to solve the *first equation*, Eq. (14.75), we write

$$f(x) = A \cos \rho(x + x_0), \quad A \neq 0.$$

A and x_0 are the two integration constants. These are determined with the help of the boundary conditions. Thus $E_z|_{\text{surface}} = 0$ implies

$$f(0) = 0, \quad f(a) = 0,$$

so that

$$\cos \rho x_0 = 0, \quad \cos \rho(a + x_o) = 0,$$

and hence

$$\rho x_0 = \pm \frac{\pi}{2}, \pm \frac{3\pi}{2}, \dots, \quad \rho(a + x_0) = \pm \frac{\pi}{2}, \pm \frac{3\pi}{2}, \dots,$$

and

$$\rho = \frac{n\pi}{a}, \quad n = 0, \pm 1, \pm 2, \dots. \tag{14.77}$$

(If we had instead of $\rho^2 \to -\rho^2$ and hence $f \propto \cosh \rho(x+x_0)$ or $\sinh \rho(x+x_0)$, the boundary conditions could not be satisfied in a nontrivial way). We conclude that

$$\begin{aligned}
f(x) &= A \cos \rho(x + x_0) \\
&= A(\cos \rho x \underbrace{\cos \rho x_0}_{0} - \sin \underbrace{\rho}_{\frac{n\pi}{a}} x \sin \rho x_0) \\
&= A' \sin \left(\frac{n\pi x}{a} \right), \quad n = 0, \pm 1, \pm 2, \pm 3, \dots. \tag{14.78}
\end{aligned}$$

We now consider the *second equation*, Eq. (14.76). Proceeding as in the above case we obtain

$$g(y) = B \cos \sqrt{\gamma^2 - \rho^2}(y + y_0), \quad B \neq 0, \tag{14.79}$$

and

$$E_z|_{\text{surface}} = 0 \quad \text{implies} \quad g(0) = 0, \quad g(b) = 0,$$

i.e.

$$\begin{aligned}
\cos \sqrt{\gamma^2 - \rho^2} y_0 = 0 &= \cos \sqrt{\gamma^2 - \rho^2}(b + y_0) \\
&= -\sin \sqrt{\gamma^2 - \rho^2} b \sin \sqrt{\gamma^2 - \rho^2} y_0,
\end{aligned}$$

so that

$$\sqrt{\gamma^2 - \rho^2} y_0 = \frac{\pi}{2}, \frac{3\pi}{2}, \dots, \quad \sqrt{\gamma^2 - \rho^2} b = m\pi,$$

and

$$g(y) = B' \sin\left(\frac{m\pi y}{b}\right), \quad m = 0, \pm 1, \ldots. \tag{14.80}$$

Summarising we find: *TM modes have the eigensolutions* of Eq. (14.72) given by

$$E_z(x, y) = C' \sin\left(\frac{n\pi x}{a}\right) \sin\left(\frac{m\pi y}{b}\right) \tag{14.81a}$$

with

$$\gamma^2 - \rho^2 = \frac{m^2 \pi^2}{b^2}, \quad \rho = \frac{n\pi}{a}. \tag{14.81b}$$

From this we obtain the *eigenvalue relation*

$$\gamma^2 = \gamma_\lambda^2, \quad \gamma_\lambda^2 \equiv \frac{m^2 \pi^2}{b^2} + \frac{n^2 \pi^2}{a^2}, \quad \lambda \equiv \{m, n\}. \tag{14.82}$$

The x- and y-components of $\mathbf{E}$ can now be obtained from E_z with the help of Eq. (14.20), or one can obtain these in some other way. Similar considerations apply in the case of the components of $\mathbf{B}$.

We saw that the following conditions hold at the boundary:

$$B_n^{(1)} = B_n^{(2)}, \quad E_t^{(1)} = E_t^{(2)},$$

i.e. for the fields from inside the wave guide

$$(\mathbf{n} \cdot \mathbf{B})_{\text{surface}} = 0, \quad (\mathbf{n} \times \mathbf{E})_{\text{surface}} = 0.$$

We thus obtain
(a) for planes $x = 0, a$, $\mathbf{n}$ parallel and antiparallel to $\mathbf{e}_x$:

$$B_x = 0, E_y = 0 = E_z, \quad \text{i.e. } E_x \neq 0;$$

(b) for planes $y = 0, b$, $\mathbf{n}$ parallel and antiparallel to $\mathbf{e}_y$:

$$B_y = 0, E_x = 0 = E_z, \quad \text{i.e. } E_y \neq 0.$$

We assume

$$\mathbf{E}(x, y, z, t) = \mathbf{E}(x, y)e^{i(kz - \omega t)}, \quad \mathbf{B}(x, y, z, t) = \mathbf{B}(x, y)e^{i(kz - \omega t)}.$$

The equations to be solved are (for $\mu = \mu_0, \epsilon = \epsilon_0, \epsilon_0 \mu_0 = 1/c^2$)

$$\left(\nabla_\perp^2 + \frac{\omega^2}{c^2} - k^2\right) \begin{cases} \mathbf{E}(x, y) = 0, \\ \mathbf{B}(x, y) = 0, \end{cases} \text{ with } E_z|_{\text{surface}} = 0, \left.\frac{\partial B_z}{\partial n}\right|_{\text{surface}} = 0.$$
$$\tag{14.83}$$

As an example we consider E_x (other components can be dealt with similarly). We therefore select the equation

$$\left(\nabla_\perp^2 + \frac{\omega^2}{c^2} - k^2\right) E_x(x, y) = 0. \tag{14.84}$$

The variables x, y can be separated, i.e. we put

$$E_x \propto (\sin \rho x \text{ or } \cos \rho x)(y-\text{part}),$$

so that

$$\left(\frac{\partial^2}{\partial y^2} - \rho^2 + \frac{\omega^2}{c^2} - k^2\right)(y-\text{part}) = 0.$$

Since E_x is nonzero for $x = 0, a$ (see boundary conditions), we choose $E_x \propto \cos \rho x, \rho = m\pi/a$. Since, however, E_x has to vanish for $y = 0, b$ (see boundary conditions), we choose for the y-part

$$\sin\left(\frac{y\pi n}{b}\right), \quad n = 0, 1, 2, \ldots,$$

so that

$$E_x = \alpha \cos\left(\frac{m\pi x}{a}\right) \sin\left(\frac{n\pi y}{b}\right). \tag{14.85}$$

We obtain the same expression with Eq. (14.20) from E_z. For E_x to satisfy Eq. (14.84), we must have

$$-\left(\frac{m^2\pi^2}{a^2} + \frac{n^2\pi^2}{b^2}\right) + \frac{\omega^2}{c^2} - k^2 = 0.$$

For the wave $\mathbf{E}(x, y, z, t) \propto e^{ikz}$ to propagate in the direction of z, we must have $k^2 > 0$, i.e. $\mathbf{E}$ must be periodic in z. This means

$$k^2 > 0 : \frac{\omega^2}{c^2} \geq \frac{\omega_0^2}{c^2}, \quad \omega_0 = c\sqrt{\frac{m^2\pi^2}{a^2} + \frac{n^2\pi^2}{b^2}}. \tag{14.86}$$

We see: ω must be larger than some *limiting frequency* ω_0. Proceeding similarly with all components $E_x, E_y, E_z, B_x, B_y, B_z$, we obtain the following expressions

$$\left.\begin{array}{l} E_x = \alpha \cos(\frac{m\pi x}{a}) \sin(\frac{n\pi y}{b}) \\ E_y = \beta \sin(\frac{m\pi x}{a}) \cos(\frac{n\pi y}{b}) \\ E_z = \gamma \sin(\frac{m\pi x}{a}) \sin(\frac{n\pi y}{b}) \\ B_x = \alpha' \sin(\frac{m\pi x}{a}) \cos(\frac{n\pi y}{b}) \\ B_y = \beta' \cos(\frac{m\pi x}{a}) \sin(\frac{n\pi y}{b}) \\ B_z = \gamma' \cos(\frac{m\pi x}{a}) \cos(\frac{n\pi y}{b}) \end{array}\right\} e^{i(kz-\omega t)}. \tag{14.87}$$

We refer to these expressions again later.[†]

TM waves have everywhere $B_z = 0$, i.e. $\gamma' = 0$. We deduce from Eq. (14.87) that these waves are different from zero if (to be shown) $m \neq 0, n \neq 0$. In order to see this, we recall: The expressions for $\mathbf{E}, \mathbf{B}$ must satisfy Maxwell's equations. Hence we consider these:

(a) $\boldsymbol{\nabla} \cdot \mathbf{E} = 0$ yields (apart from the factor $\sin(n\pi x/a)\sin(n\pi y/b)$)

$$-\alpha \frac{m\pi}{a} - \beta \frac{n\pi}{b} + ik\gamma = 0; \qquad (14.88)$$

(b) $\boldsymbol{\nabla} \cdot \mathbf{B} = 0$ yields

$$\alpha' \frac{m\pi}{a} + \beta' \frac{n\pi}{b} + ik\gamma' = 0; \qquad (14.89)$$

(c) $\boldsymbol{\nabla} \times \mathbf{E} = i\omega \mathbf{B}$ yields [‡]

$$\alpha' i\omega = \gamma \frac{n\pi}{b} - ik\beta,$$
$$\beta' i\omega = ik\alpha - \gamma \frac{m\pi}{a},$$
$$\gamma' i\omega = \beta \frac{m\pi}{a} - \alpha \frac{n\pi}{b}; \qquad (14.90)$$

(d) $\mu \boldsymbol{\nabla} \times \mathbf{H} = -i\omega \epsilon \mu \mathbf{E}$ yields

$$0 = i\omega\epsilon\mu\alpha - \gamma' \frac{n\pi}{b} - \beta' ik,$$
$$0 = i\omega\epsilon\mu\beta + ik\alpha' + \gamma' \frac{m\pi}{a},$$
$$0 = i\omega\epsilon\mu\gamma - \beta' \frac{m\pi}{a} + \alpha' \frac{n\pi}{b}. \qquad (14.91)$$

For TM waves $\gamma' = 0$. We show: For $m = 0, n \neq 0$ the entire field is zero. From (b) we have

$$\alpha' \frac{m\pi}{a} + \beta' \frac{n\pi}{b} = 0.$$

For $m = 0, n \neq 0 \rightarrow \beta' = 0$. It then follows from (14.91), since $\gamma' = 0, \beta' = 0 \rightarrow \alpha = 0$. Then $E_x = 0$ (since $\alpha = 0$), $E_y, E_z, B_x = 0$ (since $m = 0$), $B_y = 0$ (since $\beta' = 0$), $B_z = 0$ (since $\gamma' = 0$). Hence the entire field $= 0$ (can

---

[†]Later we write

$$\mathbf{B} = \left(\frac{\partial \Psi}{\partial y}, -\frac{\partial \Psi}{\partial x}, 0\right), \quad \mathbf{E} = \left(\frac{\partial \Pi}{\partial y}, -\frac{\partial \Pi}{\partial x}, 0\right),$$

so that according to Eq. (14.87), $\Psi \propto \sin\ldots\sin\ldots$ and $\Pi \propto \cos\ldots\cos\ldots$.

[‡]Observe that for $m = 1, n = 0$ one obtains $\alpha' = -\frac{k\beta}{\omega}, \gamma' = \frac{\beta\pi}{ia\omega}$. These expressions will be used in Example 14.2.

also be obtained faster from Eq. (14.69)). Thus there is only one limiting frequency ω_{11} which is such, that TM waves are possible only if

$$\omega > \omega_{11}, \quad \omega_{11} = c\sqrt{\frac{\pi^2}{a^2} + \frac{\pi^2}{b^2}}. \tag{14.92}$$

TE waves have everywhere $E_z = 0$, i.e. $\gamma = 0$. In this case a solution exists also in the case when either m is zero *or* n is zero. The lowest limiting frequency for $a > b, m = 1, n = 0$ is (for $m = 0, n = 0$, according to (b) $\gamma' = 0$, implies $E_x = 0, E_y = 0$, hence no field):

$$\omega_{10} = c\frac{\pi}{a} < \omega_{11}. \tag{14.93}$$

With this the fundamental wave is given by

$$E_x = 0, \; E_z = 0, \; E_y = \beta \sin\frac{\pi x}{a} e^{i(kz - \omega t)} = \frac{\beta}{2i}\left[e^{i(kz + \frac{\pi}{a}x - \omega t)} - e^{i(kz - \frac{\pi}{a}x - \omega t)} \right],$$

which is a superposition of two waves with $k_x = \pi/a$.

Previously we had defined a plane wave by the expression

$$e^{i(\mathbf{k}\cdot\mathbf{r} - \omega t)} \quad \text{for} \quad \mathbf{k}\cdot\mathbf{r} - \omega t = \text{const.}$$

The wave vector $\mathbf{k}$ is perpendicular to the plane with constant phase and points in the direction of propagation of the wave. The equations

$$y = 0, \quad kz \pm \frac{\pi}{a}x - \omega t = \text{const.} = \alpha$$

describe wave-fronts of plane waves. The gradients of the wave-fronts are therefore, see Fig. 14.5, according to the waves of these wave-fronts and at $t = 0$, i.e.

$$z = \frac{\alpha}{k} \mp \frac{\pi}{a}\frac{x}{k},$$

given by the derivative

$$\frac{dz}{dx} = \mp\frac{\pi}{ak} \equiv \mp\tan\theta$$

as indicated in Fig. 14.5. One should note here the angles ϵ of both wave-fronts with respect to the z-axis, and the angles $|\pi/2 - \epsilon|$ with respect to the x-axis. One wave-front can be looked at as the reflection of the other at the wall of the wave guide. The direction of propagation (i.e. of the normal to the wave front) is in both cases given by

$$\cos\epsilon = \pm\frac{\alpha/k}{\sqrt{(\alpha/k)^2 + (a\alpha/\pi)^2}} = \pm\frac{\pi/a}{\sqrt{(\pi/a)^2 + k^2}}.$$
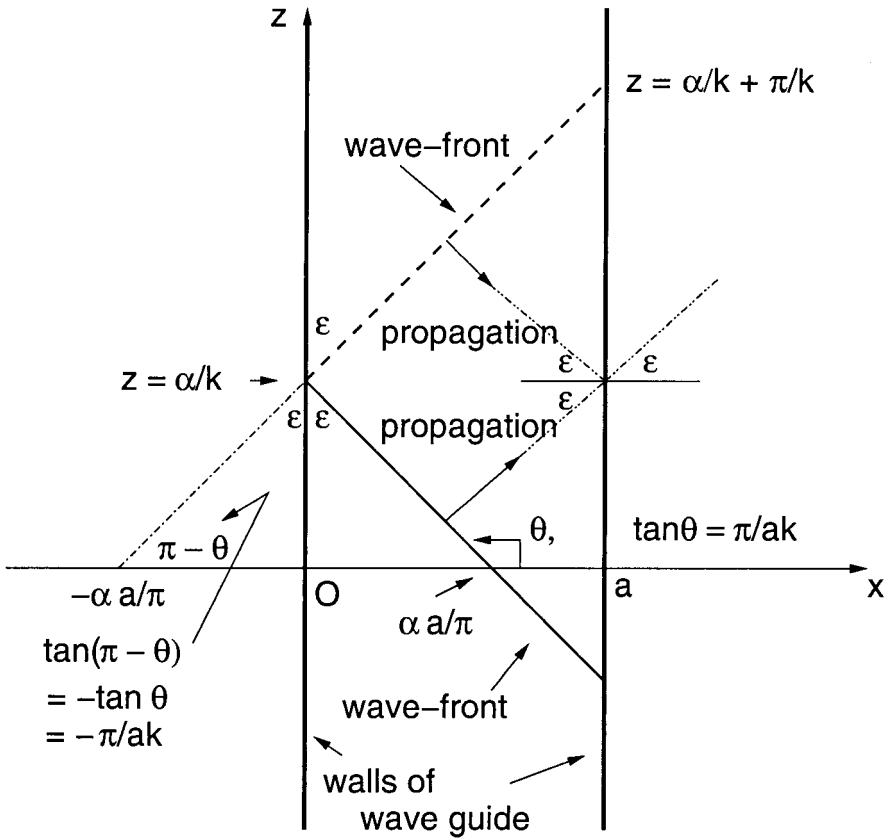
Fig. 14.5 Reflection at the walls of the wave guide.

However

$$\frac{\omega^2}{c^2} = k^2 + \frac{m^2\pi^2}{a^2} + \frac{n^2\pi^2}{b^2}.$$

In the case of the fundamental TE wave with $m = 1, n = 0$, we have

$$\frac{\omega^2}{c^2} = k^2 + \frac{\pi^2}{a^2} = k^2 + \frac{{\omega_{10}}^2}{c^2}, \quad \omega_{10}^2 = \frac{\pi^2}{a^2}c^2, \quad \frac{\pi}{a} = \frac{\omega_{10}}{c},$$

or

$$kc = \sqrt{\omega^2 - {\omega_{10}}^2}. \tag{14.94}$$

It follows that

$$\cos\epsilon = \mp\frac{\omega_{10}/c}{\omega/c} = \mp\frac{\omega_{10}}{\omega}.$$

(Later we shall need $1 - \cos^2\epsilon = 1 - (\omega_{10}/\omega)^2$). This means

for $\quad e^{i(kz \pm \frac{\pi}{a}x - \omega t)} \quad$ we have $\quad \cos\epsilon = \mp\frac{\omega_{10}}{\omega}.$

We see therefore: We can imagine the field $\mathbf{E}$ as arising from a continuous reflection of two superposed plane waves at the walls $x = 0, a$ which are incident on the walls at the angle ϵ, one wave permitting interpretation as the reflection of the other (with the correct phase difference). We note that at the critical frequency ω_{10} we have $k = 0$. In this case $\cos \epsilon = \pm 1$, i.e. $\epsilon = 0, \pm\pi, \pm 2\pi, \ldots$.

In our further exploration of wave guides with rectangular cross section we restrict ourselves to some general remarks.
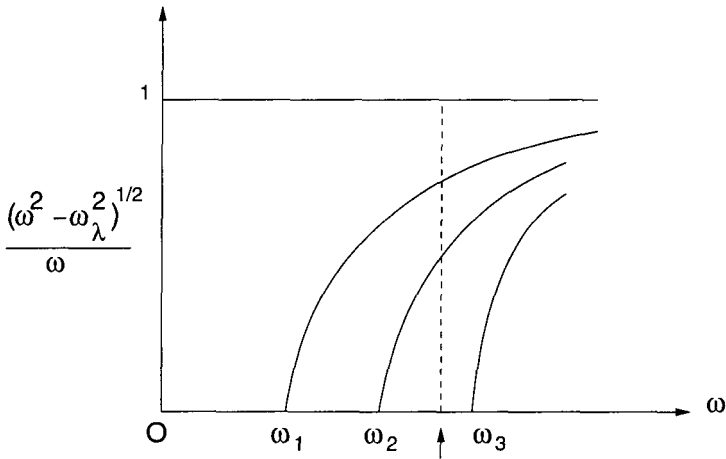

Fig. 14.6 At a given frequency ω only modes with eigenfrequency $\omega_\lambda < \omega$ are transmitted.

We let the eigenvalues of the eigenvalue problem of $E_{z\lambda}$ or $B_{z\lambda}$ be

$$\gamma^2 = \gamma_\lambda^2 > 0, \quad \lambda = 1, 2, \ldots.$$

Since $\gamma^2 = \mu\epsilon\omega^2 - k^2$, it follows that $\gamma_\lambda^2 = \mu\epsilon\omega^2 - k^2$, or

$$k^2 = \mu\epsilon\left(\omega^2 - \frac{\gamma_\lambda^2}{\mu\epsilon}\right).$$

With

$$\omega_\lambda^2 \equiv \frac{\gamma_\lambda^2}{\mu\epsilon}$$

we have

$$k = \sqrt{\mu\epsilon}\sqrt{\omega^2 - \omega_\lambda^2}. \tag{14.95}$$

In order for the waves to propagate in the wave guide proportional to $e^{i(kz-\omega t)}$ k must be real, i.e. $\omega^2 > \omega_\lambda^2$. Thus ω_λ is a *critical* or *limiting frequency*. This

is indicated in Fig. 14.6. This means that if $\omega_2 < \omega$, then $\lambda_2 > \lambda$, i.e. the wavelength must be less than the appropriate λ_2, if it is to propagate in the wave guide.

14.6.3   Wave Guides with Circular Cross Section

As another example we consider a wave guide with a circular cross section of radius $\rho = a$, as illustrated in Fig. 14.7, and made of ideal conductor material.
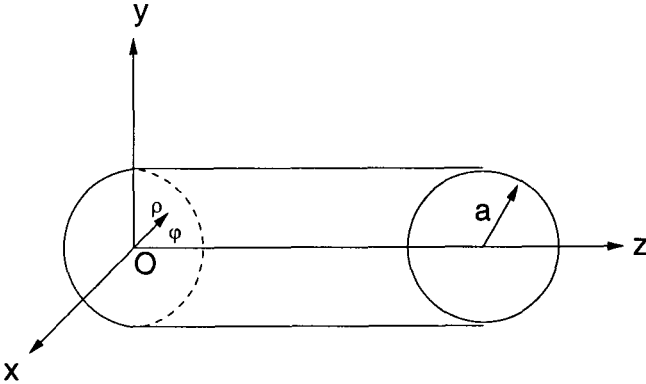

Fig. 14.7 Wave guide with circular cross section.

The boundary condition on E_z is $E_z| = 0$ for $\rho = a$, the angle φ being arbitrary. In planar polar coordinates Eq. (14.72) is

$$\left(\frac{\partial^2}{\partial \rho^2} + \frac{1}{\rho} \frac{\partial}{\partial \rho} + \frac{1}{\rho^2} \frac{\partial^2}{\partial \varphi^2} + \gamma^2 \right) E_z(\rho, \varphi) = 0. \tag{14.96}$$

We write

$$E_z(\rho, \varphi) = g(\rho) h(\varphi),$$

so that

$$\frac{1}{g(\varphi)} \left(\frac{\partial^2 g(\rho)}{\partial \rho^2} + \frac{1}{\rho} \frac{\partial g(\rho)}{\partial \rho} \right) + \frac{1}{\rho^2} \frac{1}{h(\varphi)} \frac{\partial^2 h(\varphi)}{\partial \varphi^2} + \gamma^2 = 0.$$

The variables ρ and φ can be separated by setting

$$\frac{1}{h(\varphi)} \frac{\partial^2 h(\varphi)}{\partial \varphi^2} = -\mu^2 \tag{14.97}$$

and

$$\frac{\partial^2 g(\rho)}{\partial \rho^2} + \frac{1}{\rho} \frac{\partial g(\rho)}{\partial \rho} + \left(\gamma^2 - \frac{\mu^2}{\rho^2} \right) g(\rho) = 0. \tag{14.98}$$

The solution of the first equation is

$$h(\varphi) = e^{\pm i \mu \varphi}, \quad \mu = \text{const.}$$

Clearly we must demand that $E_z(\varphi) = E_z(\varphi + 2\pi)$, i.e. $h(\varphi) = h(\varphi + 2\pi)$, and therefore we must have

$$e^{\pm i \mu 2\pi} = 1, \quad \text{i.e.} \quad \mu = m = 0, \pm 1, \pm 2, \ldots.$$

The second equation is then

$$\frac{\partial^2 g(\rho)}{\partial \rho^2} + \frac{1}{\rho} \frac{\partial g(\rho)}{\partial \rho} + \left(\gamma^2 - \frac{m^2}{\rho^2} \right) g(\rho) = 0.$$

The equation

$$\frac{\partial^2 Z(x)}{\partial x^2} + \frac{1}{x} \frac{\partial Z(x)}{\partial x} + \left(1 - \frac{m^2}{x^2} \right) Z(z) = 0 \qquad (14.99)$$

is the differential equation of the cylinder functions $J_m(x), Y_m(x)$, so that the general solution is

$$g(\rho) = A J_m(\gamma \rho) + B Y_m(\gamma \rho).$$

The solution $Y_m(x)$ is given by

$$Y_m(x) = \frac{2}{\pi} J_m(x) \ln(e^C x/2) + O(x), \quad C = \text{Euler constant},$$

the Bessel function $J_m(x)$ by

$$J_m(x) = \left(\frac{x}{2} \right)^m \sum_{l=0}^{\infty} \frac{\left(\frac{ix}{2} \right)^{2l}}{l!(m+l)!}.$$

The field $E_z(x, y)$ or $E_z(\rho, \varphi)$ must be finite for $\rho \to 0$. Since $Y_m(\gamma \rho)$ becomes infinite for $\rho \to 0$, we are compelled to impose the finiteness condition

$$B = 0, \quad m \geq 0.$$

We thus obtain for the TM modes

$$E_z(\rho, \varphi) = A J_m(\gamma \rho) e^{im\varphi}. \qquad (14.100)$$

The boundary condition $E_z = 0$ for $\rho = a$ implies the eigenvalue equation

$$J_m(\gamma a) = 0.$$

This is a transcendental equation with infinitely many roots. For the s-th positive root γa with $s \gg m$, an expansion is given in books on Special Functions, i.e.

$$\gamma a \sim \beta - \frac{\mu - 1}{8\beta} - \frac{4(\mu - 1)(7\mu - 31)}{3(8\beta)^3} + O(1/\beta^5),$$

where

$$\beta = \left(s + \frac{1}{2}m - \frac{1}{2} \right)\pi, \quad \mu = 4m^2.$$

In order to understand the reflection of the wave at the walls, we can argue in a way analogous to our treatment of the previous case (cf. after Eq. (14.93)), where we expressed the trigonometric function in terms of exponentials. This means in the present case, we have to re-express the Bessel function J_m in terms of Hankel functions $H^{(1)}, H^{(2)}$, i.e. we have to use the following formulae from books on Special Functions

$$J_m(R) = \frac{1}{2}\left[H_m^{(1)}(R) + H_m^{(2)}(R) \right]$$

with

$$H_m^{(1)}(R) = \sqrt{\frac{2}{\pi}} e^{i(R - \frac{m\pi}{2} - \frac{\pi}{4})} \left[1 + O\left(\frac{1}{R} \right) \right],$$

$$H_m^{(2)}(R) = \sqrt{\frac{2}{\pi}} e^{-i(R - \frac{m\pi}{2} - \frac{\pi}{4})} \left[1 + O\left(\frac{1}{R} \right) \right],$$

for $R \neq 0$, i.e. away from the axis of the wave guide. We do not enter into further details.

14.7 Alternative Treatment using Scalar and Vector Potentials

Our first step is to search for equations for $\mathbf{E}, \mathbf{B}$ in terms of the scalar potential ϕ and the vector potential $\mathbf{A}$. Again we assume that the interior of the wave guide is filled with some homogeneous material with electromagnetic constants ϵ, μ (e.g. those of the vacuum). Then inside the wave guide the following equations apply

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}, \quad \nabla \times \mathbf{H} = \frac{\partial \mathbf{D}}{\partial t} = \epsilon \dot{\mathbf{E}},$$

$$\mathbf{B} = \mu \mathbf{H}, \quad \mathbf{D} = \epsilon \mathbf{E}, \quad \nabla \cdot \mathbf{E} = 0, \quad \nabla \cdot \mathbf{B} = 0.$$

But, we also have

$$\mathbf{E} = -\dot{\mathbf{A}} - \nabla\phi,$$

and

$$\nabla \times \mathbf{B} = \nabla \times (\nabla \times \mathbf{A}) = \nabla(\nabla \cdot \mathbf{A}) - \nabla^2\mathbf{A}$$

and

$$\nabla \times \mathbf{B} = \mu\nabla \times \mathbf{H} = \mu\epsilon\dot{\mathbf{E}}.$$

Hence we can write

$$\mu\epsilon\dot{\mathbf{E}} = \nabla(\nabla \cdot \mathbf{A}) - \nabla^2\mathbf{A} = \mu\epsilon[-\ddot{\mathbf{A}} - \nabla\dot{\phi}]. \tag{14.101}$$

Now that we are using the vector potential $\mathbf{A}$, we still have the freedom to choose some *gauge fixing condition*. We choose again (as in our treatment of the Liénard–Wiechert potentials) the *Lorentz gauge*, i.e. we set

$$\nabla \cdot \mathbf{A} = -\mu\epsilon\dot{\phi}.$$

Then

$$\left(\nabla^2 - \mu\epsilon\frac{\partial^2}{\partial t^2}\right)\mathbf{A} = 0.$$

With the ansatz $\mathbf{A} \propto e^{-i\omega t}$ (as above for $\mathbf{E}, \mathbf{H}$) the following equation results

$$(\nabla^2 + \mu\epsilon\omega^2)\mathbf{A} = 0. \tag{14.102}$$

Hence from Eq. (14.101):

$$-i\mu\epsilon\omega\mathbf{E} = \nabla(\nabla \cdot \mathbf{A}) + \mu\epsilon\omega^2\mathbf{A} \tag{14.103}$$

with

$$\mathbf{B} = \mu\mathbf{H} = \nabla \times \mathbf{A}. \tag{14.104}$$

Equations (14.103) and (14.104) show: $\mathbf{E}, \mathbf{B}$ can be derived from $\mathbf{A}$, which is a solution of Eq. (14.102).

We consider now the TM and TE cases separately.

(a) *TM*: $B_z = 0$ everywhere.

We set

$$\mathbf{A} = (0, 0, \Psi).$$

Then

$$\mathbf{B} = \nabla \times \mathbf{A} = \begin{vmatrix} \mathbf{e}_x & \mathbf{e}_y & \mathbf{e}_z \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ 0 & 0 & \Psi \end{vmatrix} = \left(\frac{\partial\Psi}{\partial y}, -\frac{\partial\Psi}{\partial x}, 0\right)$$

and from Eq. (14.103),

$$-i\mu\epsilon\omega\mathbf{E} = \boldsymbol{\nabla}\left(\frac{\partial\Psi}{\partial z}\right) + \mu\epsilon\omega^2(0, 0, \Psi),$$

i.e.

$$-i\mu\epsilon\omega\mathbf{E} = \left(\frac{\partial^2\Psi}{\partial x\partial z}, \frac{\partial^2\Psi}{\partial y\partial z}, \frac{\partial^2\Psi}{\partial z^2} + \mu\epsilon\omega^2\Psi\right).$$

However, we still have the equation $\boldsymbol{\nabla}\cdot\mathbf{E} = 0$, i.e. from Eq. (14.103)

$$\frac{\partial}{\partial z}\left(\frac{\partial^2\Psi}{\partial x^2} + \frac{\partial^2\Psi}{\partial y^2} + \frac{\partial^2\Psi}{\partial z^2} + \mu\epsilon\omega^2\Psi\right) = 0,$$

and this is satisfied in view of Eq. (14.102), for $\mathbf{A} = (0, 0, \Psi)$, i.e.

$$\nabla^2\Psi + \mu\epsilon\omega^2\Psi = 0. \tag{14.105}$$

All fields can now be derived from the one scalar function Ψ, which is a solution of this equation. The function Ψ is frequently called *"stream function"*. The associated boundary condition is, as we saw,

$$\mathbf{E}_z|_{\text{surface}} = 0 \quad \left(\text{also} \quad \frac{\partial B_z}{\partial n}\bigg|_{\text{surface}} = 0, \text{ since } B_z = 0\right),$$

i.e. on the surface $E_{\text{tang}} = 0$. For the rectangular wave guide we have (see above)

$$\Psi \propto \sin\left(\frac{n\pi x}{a}\right)\sin\left(\frac{m\pi y}{b}\right)e^{ikz}e^{-i\omega t}.$$

(b) *TE*: $E_z = 0$ everywhere.

The relevant Maxwell equations are symmetric under the interchange

$$\mathbf{E} \leftrightarrow \mathbf{H} \quad \text{with} \quad -\omega\mu \leftrightarrow \omega\epsilon,$$

i.e. $\epsilon \leftrightarrow -\mu$. Thus in this case we have a scalar potential Π with

$$\nabla^2\Pi + \mu\epsilon\omega^2\Pi = 0, \tag{14.106}$$

where this time

$$\mathbf{E} = \left(\frac{\partial\Pi}{\partial y}, -\frac{\partial\Pi}{\partial x}, 0\right)$$

and

$$-i\mu\epsilon\omega\mathbf{H} = \left(\frac{\partial^2\Pi}{\partial x\partial z}, \frac{\partial^2\Pi}{\partial y\partial z}, \frac{\partial^2\Pi}{\partial z^2} + \mu\epsilon\omega^2\Pi\right).$$

Again all fields can be derived from the one scalar function Π which is a solution of the above equation with the boundary condition

$$\left.\frac{\partial B_z}{\partial n}\right|_{\text{surface}} = 0 \quad \text{and} \quad E_z = 0 \quad \text{everywhere.}$$

For the rectangular wave guide this means

$$\Pi \propto \cos\left(\frac{n\pi x}{a}\right) \cos\left(\frac{m\pi y}{b}\right) e^{ikz} e^{-i\omega t}.$$

14.8 Wave Velocities

We return to our considerations of rectangular wave guides. The wave or phase velocity of the TE fundamental mode with $m = 1, n = 0$ is (cf. Eq. (14.94) and thereafter)

$$\begin{aligned}
v_{\text{phase}} &= \frac{\omega}{k} = \frac{\omega c}{\sqrt{\omega^2 - \omega_{10}^2}} = \frac{c}{\sqrt{1 - (\omega_{10}/\omega)^2}} \\
&= \frac{c}{\sqrt{1 - \cos^2\epsilon}} = \frac{c}{\sin\epsilon} > c.
\end{aligned} \qquad (14.107)$$

The wavelength is $\lambda = 2\pi c/\omega$.

Next we consider the phase velocity in the general case, this means for

$$\omega_{10} \to \omega_0 = c\sqrt{\frac{m^2\pi^2}{a^2} + \frac{n^2\pi^2}{b^2}}, \quad \text{i.e.} \quad \frac{\omega^2}{c^2} = k^2 + \frac{\omega_0^2}{c^2}. \qquad (14.108)$$

In this general case the phase velocity is

$$v_{\text{phase}} = \frac{\omega}{k} = \frac{\omega}{\sqrt{\omega^2 - \omega_0^2}} c > c.$$

Thus the phase velocity is always larger than the velocity of light! In fact it becomes infinite at the critical frequency ω_0.

The velocity of energy transport is here in the case without attenuation, i.e. without loss of energy, the so-called *group velocity*. This group velocity v_g is defined by

$$v_g = \frac{d\omega}{dk} = \frac{d}{dk}\sqrt{k^2 c^2 + \omega_0^2} = \frac{kc^2}{\sqrt{k^2 c^2 + \omega_0^2}} = \frac{c\sqrt{\omega^2 - \omega_0^2}}{\omega} < c. \qquad (14.109)$$

We see that the group velocity of the wave in the wave guide is always less than the velocity of light and vanishes at the critical frequency. Finally we observe that

$$v_{\text{phase}} v_g = c^2. \qquad (14.110)$$

This is again a relation which is of wider generality than our derivation here.

14.9   Energy Transport in Wave Guides

14.9.1   The Complex Poynting Vector

In our earlier treatment of the Poynting vector we always assumed real quantities. We now wish to allow for complex fields, and complex ϵ, μ.

For the calculation of the transport of energy in wave guides we consider as a first step time averaged products of real vectors, which result from complex vectors and lead to the definition of the *complex Poynting vector*. This complex Poynting vector takes into account loss of energy due to complex ϵ and μ. Consider the real product,[§] in which we again separate the time dependence with the factor $e^{-i\omega t}$:

$$
\begin{aligned}
\mathbf{j}&(x,t) \cdot \mathbf{E}(x,t) \\
&= \Re[\mathbf{j}(x)e^{-i\omega t}] \cdot \Re[\mathbf{E}(x)e^{-i\omega t}] \\
&= \frac{1}{2}[\mathbf{j}(x)e^{-i\omega t} + \mathbf{j}^*(x)e^{i\omega t}] \cdot \frac{1}{2}[\mathbf{E}(x)e^{-i\omega t} + \mathbf{E}^*(x)e^{i\omega t}] \\
&= \frac{1}{4}[\mathbf{j}(x) \cdot \mathbf{E}(x)e^{-2i\omega t} + \mathbf{j}^*(x) \cdot \mathbf{E}^*(x)e^{2i\omega t} + \mathbf{j}(x) \cdot \mathbf{E}^*(x) + \mathbf{j}^*(x) \cdot \mathbf{E}(x)] \\
&= \frac{1}{2}\Re[\mathbf{j}(x) \cdot \mathbf{E}(x)e^{-2i\omega t} + \mathbf{j}^*(x) \cdot \mathbf{E}(x)]. \qquad (14.111)
\end{aligned}
$$

The time average of this is (averaged over one oscillation period $T = 2\pi/\omega$)

$$
\frac{1}{T}\int_0^T \mathbf{j}(x,t) \cdot \mathbf{E}(x,t)dt = \frac{1}{T}\int_0^T dt \frac{1}{2}\Re[\mathbf{j}(x) \cdot \mathbf{E}(x)e^{-2i\omega t} + \mathbf{j}^*(x) \cdot \mathbf{E}(x)]
$$

$$
= \frac{1}{2}\Re[\mathbf{j}^*(x) \cdot \mathbf{E}(x)] + \frac{1}{2}\Re\left\{\mathbf{j}(x) \cdot \mathbf{E}(x)\frac{1}{T}\int_0^T e^{-2i\omega t}dt\right\}.
$$

However

$$
\frac{1}{T}\int_0^T e^{-2i\omega t}dt = \frac{1}{-2i\omega T}[e^{-2i\omega t}]_0^{t=\frac{2\pi}{\omega}} = 0.
$$

Therefore

$$
\frac{1}{T}\int_0^T \mathbf{j}(x,t) \cdot \mathbf{E}(x,t)dt = \frac{1}{2}\Re[\mathbf{j}^*(x) \cdot \mathbf{E}(x)].
$$

Next we consider this expression integrated over a spatial volume V, i.e.

$$
\int_V dV \frac{1}{T}\int_0^T \mathbf{j}(x,t) \cdot \mathbf{E}(x,t)dt = \int_V dV \frac{1}{2}\Re[\mathbf{j}^*(x) \cdot \mathbf{E}(x)]
$$

[§]The Ohmic power is $\int dV \frac{1}{T}\int_0^T \mathbf{j} \cdot \mathbf{E}dt$ where $\mathbf{j}dV = Ids = (dq/dt)ds$.

$$= \int_V dV \frac{1}{2} \Re[\{ \nabla \times \mathbf{H}^*(x) - i\omega \mathbf{D}^*(x) \} \cdot \mathbf{E}(x)].$$

In this foregoing step we replaced $\mathbf{j}$ by the Maxwell equation

$$\nabla \times \mathbf{H} = \mathbf{j} + \frac{\partial \mathbf{D}}{\partial t} \quad \text{with} \quad \mathbf{D}(x,t) = e^{-i\omega t} \mathbf{D}(x),$$

and to go to the next line we use

$$\text{div}(\underbrace{\mathbf{u}}_{\mathbf{E}} \times \underbrace{\mathbf{v}}_{\mathbf{H}^*}) = \mathbf{v} \cdot \text{curl} \, \mathbf{u} - \mathbf{u} \cdot \text{curl} \, \mathbf{v},$$

so that we obtain

$$\int_V dV \frac{1}{T} \int_0^T \mathbf{j}(x,t) \cdot \mathbf{E}(x,t) dt = \int_V dV \frac{1}{2} \Re[\{ -\nabla \cdot (\mathbf{E}(x) \times \mathbf{H}^*(x))$$
$$+ \mathbf{H}^*(x) \cdot (\nabla \times \mathbf{E}(x)) \} - i\omega \mathbf{D}^*(x) \cdot \mathbf{E}(x)].$$

Here we use the Maxwell equation

$$\nabla \times \mathbf{E}(x,t) = -\frac{\partial \mathbf{B}(x,t)}{\partial t} = i\omega \mathbf{B}(x,t), \quad \nabla \times \mathbf{E}(x) = i\omega \mathbf{B}(x)$$

and obtain the expression

$$\int_V dV \frac{1}{T} \int_0^T \mathbf{j}(x,t) \cdot \mathbf{E}(x,t) dt = \int_V dV \frac{1}{2} \Re[-\nabla \cdot (\mathbf{E}(x) \times \mathbf{H}^*(x))$$
$$- i\omega(\mathbf{D}^*(x) \cdot \mathbf{E}(x) - \mathbf{H}^* \cdot \mathbf{B}(x))].$$

One now defines the *complex Poynting vector*

$$\mathbf{S} := \frac{1}{2}(\mathbf{E}(x) \times \mathbf{H}^*(x)) \tag{14.112}$$

and the *complex energy densities*

$$\omega_E = \frac{1}{2} \mathbf{E} \cdot \mathbf{D}^*(x), \quad \omega_M = \frac{1}{2} \mathbf{B}(x) \cdot \mathbf{H}^*(x). \tag{14.113}$$

Hence we obtain

$$\int_V dV \frac{1}{T} \int_0^T \mathbf{j}(x,t) \cdot \mathbf{E}(x,t) dt$$
$$= -\Re \left[\int_V dV \nabla \cdot \mathbf{S} + i\omega \int_V dV (\omega_E - \omega_M) \right],$$

or

$$\int_V dV \frac{1}{T} \int_0^T \mathbf{j}(x,t) \cdot \mathbf{E}(x,t) dt = \int_V dV \frac{1}{2} \Re[\mathbf{j}^*(x) \cdot \mathbf{E}(x)]$$

$$= -\Re\left[\int_{F(V)} \mathbf{S} \cdot d\mathbf{F} + i\omega \int_V dV(\omega_E - \omega_M) \right]. \tag{14.114}$$

This result is to be compared with the expression we had earlier. The minus sign in the second contribution originates from the appearance of $\mathbf{j}^*$ (instead of $\mathbf{j}$) in the preceding line. In the case of lossless conductors (i.e. with no damping) or dielectric media we have

$$\mathbf{B}(x) = \mu \mathbf{H}(x), \quad \mathbf{D}(x) = \epsilon \mathbf{E}(x), \quad \epsilon, \mu \text{ real,}$$

and the energy densities ω_E, ω_M are real. In these cases we obtain

$$\int_V dV \underbrace{\frac{1}{T} \int_0^T \mathbf{j}(x,t) \cdot \mathbf{E}(x,t) dt}_{\text{Ohmic power}} = - \int_{F(V)} \Re \mathbf{S} \cdot \mathbf{n} dF. \tag{14.115}$$

The left side of the equation represents the work (averaged over one period) done by the field $\mathbf{E}$ on the source current density $\mathbf{j}$ in the volume V. The expression $+ \int \Re \mathbf{S} \cdot \mathbf{n} dF$ represents the corresponding transport of energy through the boundary surface to outside, so that $- \int \Re \mathbf{S} \cdot \mathbf{n} dF$ is the corresponding flow into the volume V. In other words, the energy flow $- \int \Re \mathbf{S} \cdot \mathbf{n} dF$ into the volume V, corresponds to the work done by the field $\mathbf{E}$ on a corresponding current density in V. If ϵ and μ are complex, there are losses of energy which appear in the energy equivalence relation through $\Re[i\omega \int dV(\omega_E - \omega_M)]$.

14.9.2   Application of the Complex Poynting Vector

Our first considerations of energy transport assume that the wave guide is made of ideal conductor material. The primary step is to evaluate the complex Poynting vector for TM and TE waves. In the case of **TM waves** we had

$$B_z = 0 \text{ everywhere,} \quad \mathbf{B}_\perp = \frac{i\mu\epsilon\omega}{\gamma^2}(\mathbf{e}_z \times \boldsymbol{\nabla}_\perp)E_z,$$

$$\mathbf{E}_\perp = \frac{ik}{\gamma^2}\boldsymbol{\nabla}_\perp E_z, \tag{14.116}$$

with $k^2 = \mu\epsilon\omega^2 - \gamma^2, \gamma^2 \rightarrow \gamma_\lambda^2 = \mu\epsilon\omega_\lambda^2$. With this, we obtain $(\mathbf{E}(x) := \mathbf{E}(x,y))$

$$
\begin{aligned}
2\mathbf{S} &= \mathbf{E}(x) \times \mathbf{H}^*(x) \\
&= (\mathbf{E}_\perp(x), E_z(x)\mathbf{e}_z) \times (\mathbf{H}_\perp^*(x), 0) \\
&= (E_z(x)\mathbf{e}_z \times \mathbf{H}_\perp^*(x), \ \mathbf{E}_\perp(x) \times \mathbf{H}_\perp^*(x)) \\
&\equiv (2\mathbf{S}_\perp, 2S_z\mathbf{e}_z),
\end{aligned}
$$

where

$$
2\mathbf{S}_\perp = E_z(x)\mathbf{e}_z \times \mathbf{H}_\perp^*(x) \overset{(14.116)}{=} -(E_z(x)\mathbf{e}_z) \times \frac{i\epsilon\omega}{\gamma^2}(\mathbf{e}_z \times \mathbf{\nabla}_\perp)E_z
$$

$$
= E_z(x)\frac{i\epsilon\omega}{\gamma^2}\mathbf{\nabla}_\perp E_z.
$$

This contribution is, in general, purely imaginary, i.e. when ϵ, γ^2, E_z are real. The real part of $\mathbf{S}$ yields — as we have seen — the flow of energy averaged over one period of oscillation; in general $\mathbf{S}_\perp$ therefore does not contribute to this.

We restrict ourselves in the first place to wave guides made of perfect conductor material and *calculate the flow of energy in the longitudinal direction*, i.e. along the axis of the wave guide. In this case

$$
2S_z\mathbf{e}_z = \mathbf{E}_\perp(x) \times \mathbf{H}_\perp^*(x) \overset{(14.116)}{=} \frac{\epsilon\omega k}{\gamma^4}(\mathbf{\nabla}_\perp E_z) \times (\mathbf{e}_z \times \mathbf{\nabla}_\perp)E_z^*
$$

$$
= \frac{\epsilon\omega k}{\gamma^4}(\mathbf{\nabla}_\perp E_z) \cdot (\mathbf{\nabla}_\perp E_z)^*\mathbf{e}_z. \tag{14.117}
$$

The flow of energy through the cross section A of the wave guide is therefore

$$
\Re \int_A S_z\mathbf{e}_z \cdot d\mathbf{F} = \Re\frac{\epsilon\omega k}{2\gamma^4} \int_A (\mathbf{\nabla}_\perp E_z) \cdot (\mathbf{\nabla}_\perp E_z)^* dF. \tag{14.118}
$$

In Chapter 3 we used two Green's theorems of which the first was

$$
\int_V [\phi\nabla^2\psi + (\mathbf{\nabla}\phi) \cdot (\mathbf{\nabla}\psi)]dV = \oint_F \phi(\mathbf{\nabla}\psi) \cdot d\mathbf{F}.
$$

Reducing the space dimension by one and making the replacement $\phi \rightarrow \phi^*$ we obtain[¶]

$$
\int_A [\phi^*\nabla_\perp^2\psi + (\mathbf{\nabla}_\perp\phi)^* \cdot (\mathbf{\nabla}_\perp\psi)]dF = \oint_C \phi^*(\mathbf{\nabla}_\perp\psi) \cdot d\mathbf{l}.
$$

[¶]An analogous reduction of the space dimension by one can also be applied for instance to Gauss' divergence theorem, and yields $\int \mathbf{\nabla}_\perp \cdot \mathbf{A}dF = \oint_{\text{path}} \mathbf{A} \cdot d\mathbf{l}$.

Hence

$$\int_A S_z \mathbf{e}_z \cdot d\mathbf{F} = \frac{\epsilon \omega k}{2\gamma^4}\left[\oint_C \underbrace{E_z^*}_{0}(\boldsymbol{\nabla}_\perp E_z) \cdot d\mathbf{l} - \int_A E_z^* \nabla_\perp^2 E_z dF\right], \qquad (14.119)$$

where the path is to be taken along the surface. In view of the boundary condition $E_z|_{\text{surface}} = 0$ the first integral vanishes. The second integral can be rewritten with the help of Eq. (14.31), i.e.

$$(\nabla_\perp^2 + \gamma^2)E_z(x,y) = 0,$$

so that the energy transmitted by the wave guide per unit time is the power

$$\langle L \rangle = \Re \int_A S_z \mathbf{e}_z \cdot d\mathbf{F} = \Re \frac{\epsilon \omega k}{2\gamma^2} \int_A E_z^*(x,y)E_z(x,y)dF. \qquad (14.120)$$

The integral on the right is a normalisation integral of E_z. Since, as we saw,

$$k = \sqrt{\mu\epsilon}\sqrt{\omega^2 - \omega_\lambda^2}, \quad \gamma^2 = \gamma_\lambda^2 = \mu\epsilon\omega_\lambda^2, \qquad (14.121)$$

it follows that

$$\int_A S_z \mathbf{e}_z \cdot d\mathbf{F} = \frac{1}{2}\sqrt{\frac{\epsilon}{\mu}}(\frac{\omega}{\omega_\lambda})^2\sqrt{1 - \frac{\omega_\lambda^2}{\omega^2}} \int_A E_z^*(x,y)E_z(x,y)dF. \qquad (14.122)$$

For **TE waves** one obtains an analogous expression; in this case, however, the boundary condition which enters is $\partial B_z/\partial n|_{\text{surface}} = 0$.

14.9.3 Attenuation of Wave Guides ($\sigma \neq 0$)

Our next step is to consider the *attenuation* of wave guides, i.e. the loss of energy along the wave guide when the walls are made of a metal with *finite conductivity*. In the case of finite conductivity we observed earlier, the surface current density K does not exist. However, we then have the current density $\mathbf{j}$ in the non-ideal conductor, and we can define an *effective surface current* density $\mathbf{K}_{\text{eff}}$ by

$$\mathbf{K}_{\text{eff}} = \int \mathbf{j}(\rho)d\rho, \qquad (14.123)$$

where ρ is the depth of the current $\mathbf{j}(\rho)$ below the surface. We derived earlier the following expression for the Ohmic power of a conductor with conductivity σ:

$$L = \int \mathbf{j} \cdot \mathbf{E}dV. \qquad (14.124)$$

Averaging this over one period of oscillation for real ϵ and μ of the interior of the wave guide, and independent of the frequency, we obtain with (14.112)

$$\langle L \rangle = - \int_{F(V)} \Re \mathbf{S} \cdot \mathbf{n} dF = - \frac{1}{2} \int \Re \mathbf{n} \cdot \mathbf{E} \times \mathbf{H}^* dF, \qquad (14.125)$$

so that

$$\frac{d \langle L \rangle}{dF} = - \frac{1}{2} \Re [\mathbf{n} \cdot \mathbf{E} \times \mathbf{H}^*]. \qquad (14.126)$$

For $\mathbf{E}$ we have to substitute here the field beyond the (hollow) interior of the wave guide. We calculate this as follows. Neglecting the displacement current we obtain from Maxwell's two "curl" equations (for inside the metal)

$$\mathbf{\nabla} \times \mathbf{H} = \mathbf{j}, \quad \mathbf{\nabla} \times \mathbf{E} = - \frac{\partial \mathbf{B}}{\partial t}$$

with

$$\mathbf{E}, \mathbf{H} \propto e^{-i\omega t}, \quad \mathbf{B} = \mu \mathbf{H}, \qquad \mathbf{j} = \sigma \mathbf{E}$$

the relations

$$\mathbf{E} = \frac{1}{\sigma} \mathbf{\nabla} \times \mathbf{H}, \quad \mathbf{H} = \frac{1}{i\mu\omega} \mathbf{\nabla} \times \mathbf{E}. \qquad (14.127)$$

Inserting one equation into the other we obtain

$$\mathbf{H} = \frac{1}{i\omega\mu\sigma} \mathbf{\nabla} \times (\mathbf{\nabla} \times \mathbf{H}) = \frac{1}{i\omega\mu\sigma} [\mathbf{\nabla} \underbrace{(\mathbf{\nabla} \cdot \mathbf{H})}_{0} - \nabla^2 \mathbf{H}],$$

so that

$$\nabla^2 \mathbf{H} + i\omega\mu\sigma \mathbf{H} = 0. \qquad (14.128)$$

We let $-\rho\mathbf{n}$ be the coordinate pointing *into* the conductor or wall, and we make the assumption that the variation of $\mathbf{H}$ in the plane perpendicular to $\mathbf{n}$ is so slow that we have in good approximation

$$\nabla^2 \simeq \frac{\partial^2}{\partial \rho^2} \quad \text{(assumption)}.$$

We then have for the field *inside* the conductor

$$\frac{\partial^2}{\partial \rho^2} \mathbf{H} + i\omega\mu\sigma \mathbf{H} = 0$$

and (with $a = $ const. and $\sqrt{i} = (1+i)/\sqrt{2}$)

$$\mathbf{H}(x, y, z, t) = \mathbf{H}_0 e^{-i\omega t} e^{ikz} e^{\pm i \sqrt{i\omega\mu\sigma}(\rho-a)}$$
$$= \mathbf{H}_0 e^{-i\omega t} e^{ikz} e^{-(\rho-a)/\rho_0} e^{i(\rho-a)/\rho_0}, \qquad (14.129)$$

where $\mathbf{H}_0 = $ const. and

$$\rho_0 = \frac{\sqrt{2}}{\sqrt{\omega\mu\sigma}} \tag{14.130}$$

is the *skin depth*. For $\mathbf{E}$ we have to the same approximation

$$\mathbf{E} = \frac{1}{\sigma}\boldsymbol{\nabla} \times \mathbf{H} \simeq \frac{1}{\sigma}\boldsymbol{\nabla}_\rho \times \mathbf{H} \quad \parallel \quad \mathbf{e}_z \text{ for } TM. \tag{14.131}$$

We note that $\mathbf{H}_t = \mathbf{H}_0 e^{-i\omega t}e^{ikz}$ is the tangential field at (just outside) the surface of the conductor with $\rho = a$ as indicated in Fig. 14.8. Now,

$$\boldsymbol{\nabla}_\rho = -\mathbf{n}\frac{\partial}{\partial\rho} \equiv \mathbf{e}_\rho\frac{\partial}{\partial\rho}.$$

Here $\mathbf{n}$ is the unit vector pointing vertically out of the conductor. Then in the conductor (and just outside for $\rho = a$)

$$\underbrace{\mathbf{E}}_{\parallel\ \mathbf{e}_z\ \text{for TM}} = -\frac{1}{\sigma}(\mathbf{n} \times \mathbf{H})\frac{(i-1)}{\rho_0} \propto \frac{1}{\sqrt{\sigma}} \tag{14.132}$$

and
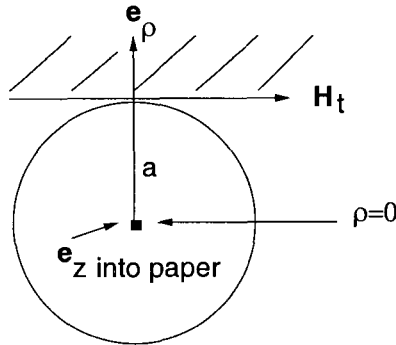
Fig. 14.8 The directions of vectors.

$$\mathbf{K}_{\text{eff}} = \int_a^\infty \mathbf{j}d\rho = \int_a^\infty \sigma\mathbf{E}d\rho = -\int_a^\infty \frac{\sigma}{\sigma\rho_0}(\mathbf{n} \times \mathbf{H})(i-1)d\rho$$

$$= -\left[\frac{1}{\rho_0}(\mathbf{n} \times \mathbf{H})\frac{(i-1)}{(i-1)}\rho_0\right]_a^\infty,$$

where in the last step one integrates over the two exponentials in Eq. (14.129) with respect to ρ leaving exponentials and hence again $\mathbf{H}$, so that with $\rho_0 = a$,

$$\mathbf{K}_{\text{eff}} = (\mathbf{n} \times \mathbf{H}_t), \quad \mathbf{H}_t = \mathbf{H}_0 e^{-i\omega t}e^{ikz}, \quad \mathbf{H}_t = \mathbf{H}(\rho = a, z, t). \tag{14.133}$$

One may observe that this relation has the same form as that for the surface current density K in the case of conductivity $\sigma = \infty$. We now have

$$
\begin{aligned}
|\mathbf{K}_{\text{eff}}|^2 &= (\mathbf{n} \times \mathbf{H}_t) \cdot (\mathbf{n} \times \mathbf{H}_t^*) \equiv \mathbf{n} \cdot [\mathbf{H}_t \times (\mathbf{n} \times \mathbf{H}_t^*)] \\
&= \mathbf{n} \cdot [(\mathbf{H}_t \cdot \mathbf{H}_t^*)\mathbf{n} - (\mathbf{H}_t \cdot \mathbf{n})\mathbf{H}_t^*] \\
&\simeq |\mathbf{H}_t|^2,
\end{aligned}
$$

since from

$$
\mathbf{H}_t = \frac{1}{i\omega\mu} \boldsymbol{\nabla} \times \mathbf{E}_t
$$

in the conductor, it follows that

$$
\mathbf{H}_t \cdot \mathbf{n} \simeq \frac{1}{i\omega\mu} \left\{ \left(-\mathbf{n}\frac{\partial}{\partial\rho} \right) \times \mathbf{E}_t \right\} \cdot \mathbf{n} = 0.
$$

We now calculate $\langle L \rangle$ by inserting in Eq. (14.126) the expression (14.132) for $\mathbf{E}$ in the conductor. We have

$$
\frac{d\langle L \rangle}{dF} = -\frac{1}{2}\Re[\mathbf{n} \cdot (\mathbf{E} \times \mathbf{H}^*)],
$$

and so

$$
\begin{aligned}
\frac{d\langle L \rangle}{dF} &\overset{(14.132)}{=} \frac{1}{2\sigma\rho_0}\Re[\ \underbrace{\mathbf{n} \cdot \{(\mathbf{n} \times \mathbf{H}) \times \mathbf{H}^*\}}_{} (i-1)] \\
&\qquad \underbrace{-\mathbf{n} \cdot \{\mathbf{H}^* \times (\mathbf{n} \times \mathbf{H})\}}_{-\mathbf{n}\cdot\{(\mathbf{H}^*\cdot\mathbf{H})\mathbf{n}-(\mathbf{H}^*\cdot\mathbf{n})\mathbf{H}\}} \\
&= \frac{1}{2\sigma\rho_0}\Re[-(i-1)|\mathbf{H}|^2] = \frac{1}{2\sigma\rho_0}|\mathbf{H}|^2,
\end{aligned}
$$

and hence

$$
\frac{d\langle L \rangle}{dF} \simeq \frac{1}{2\sigma\rho_0}|\mathbf{H}|^2. \tag{14.134}
$$

Here, as indicated in Fig. 14.9, $dF = dz\,dl$, where dl is an element of the path around the surface of the wave guide, so that (with the minus sign indicating that the expression represents a rate of loss)

$$
\begin{aligned}
\frac{d\langle L \rangle}{dz} &= -\frac{1}{2\sigma\rho_0}\oint_C |\mathbf{H}|^2 dl \\
&= -\frac{1}{2\sigma\rho_0}\oint_C |\mathbf{H}_t|^2 dl \\
&= -\oint_C \frac{1}{2\sigma\rho_0}|\mathbf{K}_{\text{eff}}|^2 dl \\
&= -\oint_C \frac{1}{2\sigma\rho_0}|\mathbf{n} \times \mathbf{H}_t|^2 \underbrace{dl}_{a\,d\varphi}. \tag{14.135}
\end{aligned}
$$

This is the loss of energy per unit time and per unit length of the wave guide as a consequence of the finite conductivity σ of its walls. For $\sigma \to \infty$ the right hand side of the attenuation rate (14.135) vanishes, as expected.
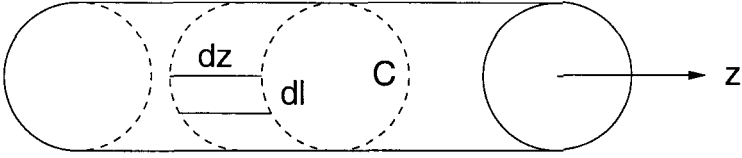

Fig. 14.9 The surface element $dF = dz dl$.

The integral is a line integral in the surface around the wave guide as indicated in Fig. 14.9. The expression for $d\langle L \rangle/dz$ involves only the value of the field $\mathbf{H}$ at the surface of the conductor.

We have seen earlier that the dispersion relation for conductors is

$$k^2 = \mu \epsilon \omega^2 \left(1 + i \frac{\sigma}{\omega \epsilon} \right). \tag{14.136}$$

We now have to distinguish between the material of the walls and that of the dielectric inside the wave guide. The walls have a high conductivity so that the skin depth is as small as possible and the wave guide is comparable to one made of ideal conductor material. On the other hand one wants the dielectric of the interior to have a conductivity as small as possible. Our earlier investigation of the behaviour of the conductivity σ was based on a simple model consideration for the calculation of the frequency dependence of σ. These considerations do not take into account for instance a possible dependence on boundary conditions or on the geometry of the conductor. In the following we shall argue that if σ is finite and large, and (as we assume) independent of the frequency, the propagation vector $\mathbf{k}$ of the electromagnetic wave develops an imaginary part due to the fact that — in this case of finite conductivity σ — the resistivity is finite so that damping is present and hence attenuation takes place. Thus we write

$$k = \sqrt{k^2} = k_R + i k_I$$

in the case of penetration into the walls of the wave guide. The propagation of the fields $\mathbf{E}, \mathbf{H} \propto e^{ikz}$ is therefore damped. The field $\mathbf{H}$ enters the power $\langle L \rangle$ quadratically. Hence we can write

$$\langle L \rangle = e^{-2k_I z} \langle L \rangle_0,$$

or

$$\frac{d\langle L\rangle}{dz} = -2k_I\langle L\rangle,$$

i.e.

$$k_I = -\frac{1}{2\langle L\rangle}\frac{d\langle L\rangle}{dz}. \tag{14.137}$$

The quantity k_I is the so-called *attenuation constant*. The intention is, of course, to achieve k_I as small as possible for an optimal use of the wave guide. If the interior of the wave guide is filled with a dielectric which has a small conductivity σ_0, this naturally also contributes to some damping, which we can take into account in the above dispersion relation. In this case we have *inside the wave guide* ("combining" (14.136) and $k^2 = \mu\epsilon\omega^2(1 - \omega_\lambda^2/\omega^2)$)

$$k^2 = \mu\epsilon\omega^2\left[1 - \frac{\omega_\lambda^2}{\omega^2}\right](1 + O(\sigma_0)). \tag{14.138}$$

In the next section we investigate an evaluation of the attenuation constant, and how this can be minimised for an optimal use of the wave guide.

14.9.4 Optimal Use of a Wave Guide

Our next objective is a model evaluation of k_I given by Eq. (14.137) for the case of a wave guide with circular cross section and filled with an ideal dielectric ($\sigma_0 = 0$ in Eq. (14.138)). This requires the calculation of $d\langle L\rangle/dz$. Assuming the conductivity of the walls of the wave guide is close to that of a perfect conductor, we can take (14.120) for the power $\langle L\rangle$. As an example we consider the case of **TM waves**. In this case we have (see Eqs. (14.69) and (14.67))

$$\mu\mathbf{H}_\perp = \mathbf{B}_\perp = \frac{i\mu\epsilon\omega}{\gamma^2}(\mathbf{e}_z \times \boldsymbol{\nabla}_\perp)E_z, \quad B_z = 0.$$

Here E_z is to be obtained as solution of the equation

$$(\boldsymbol{\nabla}_\perp^2 + \gamma^2)E_z(x, y) = 0$$

with boundary condition $E_z|_{\text{surface}} \simeq O(1/\sqrt{\sigma})$ (see Eq. (14.22)).[||] Thus our calculations will be concerned with E_z and its derivatives. In the case of the wave guide with circular cross section of radius a we had (cf. Eq. (14.100))

$$E_z(\rho, \varphi) = AJ_m(\gamma\rho)e^{im\varphi}, \quad E_z|_{\text{surface}} \propto J_m(\gamma a) \simeq 0.$$

[||]The condition $E_t^{(1)} - E_t^{(2)} = 0$ implies when $\sigma \neq \infty$, $E_t^{(1)} = O(1/\sqrt{\sigma})$. For very good conductors we can replace the right side by zero. For inclusion of such corrections see Jackson [3], Sec. 8.6, Perturbation of Boundary Conditions.

Hence we have in this case (TM: with $\mathbf{H}_z = 0$)

$$-\mathbf{n} \times \mathbf{H}_\perp = -\mathbf{n} \times \frac{i\epsilon\omega}{\gamma^2}(\mathbf{e}_z \times \boldsymbol{\nabla}_\perp)E_z$$

$$= -\frac{i\epsilon\omega}{\gamma^2}\mathbf{e}_z(\mathbf{n} \cdot \boldsymbol{\nabla}_\perp E_z). \qquad (14.139)$$

Here

$$\boldsymbol{\nabla}_\perp = \mathbf{e}_\rho\frac{\partial}{\partial\rho} + \mathbf{e}_\varphi\frac{1}{\rho}\frac{\partial}{\partial\varphi}$$

and

$$-\mathbf{n} \times \mathbf{H}_\perp = \frac{i\epsilon\omega}{\gamma^2}\mathbf{e}_z\frac{\partial E_z}{\partial\rho},$$

which is a vector along $\mathbf{e}_z$ derived from E_z (recall we have $\mathbf{H}_z = 0$ here). We can call this $\mathbf{H}^{\text{eff}}$ and insert this into Eq. (14.135) to obtain the loss of power which this gives rise to along the wave guide. Thus

$$\frac{d\langle L\rangle}{dz} \overset{(14.135)}{=} -\oint_C \frac{1}{2\sigma\rho_0}\left(\frac{\epsilon\omega}{\gamma^2}\right)^2\left|\frac{\partial E_z}{\partial\rho}\right|_{\rho=a}^2 ad\varphi$$

$$= -2\pi\frac{a(\epsilon\omega)^2}{2\sigma\rho_0\gamma^4}\left|\frac{\partial E_z}{\partial\rho}\right|_{\rho=a}^2$$

$$= -\frac{\pi a(\epsilon\omega)^2}{\sigma\rho_0\gamma^2}|A|^2\left(\frac{\partial J_m(z)}{\partial z}\right)_{z=\gamma a}^2$$

$$\equiv \tilde{A}. \qquad (14.140)$$

For the power $\langle L\rangle$ we derived the expression given by Eq. (14.120). We use this here with $dF = \rho d\varphi d\rho$ for the calculation of the main part of power arising from E_z:

$$L = \Re\int_A S_z\mathbf{e}_z \cdot d\mathbf{F}$$

$$= \Re\frac{\epsilon\omega k}{2\gamma^2}\int_A E_z{}^*(x,y)E_z(x,y)dF$$

$$= \Re\frac{\epsilon\omega k}{2\gamma^2}\int_A \rho d\varphi d\rho E_z{}^*(\rho,\varphi)E_z(\rho,\varphi)$$

$$= \Re\frac{\epsilon\omega k}{2\gamma^2}2\pi\int_{\rho=0}^a \rho d\rho|A|^2[J_m(\gamma\rho)]^2$$

$$\equiv \langle L\rangle \equiv \tilde{B}, \qquad (14.141)$$

where (see Eq. (14.138)) $\Re k \simeq \sqrt{\mu\epsilon}\sqrt{\omega^2 - \omega_\lambda^2}$, if we neglect a (generally small) boundary-dependent contribution to $\Re k$. We obtain therefore

$$
\begin{aligned}
k_I &= -\frac{1}{2\langle L\rangle}\frac{d\langle L\rangle}{dz} \equiv -\frac{1}{2}\frac{\tilde{A}}{\tilde{B}} \\
&= \frac{1}{2}\frac{\pi a(\epsilon\omega)^2}{\sigma\rho_0\gamma^2}\frac{\gamma^2}{\pi\epsilon\omega\Re k}\frac{(\partial J_m(z)/\partial z)^2_{z=\gamma a}}{\int_0^a \rho d\rho[J_m(\gamma\rho)]^2} \\
&= \frac{\epsilon\omega a}{2\sigma\rho_0\Re k}\frac{(\partial J_m(z)/\partial z)^2_{z=\gamma a}}{\int_0^a \rho d\rho[J_m(\gamma\rho)]^2}.
\end{aligned}
\tag{14.142}
$$

In order to keep the attenuation, i.e. the loss of energy, as small as possible, one wants to minimise this expression or, put differently, one wants to determine those modes for which k_I is minimal. We are therefore interested in the behaviour of this expression as a function of the frequency ω ($\gamma^2 = \mu\epsilon\omega^2 - k^2$). Thus we obtain with the substitution of

$$
\rho_0 = \rho_{0\lambda}\sqrt{\frac{\omega_\lambda}{\omega}}
$$

(since the skin depth $\rho_0 \propto 1/\sqrt{\omega}$, as we saw in Eq. (14.130)) and

$$
\gamma \to \gamma_\lambda, \quad \lambda = \{m, n\},
$$

the expression

$$
k_I = \frac{\epsilon a}{2\sigma\rho_{0\lambda}}\frac{\omega}{\sqrt{\frac{\omega_\lambda}{\omega}}}\frac{1}{\sqrt{\mu\epsilon}\sqrt{\omega^2 - \omega_\lambda^2}}f(\gamma_\lambda),
\tag{14.143}
$$

where

$$
f(\gamma_\lambda) = \left[\frac{(\partial J_m(z)/\partial z)^2_{z=\gamma a}}{\int_0^a \rho d\rho[J_m(\gamma\rho)]^2}\right]_{\gamma_\lambda}.
\tag{14.144}
$$

Thus

$$
k_I = \frac{a\sqrt{\frac{\epsilon}{\mu}}}{2\sigma\rho_{0\lambda}}\frac{\omega\sqrt{\frac{\omega}{\omega_\lambda}}}{\sqrt{\omega^2 - \omega_\lambda^2}}f(\gamma_\lambda) = \frac{a\sqrt{\frac{\epsilon}{\mu}}}{2\sigma\rho_{0\lambda}}\frac{\sqrt{\frac{\omega}{\omega_\lambda}}}{\sqrt{1 - \frac{\omega_\lambda^2}{\omega^2}}}f(\gamma_\lambda) \overset{\sigma\to\infty}{\to} 0.
$$

The best use is made of the wave guide if it is used for those modes whose eigenfrequencies are closest to the frequency at the minimum of k_I. The frequency at the minimum of the curve in Fig. 14.10 is easily determined by equating the derivative of k_I to zero,

$$
\frac{d}{d\omega}\left[\frac{\sqrt{\omega}}{\sqrt{1 - \frac{\omega_\lambda^2}{\omega^2}}}\right] = 0,
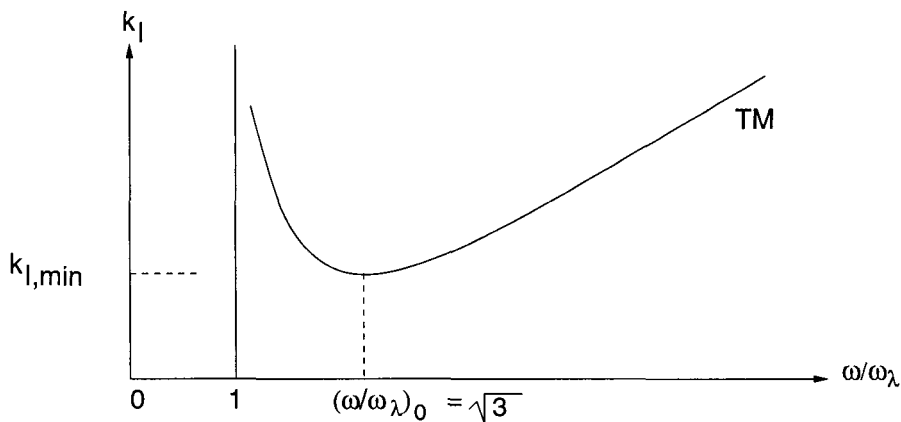$$

Fig. 14.10 The behaviour of k_I for TM modes.

i.e.

$$\left(1 - \frac{\omega_\lambda^2}{\omega^2}\right) = \omega\left(2\frac{\omega_\lambda^2}{\omega^2}\right), \quad \frac{\omega}{\omega_\lambda} = \sqrt{3}.$$

The behaviour of the attenuation constant near this minimum is shown in Fig. 14.10.

14.10   Resonators (Closed Wave Guides)

In the case of resonators or cavities we have additional boundary conditions at values of z, where the two ends of the resonator have their closures, and the boundary conditions there supply a third integer (recall that our considerations in Sec. 14.6 with the enforcement of boundary conditions, led to wave guide modes characterised by two integers). All these integers arise in analogy to quantum numbers in quantum mechanics, there each corresponding to quantisation of one degree of freedom. Here, of course, we are dealing with macroscopic physics, and the analogy is restricted to that of the mathematical eigenvalue problem. We assume that the cylindrical resonator is closed at both ends by plates made of the same conductor material as the body of the resonator. The electromagnetic waves can now also be reflected at the two ends.

We begin with **TM modes**, with time-dependence $e^{-i\omega t}$ (as before). In this case we have $\mathbf{B}_z = 0$ (everywhere), and we make the ansatz

$$\mathbf{E}_z = \psi(x, y)(A \sin kz + B \cos kz)e^{-i\omega t}\mathbf{e}_z. \qquad (14.145)$$

As before we obtain the transverse components from the relations derived earlier, i.e. Eqs. (14.20) and (14.21), and hence here, with $\gamma^2 = \epsilon\mu\omega^2 - k^2$,

from

$$\mathbf{E}_\perp = \frac{k}{\gamma^2}\boldsymbol{\nabla}_\perp\psi(x,y)(A\cos kz - B\sin kz)e^{-i\omega t},$$

$$\mathbf{B}_\perp = \frac{1}{\gamma^2}i\mu\epsilon\omega\mathbf{e}_z \times \boldsymbol{\nabla}_\perp\psi(x,y)(A\sin kz + B\cos kz)e^{-i\omega t}. \quad (14.146)$$

In view of the general boundary condition

$$E_t^{(1)} = E_t^{(2)}$$

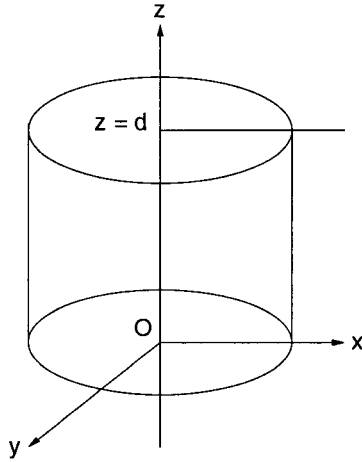we have at the ends: Inside $\mathbf{E}_\perp = 0$ for ideal conductors.

Fig. 14.11 The resonator.

Hence we have, with the geometry of the resonator as shown in Fig. 14.11,

$$\mathbf{E}_\perp(z=0) = 0 = \mathbf{E}_\perp(z=d),$$

and hence

$$A = 0 \quad \text{and} \quad \sin kd = 0, \quad \text{i.e.} \quad k = \frac{l\pi}{d}, \ l = 0, \pm 1, \pm 2, \ldots.$$

With this we have for the electric field

$$\mathbf{E}_z = B\cos\left(\frac{l\pi z}{d}\right)\psi(x,y)e^{-i\omega t}\mathbf{e}_z,$$

$$\mathbf{E}_\perp = -B\frac{l\pi}{d\gamma^2}\sin\left(\frac{l\pi z}{d}\right)\boldsymbol{\nabla}_\perp\psi(x,y)e^{-i\omega t}, \quad (14.147a)$$

and for the magnetic field

$$\mathbf{B}_z = 0,$$

$$\mathbf{B}_\perp = B\frac{i\mu\epsilon\omega}{\gamma^2}\cos\left(\frac{l\pi z}{d}\right)\mathbf{e}_z \times \boldsymbol{\nabla}_\perp\psi(x,y)e^{-i\omega t}. \qquad (14.147\mathrm{b})$$

For **TE modes** we have $\mathbf{E}_z = 0$ (everywhere), and we make the ansatz

$$\mathbf{B}_z = \psi(x,y)(A\sin kz + B\cos kz)\mathbf{e}_z e^{-i\omega t}. \qquad (14.148)$$

At the ends $\mathbf{B}_n$ is continuous (this condition $B_n^{(1)} = B_n^{(2)}$ is again of general validity), but outside the field is zero (as a consequence of the skin effect the fields do not penetrate into the metal except at extremely high frequencies). Hence we have $\mathbf{B}_z(z=0) = 0, \mathbf{B}_z(z=d) = 0$ implying

$$B = 0, \quad k = \frac{l\pi}{d}, \quad l = 0, \pm 1, \pm 2, \ldots.$$

We thus obtain for the electric field with Eq. (14.20)

$$\mathbf{E}_z = 0,$$

$$\mathbf{E}_\perp = -A\frac{i\omega}{\gamma^2}\sin\left(\frac{l\pi z}{d}\right)\mathbf{e}_z \times \boldsymbol{\nabla}_\perp\psi(x,y)e^{-i\omega t}, \qquad (14.149\mathrm{a})$$

and for the magnetic field with Eq. (14.21)

$$\mathbf{B}_z = A\psi(x,y)\sin\left(\frac{l\pi z}{d}\right)e^{-i\omega t}\mathbf{e}_z,$$

$$\mathbf{B}_\perp = A\frac{l\pi}{d\gamma^2}\cos\left(\frac{l\pi z}{d}\right)\boldsymbol{\nabla}_\perp\psi(x,y)e^{-i\omega t}. \qquad (14.149\mathrm{b})$$

Since $\mathbf{E}_z$ (TM) and $\mathbf{B}_z$ (TE) satisfy the equation

$$(\triangle_\perp + \gamma^2)\begin{Bmatrix}\mathbf{E}_z \\ \mathbf{B}_z\end{Bmatrix} = 0,$$

with boundary conditions $E_z|_{\text{surface}} = 0, (\partial B_z/\partial n)|_{\text{surface}} = 0$, it follows, that the scalar function $\psi(x,y)$ is in these cases solution of

$$(\triangle_\perp + \gamma^2)\psi(x,y) = 0$$

with the boundary condition

$$\psi|_{\text{surface}} = 0 \quad \text{or} \quad \left.\frac{\partial\psi}{\partial n}\right|_{\text{surface}} = 0,$$

where in both cases

$$\gamma^2 = \epsilon\mu\omega^2 - k^2 = \epsilon\mu\omega^2 - \left(\frac{l\pi}{d}\right)^2. \tag{14.150}$$

The equation for ψ is

$$\left(\frac{d^2}{dx^2} + \frac{d^2}{dy^2} + \gamma^2\right)\psi(x,y) = 0. \tag{14.151}$$

We now choose a *resonator of rectangular cross section* as illustrated in Fig. 14.12. The condition $\psi|_{\text{surface}} = 0$ implies that

- $\psi = 0$ on planes $x = 0, a, \ y \in [0, b], \ $ and
- $\psi = 0$ on planes $y = 0, b, \ x \in [0, a]$.

We choose the solution such that these conditions are satisfied, i.e.

$$\psi \propto \sin\left(\frac{m\pi x}{a}\right)\sin\left(\frac{n\pi y}{b}\right),$$
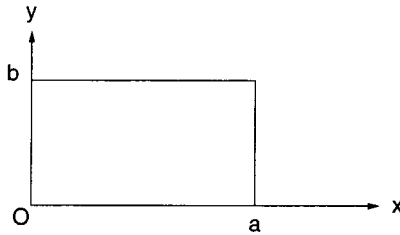
Fig. 14.12 Cross section of a rectangular resonator.

where m, n are integers. Inserting this into the equation for ψ, we obtain

$$-\left(\frac{m\pi}{a}\right)^2 - \left(\frac{n\pi}{b}\right)^2 + \gamma^2 = 0,$$

or

$$\epsilon\mu\omega^2 = \left(\frac{l\pi}{d}\right)^2 + \left(\frac{m\pi}{a}\right)^2 + \left(\frac{n\pi}{b}\right)^2. \tag{14.152}$$

We infer from this relation that the *eigenfrequencies* $\omega = \omega_{lmn}$ can be changed by changing the lengths a, b.

In the case of a *cylindrical resonator with circular cross section* of radius R as illustrated in Fig. 14.13, the function $\psi = \psi(\rho, \varphi)$ is given by the two-dimensional Laplace equation

$$(\triangle_\perp + \gamma^2)\psi(\rho, \varphi) = 0, \quad \triangle_\perp = \frac{\partial^2}{\partial\rho^2} + \frac{1}{\rho}\frac{\partial}{\partial\rho} + \frac{1}{\rho^2}\frac{\partial^2}{\partial\varphi^2}. \tag{14.153}$$

Using again the method of separation of variables and *demanding*, that

$$\psi(\rho, \varphi) = \psi(\rho, \varphi + 2m\pi),$$

we obtain

$$\psi(\rho, \varphi) = \psi(\rho)e^{im\varphi}, \quad m = 0, 1, 2, \ldots$$

with

$$\left(\frac{\partial^2}{\partial\rho^2} + \frac{1}{\rho}\frac{\partial}{\partial\rho} + \gamma^2 - \frac{m^2}{\rho^2}\right)\psi(\rho) = 0, \tag{14.154}$$

which is the Bessel differential equation. We thus have

$$\psi(\rho, \varphi) \propto J_m(\gamma\rho)e^{im\varphi} \tag{14.155}$$

and with Eq. (14.147a):

$$\mathbf{TM}: \quad \mathbf{E}_z = BJ_m(\gamma\rho)e^{im\varphi}\cos\left(\frac{l\pi z}{d}\right)e^{-i\omega t}\mathbf{e}_z. \tag{14.156}$$

The boundary condition

$$\psi(\gamma\rho)|_{\rho=R} = 0 \quad \text{implies} \quad J_m(\gamma R) = 0.$$

We let x_{mn} be the n-th zero $\neq 0$ ($n = 1, 2, 3, \ldots$) of $J_m(x)$. Then $\gamma R = x_{mn}$, i.e.

$$\gamma \equiv \gamma_{mn} = \frac{x_{mn}}{R}.$$
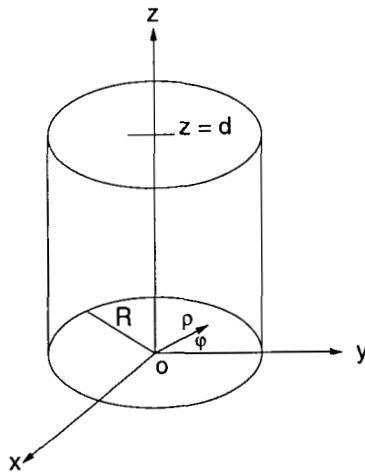
Fig. 14.13 The resonator with circular cross section.

For ω we obtain, since $\gamma^2 = \epsilon\mu\omega^2 - (l\pi/d)^2$,

$$\omega^2 = \frac{1}{\epsilon\mu}\left[\gamma^2 + \left(\frac{l\pi}{d}\right)^2\right] = \frac{1}{\epsilon\mu}\left[\left(\frac{x_{mn}}{R}\right)^2 + \left(\frac{l\pi}{d}\right)^2\right]. \tag{14.157}$$

The *lowest frequency* is ω_0:

$$\omega_0^2 = \frac{1}{\epsilon\mu}\left(\frac{x_{01}}{R}\right)^2, \quad (m = 0, n = 1, l = 0), \quad \text{where} \quad x_{01} \sim 2.4048.$$

We note, that ω_0 is independent of d. We obtain therefore for the corresponding **TM mode**:

$$\mathbf{E}_z = B J_0\left(\frac{2.4\rho}{R}\right)e^{-i\omega t}\mathbf{e}_z, \quad \mathbf{B}_\perp = \frac{i\mu\epsilon\omega_0}{\gamma_{01}^2}B\mathbf{e}_z \times \boldsymbol{\nabla}_\perp J_0\left(\frac{2.4\rho}{R}\right)e^{-i\omega t}.$$

In cylindrical coordinates

$$\boldsymbol{\nabla}_\perp = \mathbf{e}_\rho\frac{\partial}{\partial\rho} + \frac{1}{\rho}\mathbf{e}_\varphi\frac{\partial}{\partial\varphi}.$$

Hence

$$\boldsymbol{\nabla}_\perp J_0\left(\frac{2.4\rho}{R}\right) = \mathbf{e}_\rho\frac{\partial}{\partial\rho}J_0\left(\frac{2.4\rho}{R}\right),$$

so that

$$\mathbf{B}_\perp \equiv \mathbf{B}_\varphi = \frac{i\mu\epsilon\omega_0}{\gamma_{01}^2}B\frac{\partial}{\partial\rho}J_0(\gamma_{01}\rho)e^{-i\omega t}\underbrace{\mathbf{e}_\varphi}_{\mathbf{e}_z\times\mathbf{e}_\rho}, \quad \left(\gamma_{01} \simeq \frac{2.4}{R}\right).$$

However (see Tables of Special Functions)

$$\frac{d}{dx}J_m(x) = \frac{1}{2}(J_{m-1}(x) - J_{m+1}(x))$$

and $J_{-m}(x) = (-1)^m J_m(x)$, so that

$$\frac{d}{dx}J_0(x) = \frac{1}{2}(J_{-1}(x) - J_1(x)) = -J_1(x).$$

Hence

$$\mathbf{B}_\varphi = -i\frac{\mu\epsilon\omega_0}{\gamma_{01}^2}B\gamma_{01}J_1(\gamma_{01}\rho)e^{-i\omega t}\mathbf{e}_\varphi,$$

where

$$\gamma_{01} = \sqrt{\epsilon\mu}\omega_0,$$

and therefore

$$\mathbf{B}_\varphi = -i\sqrt{\mu\epsilon}BJ_1(\gamma_{01}\rho)e^{-i\omega t}\mathbf{e}_\varphi.$$

For **TE modes** we had Eqs. (14.149a) and (14.149b). These yield in the present case

$$B_z = AJ_m(\gamma_{mn}\rho)e^{im\varphi}\sin\left(\frac{l\pi z}{d}\right)e^{-i\omega t}. \tag{14.158}$$

The condition

$$\left.\frac{\partial B_z}{\partial n}\right|_{\rho=R} = \left.\frac{\partial B_z}{\partial \rho}\right|_{\rho=R} = 0$$

implies

$$\left.\frac{\partial}{\partial \rho}J_m(\gamma_{mn}\rho)\right|_{\rho=R} = 0.$$

We thus obtain the quantities γ_{mn} from the zeros of the derivative of the Bessel function. For the lowest frequency** we need the lowest root; this is ($l \neq 0$ from Eq. (14.158), so that $B_z \neq 0$)

$$J_1'(x_{11}) = 0 \quad \text{with} \quad x_{11} = 1.841,$$

so that $m = 1, n = 1$. Since

$$\omega^2 = \frac{1}{\epsilon\mu}\left[\gamma^2 + \left(\frac{l\pi}{d}\right)^2\right]$$

with $\gamma_{mn} = x_{mn}/R$ it follows that

$$\omega_{111}^2 = \frac{1}{\epsilon\mu}\left[\left(\frac{1.8}{R}\right)^2 + \left(\frac{\pi}{d}\right)^2\right]$$

($l = 1$; $B_z = 0$ for $l = 0$). We see that the fundamental frequency ω_{111} of the TE mode can be changed by changing d, the height of the cylindrical resonator (contrary to the fundamental frequency of the TM mode). Finally from Eq. (14.149b) we have

$$B_z = AJ_1(\gamma_{11}\rho)\sin\left(\frac{\pi z}{d}\right)e^{-i\omega t}e^{i\varphi}$$

(for the factor in φ see Eq. (14.155)). The resonator modes we investigated here find application for instance in accelerator technology.[††]

** The roots of $J_0'(x) = 0$ are $x = 3.8, 7.0, 10.1, \ldots$.

[††] See e.g. L. Palumbo and V. G. Vaccaro, *Wake field measurements*, Frascati-Report LNF-89/035(P)(1989). See also J. Slater, *Microwave Electronics* (D. Van Nostrand, 1950), L. C. Maier and J. C. Slater, "Field Strength Measurements in Resonance Cavities", *J. Appl. Phys.* **23** (1952) 68.

14.11   Examples

Example 14.1: **Vector potentials and wave guides**

Define in the charge-free interior of a wave guide the electric vector potential $\mathbf{G}$ by $\mathbf{E} = \nabla \times \mathbf{G}$. Similarly we have $\mathbf{B} = \nabla \times \mathbf{A}$. Find the wave equations of the vector potentials and sketch qualitatively the moduli of TE and TM modes of the two-dimensional fields.

Solution: In the charge-free space ($\rho = 0, \mathbf{j} = 0$) we can define an electric vector potential $\mathbf{G}$ by $\mathbf{E} = \nabla \times \mathbf{G}$, so that $\nabla \cdot \mathbf{E} = 0$. Maxwell's curl equations are here

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}, \quad \nabla \times \mathbf{B} = \mu \frac{\partial \mathbf{D}}{\partial t}.$$

From these we deduce

$$\mathbf{E} = -\frac{\partial \mathbf{A}}{\partial t}, \quad \mathbf{B} = \mu\epsilon \frac{\partial \mathbf{G}}{\partial t}.$$

Moreover

$$\nabla \times (\nabla \times \mathbf{G}) = -\frac{\partial}{\partial t}(\nabla \times \mathbf{A}), \; \nabla \times (\nabla \times \mathbf{A}) = \mu\epsilon \frac{\partial}{\partial t}(\nabla \times \mathbf{G}).$$

and with "*curl curl = grad div − div grad* " we obtain

$$\nabla(\nabla \cdot \mathbf{G}) - \triangle \mathbf{G} = -\left(\nabla \times \frac{\partial \mathbf{A}}{\partial t} \right) = -\frac{\partial \mathbf{B}}{\partial t} = -\mu\epsilon \frac{\partial^2 \mathbf{G}}{\partial t^2}$$

and

$$\nabla(\nabla \cdot \mathbf{A}) - \triangle \mathbf{A} = \mu\epsilon \left(\nabla \times \frac{\partial \mathbf{G}}{\partial t} \right) = \mu\epsilon \frac{\partial \mathbf{E}}{\partial t} = -\mu\epsilon \frac{\partial^2 \mathbf{A}}{\partial t^2}.$$

We choose the gauge fixing conditions

$$\nabla \cdot \mathbf{A} = 0, \quad \nabla \cdot \mathbf{G} = 0.$$

Then

$$-\triangle \mathbf{G} + \mu\epsilon \frac{\partial^2 \mathbf{G}}{\partial t^2} = 0,$$

and a similar equation holds for $\mathbf{A}$. TE modes are defined by $E_z = 0$. With $\mathbf{G} = (0, 0, \Pi)$, it follows that

$$\mathbf{E} = \nabla \times \mathbf{G} = \left(\frac{\partial \Pi}{\partial y}, -\frac{\partial \Pi}{\partial x}, 0 \right).$$

It may be noted that the potential $\mathbf{G}$ has in this case only a z-component (compare with $\mathbf{B}$ in Sec. 8.4.). As an example we take, with $m, n = 0, 1, \ldots$, as in Eq. (14.87),

$$\Pi = \cos \frac{m\pi x}{a} \cos \frac{n\pi y}{b} e^{ikz} e^{-i\omega t}, \quad \frac{\partial \Pi}{\partial y} = -\frac{n\pi}{b} \cos \frac{m\pi x}{a} \sin \frac{n\pi y}{b} e^{ikz} e^{-i\omega t},$$

$$\frac{\partial \Pi}{\partial x} = -\frac{m\pi}{a} \sin \frac{m\pi x}{a} \cos \frac{n\pi y}{b} e^{ikz} e^{-i\omega t}.$$

Lines of constant electric field strength are given by $|\mathbf{E}|^2 = $ const., i.e.

$$\left(\frac{n\pi}{b} \right)^2 \cos^2 \frac{m\pi x}{a} \sin^2 \frac{n\pi y}{b} + \left(\frac{m\pi}{a} \right)^2 \sin^2 \frac{m\pi x}{a} \cos^2 \frac{n\pi y}{b} = \text{const.}$$

The curves given by this equation are ellipses in the (x, y)-plane, as one can see by selecting specific modes. One can also see this simply by considering small values of x and y, so that the sine can be approximated by its argument. Then the Cartesian equation of an ellipse is obtained, and
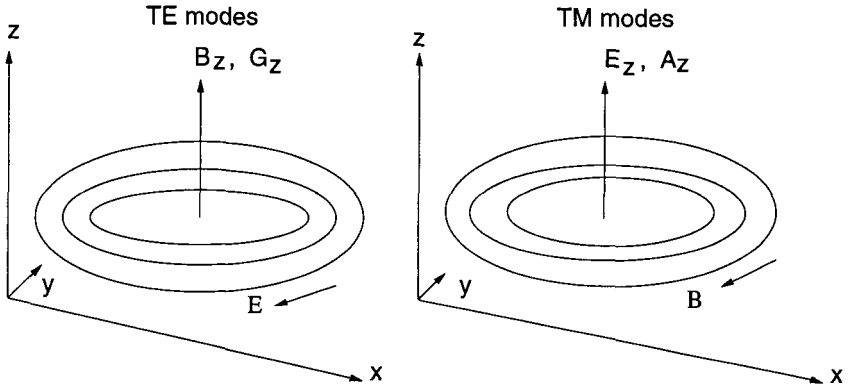
Fig. 14.14 *TE* and *TM* modes.

since lines of different $|\mathbf{E}|^2$ cannot cut, ellipses result also for other values of x and y. Analogous considerations apply to TM modes. In their case $B_z = 0$ and $\mathbf{A} = (0, 0, \Psi)$ with

$$\mathbf{B} = \left(\frac{\partial \Psi}{\partial y}, -\frac{\partial \Psi}{\partial x}, 0 \right).$$

The curves are sketched schematically in Fig. 14.14.

Example 14.2: Induced charges and currents

Consider the fundamental TE mode of a rectangular wave guide and calculate all charges and currents induced in its surfaces (the wave guide is made of ideal conductor material).

Solution: We have the equations (see summaries of formulae in Secs. 4.7 and 6.8)

$$(1). \ D_n^{(1)} - D_n^{(2)} = \sigma, \ \ \text{physical} \ \Rightarrow \Re(\epsilon_1 E_n^{(1)} - \epsilon_2 E_n^{(2)}) = \sigma,$$

$$(2). \ \mathbf{n} \times (\mathbf{H}^{(2)} - \mathbf{H}^{(1)}) = \mathbf{K},$$

from which σ and $\mathbf{K}$ are to be calculated (definition of surface current density $\mathbf{K}$ in Sec. 11.2). In Sec. 14.6 we obtained the fundamental TE wave for $m = 1$ and $n = 0$. We have therefore (cf. Eq. (14.87)):

$$\mathbf{E} = \mathbf{e}_y \beta \sin \frac{\pi x}{a} e^{i(kz - \omega t)}.$$

Since $\mathbf{E} \parallel \mathbf{e}_y$, it follows that (see Fig. 14.15 for explanation of subscripts) $\sigma_① = 0 = \sigma_③, \sigma_② = -\sigma_④$. Calculation of $\sigma_②$: Analogous to the case of the parallel plate condenser we have $D_n^{(1)} = 0$. Hence the surface charge density on surface ② is (with ϵ the dielectric constant of the inside of the wave guide and the surface normal pointing in the direction of $-\mathbf{e}_y$)

$$\sigma_② = \Re(-D_n^{(2)}) = \Re(-)(-)\epsilon\beta \sin \frac{\pi x}{a} e^{i(kz - \omega t)} = \epsilon\beta \sin \frac{\pi x}{a} \cos(kz - \omega t),$$

$$\therefore \sigma_② = \frac{\beta}{\epsilon} \sin \frac{\pi x}{a} \cos(kz - \omega t).$$

For the field $\mathbf{B}$ we have (see (14.87) and (14.90) for $m = 1, n = 0$)

$$\mathbf{B} = \beta \left[-\mathbf{e}_x \frac{k}{\omega} \sin \frac{\pi x}{a} + \mathbf{e}_z \frac{\pi}{ai\omega} \cos \frac{\pi x}{a} \right] e^{i(kz - \omega t)},$$

and we have to evaluate
$$\mathbf{n} \times (\mathbf{H}^{(2)} - \mathbf{H}^{(1)}).$$

Since for the *TE* wave (cf. Eq. (14.70))

$$\mathbf{E}_\perp = -\frac{\omega}{k}(\mathbf{e}_z \times \mathbf{B}_\perp), \text{ with } \mathbf{e}_z \nparallel \mathbf{B}_\perp,$$

we have $\mathbf{B}_\perp = 0$ for $\mathbf{E}_\perp = 0$, i.e. outside of the wave guide $\mathbf{B} = 0$, i.e. $\mathbf{H}^{(1)} = 0$. It remains to
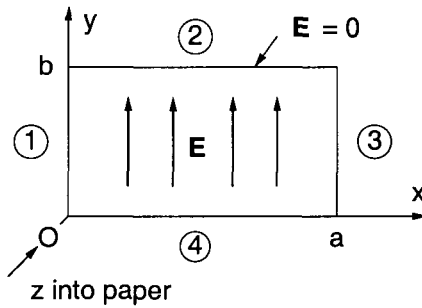

Fig. 14.15 Cross section through the rectangular wave guide.

calculate $\mathbf{n} \times \mathbf{H}$ for the respective surfaces as shown in Fig. 14.15 (therewith obtaining the surface current densities $\mathbf{K}$).

Surface ②: $\mathbf{n} = \mathbf{e}_y, y = b$.

$$\mathbf{K} = \mathbf{n} \times \mathbf{H} = \Re\left\{\frac{1}{\mu}\mathbf{e}_y \times \beta\left[-\mathbf{e}_x\frac{k}{\omega}\sin\frac{\pi x}{a} + \mathbf{e}_z\frac{\pi}{ai\omega}\cos\frac{\pi x}{a}\right]e^{i(kz-\omega t)}\right\}$$

$$= -\frac{\beta}{\mu}\left[\frac{k}{\omega}\sin\frac{\pi x}{a}\cos(kz-\omega t)\mathbf{e}_z + \frac{\pi}{a\omega}\cos\frac{\pi x}{a}\sin(kz-\omega t)\mathbf{e}_x\right].$$

Surface ①: $\mathbf{n} = \mathbf{e}_x, x = a$.

$$\mathbf{K} = \mathbf{n} \times \mathbf{H} = \Re\left\{\frac{1}{\mu}\mathbf{e}_x \times \beta\left[-\mathbf{e}_x\frac{k}{\omega}\sin\frac{\pi x}{a} + \mathbf{e}_z\frac{\pi}{ai\omega}\cos\frac{\pi x}{a}\right]e^{i(kz-\omega t)}\right\}\Bigg|_{x=a}$$

$$= \frac{\beta}{\mu}\frac{\pi}{a\omega}\sin(kz-\omega t)\mathbf{e}_y.$$

Example 14.3: Determination of modes

Consider two parallel, infinitely extended, perfectly conducting walls in the planes $y = 0, b$, as indicated in Fig. 14.16. The electric vector of the incident *TE* wave is polarised in the direction of x, i.e.

$$\mathbf{E}_0 = \mathbf{e}_x E_0^0 \exp[-i(\omega t - \mathbf{k}_0 \cdot \mathbf{r})]$$

with $\mathbf{r} = (0, y, z)$. Show that only waves with frequency

$$\omega_n = \frac{n\pi c}{b\cos\theta_0}, \quad n = 0, 1, 2, \dots$$

are propagated between the planes, where θ_0 is the angle of incidence between $\mathbf{k}_0$ and $\mathbf{r}$.

Solution: According to the geometry depicted in Fig. 14.16 the vector $\mathbf{k}_0$ is perpendicular to the plane of equal phases $\mathbf{k}_0 \cdot \mathbf{r} = \text{const}$. This means
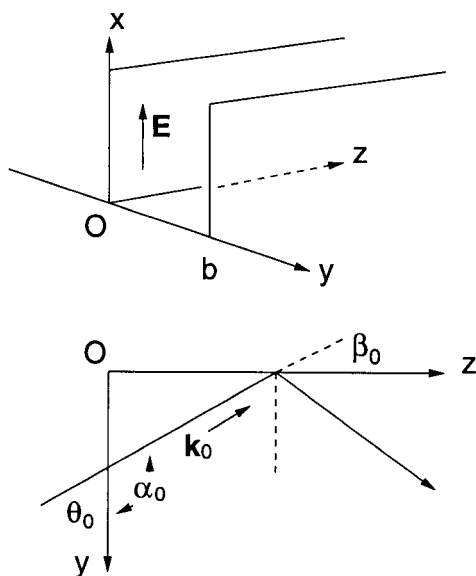
Fig. 14.16 Reflection between planes.

$$k_0 y \cos(\underbrace{\mathbf{k_0}, \mathbf{y}}_{\alpha_0 = \pi - \theta_0}) + k_0 z \cos(\underbrace{\mathbf{k_0}, \mathbf{z}}_{\beta_0 = \frac{\pi}{2} - \theta_0}) = \text{const.},$$

i.e.

$$-k_0 y \cos \theta_0 + k_0 z \sin \theta_0 = \text{const.},$$

so that

$$\mathbf{E}_0 = \mathbf{e}_x E_0^0 e^{-i\omega t} e^{ik_0(-y \cos \theta_0 + z \sin \theta_0)}.$$

Since $\mathbf{E}_0$ has to satisfy the wave equation, and there is no propagation of the wave in x, we see that its dispersion relation is

$$\frac{\omega^2}{c^2} = k_y^2 + k_z^2 = k_0^2 \cos^2 \theta_0 + k_0^2 \sin^2 \theta_0 = \mathbf{k}_0^2.$$

In vacuum $\omega = k_0 c$. (Here we have total reflection since the wall is a perfect conductor with $\sigma = \infty$, i.e. no radiation or energy penetrates into the wall). The $\mathbf{E}$ wave totally reflected from the wall, can also be written with the help of Fig. 14.16 as

$$\mathbf{E}_0'' = \mathbf{e}_x E_0'' e^{-i\omega t} e^{ik_0(y \cos \theta_0 + z \sin \theta_0)}.$$

The continuity condition of the field $\mathbf{E}$, i.e.

$$(\text{total } \mathbf{E}_{\text{tang}})|_{\text{wall}} = 0$$

($E_{\text{tang}}^{(1)} = E_{\text{tang}}^{(2)} = 0$ in the perfect conductor, see skin effect in Sec. 9.5) implies, as shown in Sec. 11.4.2 (cf. Eq. (11.52)),

$$E_0'' = -E_0^0, \quad (E_0' = 0).$$

Hence

$$\mathbf{E}_0'' = -\mathbf{e}_x E_0^0 e^{-i\omega t} e^{ik_0(y \cos \theta_0 + z \sin \theta_0)}.$$

The entire field $\mathbf{E}$ at (x, y, z, t) is then:

$$\mathbf{E} = \mathbf{E}_0 + \mathbf{E}_0'' = \mathbf{e}_x E_0^0 e^{-i\omega t} e^{ik_0 z \sin \theta_0} [e^{-ik_0 y \cos \theta_0} - e^{ik_0 y \cos \theta_0}]$$
$$= -2i\mathbf{e}_x E_0^0 e^{-i\omega t} e^{ik_0 z \sin \theta_0} \sin(k_0 y \cos \theta_0).$$

The boundary condition at the wall referred to above implies at $y = 0, b$:

$$\sin(k_0 b \cos \theta_0) = 0, \quad \therefore \quad k_0 b \cos \theta_0 = n\pi, \quad \underbrace{n = 1, 2, \dots}_{\text{mode numbers}}.$$

The dispersion relation is now seen to be

$$\frac{\omega^2}{c^2} = k_0^2 \cos^2 \theta_0 + k_0^2 \sin^2 \theta_0$$
$$= k_0^2 \sin^2 \theta_0 + \left(\frac{n\pi}{b}\right)^2$$
$$= \left(\frac{n\pi}{b \cos \theta_0}\right)^2.$$

For wavelength $\lambda \rightarrow \lambda_0 = 2\pi/k_0$ we have

$$\lambda_0 = \frac{2\pi b \cos \theta_0}{n\pi} = \frac{2b \cos \theta_0}{n}$$

and frequency

$$\omega_{0n} = k_0 c = \frac{n\pi c}{b \cos \theta_0}.$$

For a given angle of incidence θ_0 only waves with frequencies

$$\omega_{0n} = \frac{n\pi c}{b \cos \theta_0}, \quad n = 1, 2, \dots$$

are propagated between the walls. In the text we considered cylindrical wave guides with modes characterised by two integers, and closed resonators (or cavities) with modes characterised by three integers. In the above case the waves are confined only in one dimension, and hence its modes are characterised by only one integer.

This page intentionally left blank

Chapter 15

Propagation of Waves in Media

15.1 Introductory Remarks

In this chapter we introduce the concepts of signals and wave packets and demonstrate the intricate connection between the analytic properties of the Fourier transform of the wave packet (also called its spectral function) and the limitation of its velocity by the velocity of light.

15.2 Dispersion Relation: Normal and Anomalous Dispersion

In preceding chapters we derived and became familiar with the dispersion relation of a conducting medium, i.e.

$$k^2 = \omega^2 \eta \mu = \mu\epsilon\omega^2 \left(1 + i\frac{\sigma}{\omega\epsilon}\right), \qquad (15.1)$$

which also defines the generalised dielectric constant η (one should note that in general the relation is more complicated, in particular, as we saw in Chapter 9, if the frequency dependence of the conductivity σ is taken into account). We arrived at the expression (15.1) by assuming

$$\mathbf{j} = \sigma\mathbf{E}, \quad \mathbf{H} \propto e^{\mp i\omega t}e^{i\mathbf{k}\cdot\mathbf{r}},$$

together with the two Maxwell curl equations, and we obtained

$$\mathbf{k}^2\mathbf{H} = \mu\epsilon\omega^2 \left(1 + i\frac{4\pi\sigma}{\omega\epsilon}\right)\mathbf{H}.$$

The relation (15.1) is a general consequence if we assume that $\mathbf{H}$ does not vanish anywhere in space. We now want to study the dispersion relations as functions of the frequency ω.

We had also defined previously the generalised refractive index

$$n(k) \equiv p = \frac{\sqrt{\mu\eta}}{\sqrt{\mu_0\epsilon_0}}, \quad \sqrt{\mu_0\epsilon_0} = \frac{1}{c}, \quad c' = \frac{1}{\sqrt{\mu\eta}}, \quad p = \frac{c}{c'}, \tag{15.2}$$

so that

$$n(k) = \frac{ck}{\omega} \equiv \tilde{n}(\omega) \quad \text{or} \quad \omega(k) = \frac{ck}{n(k)} = \frac{ck}{\tilde{n}(\omega)}. \tag{15.3}$$

The *phase velocity* v_P and the *group velocity* v_G defined previously for real $n(k)$ or $\tilde{n}(\omega)$ are given by

$$v_P = \frac{\omega}{k} = \frac{c}{n(k)} = \frac{c}{\tilde{n}(\omega)},$$

$$v_G = \frac{d\omega}{dk} = \frac{c}{n(k)} - \frac{ck(dn(k)/dk)}{[n(k)]^2}$$

$$= \frac{c}{\tilde{n}(\omega)} - \frac{ck(d\tilde{n}(\omega)/d\omega)(d\omega/dk)}{[\tilde{n}(\omega)]^2}. \tag{15.4}$$

Solving the latter for $d\omega/dk$ we obtain

$$v_G = \frac{d\omega}{dk} = \frac{c}{\tilde{n}(\omega) + \omega(d\tilde{n}(\omega)/d\omega)}. \tag{15.5}$$

One says, there is no dispersion if $d\tilde{n}(\omega)/d\omega = 0$.

We now distinguish between two cases:

(1). The case of *normal dispersion* defined by

$$\frac{d\tilde{n}(\omega)}{d\omega} > 0, \quad \text{so that } v_G < v_P < c, \ \tilde{n}(\omega) > 1, \tag{15.6}$$

(in general $\epsilon/\epsilon_0 \geq 1$), and

(2). The case of *anomalous dispersion* defined by

$$\frac{d\tilde{n}(\omega)}{d\omega} < 0, \quad \text{for which } v_G > v_P, \tag{15.7}$$

which would imply $v_G > c$. This really means that the relation (15.5) then breaks down (see below), i.e. one has to take the real part when $\tilde{n}$ is complex!

The amplitude of a wave group moves (as one deduces from a superposition of plane waves) with the group velocity. Since the energy of a wave is determined by its amplitude, this implies that the energy of the wave is

transmitted with the group velocity. However, we shall see in the following that no energy can be transported with a velocity $> c$. In the following we consider frequency superpositions of waves with phases

$$\delta(\omega) = k(\omega)x - \omega t.$$

A *stationary phase* is given by $\delta'(\omega) = 0$, i.e.

$$\frac{dk(\omega)}{d\omega}x - t = 0, \quad \text{or} \quad \frac{d\omega(k)}{dk} = \frac{x}{t}.$$

The ratio x/t is described as *signal velocity*. In the present case this is equal to the group velocity. In general the group velocity of the dominant frequency can be treated as the signal velocity and hence as the velocity of the transport of energy. The Theory of Relativity teaches, that no physical velocity can exceed the velocity of light c (often cited as phenomenological input of the Special Theory of Relativity, but could also be considered as a consequence of the principle of relativity). In cases where

$$\frac{d\omega}{dk} = \frac{c}{n(k) + \omega(dn(k)/d\omega)} > c,$$

the signal velocity cannot be identified with $d\omega/dk$. The wave packet then has a more complicated structure, so that the expansion about a stationary phase is not possible. The aim of this expansion is to use the value k_0, for which $\delta'(\omega) = 0$ as a point around which $\omega(k)$ is to be expanded:

$$\omega(k) = \omega(k_0) + (k - k_0)\omega'(k_0) + \cdots .$$

If $\tilde{n}$ is complex, we equate

$$
\begin{aligned}
v_G := \Re\frac{d\omega}{dk} &= \Re\frac{c}{\tilde{n}(\omega) + \omega(d\tilde{n}(\omega)/d\omega)} \\
&= \Re\frac{c}{\{\Re\tilde{n}(\omega) + \omega(d\Re\tilde{n}(\omega)/d\omega)\} + i\{\Im\tilde{n}(\omega) + \omega(d\Im\tilde{n}(\omega)/d\omega)\}} \\
&= c\frac{\{\Re\tilde{n}(\omega) + \omega(d\Re\tilde{n}(\omega)/d\omega)\}}{\{\Re\tilde{n}(\omega) + \omega(d\Re\tilde{n}(\omega)/d\omega)\}^2 + \{\Im\tilde{n}(\omega) + \omega(d\Im\tilde{n}(\omega)/d\omega)\}^2} \\
&\leq c.
\end{aligned}
\tag{15.8}
$$

Since absorption effects are related to imaginary parts, we see that the validity of the relation has its explanation in the damping effects resulting from these imaginary parts. In the following we investigate the behaviour of $\tilde{n}(\omega)$ in more detail. At the end of the chapter we shall see that when $d\Re\tilde{n}(\omega)/d\omega < 0$, the imaginary part $\Im\tilde{n}(\omega)$ is particularly large.

15.3 Absorption, Causality and Analyticity

We return to the consideration of a plane wave which is vertically incident on a dispersive medium as indicated in Fig. 15.1. A superposition of such waves represents a wave train or *wave packet*. We write this superposition in vacuum $(n = 1)$

$$u_V(x,t) = \int_{-\infty}^{\infty} A(\omega)e^{i(kx-\omega t)}\,d\omega \tag{15.9}$$

(we shall see later that the Fourier representation of a real quantity depends only on positive frequencies).
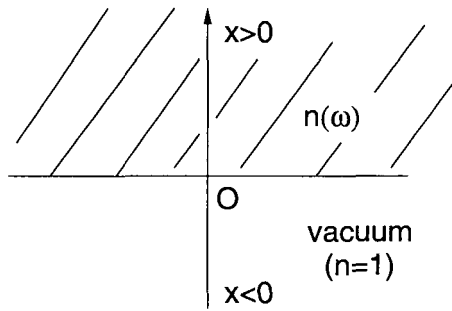
Fig. 15.1 Direction of a vertically incident wave.

As before we deduce from Maxwell's curl equations:
- *In vacuum $(x < 0)$ we have $k = \omega/c$.*
- *In the medium $(x > 0)$ we have $k(\omega) = \omega\tilde{n}(\omega)/c$.*

At the boundary plane but on the side of the vacuum

$$u_V(0,t) = \int_{-\infty}^{\infty} A(\omega)e^{-i\omega t}\,d\omega, \quad \text{i.e.} \quad A(\omega) = \frac{1}{2\pi}\int_{-\infty}^{\infty} u_V(0,t)e^{i\omega t}\,dt. \tag{15.10}$$

Every single wave is subjected to absorption and reflection at the interface as discussed in Chapter 11. The transmission coefficient defined there for $\mathbf{E}_0 \perp$ plane of incidence, i.e.

$$T = \frac{2\cos\alpha}{\cos\alpha + N\cos\beta},$$

is in the present case of a wave vertically incident

$$T = \frac{2}{1 + \tilde{n}(\omega)}, \quad (\mu = \mu_0), \tag{15.11}$$

and hence for the wave packet in the dispersive medium $(x > 0)$ we have

$$u(x,t) = \int_{-\infty}^{\infty} \left(\frac{2}{1+\tilde{n}(\omega)} \right) A(\omega) e^{i(k(\omega)x - \omega t)} \, d\omega. \tag{15.12}$$

For a further consideration of this integral we require information about the analytical properties of $A(\omega)$ and $k(\omega)$.

15.3.1 Properties of $A(\omega)$

We consider a wave packet $u(x,t)$ having a bell-shaped form, and starting from the origin $x = 0$, as indicated in Fig. 15.2. Then $u(0,t)$ has approximately the shape of

$$u(0,t) \simeq \frac{at^m}{m!} \theta(t) e^{-\epsilon t}, \quad \epsilon > 0 \tag{15.13}$$

(m integral and larger than zero).
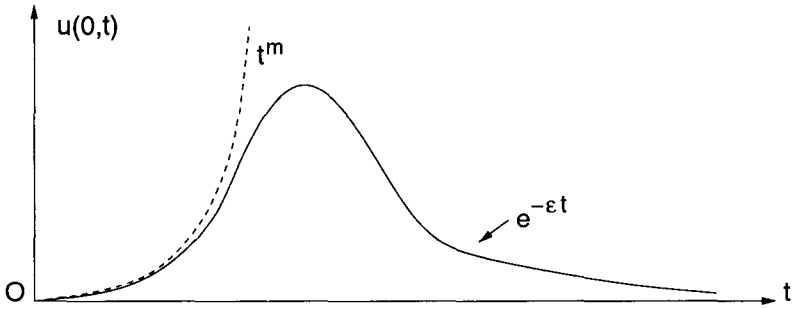
Fig. 15.2 The wave packet.

Thus the wave packet appears at the point $x = 0$ at time $t = 0$, grows to a maximum strength and ultimately disappears from $x = 0$. We have added the factor $\theta(t)$, in order to avoid the sector with $e^{-\epsilon(t \text{ negative})}$ in view of the integration we have to perform. Of course, one could achieve this also for example with

$$|t|^m e^{-\epsilon|t|}, \quad \epsilon > 0$$

or other means. On the other hand, one can argue that (15.13) has a well defined boundary with $x = 0$ at $t = 0$.

The property $u(0,t) = 0$ for $t < 0$ is described as a *causality condition*. We then obtain, from (15.12) with $x = 0$ and Fourier transformation, the

following relation

$$
\frac{2}{1+\tilde{n}(\omega)}A(\omega) = \frac{1}{2\pi}\int_{-\infty}^{\infty}u(0,t)e^{i\omega t}dt
$$

$$
= \frac{1}{2\pi}\frac{a}{m!}\int_{0}^{\infty}t^m e^{-(\epsilon-i\omega)t}dt
$$

$$
= \frac{1}{2\pi}\frac{a}{m!}\frac{m!}{(\epsilon-i\omega)^{m+1}}
$$

$$
= \frac{ai^{m+1}}{2\pi(\omega+i\epsilon)^{m+1}}, \tag{15.14}
$$

where we used the integral

$$
(z-1)! \equiv \Gamma(z) = \int_{0}^{\infty}e^{-t}t^{z-1}dt. \tag{15.15}
$$

Hence

$$
A(\omega) \propto \frac{\frac{1}{2}(1+\tilde{n}(\omega))}{(\omega+i\epsilon)^{m+1}} \overset{\text{in vacuum}}{=} \frac{1}{(\omega+i\epsilon)^{m+1}}. \tag{15.16}
$$

It is a matter of experience that the transmission of light through transparent media (i.e. those for which $\sigma = 0$) does not lead to resonance effects. We saw earlier that in the case of such media $\tilde{n}(\omega)$ is real. Hence in these cases $\tilde{n}(\omega)$ is an analytic function without poles. However, if there are resonance effects, $\tilde{n}(\omega)$ will not be everywhere an analytic function of ω, as we shall see. We infer from the above result that in that case $A(\omega)$ is, in general, an analytic function of ω with poles, i.e. a meromorphic function. The poles at

$$
\omega + i\epsilon = 0
$$

lie in the lower half of the ω-plane. We infer this from the above result. But we can see this also more generally. Consider the vacuum wave packet

$$
u_V(x,t) = \int_{-\infty}^{\infty}A(\omega)e^{i(kx-\omega t)}d\omega. \tag{15.17}
$$

In vacuum $k = \omega/c$ (this is why we can write the integral also as $\int dk$). Thus in vacuum all waves (the partial waves as well as the group) propagate with the velocity of light. According to the Special Theory of Relativity (see later) a signal or wave packet is propagated only into the *future light cone* (see Fig. 15.3), which means that for $x > ct$ the integral must vanish, i.e. we must have

$$
0 = [u_V(x,t)]_{x>ct} = \int_{-\infty}^{\infty}A(\omega)e^{i\frac{\omega}{c}(x-ct)}d\omega\theta(x-ct). \tag{15.18}
$$

We investigate this equation with the help of a contour integral in the plane of complex ω. If we choose the contour of integration C as indicated in Fig. 15.4, the infinite semicircle with $\Im\omega > 0$, and thus the factor

$$e^{i\frac{\omega}{c}(x-ct)}\theta(x-ct)$$

with exponentially decreasing behaviour, does not contribute to the integral and we have

$$0 = \int_{-\infty}^{\infty}\cdots + \int_{\frown}\cdots,$$

i.e.

$$0 = \int_{C} A(\omega)e^{i\frac{\omega}{c}(x-ct)}d\omega\theta(x-ct). \tag{15.19}$$
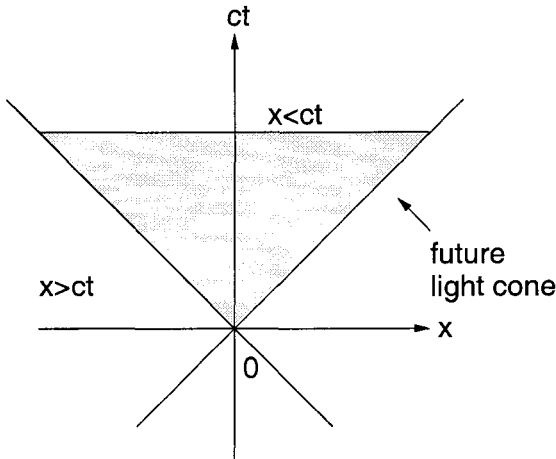
Fig. 15.3 The future light cone.

The Cauchy residue theorem now implies that $A(\omega)$ cannot possess poles in the plane $\Im\omega > 0$ (i.e. their residues would be zero). Thus $A(\omega)$ is analytic in that part of the complex plane. This means the poles of the amplitude $A(\omega)$ can only be in the lower half of the ω-plane. We obtained this result from the condition that propagation takes place only into the future light cone or time-like domain, where $x/t \leq c$ (in the following we use this result that $A(\omega)$ is a meromorphic function with poles in the lower half of the complex plane). This property of no propagation into the space-like domain depends on the other property of poles located in the lower half-plane.

We saw at the beginning that in regions of anomalous dispersion it is possible that $v_G > c$. We recall that v_G is quite generally (except when $v_G > c$) the velocity with which energy is transported, and hence a signal

with velocity x/t. As such it cannot be larger than c (according to the Special Theory of Relativity).

We return to the expression of Eq. (15.12), i.e. to

$$u(x,t) = \int_{-\infty}^{\infty} \left(\frac{2}{1+\tilde{n}(\omega)} \right) A(\omega) e^{i(k(\omega)x - \omega t)} d\omega.$$

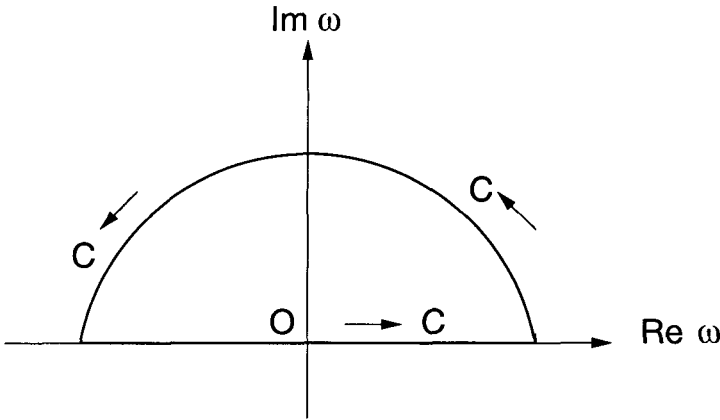In order to be able to perform the integration we have to know the analytic properties of $\tilde{n}(\omega)$.

Fig. 15.4 The integration contour C.

15.3.2   Properties of $\tilde{n}(\omega)$

In the discussion of the Drude theory we used the classical nonrelativistic equation of motion of a free electron

$$m\ddot{x} + \frac{m}{\tau}\dot{x} = eE, \quad E = E_0 e^{-i\omega t} \tag{15.20}$$

(earlier we used $E \propto e^{+i\omega t}$; hence comparisons may require some changes of signs). If the electron is not free, but bound to some atom with eigenfrequency ω_k, then the equation is in the case of a simple one-dimensional oscillator model

$$m\ddot{x} + \frac{m}{\tau}\dot{x} + m\omega_k^2 x = eE. \tag{15.21}$$

This equation takes into account that the (*almost* free conduction) electron does not follow the applied field without inertia. One should note that here we consider the electron as bound to the atom by an oscillator potential. From quantum mechanics we know that the Schrödinger equation with an

oscillator potential $V_R = ax^2, a > 0$, allows only discrete bound states (i.e. states with infinite lifetime). The additional friction term we have introduced here has its counterpart in quantum mechanics as an imaginary part of the potential

$$V_I = i\Im V \quad \text{of} \quad V = V_R + i\Im V.$$

For $\Im V < 0$ the potential V is absorptive, and violates unitarity (absorption of particles, energy into the target). A *resonance* is a state of *finite lifetime* $\tau = 1/\gamma$. It is not necessary to have $\Im V \neq 0$ in order to obtain resonances. Resonances are obtained, for instance, for screened Coulomb potentials or with potentials which permit the escape of particles to infinity through quantum tunneling. Such cases are, however, much more difficult to handle analytically, so that we restrict ourselves here for reasons of simplicity to a simple semi-classical model. Semi-classical means here, we start off classically, and then replace the frequency by discrete quantum mechanical frequencies.

Setting in Eq. (15.21):

$$x = x_0 e^{-i\omega t},$$

we have (on the right side of the following equation we should strictly have the complete Lorentz force, but we assume that $\mathbf{v} \times \mathbf{B}$ is negligible)

$$\left(-m\omega^2 - \frac{m}{\tau} i\omega + m\omega_k^2 \right) x = eE, \tag{15.22}$$

i.e. with $E(x,t) = E_0(x) e^{-i\omega t}$

$$x = \frac{eE}{m(\omega_k^2 - \omega^2 - i\gamma\omega)}, \quad (\gamma = 1/\tau). \tag{15.23}$$

For γ to represent a damping of x or E, we must have $\Im\omega < 0$, i.e. at the resonance

$$\omega = \frac{1}{2}\left[-i\gamma \pm \sqrt{4\omega_k^2 - \gamma^2} \right], \quad 4\omega_k^2 > \gamma^2,$$

and for $x \propto e^{-i\omega t}$ and $\gamma > 0$. We thus have

$$\mathbf{j} = ne\dot{x} = -\frac{ne^2 i\omega \mathbf{E}}{m(\omega_k^2 - \omega^2 - i\gamma\omega)} \equiv \sigma \mathbf{E}, \tag{15.24}$$

so that

$$\sigma = -\frac{ne^2 i\omega}{m(\omega_k^2 - \omega^2 - i\gamma\omega)} \tag{15.25}$$

and the generalised dielectric constant is

$$\eta = \epsilon + i\frac{\sigma}{\omega} = \epsilon + \frac{ne^2}{m(\omega_k^2 - \omega^2 - i\gamma\omega)}. \tag{15.26}$$

In general the electron may occupy one of many possible quantum mechanical states k. It is then necessary to replace the last equation by

$$\eta = \epsilon + \sum_k \frac{ne^2}{m} \frac{f_k}{\omega_k^2 - \omega^2 - i\gamma\omega}, \tag{15.27}$$

where the sum extends over the states of the electrons involved in the polarisation of the medium. The factor

$$\frac{e^2}{m(\omega_k^2 - \omega^2 - i\gamma\omega)}$$

is a measure of the polarisability of the medium. For strongly bound electrons $\omega_k \gg \omega$ and $\mathbf{j}$ is correspondingly small. We obtain therefore for the refractive index p the result

$$p = \sqrt{\frac{\eta}{\epsilon}} = \sqrt{1 + \frac{1}{\epsilon}\sum_k \frac{ne^2}{m} \frac{f_k}{\omega_k^2 - \omega^2 - i\gamma\omega}}. \tag{15.28}$$

The coefficients f_k are called *oscillator strengths*; they specify the probability of an electron with energy $\hbar\omega_k$ to be in state k. For atoms in vacuum ϵ is equal to ϵ_0, and the refractive index is

$$p = \sqrt{\frac{\eta}{\epsilon_0}} \simeq 1 + \frac{ne^2}{2m\epsilon_0}\sum_k \frac{f_k}{\omega_k^2 - \omega^2 - i\gamma\omega} \equiv \tilde{n}(\omega). \tag{15.29}$$

For our purposes here it suffices to consider a dielectric, which permits only one resonance energy $\hbar\omega_k = \hbar\omega_0$, so that

$$p^2(\omega) = 1 + \frac{\omega_P^2}{\omega_0^2 - \omega^2 - i\gamma\omega}, \quad \omega_P^2 \equiv \frac{ne^2 f_0}{m\epsilon_0}. \tag{15.30}$$

We set

$$p^2(\omega) = \frac{\omega_0^2 - \omega^2 - i\gamma\omega + \omega_P^2}{\omega_0^2 - \omega^2 - i\gamma\omega} \equiv \frac{(\omega - \omega_a)(\omega - \omega_b)}{(\omega - \omega_c)(\omega - \omega_d)} \equiv \tilde{n}^2(\omega). \tag{15.31}$$

Then it is easy to calculate

$$\begin{matrix} \omega_a \\ \omega_b \end{matrix} = -i\frac{\gamma}{2} \pm \sqrt{\omega_0^2 + \omega_P^2 - \frac{\gamma^2}{4}},$$

$$\begin{matrix} \omega_c \\ \omega_d \end{matrix} = -i\frac{\gamma}{2} \pm \sqrt{\omega_0^2 - \frac{\gamma^2}{4}}.$$

We observe by looking at Eq. (15.31), that the poles of $p^2(\omega)$, like those of $A(\omega)$, are located in the lower half of the ω-plane, as indicated in Fig. 15.5. We see that along

$$\Re\omega_d \leq \Re\omega \leq \Re\omega_b, \quad \Re\omega_c \leq \Re\omega \leq \Re\omega_a$$

the refractive index $p(\omega)$ is pure imaginary. We also see that
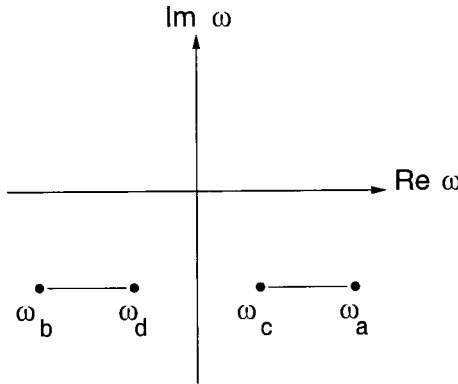
$$\lim_{\omega \to \infty} \tilde{n}(\omega) = 1.$$

Fig. 15.5 The poles of $p(\omega)$ in the ω-plane.

Next we consider the factor $1/(1+\tilde{n}(\omega))$ in the function $u(x,t)$, the latter being the wave packet in the dispersive medium. We have

$$\frac{1}{1+\tilde{n}(\omega)} = \frac{1}{1 + \sqrt{(\omega - \omega_a)(\omega - \omega_b)/(\omega - \omega_c)(\omega - \omega_d)}}$$
$$= \frac{(\omega - \omega_c)(\omega - \omega_d) - \sqrt{(\omega - \omega_a)(\omega - \omega_b)(\omega - \omega_c)(\omega - \omega_d)}}{(\omega - \omega_c)(\omega - \omega_d) - (\omega - \omega_a)(\omega - \omega_b)}.$$

The denominator is

$$(\omega - \omega_c)(\omega - \omega_d) - (\omega - \omega_a)(\omega - \omega_b)$$
$$= \omega^2 - (\omega_c + \omega_d)\omega + \omega_c\omega_d - \omega^2 + (\omega_a + \omega_b)\omega - \omega_a\omega_b$$
$$= w(\omega_a + \omega_b - \omega_c - \omega_d) + (\omega_c\omega_d - \omega_a\omega_b)$$
$$= -\frac{\gamma^2}{4} - \omega_0^2 + \frac{2\gamma^2}{4} + \omega_0^2 + \omega_P^2 - \frac{\gamma^2}{4} = \omega_P^2.$$

Hence

$$\frac{1}{1+\tilde{n}(\omega)} = \frac{(\cdots)(\cdots) - \sqrt{(\cdots)(\cdots)(\cdots)(\cdots)}}{\omega_P^2},$$

i.e. $1/(1 + \tilde{n}(\omega))$ has no pole in the ω-plane. This example verifies that in the given case, and more generally (here without proof), the factor $2/(1 + \tilde{n}(\omega))$ has no poles, i.e. is regular.

15.4 No Wave Packet with a Velocity $> c$

We are now in a position to prove that no wave packet in a dispersive medium with refractive index $\tilde{n}(\omega)$ can propagate with a velocity larger than the velocity of light. We consider the integral (see $u(x,t)$ of Eq. (15.12))

$$I = \int_{\curvearrowright} \frac{2}{1 + \tilde{n}(\omega)} A(\omega) e^{i(k(\omega)x - \omega t)} d\omega \qquad (15.32)$$

(physically the refractive index $\tilde{n}(\omega)$ can never be -1 for any value of ω), in which the contour of integration $C = \curvearrowright$ is to be taken as indicated in Fig. 15.4. We have

$$I = \int_{-\infty}^{\infty} + \int_{\curvearrowright} \cdots = \text{wave packet } u(x,t) + \int_{\curvearrowright} \cdots .$$

We know: The integrand is a meromorphic function of ω with poles only in the lower half of the ω-plane. Applying *Cauchy's residue theorem* we obtain

$$I = 2\pi i \sum \text{residues} = 0.$$

We have therefore

$$u(x,t) = -\int_{\curvearrowright} \cdots .$$

We consider now the contour integral along the infinite semicircle, i.e. for $|\omega| \sim \infty$. Since

$$k(\omega) = \frac{\omega}{c} \tilde{n}(\omega) \to \frac{\omega}{c} \text{ for } |\omega| \to \infty,$$

i.e.

$$\frac{d\omega}{dk} \to c, \ \tilde{n}(\omega) \to 1,$$

we have

$$\int_{\curvearrowright} \Rightarrow \int_{\curvearrowright} \frac{2}{2} A(\omega) e^{i\omega(\frac{x}{c} - t)} d\omega.$$

This integral vanishes exponentially in the upper half of the ω-plane, i.e. for $\Im \omega > 0$, *provided*

$$\frac{x}{c} - t > 0, \quad \text{i.e.} \quad v_{\text{signal}} = \frac{x}{t} > c.$$

This result implies that the propagation or spreading of waves into regions beyond the future light cone are excluded, or, phrased differently, in regions $x > ct$ the wave packet $u(x, t)$ vanishes.

In the next step we demonstrate, that for signal velocities

$$v_{\text{signal}} = \frac{x}{t} < c,$$

the wave packet has a definite value. This time we integrate in the lower half plane and consider the integral

$$I = \int_{\curvearrowright} \frac{2}{1 + \tilde{n}(\omega)} A(\omega) e^{i(k(\omega)x - \omega t)} d\omega = \int_{-\infty}^{\infty} \cdots + \int_{\curvearrowright} \cdots . \tag{15.33}$$

We consider

$$\int_{\curvearrowright} \cdots \Rightarrow \int_{\curvearrowright} \frac{2}{2} A(\omega) e^{i\omega(\frac{x}{c} - t)} d\omega.$$

For $x < ct$, i.e. $v_{\text{signal}} \equiv x/t < c$, and $\Im \omega < 0$ we see that

$$e^{i\omega(\frac{x}{c} - t)}$$

is exponentially decreasing for $\omega \to \infty$. This means

$$\int_{\curvearrowright} \cdots = 0.$$

Hence

$$u(x, t) = \int_{-\infty}^{\infty} \cdots = I = 2\pi i \sum \text{residues}.$$

In order to be able to perform the integration explicitly, we require, of course, a knowledge of the function $A(\omega)$.

15.5 Explanation of the Anomalous Dispersion

We return to our approximate expression for the refractive index, i.e. to the square root of expression (15.30), i.e.

$$\tilde{n} = \sqrt{\frac{\eta}{\epsilon_0}} \simeq 1 + \frac{1}{2} \frac{\omega_P^2}{\omega_0^2 - \omega^2 - i\gamma\omega}, \tag{15.34}$$

taking into account only one pole at $\Re \omega = \omega_0$. With a real denominator we have

$$\tilde{n} = \sqrt{\frac{\eta}{\epsilon_0}} = 1 + \frac{\omega_P^2}{2} \frac{\omega_0^2 - \omega^2 + i\gamma\omega}{(\omega_0^2 - \omega^2)^2 + \gamma^2\omega^2} \tag{15.35}$$

and

$$\Re\tilde{n} = 1 + \frac{\omega_P^2}{2}\frac{(\omega_0^2 - \omega^2)}{(\omega_0^2 - \omega^2)^2 + \gamma^2\omega^2} \tag{15.36}$$

and

$$\Im\tilde{n} = \frac{\omega_P^2}{2}\frac{\gamma\omega}{(\omega_0^2 - \omega^2)^2 + \gamma^2\omega^2}, \tag{15.37}$$

and (after 3 lines of calculations)

$$\frac{d\Re\tilde{n}}{d\omega} = \frac{\omega_P^2\omega[(\omega_0^2 - \omega^2)^2 - \gamma^2\omega_0^2]}{[(\omega_0^2 - \omega^2)^2 + \gamma^2\omega^2]^2}. \tag{15.38}$$

Now (for $\omega \neq 0$),

$$\frac{d\Re\tilde{n}}{d\omega} = 0, \quad \text{where} \quad (\omega_0^2 - \omega^2)^2 = \gamma^2\omega_0^2,$$

i.e.

$$\omega^2 = \omega_0^2 \mp \gamma\omega_0,$$

and hence

$$\omega \simeq \omega_0 \mp \frac{1}{2}\gamma.$$

It follows that $\Re\tilde{n}$ has extrema at these points. Using Eq. (15.35) and setting $\omega \simeq \omega_0 \mp \gamma/2$, the refractive index $\tilde{n}$ at these points is:

$$\tilde{n} = 1 + \frac{\omega_P^2}{2}\frac{(\pm\gamma\omega_0) + i\gamma\sqrt{\omega_0^2 \mp \gamma\omega_0}}{(\pm\gamma\omega_0)^2 + \gamma^2(\omega_0^2 \mp \gamma\omega_0)} \tag{15.39}$$

or

$$\Im\tilde{n} \overset{(15.37)}{\simeq} \frac{\omega_P^2}{2}\frac{\gamma\sqrt{\omega_0^2}}{2\gamma^2\omega_0^2} = \frac{\omega_P^2}{4\gamma\omega_0} \tag{15.40}$$

and

$$\Re\tilde{n} \overset{(15.39)}{\simeq} 1 + \frac{\omega_P^2}{2}\frac{(\pm\gamma\omega_0)}{2\gamma^2\omega_0^2} = 1 \pm \frac{\omega_P^2}{4\gamma\omega_0}. \tag{15.41}$$

Thus

$$|\Re\tilde{n} - 1| = |\Im\tilde{n}| \quad \text{at} \quad \omega = \omega_0 \pm \frac{\gamma}{2}. \tag{15.42}$$

We see therefore that

$$\frac{d\Re\tilde{n}}{d\omega} < 0, \quad \text{where} \quad \omega_0 - \frac{\gamma}{2} \leq \omega \leq \omega_0 + \frac{\gamma}{2}.$$

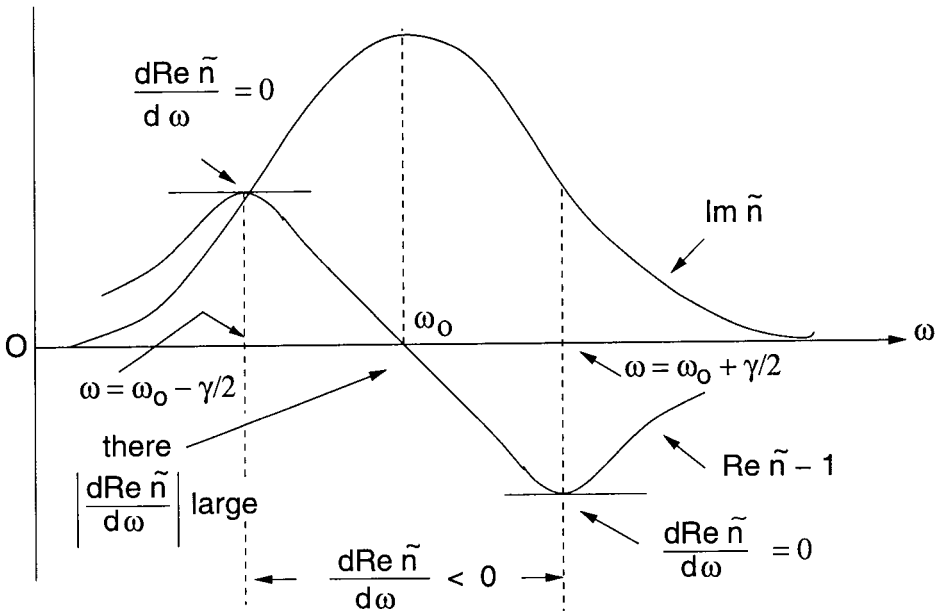In this region, see Fig. 15.6, $\Im\tilde{n}$ is particularly large (this is also the region of the maximum).

Fig. 15.6 $\Im\tilde{n}$ large where $d\Re\tilde{n}/d\omega$ is negative.

We shall see in Chapter 16, that a large imaginary part (of e.g. $\tilde{n}$) is connected with strong absorption effects. That physical velocities

$$v_G = \frac{c}{\tilde{n} + \omega(d\tilde{n}/d\omega)} > c$$

do not occur, i.e. when $d\tilde{n}/d\omega$ is negative, is now *qualitatively explained* by arguing that these mathematically arising velocities are removed by the absorptive effects.

Example 15.1: Method of stationary phase
Consider the integral

$$I = \int F(\omega)e^{i\phi(\omega)}d\omega,$$

where $F(\omega)$ is a slowly varying function of ω, whereas $\phi(\omega)$ is a rapidly varying phase with only one stationary point at ω_0. Show that approximately

$$I \simeq \sqrt{\frac{2i\pi}{\phi''(\omega_0)}}F(\omega_0)e^{i\phi(\omega_0)}.$$

Solution: In the neighbourhood of ω_0 we can write

$$\phi(\omega) = \phi(\omega_0) + \frac{1}{2}\phi''(\omega_0) + \cdots .$$

Since $F(\omega)$ varies in the neighbourhood of ω_0 only slowly, we have

$$I \simeq F(\omega_0)e^{i\phi(\omega_0)} \int_{-\infty}^{\infty} e^{\frac{i}{2}\phi''(\omega_0)(\omega-\omega_0)^2} \, d\omega.$$

The remaining integral is a Fresnel integral* (or can be evaluated with appropriate assumptions like an integral over a Gauss function) and leads to the above result.

---

*See e.g. W. Magnus and F. Oberhettinger, *Formulas and Theorems for the Functions of Mathematical Physics* (Chelsea Publ. Co., 1954) p. 96.

Chapter 16

Causality and Dispersion Relations

16.1 Introductory Remarks

In this chapter we continue the study of analytic properties of spectral functions, but concentrate on Green's functions and causality. In this connection important integral relations between real and imaginary parts of spectral functions are introduced, which are also known as dispersion relations or (in more mathematical contexts) as Hilbert transforms, and were first applied by Kramers to the dielectric susceptibility.

16.2 Cause U and Effect E

We let $U(t)$ represent a *cause*, e.g. a charge distribution which varies with time, and $E(t)$ an *effect*, e.g. an electromagnetic field. Since both are real physical quantities, we assume them to be real.

The *Fourier transforms* of the two quantities define other *representations* of $U(t)$ and $E(t)$:

$$U(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} d\omega\, e^{-i\omega t} u(\omega), \quad u(\omega) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} dt\, e^{i\omega t} U(t),$$

$$E(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} d\omega\, e^{-i\omega t} e(\omega), \quad e(\omega) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} dt\, e^{i\omega t} E(t). \quad (16.1)$$

Since we demand that U and E are real, i.e.

$$U(t) = U^*(t), \quad E(t) = E^*(t),$$

361

it follows that

$$\int_{-\infty}^{\infty} d\omega\, e^{-i\omega t} u(\omega) = \int_{-\infty}^{\infty} d\omega\, e^{i\omega t} u^*(\omega) = \int_{-\infty}^{\infty} d\omega'\, e^{-i\omega' t} u^*(-\omega')$$

$$= \int_{-\infty}^{\infty} d\omega\, e^{-i\omega t} u^*(-\omega),$$

i.e.

$$u(\omega) = u^*(-\omega), \quad e(\omega) = e^*(-\omega). \tag{16.2}$$

These relations between the integrands of the Fourier transforms or spectral functions are referred to as *"crossing relations"*. One should note that ω is assumed to be real. For $U(t)$, for instance, it follows that

$$U(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{0} d\omega\, e^{-i\omega t} u(\omega) + \frac{1}{\sqrt{2\pi}} \int_{0}^{\infty} d\omega\, e^{-i\omega t} u(\omega)$$

$$= \frac{1}{\sqrt{2\pi}} \int_{0}^{\infty} d\omega'\, e^{i\omega' t} \underbrace{u(-\omega')}_{u^*(\omega')} + \frac{1}{\sqrt{2\pi}} \int_{0}^{\infty} d\omega\, e^{-i\omega t} u(\omega)$$

$$= \frac{1}{\sqrt{2\pi}} \int_{0}^{\infty} d\omega [u(\omega) e^{-i\omega t} + u^*(\omega) e^{i\omega t}].$$

We set

$$u(\omega) = \frac{1}{2} r(\omega) e^{i\theta(\omega)}, \tag{16.3}$$

where $r(\omega), \theta(\omega)$ are real. Then

$$U(t) = \frac{1}{\sqrt{2\pi}} \int_{0}^{\infty} d\omega \left[\frac{1}{2} r(\omega) e^{i\theta(\omega)} e^{-i\omega t} + \frac{1}{2} r(\omega) e^{-i\theta(\omega)} e^{i\omega t} \right]$$

$$= \frac{1}{\sqrt{2\pi}} \int_{0}^{\infty} d\omega\, r(\omega) \cos[\omega t - \theta(\omega)], \tag{16.4}$$

i.e. the Fourier representation of $U(t)$ depends only on *positive* frequencies ($\omega > 0$). A corresponding result can be obtained for $E(t)$.

16.3 (U, E)-Linearity and Green's Functions

We let U_i be the cause of the effect E_i. Then we assume that

$$\alpha U_i + \beta U_j$$

is cause of the effect

$$\alpha E_i + \beta E_j,$$

or more generally

$$E(t) = \int_{-\infty}^{\infty} dt' G(t, t') U(t') \tag{16.5}$$

(superposition principle). Here $G(t, t')$ is a *weight* or *Green's function*. We also assume that

$$G(t, t') = G(t - t'),$$

i.e. the process connecting U and E is independent of absolute times. Then E is the inhomogeneous solution of a differential equation

$$DE(t) = U(t) \tag{16.6}$$

with

$$DG(t - t') = \delta(t - t') = \delta(t' - t). \tag{16.7}$$

We write $G(t, t')$ as $G(t - t')$ and not as $G(t' - t)$ for reasons of causality which lead us to the so-called *retarded Green's function*. Then

$$E(t) = \int_{-\infty}^{\infty} dt' G(t, t') U(t').$$

This is precisely the relation given above. Since U and E are real, G must also be real. Let $g(\omega)$ be the Fourier transform of G:

$$G(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} d\omega e^{-i\omega t} g(\omega), \quad g(\omega) = \int_{-\infty}^{\infty} G(t) e^{i\omega t} dt. \tag{16.8}$$

Then

$$\begin{aligned}
e(\omega) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} dt e^{i\omega t} E(t) \\
&= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} dt e^{i\omega t} \int_{-\infty}^{\infty} dt' G(t - t') U(t'),
\end{aligned}$$

and, inserting (16.8) for $G(t - t')$, this becomes

$$\begin{aligned}
e(\omega) &= \frac{1}{(\sqrt{2\pi})^3} \int_{-\infty}^{\infty} dt e^{i\omega t} \int_{-\infty}^{\infty} dt' U(t') \int_{-\infty}^{\infty} d\omega' e^{-i\omega'(t-t')} g(\omega') \\
&= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} dt' U(t') \int d\omega' g(\omega') e^{i\omega' t'} \underbrace{\frac{1}{2\pi} \int_{-\infty}^{\infty} dt e^{i(\omega - \omega')t}}_{\delta(\omega - \omega')}
\end{aligned}$$

and hence

$$\begin{aligned}
e(\omega) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} dt' U(t') g(\omega) e^{i\omega t'} \\
&= g(\omega) u(\omega), \tag{16.9}
\end{aligned}$$

i.e.

$$e(\omega) = g(\omega)u(\omega), \quad e^*(-\omega) = g^*(-\omega)u^*(-\omega). \tag{16.10}$$

These relations are called *convolution theorems*.

Since

$$u(\omega) = u^*(-\omega), \quad e(\omega) = e^*(-\omega),$$

it follows that

$$e(\omega) = g^*(-\omega)u(\omega), \quad \text{i.e.} \quad g(\omega) = g^*(-\omega). \tag{16.11}$$

The latter of these implies again that $G(t)$ is real (see remarks at the beginning).

A *mono-frequent* cause $U_0 e^{-i\omega t}$ leads to a mono-frequent effect:

$$
\begin{aligned}
E(t) &= \int_{-\infty}^{\infty} dt'\, G(t - t')U(t') = \int_{-\infty}^{\infty} dt'\, G(t - t')U_0 e^{-i\omega t'} \\
&= U_0 e^{-i\omega t} \int_{-\infty}^{\infty} dt'\, G(t - t') e^{i\omega(t - t')} \\
&= U_0 e^{-i\omega t} \int_{-\infty}^{\infty} \underbrace{d(t' - t)}_{-d(t - t')}\, G(t - t') e^{i\omega(t - t')},
\end{aligned}
$$

and hence

$$
\begin{aligned}
E(t) &= U_0 e^{-i\omega t} \int_{\infty}^{-\infty} (-)d(t - t') G(t - t') e^{i\omega(t - t')} \\
&= U_0 e^{-i\omega t} g(\omega) \equiv E_0 e^{-i\omega t}, \tag{16.12}
\end{aligned}
$$

where $E_0 = U_0 g(\omega)$.

16.4   Causality

The concept of causality expresses the fundamental idea that an effect $E(t)$ at time t can only have and depend on causes $U(t)$ prior to time t. Since

$$E(t) = \int_{-\infty}^{\infty} dt'\, G(t - t')U(t'),$$

(G real for U and E real), this means $E(t)$ can only depend on such $U(t')$, for which $t' < t$. More concretely

$$G(t - t') = 0 \quad \text{for} \quad t - t' < 0,$$

i.e. $G(t - t') \propto \theta(t - t')$, i.e. only such values of the cause U enter the effect $E(t)$ at time t, which appeared before the instant of time t, at which we are interested in the effect $E(t)$. This time ordering is described as *causality*. This causality is also described as *macrocausality* in contrast to *microcausality* or locality, which is a quantum theoretical concept concerning operators representing observables and implies that such observables commute if their separation is spacelike (because then they cannot exchange signals since no signal can travel with a velocity greater than that of light). The Green's function obeying macrocausality is the *retarded Green's function*. Consider once again a typical classical equation of motion

$$\ddot{x} + 2\gamma\dot{x} + \omega_0^2 x = f(t).$$

The friction term violates the time-reversal invariance of the simple Newton equation

$$\ddot{x} + \omega_0^2 x = 0$$

(which is plausible, because a process involving friction, in which therefore energy is dissipated as heat, is irreversible). For $\gamma > 0$, the direction of the friction force $2\gamma\dot{x}$ is opposite to that of the motion since

$$\ddot{x} + \omega_0^2 x = -2\gamma\dot{x} + f(t).$$

The Green's function of this problem is, however, *uniquely defined by the friction term*, because with

$$\left(\frac{d^2}{dt^2} + 2\gamma\frac{d}{dt} + \omega_0^2\right)G(t, t') = \delta(t - t')$$

we obtain

$$G(t - t') = \frac{1}{2\pi}\int_{-\infty}^{\infty} d\omega \frac{e^{-i\omega(t-t')}}{\omega_0^2 - \omega^2 - 2i\gamma\omega}. \tag{16.13}$$

The poles of the integrand are located in the domain $\Im\omega < 0$, and are given by

$$\omega = -i\gamma \pm \sqrt{\omega_0^2 - \gamma^2}, \quad \gamma > 0.$$

In this way the integral with $\gamma \neq 0$ and the condition $G = $ finite is uniquely defined. Since there are no poles in the upper half-plane

$$\int_{\curvearrowright} \cdots = 2\pi i \sum \text{residues} = 0,$$

and it follows that

$$\int_{-\infty}^{\infty} \cdots = -\int_{\curvearrowright} \cdots = 0 \text{ for } t - t' < 0,$$

i.e.

$$G(t - t') = 0 \text{ for } t - t' < 0$$

and

$$G(t - t') \neq 0 \text{ for } t - t' > 0.$$

This finite Green's function is the retarded Green's function. If

$$G(t - t') = 0 \text{ for } t - t' < 0,$$

we have

$$x(t) = \int_{-\infty}^{\infty} G(t - t')f(t')dt'$$

$$= \int_{-\infty}^{t} G(t - t')f(t')dt'. \tag{16.14}$$

The solution of the homogeneous equation which could be added to give the complete solution (with $x \to 0$ for $t \to \infty$), i.e. the solution of

$$\ddot{x} + 2\gamma\dot{x} + \omega_0^2 x = 0,$$

is

$$x_0(t) = \left(ae^{i\sqrt{\omega_0^2 - \gamma^2}\, t} + be^{-i\sqrt{\omega_0^2 - \gamma^2}\, t} \right) e^{-\gamma t}.$$

This diverges for $t \to -\infty$. To ensure that $x(t)$ is finite for all times t, we have to choose $a = 0 = b$, i.e. a contribution of the homogeneous equation to the solution is excluded by boundary conditions. In the case $\gamma \neq 0$ the integration contour is automatically fixed by the requirement of G, x to be finite. In the case $\gamma = 0$, however, the contour must be specified separately, because then the poles $\omega = \pm\omega_0$ lie on the real axis (in the above simplified model). We consider this case by selecting a pole at $\omega = \omega_0 - i\epsilon, \epsilon > 0$, with $\epsilon \to 0$.

16.5   Causality and Analyticity

We return to our original consideration and now permit *complex values of* ω. Since

$$g(\omega) = \int_{-\infty}^{\infty} dt\, G(t)e^{i\omega t}, \quad G(t) \propto \theta(t), \tag{16.15}$$

it follows that

$$g(\omega) = \int_{0}^{\infty} dt\, G(t)e^{i\omega t},$$

or (since $G(t)$ is real, see above)

$$g^*(\omega) = \int_0^\infty dt G(t) e^{-i\omega t} \neq g^*(\omega^*).$$

We now set

$$\omega = \omega_R + i\omega_I.$$

Then

$$g(\omega) \equiv g(\omega_R + i\omega_I) = \int_0^\infty dt G(t) e^{i\omega_R t} e^{-\omega_I t}$$

and

$$g^*(\omega) = g^*(\omega_R + i\omega_I) = \int_0^\infty dt G(t) e^{-i\omega_R t} e^{\omega_I t} = g(-\omega).$$

We assume that

$$G(t) \sim \text{ finite for } t \to \infty.$$

Then there exists a $g(\omega)$ for $\omega_I > 0$ and $g^*(\omega)$ for $\omega_I < 0$. The derivatives

$$\frac{d}{d\omega} g(\omega) \quad \text{and} \quad \frac{d}{d\omega} g^*(\omega)$$

also exist at every point with respectively $\omega_I > 0$ and $\omega_I < 0$. Hence $g(\omega)$ is analytic in the half-plane $\omega_I > 0$, and $g^*(\omega)$ is analytic in the half-plane $\omega_I < 0$. We see therefore that we can infer from the causality condition

$$G(t - t') = 0 \text{ for } t - t' < 0$$

the analyticity of the Fourier transforms $g(\omega)$ and $g^*(\omega)$ in the domains $\omega_I > 0$ and $\omega < 0$ respectively.

Example 16.1: The Cauchy–Riemann equations
Starting from the integral

$$g(\omega) = \int_0^\infty dt G(t) e^{i\omega t}, \quad \omega = \omega_R + i\omega_I,$$

examine whether

$$g(\omega) = g_R(\omega_R, \omega_I) + i g_I(\omega_R, \omega_I) \equiv u + iv$$

satisfies the Cauchy–Riemann equations, i.e.

$$\frac{\partial}{\partial x} g_R = \frac{\partial}{\partial y} g_I, \quad \frac{\partial}{\partial x} g_I = -\frac{\partial}{\partial y} g_R,$$

or $u_x = v_y, u_y = -v_x$.

Solution: The solution is trivial. See also Sec. 3.7.

16.6 Principal Value Integrals and Dispersion Relations

From the fact that the spectral functions $g(\omega)$ and $g^*(\omega)$ possess simple poles in one half of the ω-plane and are analytic in the other, one can derive relations between their real and imaginary parts called *dispersion relations*.

We consider the following integral taken along the infinite semicircle in the domain $\Im\omega' > 0$,

$$I = \int_{\frown} d\omega' \frac{g(\omega')}{\omega' - \omega + i\epsilon}, \quad \epsilon > 0, \quad \omega \text{ real}. \tag{16.16}$$

With $\epsilon > 0$, which displaces the pole underneath the real axis, we could exclude it from within a contour along the real axis and then back along the semicircle at infinity. We can achieve the same effect, however, with a small semicircular distortion of the contour around the pole on the real axis as indicated in Fig. 16.1.
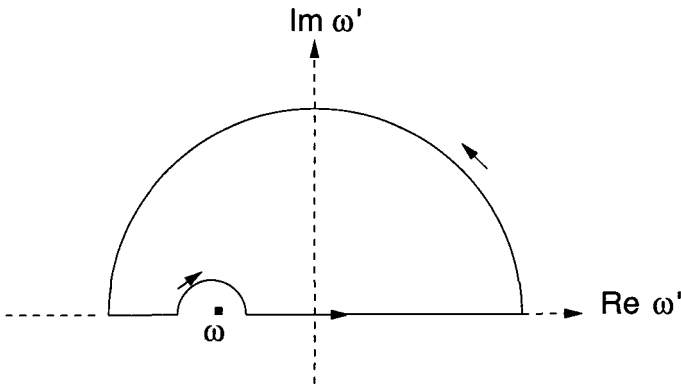
Fig. 16.1   The contour around the point ω in the complex ω'-plane.

In the upper half-plane $\Im\omega' > 0$, as we saw, $g(\omega')$ is analytic. Thus *if* $g(\omega')$ decreases sufficiently fast towards zero for $|\omega'| \to \infty, \Im\omega' > 0$, so that $\int_{\frown} \cdots = 0$, then

$$\int_{\frown} d\omega' \cdots = 2\pi i \sum \text{residues} = 0$$

and

$$\int_{-\infty}^{\infty} d\omega' \cdots = 0,$$

where the integration is meant to be taken along the contour shown in

Fig. 16.1 from $-\infty$ to $+\infty$, i.e.

$$
\begin{aligned}
0 &= \int_{-\infty}^{\infty} d\omega' \frac{g(\omega')}{\omega' - \omega} \\
&= P \int d\omega' \frac{g(\omega')}{\omega' - \omega} \quad + \underbrace{\int_{\frown} d\omega' \frac{g(\omega')}{\omega' - \omega}}_{\substack{\text{infinitesimal} \\ \text{semicircle around } \omega' = \omega}},
\end{aligned}
\tag{16.17}
$$

where

$$
P \int
$$

means *"principal value"* of the integral (which will be explained in detail below, see Eq. (16.32)). We note here that

$$
P \int d\omega' \frac{g(\omega')}{\omega' - \omega} = \lim_{\epsilon \to 0} \left(\int_{-\infty}^{\omega - \epsilon} \cdots + \int_{\omega + \epsilon}^{\infty} \cdots \right),
$$

and we shall return to this expression later. With

$$
\omega' - \omega = \rho e^{i\varphi}, \quad d\omega' = i\rho e^{i\varphi} d\varphi,
$$

we have

$$
\underbrace{\int_{\frown} d\omega' \frac{g(\omega')}{\omega' - \omega}}_{\substack{\text{infinitesimal} \\ \text{semicircle around } \omega' = \omega}} = \lim_{\rho \to 0} \left(ig(\omega) \int_{+\pi}^{0} \frac{e^{i\varphi} \rho \, d\varphi}{\rho e^{i\varphi}} \right)
$$

$$
= ig(\omega) \int_{\pi}^{0} d\varphi = -ig(\omega)\pi.
$$

Hence

$$
g(\omega) = \frac{1}{i\pi} P \int_{-\infty}^{\infty} d\omega' \frac{g(\omega')}{\omega' - \omega},
\tag{16.18}
$$

or

$$
\begin{aligned}
g(-\omega) &= \frac{1}{i\pi} P \int_{-\infty}^{\infty} d\omega' \frac{g(\omega')}{\omega' + \omega} = \frac{1}{i\pi} P \int_{\infty}^{-\infty} d\omega' \frac{g(-\omega')}{\omega' - \omega} \\
&= -\frac{1}{i\pi} P \int_{-\infty}^{\infty} d\omega' \frac{g(-\omega')}{\omega' - \omega}.
\end{aligned}
$$

Comparison with

$$
g^*(\omega) = -\frac{1}{i\pi} P \int_{-\infty}^{\infty} d\omega' \frac{g^*(\omega')}{\omega' - \omega}
$$

verifies the relation obtained previously

$$g(-\omega) = g^*(\omega). \tag{16.19}$$

The right side of the integral relations we thus obtained contains the factor "i". Taking real and imaginary parts of Eq. (16.18) we obtain:

$$\Re g(\omega) = \frac{1}{\pi} P \int_{-\infty}^{\infty} d\omega' \frac{\Im g(\omega')}{\omega' - \omega},$$

$$\Im g(\omega) = -\frac{1}{\pi} P \int_{-\infty}^{\infty} d\omega' \frac{\Re g(\omega')}{\omega' - \omega}. \tag{16.20}$$

We also have

$$\Re g^*(\omega) = -\frac{1}{\pi} P \int_{-\infty}^{\infty} d\omega' \frac{\Im g^*(\omega')}{\omega' - \omega}, \quad \Im g^*(\omega) = \frac{1}{\pi} P \int_{-\infty}^{\infty} d\omega' \frac{\Re g^*(\omega')}{\omega' - \omega}.$$

These relations are the promised *dispersion relations*. They are also called *Hilbert transforms*.

We can rewrite the dispersion relations in a form involving only positive frequencies. To achieve this, we write

$$\Re g(\omega) = -\frac{1}{\pi} P \int_{-\infty}^{0} d\omega' \frac{\Im g(\omega')}{\omega' - \omega} - \frac{1}{\pi} P \int_{0}^{\infty} d\omega' \frac{\Im g(\omega')}{\omega' - \omega}$$

$$= \frac{1}{\pi} P \int_{0}^{\infty} d\omega' \frac{\Im g(-\omega')}{-\omega' - \omega} + \frac{1}{\pi} P \int_{0}^{\infty} d\omega' \frac{\Im g(\omega')}{\omega' - \omega},$$

(of course, the pole is contained only in one integral; in the other the principal value integral is an ordinary integral). Since

$$g(\omega) = g^*(-\omega), \quad g(-\omega) = g^*(\omega),$$

we have

$$\Re g(\omega) + i \Im g(\omega) = \Re g(-\omega) - i \Im g(-\omega),$$

i.e.

$$\Re g(\omega) = \Re g(-\omega), \quad \Im g(\omega) = -\Im g(-\omega). \tag{16.21}$$

Hence

$$\Re g(\omega) = -\frac{1}{\pi} P \int_{0}^{\infty} d\omega' \frac{\Im g(\omega')}{-\omega' - \omega} + \frac{1}{\pi} P \int_{0}^{\infty} d\omega' \frac{\Im g(\omega')}{\omega' - \omega}$$

$$= \frac{1}{\pi} P \int_{0}^{\infty} d\omega' \Im g(\omega') \left[\frac{1}{\omega' + \omega} + \frac{1}{\omega' - \omega} \right]$$

$$= \frac{1}{\pi} P \int_{0}^{\infty} \Im g(\omega') d\omega' \frac{2\omega'}{\omega'^2 - \omega^2}$$

$$= \frac{2}{\pi} P \int_{0}^{\infty} d\omega' \omega' \frac{\Im g(\omega')}{\omega'^2 - \omega^2}. \tag{16.22}$$

Similarly

$$
\Im g(\omega) = -\frac{1}{\pi} P \int_{-\infty}^{\infty} d\omega' \frac{\Re g(\omega')}{\omega' - \omega}
$$

$$
= -\frac{1}{\pi} P \int_{-\infty}^{0} d\omega' \frac{\Re g(\omega')}{\omega' - \omega} - \frac{1}{\pi} P \int_{0}^{\infty} d\omega' \frac{\Re g(\omega')}{\omega' - \omega}
$$

$$
= -\frac{1}{\pi} P \int_{0}^{\infty} d\omega' \frac{\Re g(-\omega')}{-\omega' - \omega} - \frac{1}{\pi} P \int_{0}^{\infty} d\omega' \frac{\Re g(\omega')}{\omega' - \omega},
$$

and using Eq. (16.21), this becomes

$$
\Im g(\omega) = -\frac{1}{\pi} P \int_{0}^{\infty} d\omega' \frac{\Re g(\omega')}{-\omega' - \omega} - \frac{1}{\pi} P \int_{0}^{\infty} d\omega' \frac{\Re g(\omega')}{\omega' - \omega}
$$

$$
= -\frac{1}{\pi} P \int_{0}^{\infty} d\omega' \Re g(\omega') \left[-\frac{1}{\omega' + \omega} + \frac{1}{\omega' - \omega} \right]
$$

$$
= -\frac{1}{\pi} P \int_{0}^{\infty} d\omega' \Re g(\omega') \frac{(\omega' + \omega) - (\omega' - \omega)}{(\omega'^2 - \omega^2)}
$$

$$
= -\frac{2\omega}{\pi} P \int_{0}^{\infty} d\omega' \frac{\Re g(\omega')}{\omega'^2 - \omega^2}. \tag{16.23}
$$

Dispersion relations of this kind were first formulated by Kramers and Kronig (1926),[*] and, in fact, *in application to the dielectric susceptibility $\chi_e(\omega)$,* which we introduced in Chapter 4, i.e.

$$
\chi_e(\omega) = \frac{\epsilon - \epsilon_0}{\epsilon_0} = \frac{\Re\epsilon(\omega) + i\Im\epsilon(\omega) - \epsilon_0}{\epsilon_0}. \tag{16.24}
$$

We obtained the dispersion relations for $g(\omega)$ from the assumption that this function is analytic (i.e. free of poles) in the upper half of the complex ω-plane. We saw earlier that the refractive index or rather its square or the dielectric constant have the same property. For this reason one can write down dispersion relations for this quantity. Thus, from the dispersion relation in terms of positive frequencies (by identifying $g(\omega)$ with $\epsilon(\omega) - \epsilon_0$), we obtain

$$
\Re\epsilon(\omega) - \epsilon_0 = \frac{2}{\pi} P \int_{0}^{\infty} dx \frac{x\Im\epsilon(x)}{x^2 - \omega^2},
$$

$$
\Im\epsilon(\omega) = -\frac{2\omega}{\pi} P \int_{0}^{\infty} dx \frac{\Re\epsilon(x) - \epsilon_0}{x^2 - \omega^2}. \tag{16.25}
$$

We consider these relations in some applications in the next section.

---

[*]H. A. Kramers: *Collected Scientific Papers*, p. 333; R. Kronig, *J. Opt. Soc. Am.* **12** (1926) 547. See also J. Hilgevoord: *Dispersion Relations and Causal Description* (North-Holland Publ. Co., Amsterdam, 1962).

16.7   Absorption: Special Cases

We consider three cases.

(a)   The case $\Im\epsilon(\omega) = 0$ for all frequencies ω is the *case without absorption*. Recall: $\mathbf{E}, \mathbf{H} \propto e^{i(k(\omega)x - \omega t)}, k(\omega) = (\omega/c)\tilde{n}(\omega), \tilde{n}(\omega) = \sqrt{\epsilon\mu}/\sqrt{\epsilon_0\mu_0}$. If $\epsilon(\omega)$ contains an imaginary part, then also $\tilde{n}(\omega)$ and $k(\omega)$, and hence $\mathbf{E}, \mathbf{H}$ have exponentially decreasing amplitudes, i.e. there is damping. If $\Im\epsilon(\omega) = 0$, then $\Re\epsilon(\omega) = \epsilon_0$, and

$$v_{\text{phase}} = \frac{\omega}{k} = \frac{c\sqrt{\mu_0\epsilon_0}}{\sqrt{\mu\epsilon}} = \frac{1}{\sqrt{\mu\epsilon}} = c'.$$

Thus there is no dispersion.[†] Effectively the dispersion relations say: If there is dispersion in some frequency region, then (in general) there is absorption in some other region.
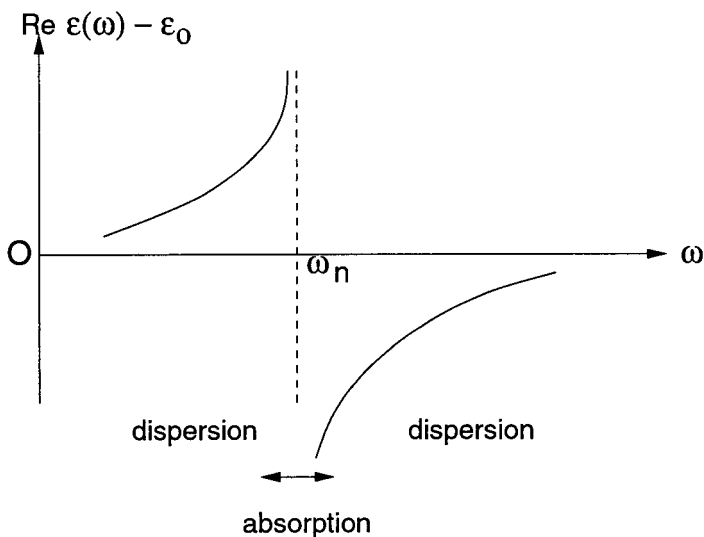
Fig. 16.2 Typical behaviour of $\Re\epsilon$ for resonance absorption.

(b)   The case $\Im\epsilon(\omega) = \kappa\delta(\omega - \omega_n), \kappa > 0, 0 < \omega_n < \infty$. This case is called the case of the *"zero width resonance"* and corresponds to that of absorption

---

[†]There is no dispersion when

$$\frac{d\tilde{n}(\omega)}{d\omega} = \frac{d}{d\omega}\left(\frac{ck}{\omega}\right) = 0.$$

in a very narrow frequency region around $\omega = \omega_n$. Then for $\omega \neq \omega_n$:

$$
\begin{aligned}
\Re\epsilon(\omega) - \epsilon_0 &= \frac{2}{\pi} P \int_0^\infty x dx \frac{\Im\epsilon(x)}{x^2 - \omega^2} \\
&= \frac{2}{\pi} P \int_0^\infty x dx \frac{\kappa\delta(x - \omega_n)}{x^2 - \omega^2} \\
&\simeq \frac{2}{\pi} \frac{\kappa\omega_n}{\omega_n^2 - \omega^2}.
\end{aligned}
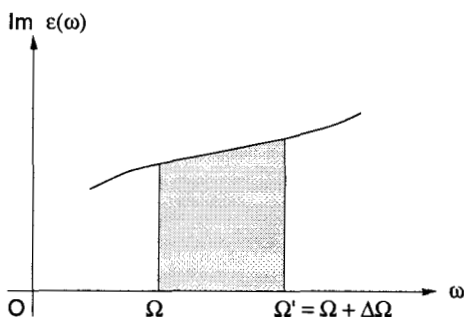\tag{16.26}
$$

This behaviour is illustrated in Fig. 16.2.

Fig. 16.3 Absorption in a finite frequency domain.

(c) The third case we consider is that with $\Im\epsilon(x) = \tau(x)\theta(x - \Omega)\theta(\Omega' - x)$, i.e. absorption in the domain $0 < \Omega < x < \Omega'$ with $\tau(x)$ regular in this domain. Here

$$
\Re\epsilon(\omega) - \epsilon_0 = \frac{2}{\pi} \int_\Omega^{\Omega'} dx x \frac{\tau(x)}{x^2 - \omega^2}.
$$

We consider two cases here:

(i) If $\omega \gg \Omega', \Omega$, we have approximately

$$
\Re\epsilon(\omega) - \epsilon_0 \simeq \frac{2}{\pi} \int_\Omega^{\Omega'} dx x \frac{\tau(x)}{-\omega^2}.
$$

If Ω', Ω are close together (a narrow band), we can write $\Omega' = \Omega + \Delta\Omega \ll \omega$ and

$$
\Re\epsilon(\omega) - \epsilon_0 \simeq -\frac{2}{\pi} \frac{\Omega}{\omega^2} \int_\Omega^{\Omega + \Delta\Omega} dx \tau(x),
$$

i.e.

$$
\int_0^\infty dx \Im\epsilon(x) = -\int_\Omega^{\Omega + \Delta\Omega} dx \tau(x) \simeq \frac{\pi}{2\Omega}\left\{\omega^2(\Re\epsilon(\omega) - \epsilon_0)\right\}.
$$

The left side represents the area under the curve $\Im\epsilon(x)$. The right side, which is valid for $\omega \gg \Omega', \Omega$, i.e. for

$$\lim_{\omega\to\infty}\left\{-\frac{\pi\omega^2}{2\Omega}(\Re\epsilon(\omega)-\epsilon_0)\right\},$$

represents the limiting value at high frequencies. Relations of this type are called *sum rules*. For instance, in our model consideration (see Eq. (15.27) for only one resonance):

$$\Re\epsilon(\omega)-\epsilon_0 = \frac{ne^2}{m}\frac{\omega_0^2-\omega^2}{(\omega_0^2-\omega^2)^2+\gamma^2\omega^2}f_0 \overset{\omega\to\infty}{\simeq} -\frac{ne^2}{m}\frac{f_0}{\omega^2}.$$

Hence

$$\int_0^\infty dx\Im\epsilon(x) = -\int_\Omega^{\Omega+\Delta\Omega} dx\tau(x) \approx -\frac{\pi}{2\Omega}\frac{ne^2}{m}f_0. \tag{16.27}$$

(ii) If, on the other hand, we have $\omega \ll \Omega, \Omega' = \Omega + \Delta\Omega$, then

$$\Re\epsilon(\omega)-\epsilon_0 \approx \frac{2}{\pi}\int_\Omega^{\Omega'} dxx\frac{\tau(x)}{x^2} \approx \frac{2}{\pi}\frac{1}{\Omega}\int_\Omega^{\Omega'} dx\tau(x) \approx -\frac{2}{\pi}\frac{1}{\Omega}\int_0^\infty \Im\epsilon(x).$$

This relation relates the area under the curve of $\Im\epsilon(x)$ to $\Re\epsilon(\omega)-\epsilon_0$ for $\omega \ll \Omega, \Omega'$, i.e. for $\omega \to 0$. In the above example

$$\lim_{\omega\to 0}\left\{\Re\epsilon(\omega)-\epsilon_0\right\} \simeq \frac{ne^2}{m}\frac{1}{\omega_0^2}f_0,$$

so that

$$-\int_0^\infty dx\Im\epsilon(x) \approx \frac{\pi}{2}\frac{\Omega}{\omega_0^2}\frac{ne^2}{m}f_0. \tag{16.28}$$

Thus one can infer from the behaviour of $\epsilon(x)$ the frequency ω_0. Since $\omega_0 \simeq \Omega$ (see Fig. 15.6), the expressions (16.27) and (16.28) agree with each other.

16.8   Comments on Principal Values

The principal value of an integral can also be defined by [‡]

$$P\int_{-\infty}^\infty f(\omega)\frac{d\omega}{\omega} = \lim_{\epsilon\to 0}\frac{1}{2}\left[\int_{-\infty}^\infty f(\omega)\frac{d\omega}{\omega-i\epsilon} + \int_{-\infty}^\infty f(\omega)\frac{d\omega}{\omega+i\epsilon}\right]. \tag{16.29}$$

[‡]See Eq. (16.17).

Here $f(\omega)$ must be a function which is analytic in the region of integration. The relation (16.29) is based on an identity, which is of considerable usefulness in many applications. This identity is *formally written*

$$\lim_{\epsilon \to 0} \left(\frac{1}{\omega \pm i\epsilon} \right) = P\left(\frac{1}{\omega} \right) \mp i\pi\delta(\omega). \tag{16.30}$$

The ϵ-prescription on the left side is to be understood as saying that with $\epsilon > 0$ the pole is to be excluded from the region enclosed by the contour. It is with this ϵ prescription that the expression on the left is to be understood as equal to the expression on the right. The identity implies the relation

$$\lim_{\epsilon \to 0} \int_{-\infty}^{\infty} \frac{f(\omega)d\omega}{\omega \pm i\epsilon} = P \int_{-\infty}^{\infty} \frac{f(\omega)d\omega}{\omega} \mp i\pi f(0), \tag{16.31}$$

in which $f(\omega)$ is an analytic function in the domain of integration. In order to prove the identity, we recall the original definition of the principal value integral, i.e.

$$P \int_{-\infty}^{\infty} f(\omega) \frac{d\omega}{\omega} = \lim_{\epsilon \to 0} \left[\int_{-\infty}^{-\epsilon} f(\omega) \frac{d\omega}{\omega} + \int_{\epsilon}^{\infty} f(\omega) \frac{d\omega}{\omega} \right]. \tag{16.32}$$

Consider the term on the left side of the identity Eq. (16.30):

$$\frac{1}{\omega \pm i\epsilon} = \frac{\omega \mp i\epsilon}{\omega^2 + \epsilon^2} = \frac{\omega}{\omega^2 + \epsilon^2} \mp \frac{i\epsilon}{\omega^2 + \epsilon^2}.$$

In the limit $\epsilon \to 0$ we obtain with the help of a definition of the delta function (cf. Eq. (2.5)), i.e.

$$\delta(x) = \lim_{\epsilon \to 0} \frac{1}{\pi} \frac{\epsilon}{(\epsilon^2 + x^2)}$$

the relations

$$\lim_{\epsilon \to 0} \frac{1}{\omega \pm i\epsilon} = \lim_{\epsilon \to 0} \frac{\omega}{\omega^2 + \epsilon^2} \mp i\pi\delta(\omega).$$

The first term on the right of this identity is $1/\omega$ for $\omega \neq 0$. Multiplying this relation by $f(\omega)$ and integrating with respect to ω, we obtain

$$\lim_{\epsilon \to 0} \int_{-\infty}^{\infty} \frac{f(\omega)d\omega}{\omega \pm i\epsilon} = \lim_{\epsilon \to 0} \int_{-\infty}^{\infty} \frac{\omega f(\omega)d\omega}{\omega^2 + \epsilon^2} \mp i\pi f(0)$$

$$= \lim_{\epsilon \to 0} \left[\int_{-\infty}^{-\epsilon} \frac{f(\omega)d\omega}{\omega} + \int_{\epsilon}^{\infty} \frac{f(\omega)d\omega}{\omega} \right]$$

$$+ \lim_{\epsilon \to 0} \int_{-\epsilon}^{\epsilon} \frac{\omega f(\omega)d\omega}{\omega^2 + \epsilon^2} \mp i\pi f(0)$$

$$= P \int_{-\infty}^{\infty} \frac{f(\omega)d\omega}{\omega} \mp i\pi f(0).$$

The integral from $-\epsilon$ to ϵ vanishes, since

$$\lim_{\epsilon \to 0} \int_{-\epsilon}^{\epsilon} \frac{\omega f(\omega) d\omega}{\omega^2 + \epsilon^2} = \lim_{\epsilon \to 0} f(0) \int_{-\epsilon}^{\epsilon} \frac{\omega d\omega}{\omega^2 + \epsilon^2} = 0,$$

since the integrand is an odd function of ω. We have thus verified the formal relation (16.30). Equation (16.29) now follows from the addition of both cases, provided the contributions of the integrals around the infinite semicircles vanish.

The relation with the upper sign is more commonly used than the other because it is applicable when the function $f(\omega)$ is analytic in the upper half-plane, i.e. the relation

$$\lim_{\epsilon \to 0} \int_{-\infty}^{\infty} \frac{f(\omega) d\omega}{\omega + i\epsilon} = P \int_{-\infty}^{\infty} \frac{f(\omega) d\omega}{\omega} - i\pi f(0).$$

If

$$\int_{-\infty}^{\infty} \cdots = \int_{\curvearrowright} \cdots = 2\pi i \sum \text{residues} = 0,$$

it follows that

$$P \int_{-\infty}^{\infty} \frac{f(\omega) d\omega}{\omega} = i\pi f(0).$$

in agreement with $f(\omega) \equiv g(\omega)$ in Eq. (16.18).

Example 16.2: Evaluation of a principal value integral

Evaluate with the help of the original definition of the principal value integral the following integral

$$P \int_{-\infty}^{\infty} \frac{\sin x' dx'}{x' - x}.$$

Solution: We separate the sine into its exponentials, so that

$$P \int_{-\infty}^{\infty} \frac{\sin x' dx'}{x' - x} = \frac{1}{2i} P \int_{-\infty}^{\infty} \frac{e^{ix'} dx'}{x' - x} - \frac{1}{2i} P \int_{-\infty}^{\infty} \frac{e^{-ix'} dx'}{x' - x}.$$

We consider the following integral along the contour C_+ in the complex x'-plane shown in Fig. 16.4. The contour C_+ excludes the pole at x. Hence $\int_{C_+} = 0$. Moreover $\int_{\curvearrowright} = 0$ (along the infinite semicircle), so that

$$P \int_{-\infty}^{\infty} + \int_{\curvearrowright} = 0,$$

where the second contribution represents the integral along the small semicircle around the point x of $x - \epsilon$ to $x + \epsilon$. It follows therefore that

$$\frac{1}{2i} P \int_{-\infty}^{\infty} \frac{e^{ix'} dx'}{x' - x} = -\frac{1}{2i} \int_{x-\epsilon}^{x+\epsilon} \frac{e^{ix'} dx'}{x' - x}.$$

With

$$x' - x = r e^{i\theta}, \quad \frac{dx'}{x' - x} = id\theta$$

we obtain

$$\frac{1}{2i} P \int_{-\infty}^{\infty} \frac{e^{ix'} dx'}{x' - x} = -\frac{1}{2i} \int_{\pi}^{0} id\theta e^{ix} = \frac{\pi e^{ix}}{2}.$$

Similarly we obtain for the contour C_- in Fig. 16.4 the result

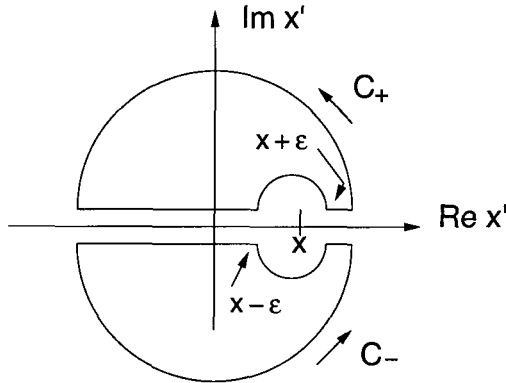$$\frac{1}{2i} \int_{C_-} \frac{e^{-ix'} dx'}{x' - x} = 0.$$

Fig. 16.4 The integration contours C_+ and C_-.

It follows that

$$\frac{1}{2i} P \int_{-\infty}^{\infty} \frac{e^{-ix'} dx'}{x' - x} = -\frac{1}{2i} \int_{x-\epsilon}^{x+\epsilon} \frac{e^{-ix'} dx'}{x' - x} = -\frac{1}{2i} \int_{-\pi}^{0} id\theta e^{-ix} = -\frac{\pi e^{-ix}}{2}.$$

Hence we obtain the result

$$P \int_{-\infty}^{\infty} \frac{\sin x' dx'}{x' - x} = \frac{\pi}{2}[e^{ix} + e^{-ix}] = \pi \cos x.$$

Example 16.3: Principal value integral[§]
Evaluate similar to the above integral

$$P \int_{a}^{\infty} \frac{\sin \sqrt{x'} dx'}{x' - x}.$$

Solution (with no derivation): For $a = 0, x > 0 : \pi \cos \sqrt{x}$; for $a = 0, x > 0(x \to -x) : e^{-\sqrt{x}}$.

[§]A table of principal value integrals of a similar kind was distributed by H. Carprasse, *An elementary method to calculate certain principal value integrals*, Universite de Liege, Belgium (1973), unpublished.

16.9   Subtracted Dispersion Relations

We obtained the dispersion relation

$$P \int_{-\infty}^{\infty} d\omega' \frac{g(\omega')}{\omega' - \omega} = i\pi g(\omega).$$

In the derivation we had assumed that

$$\int_{\frown} d\omega' \frac{g(\omega')}{\omega' - \omega} = 0$$

for $\omega' = Re^{i\theta}, R \to \infty$. It is possible, however, that this condition is not satisfied, e.g. it might happen that $g(\omega')$ approaches a nonzero constant g_0 for $|\omega'| \to \infty$. In the latter case, the integral implies

$$\int_{\frown} d\omega' \frac{g(\omega')}{\omega' - \omega} = g_0 \left(\lim_{R \to \infty} \int_{\frown} \frac{d\omega'}{\omega'} \right) = g_0 \int_0^\pi i d\theta = i\pi g_0 \neq 0.$$

We then write

$$G(\omega) = g(\omega) - g_0,$$

so that

$$\lim_{|\omega| \to \infty} G(\omega) = 0,$$

and the dispersion relations derived previously now apply to $G(\omega)$. If

$$g(\omega) \stackrel{|\omega| \to \infty}{\longrightarrow} g_1 \omega + g_0,$$

we set

$$G(\omega) = g(\omega) - (g_1 \omega + g_0)$$

and so on. The constants $g_0, g_1, \ldots$ are called *subtraction constants* and the dispersion relations are called *subtracted dispersion relations*.

Dispersion relations find important application in theories of the scattering of particles. Starting from known facts of quantum theory that bound states and resonances appear as poles of scattering amplitudes in definite variables, and that scattering states appear with branch cuts of the scattering amplitude, it is possible to construct expressions for such amplitudes in the form of dispersion relations. For instance, in the case of scattering of a particle of mass m by a Coulomb potential: The entire energy spectrum in $E = \hbar^2 k^2 / 2m$ consists of:

(a) *discrete bound states* with $E_n \propto -k_n^2, n = 0, 1, \ldots$, in the domain $E < 0$, and

(b) *continuous (scattering) states* in the domain $E > 0$ starting from $E = 0$ (as one finds, for instance, by solving the Schrödinger equation for the Coulomb potential).

The Coulomb problem possesses a branch point of the square-root type (2 Riemann sheets) at $k^2 = 0$, as one can infer from the quantum mechanical treatment (e.g. of the Coulomb phase). The *discontinuity of the function* at the cut from $k^2 = 0$ to ∞ is defined by $(f(k^{2*}) = f^*(k^2))$:

$$\frac{1}{2i}[f(k^2 + i\epsilon) - f(k^2 - i\epsilon)] = \Im f(k^2). \qquad (16.33)$$

For instance, $f(z) = z^{1/2}$ has a cut along the real axis from $z = 0$ to ∞. On the upper Riemann sheet

$$f(z) = |z^{1/2}|e^{i\theta/2}$$

and on the lower

$$f(z) = -|z^{1/2}|e^{i\theta/2}.$$

The discontinuity on the upper sheet $(\epsilon = x\theta)$ is

$$= \frac{1}{2i}\{f(z + i\epsilon) - f(z - i\epsilon)\} = |z^{1/2}|\frac{e^{i\theta/2} - e^{-i\theta/2}}{2i}$$

$$= |z^{1/2}|\sin(\theta/2) = \Im f(z) \equiv \text{discontinuity}.$$
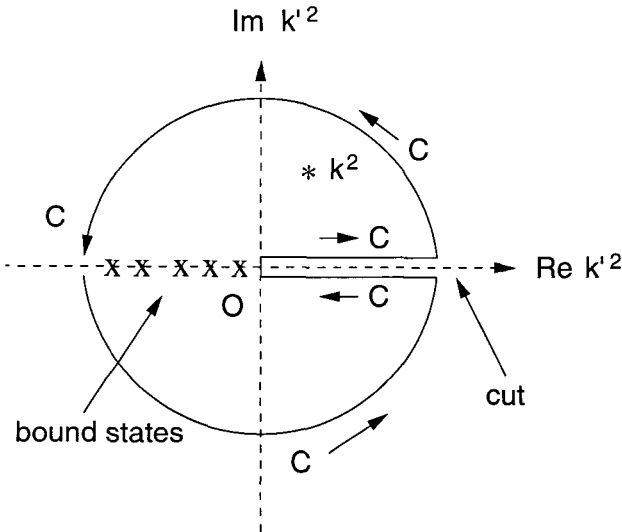
Fig. 16.5 Integration contour around the cut.

A scattering amplitude does not only depend on $E \propto k^2$, but also on the scattering angle θ, or some other corresponding variable. Thus, e.g. $f = f(k^2, \theta)$. Then (for k'^2 an arbitrary point in the k^2-plane)

$$\frac{1}{\pi} \int_0^\infty \frac{\Im f(k^2, \theta) dk^2}{k^2 - k'^2} = \frac{1}{2\pi i} \int_0^\infty dk^2 \frac{f(k^2 + i\epsilon, \theta) - f(k^2 - i\epsilon, \theta)}{k^2 - k'^2}.$$

This is an integral which circumvents the cut as indicated in Fig. 16.4. Consider the integral

$$\frac{1}{2\pi i} \int_C \frac{f(k^2, \theta)}{k^2 - k'^2} dk^2 \tag{16.34}$$

along the contour extended to infinity as in Fig. 16.4. Let us assume

$$\lim_{|k^2| \to \infty} f(k^2, \theta) = 0,$$

so that the contribution of the circle at infinity is zero. Let us also assume that $f(k^2, \theta)$ has simple poles at $k^2 = -k_n^2, n = 1, 2, \ldots$. Then, Cauchy's residue theorem tells us that

$$\frac{1}{\pi} \int_0^\infty \frac{f(k^2, \theta)}{k^2 - k'^2} dk^2 = \frac{1}{2\pi i} \int_C \frac{f(k^2, \theta)}{k^2 - k'^2} dk^2$$

$$= f(k'^2, \theta) + \sum_n \frac{g(k_n^2, \theta)}{k^2 - k'^2},$$

where

$$g(k_n^2, \theta) = \lim_{k^2 \to k_n^2} (k^2 - k_n^2) f(k^2, \theta)$$

(residue of the simple pole of f at k_n^2). Hence we obtain

$$f(k^2, \theta) = -\sum_n \frac{g(k_n^2, \theta)}{k_n^2 - k^2} + \frac{1}{\pi} \int_0^\infty dk'^2 \frac{\Im f(k'^2, \theta)}{k'^2 - k^2} \tag{16.35}$$

We have thus constructed an integral representation of the scattering amplitude with no explicit use of the potential; instead we have made use of knowledge of the spectrum obtained by solving the Schrödinger equation for the potential. The integral representation is also a dispersion relation, since for k^2 real the function $f(k^2, \theta)$ on the left is also real, i.e. the real part. One may note that here the integral was written down with our knowledge of one branch point, which characterises the onset of scattering states. The integral itself has the meaning of a principal value integral.

Chapter 17

Covariant Formulation of Electrodynamics

17.1 Introductory Remarks

In the following we consider first transformations from one frame of reference to another which moves with uniform velocity relative to the first; this is the topic of the Special Theory of Relativity. The significance and necessity of such considerations is immediately apparent if one recalls the field of an electric charge in its rest frame and then visualises this from a moving frame: In the first case one has only the static electric field, but in the second — in view of the motion of the charge — one observes also a magnetic field. It therefore becomes necessary to formulate Maxwell's equations, and more generally all laws of physics, independent of the respective reference frame, and this means in *covariant formulation*.

17.2 The Special Theory of Relativity

17.2.1 Introduction

We recapitulate first some aspects of the Special Theory of Relativity, which unifies Maxwell's electrodynamics with mechanics. The Special Theory of Relativity is based on the following two important principles or postulates of Einstein:

(1) *Einstein's principle of relativity* (1905). This principle says: The laws which describe the change of the state of a physical system are not affected by choosing one or another frame of reference which are related to each other by uniform translational motion (i.e. with constant relative velocity; systems

which are accelerated to each other are treated in the General Theory of Relativity).

Reference frames in uniform translational motion to each other (uniform motion meaning moving with constant velocity) are also called *inertial frames*. These are local reference frames in which free motion (meaning potential or applied force equal to zero) is weightless (i.e. without gravitation, i.e. no planets or other massive bodies nearby, hence the *Special* Theory of Relativity), and for which Newton's equation therefore implies*

$$m\ddot{x} = 0, \quad \therefore \quad m\dot{x} = \text{const.}, \quad \dot{x} = \frac{\text{const.}}{m}.$$

(2) *The velocity of light c is independent of the source* and is in vacuum and in the presence of electromagnetic fields the constant c in vacuum for all observers in inertial frames, as one would expect on the basis of the connection $c = 1/\sqrt{\epsilon_0 \mu_0}$. (This means that light is a field with the characteristics of waves and thus does not conform with classical corpuscular ideas).

We recapitulate briefly the development which led to these postulates. The first postulate can also be formulated as saying: It is not possible to determine a so-called "absolute velocity" (neither with electromagnetic nor with optical methods). What initiated these considerations? We recall that Newton's laws are based on the idea of an inertial frame, i.e. an unaccelerated, not rotating frame referred to an *absolute space*. This absolute space was assumed to be immovable of its own nature and with no reference to anything "exterior" (meaning that in the case of an acceleration in space, the apparent acceleration of this space — viewed from the accelerated object — is not a consequence of Newton's laws but a consequence of our acceleration relative to the absolute space, the so-called absolute acceleration). The concept of an absolute space thus offers a criterion for an absolute acceleration ($\ddot{x} = \text{const.}$), but not for the concept of an absolute velocity or a state of absolute rest. In fact, an absolute velocity cannot be determined, as the first of the above postulates of Einstein implies (see also the Michelson–Morley experiment below).

In preceding chapters we had introduced c as a constant which represents effectively the ratio between electric and magnetic units. Later we observed, that Maxwell's equations are wave equations for the fields $\mathbf{E}, \mathbf{H}$ with the phase velocity c in vacuum. More than a century ago numerous observations had led to the conclusion that c is the velocity of light. One then had to face the problem of understanding the propagation of electromagnetic waves in space, and it seemed suggestive to imagine their propagation similar to

---

*The rest frame of the particle is the frame with $\dot{x} = 0$; a particle with mass zero thus has no rest frame.

that of sound waves. Sound waves are vibrations in a medium like air. It was postulated therefore, that electromagnetic waves are propagated in a so-called "*aether*", which was assumed to be a space-filling medium introduced solely for the purpose of explaining the propagation of electromagnetic waves. c would then be the velocity of light with respect to this "*aether*", which was assumed to be stationary. Referred to a frame which moves with velocity u through the (stationary) *aether*, the velocity of light would — according to Galilean arguments — be $c' = c - u$ with respect to the moving frame. Then Maxwell's equations would have to be different in different reference frames in order to yield different velocities of light. One particular reference frame, that of the *aether*, would play a special role. This consideration is analogous to that of the velocity of sound waves relative to a train with velocity u, i.e.

$$v_{\text{relative train}} = v_{\text{in air}} - u_{\text{train}}.$$

It became necessary therefore to verify the possible existence of the *aether*, e.g. by measuring the absolute velocity v of the Earth in this *aether*. If the Earth moves through this *aether*, which was assumed to be stationary, there would have to be an observable "*aether wind*" in the opposite direction. The experiment of Michelson and Morley[†] was expected to allow the observation of this *wind*, i.e. its (absolute) velocity. Instead of considering the setup of this optical experiment, we consider the analogous case of a swimmer (velocity c in stationary water) in a river with current velocity $v \neq 0$. This consideration is simple and makes the issues solved by Einstein evident.[‡] The geometry is depicted in Fig. 17.1. The distances SR, ST in Fig. 17.1 are taken to be equal, i.e.

$$RS = ST = a.$$

Then the time the swimmer needs in order to swim from S to T and from T back to S (in the following we use the subscript "l" for "longitudinal") is

$$t_l = \frac{a}{c - v} + \frac{a}{c + v} = \frac{2ac}{c^2 - v^2}.$$

In order to reach R from S, the swimmer must swim in the direction of R'. The time he needs in order to swim from S to R is the same as the time he needs in order to swim from R to S, i.e.

$$t_\perp = 2 \frac{a}{\sqrt{c^2 - v^2}}.$$

[†]A. A. Michelson and E. W. Morley, *Am. J. Sci.* **34** (1887) 333; reprinted in *Relativity Theory: Its Origin and Impact on Modern Thought*, ed. L. P. Williams (Wiley, 1968). See also their paper in *Phil. Mag.* **24** (1887).

[‡]The swimmer analogy is also considered in L. R. Lieber, *The Einstein Theory of Relativity* (Dobson, 1949).
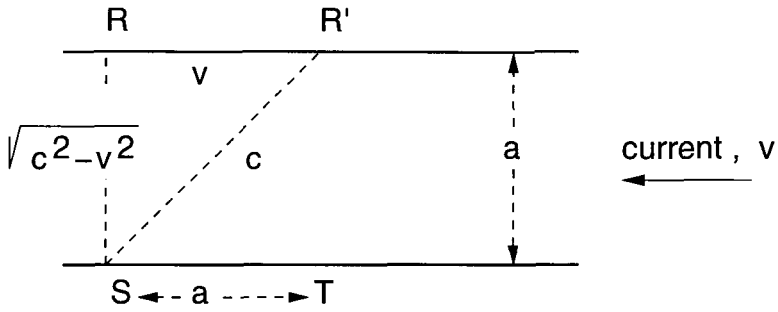
Fig. 17.1 The swimmer analogy to the experiment
of Michelson and Morley.

In the case of the swimmer (see below)

$$t_l > t_\perp \quad \text{for} \quad v < c,$$

since

$$
\begin{aligned}
t_l - t_\perp &= \frac{2ac}{c^2 - v^2} - 2\frac{a}{\sqrt{c^2 - v^2}} \\
&= 2\frac{a}{\sqrt{c^2 - v^2}}\left[\frac{c}{\sqrt{c^2 - v^2}} - 1\right] \\
&= 2\frac{a}{\sqrt{c^2 - v^2}}\left[\frac{1}{\sqrt{1 - \frac{v^2}{c^2}}} - 1\right] \\
&> 0.
\end{aligned}
\tag{17.1}
$$

In the Michelson–Morley experiment c is the velocity of light referred to the *aether*, and v the velocity of the Earth relative to the *aether*. The idea was, to measure $t_l, t_\perp$, and to determine v from $t_l - t_\perp$. The experiments of Michelson and Morley used light and mirrors and for the measurements interferometers of a high degree of precision. The results showed no difference between t_l and $t_\perp$, i.e. one had to conclude from the experiments that $t_l = t_\perp (v \neq 0)$, i.e.

$$t_\perp = 2\frac{a}{\sqrt{c^2 - v^2}} \overset{?}{=} \frac{2ac}{c^2 - v^2} = t_l.$$

Since it was required that $v \neq 0$, the next problem was to explain this result. Lorentz observed that both expressions can be made to agree for $v \neq 0$, if one assumes that the quantity a in t_l had to be replaced by

$$\frac{a\sqrt{c^2 - v^2}}{c} < a,$$

i.e. that parallel to v matter shrinks (as a consequence of its electric prop-
erties). This was the so-called "Lorentz-contraction" hypothesis. Since all
measuring rods would also shrink in the same way, the effect would not be
observable. This kind of adjustment of explanations was, however, rejected
in particular by Poincaré. Lorentz (1904) therefore undertook a deeper inves-
tigation which led him to the transformation which today carries his name,
however, without the appropriate explanation which was due to Einstein.
The transformation is written

$$x' = \gamma(x - vt), \quad y' = y, \quad z' = z,$$
$$t' = \gamma\left(t - \frac{vx}{c^2}\right), \quad \gamma = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}}. \tag{17.2}$$

Here x is the position coordinate measured in a frame fixed in the aether,
and x' that in a frame fixed on Earth at a time t, i.e. instead of

$$t = \frac{a}{c - v}, \quad (c - v)t = a = \underbrace{x'}_{\text{Earth}}$$

one has

$$t = \frac{x'}{c - v}$$

(in the frame fixed on Earth light travels in time t the distance $a \equiv x'$), so
that

$$x' = ct - vt = \underbrace{x}_{\text{aether}} -vt.$$

This expression assumes that the velocity of light c is independent of the
motion of the source. Consider again the case of the train T (with velocity
v with respect to the Earth), from which a whistle sends out a sound wave
with velocity u as indicated in Fig. 17.2. Here

$$x' = (u - v)t = x - vt,$$

i.e. u plays the role of c and is independent of the motion of the source. One
might think that by measuring x, x', t the velocity v could be determined as
an absolute velocity in contradiction to the result of the Michelson–Morley
experiment. Hence a contradiction? This apparent contradiction was re-
solved by Einstein: Every observer has his own clock, and in order to compare
times, light signals have to be exchanged. In the above $x = ct$ is the distance
travelled by light in time t (in the *aether*). The contraction hypothesis now
says that x' is to be replaced by
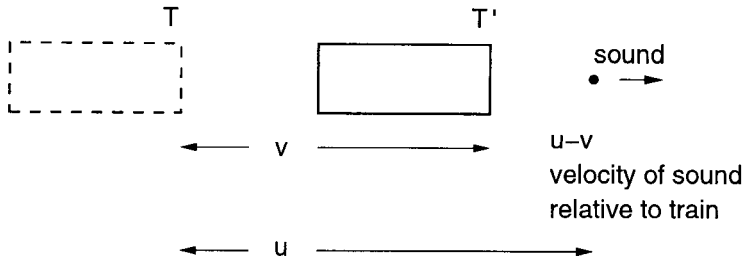
$$\sqrt{1 - \frac{v^2}{c^2}}x',$$

Fig. 17.2 Velocities of sound and the train.

so that

$$x' = \frac{x - vt}{\sqrt{1 - \frac{v^2}{c^2}}}.$$

Lorentz found that with

$$x' = \gamma(x - vt)$$

also a time t' given by

$$t' = \gamma\left(t - \frac{vx}{c^2}\right)$$

had to be introduced. But what is t'? Lorentz looked at t' as an artificial time, which had to be introduced for mathematical reasons. It became clear, therefore, that the fundamental concepts of length and time (also mass) had to be reexamined. This was done and achieved by Einstein in 1905. The statement of relativity, that for instance a material bullet fired with velocity u from a train, which is moving with velocity v, has the velocity u whether the train is moving or not, since

$$v_{\text{rel. train}} = (v + u) - v = u$$

was known before Einstein. Einstein's original intention was to extend this statement of classical mechanics to emitted light signals and hence to electrodynamics. In this Einstein realised that the laws of physics always have the same form, but are expressed in different reference frames (i.e. those of the respective observers). One consequence was a correction of Newtonian mechanics.

17.2.2 Einstein's Interpretation of Lorentz Transformations

The Lorentz transformation[§] from (x, y, z, t) to (x', y', z', t') thus describes the transformation from reference frame K of one observer to the frame K' of another observer, i.e.

$$x' = \gamma(x - vt), \quad y' = y, \quad z' = z, \quad t' = \gamma\left(t - \frac{vx}{c^2}\right). \qquad (17.3)$$

The Lorentz contraction therefore is not a real contraction (as Lorentz had suspected), but only an apparent shrinkage, which results from the difference of measurements made by observers in uniform relative motion to each other. Moreover, t' is simply the time measured in frame K'.

The reference to electrodynamics becomes particularly evident if we consider the propagation of light signals. If a bullet with velocity u is fired from a train moving with velocity v in the same direction, the velocity of the bullet with respect to a frame fixed on Earth is $u + v$, because the bullet as a travelling object had the velocity v before it was fired. If in a similar way a sound signal is sent out from a whistle on the train, i.e. with velocity u, then its velocity relative to the frame on Earth is u, i.e. less than $u + v$ (the reason being the different nature of the bullet and sound). The velocity of the sound wave depends only on the medium (i.e. the air), but not on the velocity of the source. Applied to light, this is the contents of the second postulate of above, as follows also from $c = 1/\sqrt{\epsilon_0 \mu_0}$. We deduce from these postulates now that for light observed in frame K:

$$x = ct \qquad (17.4)$$

is the distance travelled by the signal in time t. But then we must have (if the velocity of light is independent of the motion of the source) that

$$x' = ct' \qquad (17.5)$$

with the same c. We can infer from the Lorentz transformation that both equations, (17.4) and (17.5), are compatible with each other, since with $x = ct$ we have

$$x' = \gamma(c - v)t,$$

[§]The transformations today known as Lorentz transformations had already been formulated and published in 1887 by W. Voigt (*Gött. Nachr.* 1887, p. 45), however in relation to sound waves. In 1909 Lorentz wrote: *"which to my regret has escaped my notice all these years"* – see e.g. O'Rahilly [6], Vol. I, p. 325. In the chapter with title *"Voigt"* O'Rahilly attempts a critical analysis of these transformations.

so that

$$t' = \gamma\left(t - \frac{vct}{c^2}\right) = \frac{\gamma(c-v)t}{c} = \frac{x'}{c}$$

and hence

$$x' = ct',$$

i.e. light has the same velocity in every frame of reference, i.e. the observed velocity of light is independent of the motion of the observer (contrary to the claim of the *aether* theory, in which c would be the velocity of light with respect to the *aether* and hence the velocity of light observed by an observer travelling with velocity v in *aether* $c - v$, in which case the observed velocity of light would depend on the velocity of the observer).

Finally we make the following observation. Solving the equations

$$x' = \gamma(x - vt), \quad t' = \gamma\left(t - \frac{vx}{c^2}\right)$$

for x and t, we obtain

$$x = \gamma(x' + vt'), \quad t = \gamma\left(t' + \frac{vx'}{c^2}\right),$$

completely in agreement with the expectation that the frame with unprimed coordinates moves with velocity $-v$ away from the frame with primed coordinates if the frame with primed coordinates moves with velocity v away from the frame with unprimed coordinates. Einstein's principle of relativity says that the laws of physics must transform "*covariantly*" with respect to Lorentz transformations, i.e. independent of the respective reference frame in uniform translational motion.¶

17.3   Minkowski Space

We now consider Lorentz transformations in general. If one adds to the Lorentz transformations the simple translations in space and time, one obtains the *inhomogeneous Lorentz transformations*, which are also called *transformations of the Poincaré group* $\mathcal{P}$. A vector in *four-dimensional spacetime* is written in the so-called *covariant form*, which is characterised by *lower* indices, as

$$x_\mu, \quad \mu = 0, 1, 2, 3.$$

¶ "Covariant" means "co-varying", i.e. "transforming in the same way". "Covariance" can also mean independent of a specific coordinate frame, i.e. applicable to one as for any other.

The path element ds in this four-dimensional so-called *Minkowski space* $\mathbb{M}_4$ is with

$$x_0 := ct, \quad x_1 := -x, \quad x_2 := -y, \quad ; x_3 := -z$$

given by

$$(ds)^2 \equiv ds^2 = (dx_0)^2 - \sum_{i=1}^{3} (dx_i)^2. \tag{17.6}$$

We are therefore using the *metric* (or so-called *metric tensor*) of this space $\mathbb{M}_4$ given by:

$$(\eta_{\mu\nu}) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \equiv (\eta^{\mu\nu}), \quad \det(\eta_{\mu\nu}) \neq 0. \tag{17.7}$$

On the other hand, the unit matrix is defined by

$$(\mathbb{1}) = (\delta_\mu{}^\rho) \equiv (\eta_\mu{}^\rho), \quad \eta_\mu{}^\rho \eta_{\rho\nu} = \eta_{\mu\nu}, \tag{17.8}$$

so that

$$\delta_\mu{}^\rho \eta_{\rho\nu} \equiv \eta_{\mu\nu}, \quad \eta_{\mu\nu} \eta^{\nu\rho} = \delta_\mu{}^\rho. \tag{17.9}$$

The so-called *contravariant* vector x^μ with upper indices is then defined by

$$x^\mu = \eta^{\mu\nu} x_\nu,$$

so that in contravariant form

$$dx^\mu = \eta^{\mu\nu} dx_\nu,$$

or (in self-evident notation)

$$(dx^\mu) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} dx_0 \\ -dx_1 \\ -dx_2 \\ -dx_3 \end{pmatrix} = \begin{pmatrix} dx_0 \\ dx_i \end{pmatrix}$$

and

$$dx_\mu dx^\mu = (dx_0, -dx_i) \begin{pmatrix} dx_0 \\ dx_i \end{pmatrix} = (dx_0)^2 - (dx_i)^2 = (ds)^2.$$

With $\eta^{\mu\nu}$ we can "raise" or "lower" indices, and indices of the same type (one upper, one lower) are summed over (if not explicitly stated, this is understood, and called Einstein's summation convention). We shall become

familiar in the following with manipulations of this type. A quantity which
is such that all indices are summed over is a *Lorentz invariant*, e.g.

$$dx^\mu dx_\mu = (ds)^2$$

(the disappearance of the indices on the right side is described as their "con-
traction").

For later purposes we deviate a little and devote the remainder of this
section to the *transformation from Cartesian to spherical polar coordinates*
$(x_\mu) := (ct, r, \theta, \varphi)$. The square of the distance $(ds)^2$ between two neighbour-
ing spacetime points $P(ct, \mathbf{r})$ and $P(ct + cdt, \mathbf{r} + \mathbf{dr})$ with

$$x_0 = ct, \ x = r\sin\theta\cos\varphi, \ y = r\sin\theta\sin\varphi, \ z = r\cos\theta,$$

is, as we can also see geometrically,

$$ds^2 = (cdt)^2 - (dr)^2 - (rd\theta)^2 - r^2\sin^2\theta(d\varphi)^2 \equiv (cdt)^2 - dr^2 - r^2 d\Omega_2^2.$$

This means the metric $\tilde{\eta}$ is given by

$$(\tilde{\eta}) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -r^2 & 0 \\ 0 & 0 & 0 & -r^2\sin^2\theta \end{pmatrix}, \tag{17.10}$$

since then

$$(ds)^2 = \begin{pmatrix} cdt & dr & d\theta & d\varphi \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -r^2 & 0 \\ 0 & 0 & 0 & -r^2\sin^2\theta \end{pmatrix} \begin{pmatrix} cdt \\ dr \\ d\theta \\ d\varphi \end{pmatrix},$$

and so

$$(ds)^2 = \begin{pmatrix} cdt & dr & d\theta & d\varphi \end{pmatrix} \begin{pmatrix} cdt \\ -dr \\ -r^2 d\theta \\ -r^2\sin^2\theta d\varphi \end{pmatrix}$$

$$= (cdt)^2 - (dr)^2 - (rd\theta)^2 - r^2\sin^2\theta(d\varphi)^2. \tag{17.11}$$

We can obtain the metric $(\tilde{\eta})$ also as follows. We start from the transforma-
tion

$$dx'_\mu = A_\mu{}^\kappa dx_\kappa$$

and set

$$(x'_\mu) := (x^0, x, y, z) \quad \text{(Cartesian spatial coordinates)},$$
$$(x_\mu) := (x^0, r, \theta, \varphi) \quad \text{(spherical spatial coordinates)},$$

so that

$$(A_\mu{}^\kappa) = \begin{pmatrix} \frac{\partial x^0}{\partial x^0} & \frac{\partial x^0}{\partial r} & \frac{\partial x^0}{\partial \theta} & \frac{\partial x^0}{\partial \varphi} \\ \frac{\partial x}{\partial x^0} & \frac{\partial x}{\partial r} & \frac{\partial x}{\partial \theta} & \frac{\partial x}{\partial \varphi} \\ \frac{\partial y}{\partial x^0} & \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \theta} & \frac{\partial y}{\partial \varphi} \\ \frac{\partial z}{\partial x^0} & \frac{\partial z}{\partial r} & \frac{\partial z}{\partial \theta} & \frac{\partial z}{\partial \varphi} \end{pmatrix}. \tag{17.12}$$

Then

$$\begin{aligned} dx'^\mu dx'_\mu &= \eta^{\mu\nu} dx'_\nu dx'_\mu \\ &= \eta^{\mu\nu} A_\nu{}^\rho dx_\rho A_\mu{}^\kappa dx_\kappa \\ &= (A^{T\rho}{}_\nu \eta^{T\nu\mu} A_\mu{}^\kappa) dx_\rho dx_\kappa \\ &\equiv \tilde{\eta}^{\rho\kappa} dx_\rho dx_\kappa, \end{aligned} \tag{17.13}$$

where

$$\tilde{\eta}^{\rho\kappa} = A^{T\rho}{}_\nu \eta^{T\nu\mu} A_\mu{}^\kappa$$

and explicitly

$$(\tilde{\eta}^{\rho\kappa}) = \begin{pmatrix} \frac{\partial x^0}{\partial x^0} & \frac{\partial x}{\partial x^0} & \frac{\partial y}{\partial x^0} & \frac{\partial z}{\partial x^0} \\ \frac{\partial x^0}{\partial r} & \frac{\partial x}{\partial r} & \frac{\partial y}{\partial r} & \frac{\partial z}{\partial r} \\ \frac{\partial x^0}{\partial \theta} & \frac{\partial x}{\partial \theta} & \frac{\partial y}{\partial \theta} & \frac{\partial z}{\partial \theta} \\ \frac{\partial x^0}{\partial \varphi} & \frac{\partial x}{\partial \varphi} & \frac{\partial y}{\partial \varphi} & \frac{\partial z}{\partial \varphi} \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$$

$$\begin{pmatrix} \frac{\partial x^0}{\partial x^0} & \frac{\partial x^0}{\partial r} & \frac{\partial x^0}{\partial \theta} & \frac{\partial x^0}{\partial \varphi} \\ \frac{\partial x}{\partial x^0} & \frac{\partial x}{\partial r} & \frac{\partial x}{\partial \theta} & \frac{\partial x}{\partial \varphi} \\ \frac{\partial y}{\partial x^0} & \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \theta} & \frac{\partial y}{\partial \varphi} \\ \frac{\partial z}{\partial x^0} & \frac{\partial z}{\partial r} & \frac{\partial z}{\partial \theta} & \frac{\partial z}{\partial \varphi} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \frac{x}{r} & \frac{y}{r} & \frac{z}{r} \\ 0 & \frac{xz}{\sqrt{r^2-z^2}} & \frac{yz}{\sqrt{r^2-z^2}} & -\sqrt{r^2-z^2} \\ 0 & -y & x & 0 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \frac{x}{r} & \frac{xz}{\sqrt{r^2-z^2}} & -y \\ 0 & \frac{y}{r} & \frac{yz}{\sqrt{r^2-z^2}} & x \\ 0 & \frac{z}{r} & -\sqrt{r^2-z^2} & 0 \end{pmatrix}$$

$$= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -r^2 & 0 \\ 0 & 0 & 0 & -r^2 \sin^2\theta \end{pmatrix}.$$

17.4   The 10-Parametric Poincaré Group

The 10-parametric Poincaré group consists of the following transformations:

(1) *Translations* (4 independent parameters a_μ)

$$T_4 : \quad x'_\mu = x_\mu + a_\mu, \quad a_\mu \text{ real.}$$

The infinitesimal transformations are given by

$$\delta x_\mu = x'_\mu - x_\mu = \alpha_\mu \quad \text{infinitesimal.}$$

(2) *The proper homogeneous Lorentz transformations* (6 independent parameters)[‖]

$$L_6^+ : \quad x'_\mu = l_\mu{}^\nu x_\nu,$$

$$x'^\mu = \eta^{\mu\nu} l_\nu{}^\rho x_\rho = l^{\mu\rho} x_\rho$$

with the orthogonality condition

$$l^T{}_\mu{}^\kappa l_\kappa{}^\rho = \delta_\mu{}^\rho, \quad \underbrace{\det(l_\mu{}^\nu) = +1}_{\text{therefore } L^+}, \quad l_\mu{}^\nu \text{ real} \tag{17.14}$$

from the condition of invariance of $x^\mu x_\mu$ under Lorentz transformations, i.e. from

$$x'^\mu x'_\mu = x^\mu x_\mu.$$

The *transpose index* T is defined by

$$l^{T\kappa\nu} := l^{\nu\kappa}, \quad l^T{}_\mu{}^\kappa := l^\kappa{}_\mu \ne l_\mu{}^\kappa. \tag{17.15}$$

The matrices (l) define a 4×4 representation of the non-compact Lorentz group $SO(3,1)$; the number "3" refers to the three rotation angles with compact domains of validity, and the number "1" to the non-compact parameter of a Lorentz transformation. Strictly speaking the vectors x^μ, x_μ belong to different spaces T, T^*. But since a metric is defined on the 4-dimensional manifold $\mathbb{R}^4$ (the Minkowski metric η, which together with $\mathbb{R}^4$ defines the Minkowski space $\mathbb{M}_4$), it suffices to focus on only one space. The covariant

---

[‖] The real matrix l consists of $4 \times 4 = 16$ real elements or parameters, which are restricted by the orthogonality condition in Minkowski space, $l^T l = \mathbb{1}_{4 \times 4}$, i.e. by 4 (diagonal-) + 6 (off-diagonal-) = 10 conditions. Thus there remain 16 minus 10 = 6 independent parameters corresponding to 3 spatial rotations and 3 Lorentz transformations (and hence velocities) in these directions.

components of a contravariant vector $x^\nu \in T$ then follow from $x_\mu = \eta_{\mu\nu} x^\nu$. Verifying in detail:

$$
\begin{aligned}
x'^\mu x'_\mu &= \eta^{\mu\nu} l_\nu{}^\rho l_\mu{}^\kappa x_\rho x_\kappa = (\eta^{\mu\nu} l_\mu{}^\kappa)(l_\nu{}^\rho \eta_{\rho\epsilon}) x^\epsilon x_\kappa \\
&= l^{\nu\kappa} l_{\nu\epsilon} x^\epsilon x_\kappa = l^{T\kappa\nu} l_{\nu\epsilon} x^\epsilon x_\kappa \\
&= \eta^{\kappa\lambda} l^T{}_\lambda{}^\nu l_\nu{}^\omega \eta_{\omega\epsilon} x^\epsilon x_\kappa = \eta^{\kappa\lambda} \delta_\lambda{}^\omega \eta_{\omega\epsilon} x^\epsilon x_\kappa \\
&= \delta^\kappa{}_\epsilon x^\epsilon x_\kappa \quad (\eta^\kappa{}_\epsilon \equiv \delta^\kappa{}_\epsilon) \\
&= x^\kappa x_\kappa.
\end{aligned}
$$

A special case is the case of motion of the reference frame system $K'(x'_\mu)$ with constant velocity v along the x-axis of K. We set

$$
\beta = \frac{v}{c}, \quad \gamma = \frac{1}{\sqrt{1-\beta^2}}. \tag{17.16}
$$

The Lorentz transformation given earlier is now in contravariant and covariant forms respectively (with $x^1 \equiv x, x^0 \equiv ct$):

$$
\begin{aligned}
x'^0 &= \gamma(x^0 - \beta x^1), \\
x'^1 &= \gamma(x^1 - \beta x^0), \\
x'^2 &= x^2, \\
x'^3 &= x^3,
\end{aligned} \tag{17.17a}
$$

and

$$
\begin{aligned}
x'_0 &= \gamma(x_0 + \beta x_1), \\
x'_1 &= \gamma(x_1 + \beta x_0), \\
x'_2 &= x_2, \\
x'_3 &= x_3.
\end{aligned} \tag{17.17b}
$$

In this special case $(l_\mu{}^\nu)$ is given by (since $x'_\mu = l_\mu{}^\nu x_\nu$)

$$
(l_\mu{}^\nu) = \begin{pmatrix}
\frac{\partial x'_0}{\partial x_0} & \frac{\partial x'_0}{\partial x_1} & \frac{\partial x'_0}{\partial x_2} & \frac{\partial x'_0}{\partial x_3} \\
\frac{\partial x'_1}{\partial x_0} & \frac{\partial x'_1}{\partial x_1} & \frac{\partial x'_1}{\partial x_2} & \frac{\partial x'_1}{\partial x_3} \\
\frac{\partial x'_2}{\partial x_0} & \frac{\partial x'_2}{\partial x_1} & \frac{\partial x'_2}{\partial x_2} & \frac{\partial x'_2}{\partial x_3} \\
\frac{\partial x'_3}{\partial x_0} & \frac{\partial x'_3}{\partial x_1} & \frac{\partial x'_3}{\partial x_2} & \frac{\partial x'_3}{\partial x_3}
\end{pmatrix} \equiv \begin{pmatrix}
\gamma & \gamma\beta & 0 & 0 \\
\gamma\beta & \gamma & 0 & 0 \\
0 & 0 & 1 & 0 \\
0 & 0 & 0 & 1
\end{pmatrix}. \tag{17.18}
$$

Then, as expected,

$$
\det(l_\mu{}^\nu) = \gamma^2 - \gamma^2\beta^2 = \gamma^2(1-\beta^2) = +1.
$$

Moreover from

$$x'_\mu = l_\mu{}^\nu x_\nu,$$

i.e. the matrix relation we assumed above, i.e.

$$(x'_\mu) = \begin{pmatrix} ct' \\ -x' \\ -y' \\ -z' \end{pmatrix} = \begin{pmatrix} \gamma & \gamma\beta & 0 & 0 \\ \gamma\beta & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} ct \\ -x \\ -y \\ -z \end{pmatrix},$$

we obtain

$$ct' = \gamma ct - \gamma\beta x, \quad t' = \frac{t - \frac{vx}{c^2}}{\sqrt{1 - \beta^2}}$$

and

$$-x' = \gamma\beta ct - \gamma x, \quad x' = \frac{x - vt}{\sqrt{1 - \beta^2}}.$$

We now compute, also for later reference purposes,

$$(l_{\mu\nu}) = (l_\mu{}^\rho \eta_{\rho\nu}) = \begin{pmatrix} \gamma & \gamma\beta & 0 & 0 \\ \gamma\beta & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$$

$$= \begin{pmatrix} \gamma & -\gamma\beta & 0 & 0 \\ \gamma\beta & -\gamma & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}.$$

Moreover

$$(l^{\nu\rho}) = (\eta^{\nu\mu} l_\mu{}^\rho) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} \gamma & \gamma\beta & 0 & 0 \\ \gamma\beta & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

$$= \begin{pmatrix} \gamma & \gamma\beta & 0 & 0 \\ -\gamma\beta & -\gamma & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$$

and

$$(l^\nu{}_\mu) = (l^{\nu\rho} \eta_{\rho\mu}) = \begin{pmatrix} \gamma & -\gamma\beta & 0 & 0 \\ -\gamma\beta & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

We defined above[**]

$$l^T{}_\mu{}^\nu := l^\nu{}_\mu \neq l_\mu{}^\nu,$$ (17.19)

since (this verifies the orthogonality)

$$(l^T{}_\mu{}^\nu)(l_\nu{}^\rho) = (l^\nu{}_\mu)(l_\nu{}^\rho)$$

$$= \begin{pmatrix} \gamma & -\gamma\beta & 0 & 0 \\ -\gamma\beta & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \gamma & \gamma\beta & 0 & 0 \\ \gamma\beta & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

$$= \begin{pmatrix} \gamma^2 - \gamma^2\beta^2 & 0 & 0 & 0 \\ 0 & -\gamma^2\beta^2 + \gamma^2 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} = (\mathbb{1}).$$

In applications it is sometimes useful to use yet another formulation of the Lorentz transformation. We set

$$\cosh\phi := \frac{1}{\sqrt{1-\beta^2}}, \qquad \beta = \frac{v}{c}.$$ (17.20)

Since

$$\sinh^2\phi = \cosh^2\phi - 1 = -1 + \frac{1}{1-\beta^2} = \frac{\beta^2}{1-\beta^2},$$

we have

$$\sinh\phi = \frac{\beta}{\sqrt{1-\beta^2}}.$$

Then the equations

$$x' = \frac{x - vt}{\sqrt{1-\beta^2}}, \qquad t' = \frac{t - (vx/c^2)}{\sqrt{1-\beta^2}}$$

can be written

$$x' = x\cosh\phi - ct\sinh\phi, \qquad ct' = ct\cosh\phi - x\sinh\phi,$$ (17.21)

i.e.

$$(x'^\mu) = (A)(x^\mu)$$

[**]This — perhaps not very appealing — definition of indices (because of their skew positioning) agrees with $(l^{\nu\rho})^T = (l^{\rho\nu})$, since

$$(l^{\nu\rho})^T = (\eta^{\nu\mu} l_\mu{}^\rho)^T = (l^T{}_\mu{}^\rho)(\eta^{T\nu\mu}) = l^\rho{}_\mu \eta^{\mu\nu} = (l^{\rho\nu})$$

(also on multiplying out).

and

$$\begin{pmatrix} x' \\ y' \\ z' \\ ct' \end{pmatrix} = \begin{pmatrix} \cosh\phi & 0 & 0 & -\sinh\phi \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -\sinh\phi & 0 & 0 & \cosh\phi \end{pmatrix} \begin{pmatrix} x \\ y \\ z \\ ct \end{pmatrix}$$

with

$$\det(A) = \cosh^2\phi - \sinh^2\phi = 1.$$

This form of the transformation reminds one of rotations, however, with imaginary angles, which make $1/\sqrt{1-\beta^2}$ noncompact.

17.4.1   Covariant and Contravariant Derivatives

The (homogeneous) Lorentz transformations describe "rotations" in four-dimensional Minkowski space ("rotations" in the sense mentioned earlier). We now define contravariant and covariant derivatives as follows:[*]

$$(\partial^\mu) \equiv \left(\frac{\partial}{\partial x_\mu}\right) = \left(\frac{\partial}{\partial x_0}, \underbrace{-\nabla}_{\partial^i = \frac{\partial}{\partial x_i}}\right),$$

$$(\partial_\mu) \equiv \left(\frac{\partial}{\partial x^\mu}\right) = \left(\frac{\partial}{\partial x^0}, \underbrace{\nabla}_{\partial_i = \frac{\partial}{\partial x^i}}\right). \tag{17.22}$$

With these definitions we have

$$\partial^\mu \partial_\mu = \frac{\partial^2}{\partial x_0{}^2} - \nabla^2 \equiv -\Box. \tag{17.23}$$

The *chain rule* applies:

$$(\partial^\mu) \equiv \left(\frac{\partial}{\partial x_\mu}\right) = \underbrace{\frac{\partial x'^\rho}{\partial x_\mu}\frac{\partial}{\partial x'^\rho} = \eta^{\rho\kappa} l_\kappa{}^\mu \frac{\partial}{\partial x'^\rho}}_{x'^\rho = \eta^{\rho\kappa} x'_\kappa,\; x'_\kappa = l_\kappa{}^\nu x_\nu} = l^{\rho\mu}\partial'_\rho. \tag{17.24}$$

Hence

$$\partial^\mu = l^{\rho\mu}\partial'_\rho$$

and

$$\eta_{\nu\mu}\partial^\mu = \eta_{\nu\mu}l^{\rho\mu}\partial'_\rho,$$

so that

$$\partial_\nu = \partial'_\rho l^\rho{}_\nu. \tag{17.25}$$

---

[*]Note the reason: $\partial^\mu(x^2) \equiv \partial^\mu(x^\rho x_\rho) = 2x^\mu$.

We can show that this relation is identically satisfied, if ∂'_ρ transforms like the vector x'_ρ, because then

$$\partial'_\rho l^\rho{}_\nu = l^\rho{}_\nu \partial'_\rho = l^\rho{}_\nu l_\rho{}^\kappa \partial_\kappa \overset{(17.19)}{=} l^T{}_\nu{}^\rho l_\rho{}^\kappa \partial_\kappa \overset{(17.14)}{=} \delta_\nu{}^\kappa \partial_\kappa = \partial_\nu.$$

This means the derivatives transform like the corresponding vectors.

17.5 Construction of the Field Tensor $F_{\mu\nu}$

One can show that for instance the four-divergence $\partial_\mu A^\mu$ and the d'Alembert operator $\Box \equiv \partial^\mu \partial_\mu$ are *Lorentz scalars, Lorentz invariants*. Since $\det(l_\mu{}^\nu) = 1$, it follows that the transformation of the *volume element in Minkowski space* is (cf. Eq. (17.18) and thereafter)

$$dx'_0 dx'_1 dx'_2 dx'_3 = \begin{vmatrix} \dfrac{\partial x'_0}{\partial x_0} & \dfrac{\partial x'_0}{\partial x_1} & \dfrac{\partial x'_0}{\partial x_2} & \dfrac{\partial x'_0}{\partial x_3} \\[2mm] \dfrac{\partial x'_1}{\partial x_0} & \dfrac{\partial x'_1}{\partial x_1} & \dfrac{\partial x'_1}{\partial x_2} & \dfrac{\partial x'_1}{\partial x_3} \\[2mm] \dfrac{\partial x'_2}{\partial x_0} & \dfrac{\partial x'_2}{\partial x_1} & \dfrac{\partial x'_2}{\partial x_2} & \dfrac{\partial x'_2}{\partial x_3} \\[2mm] \dfrac{\partial x'_3}{\partial x_0} & \dfrac{\partial x'_3}{\partial x_1} & \dfrac{\partial x'_3}{\partial x_2} & \dfrac{\partial x'_3}{\partial x_3} \end{vmatrix} dx_0 dx_1 dx_2 dx_3$$

$$= dx_0 dx_1 dx_2 dx_3, \tag{17.26}$$

i.e. *the volume element is invariant.*

In order to achieve the covariant formulation of Maxwell's equations, we consider first the *continuity equation*, i.e.

$$\nabla \cdot \mathbf{j} + \frac{\partial \rho}{\partial t} = 0. \tag{17.27}$$

We can write this now (with $(\partial_\mu) = (\partial_0, \partial_i)$)

$$\partial_\mu J^\mu \equiv \frac{\partial J^\mu}{\partial x^\mu} = 0, \tag{17.28}$$

Here

$$J^\mu = (c\rho, \mathbf{j}) \tag{17.29}$$

is a four-vector (whether here $+\mathbf{j}$ or $-\mathbf{j}$, is a question of the direction of the current density $\mathbf{j} = \rho d\mathbf{x}/dt$). The *equations of the potentials*

$$\triangle \mathbf{A} - \frac{1}{c^2} \frac{\partial^2 \mathbf{A}}{\partial t^2} = -\mu_0 \mathbf{j},$$

$$\triangle \phi - \frac{1}{c^2} \frac{\partial^2 \phi}{\partial t^2} = -\frac{1}{\epsilon_0} \rho = -\mu_0 c^2 \rho, \tag{17.30}$$

together with the *Lorentz gauge*

$$\boldsymbol{\nabla} \cdot \mathbf{A} + \frac{1}{c^2}\frac{\partial \phi}{\partial t} = 0$$

and

$$(A^\mu) = \left(\frac{1}{c}\phi, \mathbf{A}\right) \equiv (A^0, \mathbf{A}), \tag{17.31}$$

can be written in the form

$$\partial^\rho \partial_\rho A^\mu = \mu_0 J^\mu, \quad -\square \equiv \partial^\mu \partial_\mu = \frac{1}{c^2}\frac{\partial^2}{\partial t^2} - \triangle, \tag{17.32}$$

with

$$\partial_\mu A^\mu = 0.$$

The *field strengths* $\mathbf{E}$ and $\mathbf{B}$ are given by

$$\mathbf{E} := (E_1, E_2, E_3) = -\boldsymbol{\nabla}\phi - \frac{\partial \mathbf{A}}{\partial t} = -c\left(\boldsymbol{\nabla} A^0 + \frac{\partial \mathbf{A}}{\partial x_0}\right),$$
$$\mathbf{B} := (B_1, B_2, B_3) = \boldsymbol{\nabla} \times \mathbf{A}, \tag{17.33}$$

and so, recalling Eqs. (17.22) and (17.31), by

$$\frac{1}{c}E_1 = -\frac{\partial A_0}{\partial x^1} - \frac{\partial A^1}{\partial x_0} = -\partial^0 A^1 + \partial^1 A^0,$$

$$\frac{1}{c}E_2 = -\frac{\partial A_0}{\partial x^2} - \frac{\partial A^2}{\partial x_0} = -\partial^0 A^2 + \partial^2 A^0,$$

$$\frac{1}{c}E_3 = -\frac{\partial A_0}{\partial x^3} - \frac{\partial A^3}{\partial x_0} = -\partial^0 A^3 + \partial^3 A^0,$$

$$B_1 = \frac{\partial A^3}{\partial x^2} - \frac{\partial A^2}{\partial x^3} = -(\partial^2 A^3 - \partial^3 A^2),$$

$$B_2 = \frac{\partial A^1}{\partial x^3} - \frac{\partial A^3}{\partial x^1} = -(\partial^3 A^1 - \partial^1 A^3),$$

$$B_3 = \frac{\partial A^2}{\partial x^1} - \frac{\partial A^1}{\partial x^2} = -(\partial^1 A^2 - \partial^2 A^1). \tag{17.34}$$

The similar form of the expressions for E_i/c and B_j suggests the *introduction of the field tensor* $F_{\mu\nu}$, i.e. an antisymmetric tensor of the second rank with 6 independent components, i.e.

$$F^{\mu\nu} \equiv \frac{\partial A^\nu}{\partial x_\mu} - \frac{\partial A^\mu}{\partial x_\nu} \equiv \partial^\mu A^\nu - \partial^\nu A^\mu$$

$$= \begin{pmatrix} 0 & -\frac{E_1}{c} & -\frac{E_2}{c} & -\frac{E_3}{c} \\ \frac{E_1}{c} & 0 & -B_3 & B_2 \\ \frac{E_2}{c} & B_3 & 0 & -B_1 \\ \frac{E_3}{c} & -B_2 & B_1 & 0 \end{pmatrix}. \tag{17.35}$$

One should note that here $\mathbf{E}/c$ and $\mathbf{B}$ appear, and not $\mathbf{E}/c$ and $\mathbf{H}$. The corresponding tensor with lower indices is obtained as

$$F_{\mu\nu} = \eta_{\mu\rho} F^{\rho\gamma} \eta_{\gamma\nu}$$

$$= \begin{pmatrix} 0 & \frac{E_1}{c} & \frac{E_2}{c} & \frac{E_3}{c} \\ -\frac{E_1}{c} & 0 & -B_3 & B_2 \\ -\frac{E_2}{c} & B_3 & 0 & -B_1 \\ -\frac{E_3}{c} & -B_2 & B_1 & 0 \end{pmatrix} = -F_{\nu\mu}. \qquad (17.36)$$

$F_{\mu\nu}$ is also an antisymmetric tensor. Furthermore we observe: $F_{\mu\nu}$ follows from $F^{\mu\nu}$ by the substitution $\mathbf{E} \to -\mathbf{E}$. Another quantity which is needed is the so-called *dual field tensor* $\mathcal{F}^{\mu\nu}$ defined by[†]

$$\mathcal{F}^{\mu\nu} \equiv {}^{*}F^{\mu\nu} = \frac{1}{2}\epsilon^{\mu\nu\rho\sigma} F_{\rho\sigma}$$

$$= \begin{pmatrix} 0 & -B_1 & -B_2 & -B_3 \\ B_1 & 0 & \frac{E_3}{c} & -\frac{E_2}{c} \\ B_2 & -\frac{E_3}{c} & 0 & \frac{E_1}{c} \\ B_3 & \frac{E_2}{c} & -\frac{E_1}{c} & 0 \end{pmatrix}, \qquad (17.37)$$

where

$$\epsilon^{\mu\nu\rho\sigma} = +1 \quad \text{for} \quad \mu\nu\rho\sigma \quad \text{an even permutation of } 0,1,2,3,$$
$$= -1 \quad \text{for} \quad \mu\nu\rho\sigma \quad \text{an odd permutation of } 0,1,2,3,$$
$$= 0, \quad \text{if two indices are equal.} \qquad (17.38)$$

(Note: *"Self-duality"* means ${}^{*}F^{\mu\nu} = F^{\mu\nu}$, *anti-self-duality* ${}^{*}F^{\mu\nu} = -F^{\mu\nu}$, neither of which apply here.)   The dual field tensor $\mathcal{F}^{\mu\nu}$ is obtained from $F^{\mu\nu}$ by the substitution

$$\frac{1}{c}\mathbf{E} \to \mathbf{B}, \quad \mathbf{B} \to -\frac{1}{c}\mathbf{E}.$$

We see: In four-form notation the electromagnetic field is no longer described by two separate vectors, but by one antisymmetric tensor of the second rank. In a Lorentz transformation, i.e. transformation from the reference frame of

---

[†]Note that in 3 dimensions (with $i,j,k = 1,2,3$, $g = \det(g_{ij}) = 1$) the dual of the 2-form F_{ij} is the 1-form B^k, i.e.

$$B^i = \frac{1}{2\sqrt{g}}\epsilon^{ijk} F_{jk},$$

and so $B^1 = F_{23}$, $B^2 = F_{31}$ and $B^3 = F_{12}$.

one observer to that of another observer moving relative to the first with uniform velocity, the components E_i, B_i are mixed. Thus, e.g. the pure field $\mathbf{E}/c$ of a static charge in one frame becomes a mixture of $\mathbf{E}'/c$ and $\mathbf{B}'$ in the other frame.
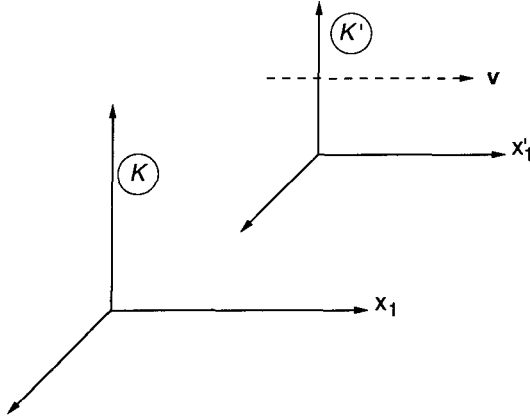
Fig. 17.3 Reference frames K, K' in uniform translational
motion with respect to each other.

The Lorentz transformation of the field tensor $F_{\mu\nu}$ is given by the transformation of a second rank tensor (see also below):

$$F'_{\mu\nu} = l_\mu{}^\sigma l_\nu{}^\rho F_{\sigma\rho} = l_\mu{}^\sigma F_{\sigma\rho} \underbrace{(l^{T\rho}{}_\nu)}_{=l_\nu{}^\rho} \tag{17.39}$$

with the special Lorentz transformation

$$(l_\mu{}^\nu) \overset{(17.18)}{=} \begin{pmatrix} \gamma & \gamma\beta & 0 & 0 \\ \gamma\beta & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{17.40}$$

called a *"boost"* along the x_1-axis with velocity $c\beta$ from reference frame $K(E, B)$ to reference frame $K'(E', B')$ indicated in Fig. 17.3.

The relation (17.39) defines a tensor of the second rank. That the quantity defined by the expression

$$F_{\mu\nu}(x) := \partial_\mu A_\nu(x) - \partial_\nu A_\mu(x)$$

is a covariant tensor of the second rank, can be shown by starting from the relation which defines a covariant vector (i.e. tensor of the first rank), and that

means the way this transforms under the appropriate transformations, in the present case of Lorentz transformations. Thus we recall the transformation of a covariant vector $A_\mu(x)$:

$$A_\nu(x') = \frac{\partial x^\mu}{\partial x'^\nu} A_\mu(x). \tag{17.41}$$

In order to verify Eq. (17.39), we have to show that with

$$x'_\mu = l_\mu{}^\nu x_\nu \tag{17.42}$$

the relation (17.41) implies

$$A_\mu(x') = l_\mu{}^\nu A_\nu(x). \tag{17.43}$$

Since $l^T{}_\mu{}^\kappa l_\kappa{}^\rho = \delta_\mu{}^\rho$, we deduce from (17.42):

$$x_\mu = l^T{}_\mu{}^\kappa x'_\kappa,$$

or

$$x^\mu = \eta^{\mu\rho} l^T{}_\rho{}^\kappa x'_\kappa = \eta^{\mu\rho} l^T{}_\rho{}^\kappa \eta_{\kappa\nu} x'^\nu,$$

so that

$$\frac{\partial x^\mu}{\partial x'^\nu} = \eta^{\mu\rho} l^T{}_\rho{}^\kappa \eta_{\kappa\nu} = l^\mu{}_\nu{}^T = l_\nu{}^\mu \tag{17.44}$$

and hence

$$\begin{aligned}
\frac{\partial x^\mu}{\partial x'^\nu} A_\mu(x) &= \eta^{\mu\rho} l^T{}_\rho{}^\kappa \eta_{\kappa\nu} A_\mu(x) \\
&= A^\rho(x) l^T{}_\rho{}^\kappa \eta_{\kappa\nu} \\
&= \eta_{\nu\kappa} l^\kappa{}_\rho A^\rho(x) \\
&= l_{\nu\rho} \eta^{\rho\kappa} A_\kappa(x) = l_\nu{}^\kappa A_\kappa(x),
\end{aligned}$$

i.e. Eq. (17.41) implies

$$A_\nu(x') = l_\nu{}^\kappa A_\kappa(x),$$

which had to be shown.

We now consider the covariant field strength tensor

$$
\begin{aligned}
F_{\mu\nu}(x') &\equiv \partial_{[\mu}A_{\nu]}(x') \\[2mm]
&= \frac{\partial}{\partial x'^\mu}A_\nu(x') - \frac{\partial}{\partial x'^\nu}A_\mu(x') \\[2mm]
&\overset{(17.41)}{=} \frac{\partial}{\partial x'^\mu}\left[A_\beta(x)\frac{\partial x^\beta}{\partial x'^\nu}\right] - \frac{\partial}{\partial x'^\nu}\left[A_\beta(x)\frac{\partial x^\beta}{\partial x'^\mu}\right] \\[2mm]
&= \left(\frac{\partial A_\beta(x)}{\partial x^\gamma}\frac{\partial x^\gamma}{\partial x'^\mu}\right)\frac{\partial x^\beta}{\partial x'^\nu} + A_\beta(x)\frac{\partial^2 x^\beta}{\partial x'^\mu \partial'^\nu} \\[2mm]
&\quad - \left(\frac{\partial A_\beta(x)}{\partial x^\gamma}\frac{\partial x^\gamma}{\partial x'^\nu}\right)\frac{\partial x^\beta}{\partial x'^\mu} - A_\beta(x)\frac{\partial^2 x^\beta}{\partial x'^\nu \partial x'^\mu} \\[2mm]
&= \frac{\partial x^\gamma}{\partial x'^\mu}\frac{\partial x^\beta}{\partial x'^\nu}\partial_{[\gamma}A_{\beta]}(x) \\[2mm]
&\overset{(17.44)}{=} l_\mu{}^\gamma l_\nu{}^\beta \partial_{[\gamma}A_{\beta]}(x)
\end{aligned}
\tag{17.45}
$$

in agreement with Eq. (17.39).

In multiplying matrices we have to keep in mind that Lorentz indices disappear through contraction. Consider first

$$
(l_\mu{}^\sigma F_{\sigma\rho}) \overset{(17.18)}{=}
\begin{pmatrix}
\gamma & \gamma\beta & 0 & 0 \\
\gamma\beta & \gamma & 0 & 0 \\
0 & 0 & 1 & 0 \\
0 & 0 & 0 & 1
\end{pmatrix}
\begin{pmatrix}
0 & \frac{E_1}{c} & \frac{E_2}{c} & \frac{E_3}{c} \\
-\frac{E_1}{c} & 0 & -B_3 & B_2 \\
-\frac{E_2}{c} & B_3 & 0 & -B_1 \\
-\frac{E_3}{c} & -B_2 & B_1 & 0
\end{pmatrix}
$$

$$
=
\begin{pmatrix}
-\frac{\gamma\beta E_1}{c} & \frac{\gamma E_1}{c} & \frac{\gamma E_2}{c} - \gamma\beta B_3 & \frac{\gamma E_3}{c} + \gamma\beta B_2 \\
-\frac{\gamma E_1}{c} & \frac{\gamma\beta E_1}{c} & \frac{\gamma\beta E_2}{c} - \gamma B_3 & \frac{\gamma\beta E_3}{c} + \gamma B_2 \\
-\frac{E_2}{c} & B_3 & 0 & -B_1 \\
-\frac{E_3}{c} & -B_2 & B_1 & 0
\end{pmatrix}
$$

$$
\equiv (L)_{\mu\rho}.
\tag{17.46}
$$

Next we consider

$$(L)_{\mu\rho}(l^T)^\rho{}_\nu$$

$$= \begin{pmatrix} -\frac{\gamma\beta E_1}{c} & \frac{\gamma E_1}{c} & \frac{\gamma E_2}{c} - \gamma\beta B_3 & \frac{\gamma E_3}{c} + \gamma\beta B_2 \\ -\frac{\gamma E_1}{c} & \frac{\gamma\beta E_1}{c} & \frac{\gamma\beta E_2}{c} - \gamma B_3 & \frac{\gamma\beta E_3}{c} + \gamma B_2 \\ -\frac{E_2}{c} & B_3 & 0 & -B_1 \\ -\frac{E_3}{c} & -B_2 & B_1 & 0 \end{pmatrix} \begin{pmatrix} \gamma & \gamma\beta & 0 & 0 \\ \gamma\beta & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

$$= \begin{pmatrix} 0 & \frac{E_1}{c} & \frac{\gamma E_2}{c} - \gamma\beta B_3 & \frac{\gamma E_3}{c} + \gamma\beta B_2 \\ -\frac{E_1}{c} & 0 & \frac{\gamma\beta E_2}{c} - \gamma B_3 & \frac{\gamma\beta E_3}{c} + \gamma B_2 \\ -\frac{E_2}{c} + \gamma\beta B_3 & -\frac{\gamma\beta E_3}{c} + \gamma B_3 & 0 & -B_1 \\ -\frac{\gamma E_3}{c} - \gamma\beta B_2 & -\frac{\gamma\beta E_3}{c} - \gamma B_2 & B_1 & 0 \end{pmatrix}$$

$$\tag{17.47}$$

This expression, the right hand side of Eq. (17.39), is to be identified with the left hand side of Eq. (17.39), i.e.

$$F'_{\mu\nu} = \begin{pmatrix} 0 & \frac{E'_1}{c} & \frac{E'_2}{c} & \frac{E'_3}{c} \\ -\frac{E'_1}{c} & 0 & -B'_3 & B'_2 \\ -\frac{E'_2}{c} & B'_3 & 0 & -B'_1 \\ -\frac{E'_3}{c} & -B'_2 & B'_1 & 0 \end{pmatrix},$$

so that we obtain the following transformation equations:

$$\begin{array}{ll} E'_1 = E_1, & cB'_1 = cB_1, \\ E'_2 = \gamma(E_2 - c\beta B_3), & cB'_2 = \gamma(cB_2 + \beta E_3), \\ E'_3 = \gamma(E_3 + c\beta B_2), & cB'_3 = \gamma(cB_3 - \beta E_2). \end{array} \tag{17.48}$$

It can be shown,[‡] that in the general case (i.e. for the constant velocity $\mathbf{v}$ with arbitrary orientation) these equations are:

$$\mathbf{E}' = \gamma(\mathbf{E} + \mathbf{v} \times \mathbf{B}) - \frac{\gamma^2}{c^2(\gamma+1)}\mathbf{v}(\mathbf{v} \cdot \mathbf{E}),$$

$$c\mathbf{B}' = \gamma\left(c\mathbf{B} - \frac{\mathbf{v} \times \mathbf{E}}{c}\right) - \frac{\gamma^2}{c(\gamma+1)}\mathbf{v}(\mathbf{v} \cdot \mathbf{B}). \tag{17.49}$$

For instance for E'_1 with $\mathbf{v}$ parallel to $\mathbf{x}_1$, and

$$\beta = v/c \quad \text{and} \quad \beta^2 = (\gamma^2 - 1)/\gamma^2,$$

[‡]See e.g. also Jackson [3].

one obtains the relations:

$$E'_1 = \gamma(E_1 + 0) - \frac{\gamma^2}{\gamma+1}\beta^2 E_1 = \frac{\gamma^2 + \gamma - \gamma^2\beta^2}{\gamma+1}E_1 = \frac{1+\gamma}{\gamma+1}E_1 = E_1$$

We observe again: **E** and **B** do not exist independently of each other. A purely electric or purely magnetic field in the reference frame of one observer appears as a mixture of both types of fields in the reference frame of another observer moving relative to the first with uniform velocity **v**. Thus the subdivision of the electromagnetic field into electric and magnetic components has no fundamental meaning. Which field components appear depends on the choice of reference frame of the observer. We also see that **E** and **B** are the fundamental quantities and not **E** and **H**.

17.6   Transformation from Rest Frame to Inertial Frame

As an example we consider the fields observed by an observer in reference frame K, when a point charge q passes him in rectilinear motion with uniform velocity **v**. This example is always impressive, because one starts from the simple Coulomb potential in the rest frame of the charge, then transforms to the other frame and ends up with the law of Biot–Savart in the nonrelativistic limit!

We let K' be the rest frame of the charge. In this case the transformation equations of the fields are inverse to those given above. We obtain the equations we need by making in the above the substitutions

$$\mathbf{E}', \mathbf{B}' \leftrightarrow \mathbf{E}, \mathbf{B} \qquad \text{and} \qquad \mathbf{v} \rightarrow -\mathbf{v}.$$

We obtain (since K moves away from K' with velocity $-\mathbf{v}$)

$$\begin{array}{ll}
E_1 = E'_1, & cB_1 = cB'_1, \\
E_2 = \gamma(E'_2 + c\beta B'_3), & cB_2 = \gamma(cB'_2 - \beta E'_3), \qquad (17.50) \\
E_3 = \gamma(E'_3 - c\beta B'_2), & cB_3 = \gamma(cB'_3 + \beta E'_2).
\end{array}$$

Of course, these equations also follow from the ones given earlier by solving for E_i, B_i. We let b be the shortest distance of the charge q from the observer at point P in K as indicated in Fig. 17.4. At time $t = t' = 0$ we let the origin O' be at O and charge q at the shortest distance from P. The coordinates of P as seen from K' are therefore according to Fig. 17.4:

$$x'_1 = -vt', \qquad x'_2 = b, \qquad x'_3 = 0,$$

where

$$r' = \sqrt{{x'_1}^2 + {x'_2}^2 + {x'_3}^2} = \sqrt{b^2 + (vt')^2}.$$

We ask: What is r', containing t', expressed in terms of the coordinates of $K(x_1, x_2, x_3, t)$? Thus we require $t' = t'(x_1, x_2, x_3, t)$. This follows from the Lorentz transformation in the direction of x_1:

$$t' = \gamma\left(t - \frac{vx_1}{c^2}\right) \quad \text{or} \quad t = \gamma\left(t' + \frac{vx'_1}{c^2}\right).$$
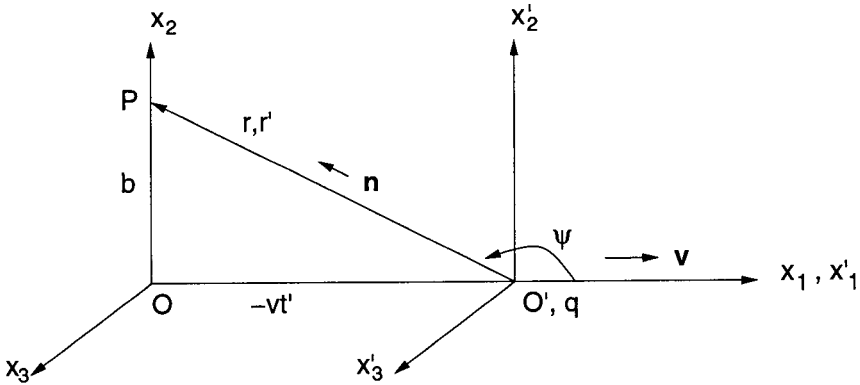

Fig. 17.4 Reference frames K, K' with charge q at the origin O' in K'.

Since we consider P fixed in K, we have $x_1 = 0$. This means

$$t' = \gamma t \quad \text{(and } x'_1 = -\gamma vt = -vt').$$

In the rest frame K' of the charge q the fields at the point P are:

$$E'_1 = -\frac{q}{4\pi\epsilon_0 r'^2}\left(\frac{vt'}{r'}\right) = -\frac{qvt'}{4\pi\epsilon_0 r'^3},$$

where $E'_1 = \mathbf{E}_1 \cdot \mathbf{e}_{x_1}$ and (cf. Fig. 17.4) $-vt'/r' = \cos\psi$. Similarly we have

$$E'_2 = \frac{q}{r'^2}\frac{\cos(\frac{\pi}{2} - \overline{\pi - \psi})}{4\pi\epsilon_0} = \frac{qb}{r'^3}\frac{1}{4\pi\epsilon_0}, \qquad E'_3 = 0 \quad \left(\propto \frac{x'_3}{r'^3}\right),$$

the latter would be the tangential component of a radial field. Since the charge is at rest in K', we also have

$$B'_1 = 0, \quad B'_2 = 0, \quad B'_3 = 0.$$

We now express these quantities in terms of the coordinates of K, i.e. with

$$r' = \sqrt{b^2 + (vt')^2} = \sqrt{b^2 + v^2\gamma^2 t^2}, \quad t' = \gamma t,$$

we obtain

$$E_1' = -\frac{q\gamma vt}{(b^2 + \gamma^2 v^2 t^2)^{3/2}}\frac{1}{4\pi\epsilon_0},$$

$$E_2' = \frac{qb}{(b^2 + \gamma^2 v^2 t^2)^{3/2}}\frac{1}{4\pi\epsilon_0}, \qquad (17.51)$$

and the others vanish. Substituting these expressions into the transformation equations (17.50), we obtain

$$\left.\begin{array}{l} E_1 = E_1' = -q\gamma vt/4\pi\epsilon_0(b^2 + \gamma^2 v^2 t^2)^{3/2}, \\ E_2 = \gamma E_2' = q\gamma b/4\pi\epsilon_0(b^2 + \gamma^2 v^2 t^2)^{3/2}, \end{array}\right\} \frac{E_1}{E_2} = -\frac{vt}{b},$$
$$E_3 = 0,$$
$$B_1 = B_1' = 0, \qquad (17.52)$$
$$B_2 = 0,$$
$$cB_3 = \gamma\beta E_2' = \beta E_2.$$

Thus the observer in K observes a magnetic field. The *relativistic effects* become particularly evident, if $v \to c$. First we see that in the *nonrelativistic limit* $\gamma \to 1$ $(\gamma = 1/\sqrt{1 - \beta^2})$:

$$cB_3 = \beta E_2 = \frac{q\gamma\beta b}{(b^2 + \gamma^2 v^2 t^2)^{3/2}}\frac{1}{4\pi\epsilon_0},$$

$$\underset{\gamma\to 1}{\simeq} \frac{q\beta b}{(b^2 + v^2 t^2)^{3/2}}\frac{1}{4\pi\epsilon_0},$$

which can be rewritten as

$$cB_3 = \frac{q}{c}\frac{vb}{r^3}\frac{1}{4\pi\epsilon_0} = \frac{q}{c}\frac{vr\sin(\pi - \psi)}{r^3}\frac{1}{4\pi\epsilon_0}$$

$$= \frac{q}{c}\frac{vr\sin\psi}{r^3}\frac{1}{4\pi\epsilon_0}$$

$$= \frac{q}{c}\frac{vr\sin(\mathbf{v}, \mathbf{r})}{r^3}\frac{1}{4\pi\epsilon_0},$$

$$c\mathbf{B} = q\frac{\mathbf{v}\times\mathbf{r}}{r^3}\frac{c\mu_0}{4\pi} \qquad (17.53)$$

with $c^2 = 1/\epsilon_0\mu_0$. Here, $-\mathbf{r}$ is the vector $\overrightarrow{PO'}$ described earlier as $\mathbf{r}$. We recognise the result (when divided by c) as the *law of Biot–Savart* for the magnetic field of a moving charge, i.e.

$$d\mathbf{B} = \frac{\mu_0 I\,d\mathbf{s}(\mathbf{r}') \times (\mathbf{r} - \mathbf{r}')}{4\pi|\mathbf{r} - \mathbf{r}'|^3}, \quad I = \frac{dq}{dt}, \quad \mathbf{v} = \frac{d\mathbf{s}}{dt}.$$

Thus the transformation laws of the Special Theory of Relativity supply not only relativistic corrections, but also the Biot–Savart law. However, it is not correct to conclude, it would be possible to derive the magnetic field and magnetic force, as well as the full set of Maxwell equations from the Coulomb law and Special Relativity. We referred to this already in Chapter 1.[§]

Next we have a closer look at *the longitudinal field E_1*. Considered as a function of vt this has the form shown in Fig. 17.5 for v very small ($\beta \simeq 0$) and very large ($\beta \lesssim 1$). The expression is

$$E_1 = -\frac{q\gamma vt}{4\pi\epsilon_0(b^2 + \gamma^2 v^2 t^2)^{3/2}}. \tag{17.54}$$
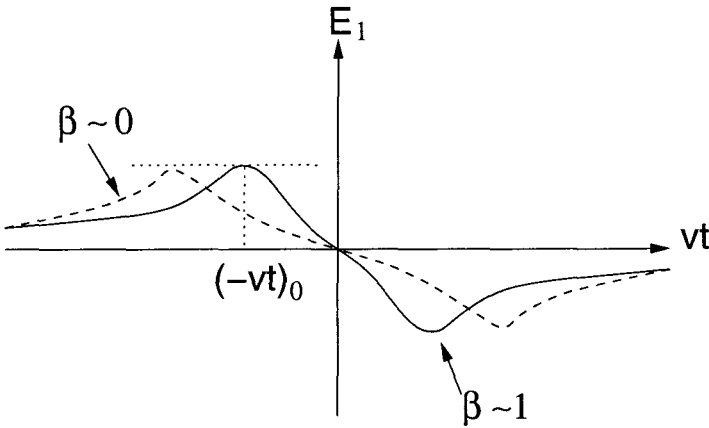
Fig. 17.5 E_1 for limiting cases of β.

The maxima and/or minima are given by

$$\frac{d}{dvt}E_1 = \frac{d}{dvt}\left[-\frac{q\gamma vt}{4\pi\epsilon_0(b^2 + \gamma^2 v^2 t^2)^{3/2}}\right] = 0,$$

i.e.

$$\frac{1}{(b^2 + \gamma^2 v^2 t^2)^{3/2}} - \frac{3}{2}\frac{2\gamma^2 v^2 t^2}{(b^2 + \gamma^2 v^2 t^2)^{5/2}} = 0,$$

or

$$b^2 + \gamma^2 v^2 t^2 - 3\gamma^2 v^2 t^2 = 0,$$

or

$$(vt)_0 = \pm\frac{1}{\sqrt{2}}\frac{b}{\gamma}, \quad \gamma^2 = \frac{1}{1 - \beta^2}.$$

[§]See e.g. Jackson [3], Sec. 12.2, p. 578.

For $\beta \lesssim 1$, we see that $(vt)_0$ is small; for $\beta \approx 0$, we see $(vt)_0 \sim \pm b/\sqrt{2} \sim$ is large, as shown in Fig. 17.5. The value of E_1 at $(vt)_0 \sim \pm b/(\sqrt{2}\gamma)$ is

$$(E_1)_0 = \mp \sqrt{\frac{4}{27}} \frac{q}{b^2} \frac{1}{4\pi\epsilon_0}.$$
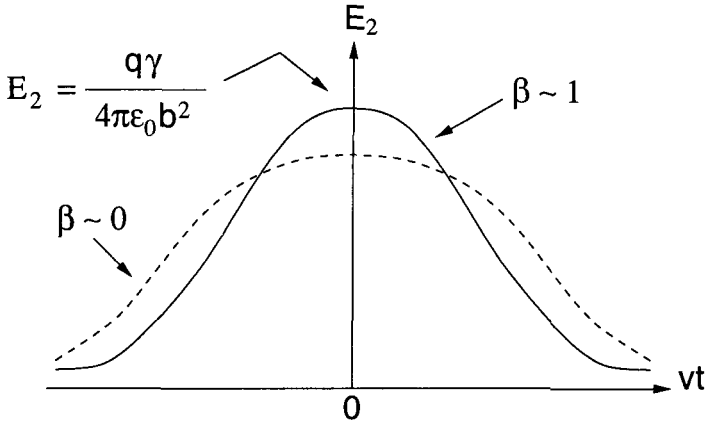
Fig. 17.6 E_2 in limits of β small and large.

This expression is independent of v; hence the maxima have the same height in Fig. 17.5. We see, that E_1 of Eq. (17.54) is an odd function of vt. In an observation in the time interval $\triangle t = (-t_0, t_0)$ the effects average out and E_1 is not observed at P. The observer at P therefore observes practically only the transverse field of the components $E_2, cB_3 = \beta E_2$, which for $\beta \approx 1$ are not only mutually transverse but also almost equal in modulus. If we look at the field E_2 as a function of vt, we see that this has the following form:

$$E_2 = \frac{q\gamma b}{4\pi\epsilon_0(b^2 + \gamma^2 v^2 t^2)^{3/2}}. \tag{17.55}$$

For the limiting cases the behaviour is sketched in Fig. 17.6. We see that E_2 has the shape of a "pulse". Thus the field of a charged particle moving at relativistic speed is equivalent to that of an electromagnetic pulse. This principle is exploited in attempts to observe heavy magnetic poles. In such experiments an atom is used to feel the effect of a passing magnetic pole (see also Chapter 21).

Finally we consider the spatial distribution of the fields with reference to the instantaneous position of the charge. With $\mathbf{E} = (E_1, E_2)$ we have the ratio

$$\frac{E_1}{E_2} = -\frac{vt}{b}.$$

From this and the force diagram we infer that the electric field has the direction of the vector $\mathbf{n}$ along r, with $\mathbf{E} \parallel \mathbf{n}$ and (see Fig. 17.7),

$$\cos \psi = \frac{\mathbf{n} \cdot \mathbf{v}}{v} = -\frac{vt}{r}.$$
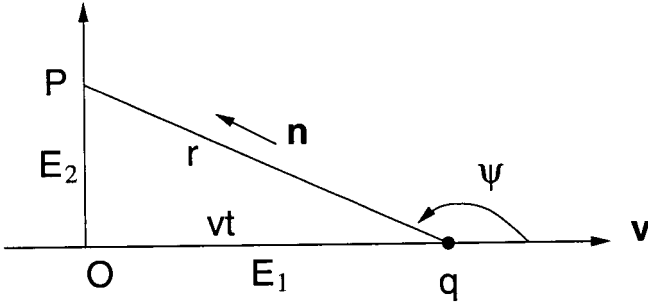
Fig. 17.7 The direction of the electric field.

But now

$$\sin \psi = \frac{b}{r}$$

and $(vt)^2 + b^2 = r^2$, so that

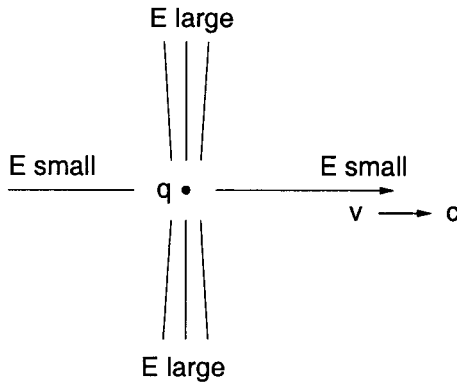$$1 - \beta^2 \sin^2 \psi = 1 - \beta^2 \frac{b^2}{r^2} = \frac{r^2 - \beta^2 b^2}{r^2} = \frac{b^2 + (vt)^2 - \beta^2 b^2}{r^2},$$

and hence

$$\begin{aligned}
1 - \beta^2 \sin^2 \psi &= \frac{b^2(1 - \beta^2) + (vt)^2}{r^2} \\
&= \frac{b^2/\gamma^2 + (vt)^2}{r^2} \\
&= \frac{b^2 + \gamma^2(vt)^2}{\gamma^2 r^2}.
\end{aligned} \tag{17.56}$$

Thus

$$\begin{aligned}
\mathbf{E} = (E_1, E_2) &= \frac{q\gamma(-vt, b)}{4\pi\epsilon_0 (b^2 + \gamma^2(vt)^2)^{3/2}} \\
&= \frac{q\gamma\mathbf{r}}{4\pi\epsilon_0 (\gamma^2 r^2)^{3/2}(1 - \beta^2 \sin^2 \psi)^{3/2}} \\
&= \frac{q\mathbf{r}}{4\pi\epsilon_0 (\gamma^2 r^3)(1 - \beta^2 \sin^2 \psi)^{3/2}}.
\end{aligned} \tag{17.57}$$
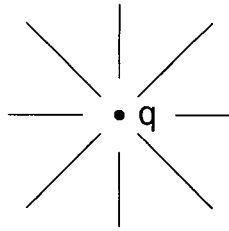
Fig. 17.8 The isotropic field for $\beta \sim 0$.

Thus the radial field has an isotropic distribution only for $\beta \sim 0$, i.e. as illustrated in Fig. 17.8. At $\psi = 0$ and π, the field $\mathbf{E}$ is

Fig. 17.9 The electric field for $v \sim c$.

$$\mathbf{E} = \frac{q\mathbf{r}}{4\pi\epsilon_0\gamma^2 r^3} = \frac{(1-\beta^2)q\mathbf{r}}{4\pi\epsilon_0 r^3} \to 0 \quad \text{for} \quad v \to c.$$

At $\psi = \pm\frac{\pi}{2}$, the field $\mathbf{E}$ is

$$\mathbf{E} = \frac{q\gamma^3\mathbf{r}}{4\pi\epsilon_0\gamma^2 r^3} = \frac{1}{4\pi\epsilon_0\sqrt{1-\beta^2}}\frac{q\mathbf{r}}{r^3} \to \infty \quad \text{for} \quad v \to c.$$

We see therefore: Whereas the electric field is isotropic in the case of the static (or almost static) charge, this is no longer the case for the field of a charge moving with velocity v; in the latter case the longitudinal component of the electric field (i.e. along the direction of motion of the particle) is smaller than the transverse component, and this effect becomes more and more pronounced as the velocity approaches that of light. This is illustrated in Fig. 17.9.

17.7 Covariant Form of Maxwell's Equations

Our next step is to reexpress Maxwell's equations in terms of the field strength tensor $F_{\mu\nu}$. Later we shall derive these equations from an action integral. Maxwell's equations as we obtained them are

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

$$\boldsymbol{\nabla} \times \mathbf{B} - \frac{1}{c^2}\frac{\partial \mathbf{E}}{\partial t} = \mu_0 \mathbf{j} \quad \text{(Ampère–Maxwell)},$$

$$\boldsymbol{\nabla} \cdot \mathbf{B} = 0 \quad \text{(no magnetic poles)},$$

$$\boldsymbol{\nabla} \times \mathbf{E} + \frac{\partial \mathbf{B}}{\partial t} = 0 \quad \text{(Faraday)}. \tag{17.58}$$

We now demonstrate that we obtain these equations by forming $\partial_\alpha F^{\alpha\beta}$ and $\partial_\alpha \mathcal{F}^{\alpha\beta}$, where $\mathcal{F}^{\alpha\beta}$ is the dual tensor defined by Eq. (17.37). In fact, we obtain (to be verified below):

$$\partial_\alpha F^{\alpha\beta} = \mu_0 J^\beta,$$

$$\partial_\alpha \mathcal{F}^{\alpha\beta} = 0, \tag{17.59}$$

for J^β defined below. We now verify these equations. Considering the left hand side of the first equation, we have

$$\partial_\alpha F^{\alpha\beta} = \left(\frac{\partial}{\partial x^0}, \boldsymbol{\nabla}\right)
\begin{pmatrix}
0 & -\frac{E_1}{c} & -\frac{E_2}{c} & -\frac{E_3}{c} \\
\frac{E_1}{c} & 0 & -B_3 & B_2 \\
\frac{E_2}{c} & B_3 & 0 & -B_1 \\
\frac{E_3}{c} & -B_2 & B_1 & 0
\end{pmatrix},$$

and hence

$$\partial_\alpha F^{\alpha\beta} = \left(\frac{1}{c}\boldsymbol{\nabla}\cdot\mathbf{E}, -\frac{1}{c}\frac{\partial E_1}{\partial x^0} + \frac{\partial B_3}{\partial y} - \frac{\partial B_2}{\partial z},\right.$$

$$-\frac{1}{c}\frac{E_2}{\partial x^0} - \frac{\partial B_3}{\partial x} + \frac{\partial B_1}{\partial z}, -\frac{1}{c}\frac{\partial E_3}{\partial x^0} + \frac{\partial B_2}{\partial x} - \frac{\partial B_1}{\partial y}\right)$$

$$= \left(\frac{1}{c}\boldsymbol{\nabla}\cdot\mathbf{E}, -\frac{1}{c}\frac{\partial \mathbf{E}}{\partial x^0} + \boldsymbol{\nabla}\times\mathbf{B}\right).$$

Setting this equal to

$$\mu_0 J^\beta = \mu_0(c\rho, \mathbf{j}),$$

and identifying components, we obtain (with $\mu_0\epsilon_0 c^2 = 1$)

$$\boldsymbol{\nabla}\cdot\mathbf{E} = \frac{\rho}{\epsilon_0}, \quad -\frac{1}{c}\frac{\partial \mathbf{E}}{\partial x^0} + \boldsymbol{\nabla}\times\mathbf{B} = \mu_0\mathbf{j},$$

in agreement with the first two of the above Maxwell equations.

Similarly

$$\partial_\alpha \mathcal{F}^{\alpha\beta} = \left(\frac{\partial}{\partial x^0}, \boldsymbol{\nabla}\right) \begin{pmatrix} 0 & -B_1 & -B_2 & -B_3 \\ B_1 & 0 & \frac{E_3}{c} & -\frac{E_2}{c} \\ B_2 & -\frac{E_3}{c} & 0 & \frac{E_1}{c} \\ B_3 & \frac{E_2}{c} & -\frac{E_1}{c} & 0 \end{pmatrix}$$

and hence

$$\partial_\alpha \mathcal{F}^{\alpha\beta} = \left(\boldsymbol{\nabla}\cdot\mathbf{B}, -\frac{\partial B_1}{\partial x^0} - \frac{1}{c}\frac{\partial E_3}{\partial y} + \frac{1}{c}\frac{\partial E_2}{\partial z},\right.$$
$$\left.-\frac{\partial B_2}{\partial x^0} + \frac{1}{c}\frac{\partial E_3}{\partial x} - \frac{1}{c}\frac{\partial E_1}{\partial z}, -\frac{\partial B_3}{\partial x^0} - \frac{1}{c}\frac{\partial E_2}{\partial x} + \frac{1}{c}\frac{\partial E_1}{\partial y}\right)$$
$$= \left(\boldsymbol{\nabla}\cdot\mathbf{B}, -\frac{\partial \mathbf{B}}{\partial x^0} - \frac{1}{c}\boldsymbol{\nabla}\times\mathbf{E}\right).$$

Equating this to zero, we obtain the other two Maxwell equations. The second of Eqs. (17.59) can also be written:

$$\partial_\alpha \epsilon^{\alpha\beta\rho\sigma} F_{\rho\sigma} = 0.$$

Using the total antisymmetry of the Levi–Civita symbol the left side can be written:

$$\epsilon^{\alpha\beta\rho\sigma}\partial_\alpha F_{\rho\sigma} = -\epsilon^{\beta\alpha\rho\sigma}\partial_\alpha F_{\rho\sigma} = -\frac{\epsilon^{\beta\alpha\rho\sigma}}{3}[\partial_\alpha F_{\rho\sigma} + \partial_\rho F_{\sigma\alpha} + \partial_\sigma F_{\alpha\rho}]$$

$$= -\frac{\epsilon^{\beta\alpha\rho\sigma}}{3}[\partial_\alpha\partial_\rho A_\sigma - \partial_\alpha\partial_\sigma A_\rho + \partial_\rho\partial_\sigma A_\alpha - \partial_\rho\partial_\alpha A_\sigma + \partial_\sigma\partial_\alpha A_\rho - \partial_\sigma\partial_\rho A_\alpha] = 0.$$

Solely from the antisymmetry

$$F_{\mu\nu} = -F_{\nu\mu}$$

we obtain

$$\partial_\alpha F_{\beta\gamma} + \partial_\beta F_{\gamma\alpha} + \partial_\gamma F_{\alpha\beta} = 0. \tag{17.60}$$

This equation is automatically valid for the antisymmetric tensor $F_{\mu\nu}$ and represents an identity which is called *Jacobi identity* or *Bianchi identity*. The equations $\partial_\alpha \mathcal{F}^{\alpha\beta} = 0$ follow from this, and are therefore sometimes referred to as examples of equations of motion which are not derived as Euler–Lagrange equations. Equations (17.59) are Maxwell's equations in covariant form, i.e. in form-invariant formulation, which means they transform under Lorentz transformations like Lorentz or Minkowski 4-vectors.

17.8 Covariantised Newton Equation of a Charged Particle

Our objective now is to reexpress the equation of motion of a Newtonian particle with charge q in the presence of an electromagnetic field (i.e. the Lorentz force) in covariant form, i.e. the equation we encountered several times earlier,

$$\frac{d\mathbf{p}}{dt} = q(\mathbf{E} + \mathbf{v} \times \mathbf{B}).$$ (17.61)

One should note that the left side of this equation consists of the vector $\mathbf{p}$ and the 0-component or d/dt of the 4-vector ∂_μ. It is therefore necessary to change something in order to achieve the same transformation property on the left of the equation as on the right. This is done by replacing t by an invariant quantity. In order to generalise the *Newton equation*

$$\frac{d}{dt}\mathbf{p} = \mathbf{F}, \quad \frac{d}{dt}p^i = F^i$$

relativistically, one introduces therefore the so-called *eigentime* τ, defined by

$$(ds)^2 = c^2(dt)^2 - (d\mathbf{x})^2$$
$$= c^2(dt)^2\left[1 - \frac{1}{c^2}\left(\frac{d\mathbf{x}}{dt}\right)^2\right]$$
$$= c^2(d\tau)^2.$$ (17.62)

τ is the time, measured by a clock fixed in the particle, i.e. in its rest frame. This is the reason that τ is called the *eigentime of the particle*. With

$$\mathbf{v} = \frac{d\mathbf{x}}{dt}$$

(which is constant for inertial frames), this is the important relation which describes the so-called *"clock paradox"*, i.e.*

$$d\tau = \frac{dt}{\gamma} = \sqrt{1 - \beta^2}dt, \quad \gamma^2 = \frac{1}{1 - \beta^2}.$$ (17.63)

This relationship expresses, so to speak, that a stationary clock runs faster than a moving clock, a phenomenon known as *"time dilation"* (cf. the

---

*For curvilinear coordinate systems the corresponding equation is $d\tau = \sqrt{g^{\rho\kappa}\,dx_\rho dx_\kappa}/c$.

discussion in Jackson [3] with detailed explanation of experimental verifica-
tions). Since $(ds)^2$ is Lorentz invariant, then also $(d\tau)^2$ or $\sqrt{(d\tau)^2} = d\tau$.
This suggests one to generalise the Newton equation as ($\alpha = 0, 1, 2, 3$)

$$\frac{d}{d\tau}(mu_\alpha) = K_\alpha \equiv \text{Minkowski force}$$

(definition). Here u_α is the *4-velocity* (also to be understood as tangential
vector with respect to the Minkowski manifold)

$$u^\alpha = \frac{dx^\alpha}{d\tau} \tag{17.64}$$

with (cf. (17.63))

$$u_i := -\frac{dx_i}{\sqrt{1-\beta^2}dt} = -\frac{v_i}{\sqrt{1-\beta^2}} \tag{17.65}$$

and

$$u_0 = \frac{dx_0}{d\tau} = \frac{cdt}{\sqrt{1-\beta^2}dt} = \frac{c}{\sqrt{1-\beta^2}}. \tag{17.66}$$

The *4-form equation* can now be written

$$\frac{d(mu_\alpha)}{\sqrt{1-\beta^2}dt} = K_\alpha. \tag{17.67}$$

The spatial part implies

$$-\frac{d}{dt}\left(m\frac{v_i}{\sqrt{1-\beta^2}}\right) = \sqrt{1-\beta^2}K_i.$$

Momentum conservation for $K_i = 0$ implies the identification

$$\text{momentum}: \quad p_i = \frac{mv_i}{\sqrt{1-\beta^2}}. \tag{17.68}$$

Hence also

$$F_i := -\sqrt{1-\beta^2}K_i. \tag{17.69}$$

The significance of the time component of the 4-form equation (17.67) can
be seen as follows. We multiply the equation by u^α:

$$u^\alpha\frac{d}{d\tau}(mu_\alpha) = u^\alpha K_\alpha = \frac{d}{d\tau}\left(\frac{1}{2}mu^\alpha u_\alpha\right).$$

However

$$u^\alpha u_\alpha = (u_0)^2 - (u_i)^2 = \frac{c^2}{1-\beta^2} - \frac{v^2}{1-\beta^2} = \frac{c^2 - v^2}{1-\beta^2} = c^2. \tag{17.70}$$

Hence

$$0 = u^\alpha K_\alpha, \quad \text{i.e.} \quad u_0 K_0 = u_i K_i,$$

or

$$\frac{c}{\sqrt{1-\beta^2}} K_0 = \frac{v_i}{\sqrt{1-\beta^2}} \frac{F_i}{\sqrt{1-\beta^2}},$$

i.e.

$$K_0 = \frac{1}{c} \frac{\mathbf{v} \cdot \mathbf{F}}{\sqrt{1-\beta^2}}. \tag{17.71}$$

Thus the fourth component of Eq. (17.67) becomes:

$$\frac{d(mc/\sqrt{1-\beta^2})}{\sqrt{1-\beta^2}dt} = \frac{1}{c} \frac{\mathbf{v} \cdot \mathbf{F}}{\sqrt{1-\beta^2}},$$

i.e.

$$\frac{d}{dt}\left(\frac{mc^2}{\sqrt{1-\beta^2}}\right) = \mathbf{v} \cdot \mathbf{F}, \tag{17.72}$$

i.e.

$$\frac{dT}{dt} = \mathbf{v} \cdot \mathbf{F}, \quad T = \frac{mc^2}{\sqrt{1-\beta^2}}. \tag{17.73}$$

Since $\mathbf{v} \cdot \mathbf{F}$ = work per unit time (or power), T is the *total energy*, i.e.

$$T = \frac{mc^2}{\sqrt{1-\beta^2}} = mc^2 + \frac{1}{2}mv^2 + O\left(\frac{1}{c^2}\right). \tag{17.74}$$

Then

$$p_\alpha = mu_\alpha \tag{17.75}$$

with

$$p_0 = mu_0 = \frac{mc}{\sqrt{1-\beta^2}} = \frac{T}{c}. \tag{17.76}$$

For $v \ll c$, Eq. (17.73) reproduces Newton's equation of motion:

$$\frac{d}{dt}\left(\frac{1}{2}m\mathbf{v}^2\right) = \mathbf{v} \cdot \mathbf{F},$$

or

$$\frac{d}{dt}(m\mathbf{v}) = \mathbf{F}.$$

From Eqs. (17.75) and (17.70), we obtain in addition

$$p^\alpha p_\alpha = m^2 u^\alpha u_\alpha = m^2 c^2, \tag{17.77}$$

and so

$$p_0^2 - \mathbf{p}^2 = m^2 c^2,$$

or

$$\left(\frac{T}{c}\right)^2 = \mathbf{p}^2 + m^2 c^2. \tag{17.78}$$

For $\mathbf{p} = 0$, $T \equiv E$ this entails the well known *Einstein formula* $E = mc^2$.

Finally we consider the Minkowskian generalisation of the Lorentz force:

$$K^\alpha \equiv qF^{\alpha\beta}u_\beta$$

$$= q\begin{pmatrix} 0 & -\frac{E_1}{c} & -\frac{E_2}{c} & -\frac{E_3}{c} \\ \frac{E_1}{c} & 0 & -B_3 & B_2 \\ \frac{E_2}{c} & B_3 & 0 & -B_1 \\ \frac{E_3}{c} & -B_2 & B_1 & 0 \end{pmatrix}\begin{pmatrix} \frac{c}{\sqrt{1-\beta^2}} \\ \frac{-v_x}{\sqrt{1-\beta^2}} \\ \frac{-v_y}{\sqrt{1-\beta^2}} \\ \frac{-v_z}{\sqrt{1-\beta^2}} \end{pmatrix}$$

and so

$$K^\alpha = \frac{q}{\sqrt{1-\beta^2}}\begin{pmatrix} \frac{\mathbf{E}\cdot\mathbf{v}}{c} \\ E_1 + v_y B_3 - v_z B_2 \\ E_2 + v_z B_1 - v_z B_3 \\ E_3 + v_x B_2 - v_y B_1 \end{pmatrix}$$

$$= \frac{q}{\sqrt{1-\beta^2}}\begin{pmatrix} \frac{\mathbf{E}\cdot\mathbf{v}}{c} \\ \mathbf{E} + \mathbf{v} \times \mathbf{B} \end{pmatrix}. \tag{17.79}$$

The equation

$$\frac{dp^\alpha}{d\tau} = \frac{d(mu^\alpha)}{d\tau} = K^\alpha$$

together with

$$\frac{dp^\alpha}{d\tau} = \frac{1}{\sqrt{1-\beta^2}}\frac{d}{dt}(p_0, \mathbf{p}) = \frac{1}{\sqrt{1-\beta^2}}\frac{d}{dt}\left(p_0, \frac{m\mathbf{v}}{\sqrt{1-\beta^2}}\right)$$

on the left hand side and (17.79) on the right, yields

$$\frac{d\mathbf{p}}{dt} \equiv \frac{d}{dt}\left(\frac{m\mathbf{v}}{\sqrt{1-\beta^2}}\right) = q(\mathbf{E} + \mathbf{v} \times \mathbf{B}),$$

$$\frac{dp_0}{dt} = q\frac{\mathbf{E}\cdot\mathbf{v}}{c}, \quad \text{i.e.} \quad \frac{d}{dt}\left(\frac{T}{c}\right) = q\frac{\mathbf{E}\cdot\mathbf{v}}{c},$$

and so

$$\underbrace{\frac{dT}{dt}}_{\text{power}} = q\underbrace{\frac{\mathbf{E}\cdot d\mathbf{x}}{dt}}_{\text{work/time}}. \tag{17.80}$$

We see that the 4-form

$$K^\alpha = qF^{\alpha\beta}u_\beta$$

is the Lorentz force supplemented by the power as the additional fourth component.

In the case of the macroscopic Maxwell equations we have to distinguish between $(\mathbf{E}, \mathbf{B})$ and $(\mathbf{D}, \mathbf{H})$. Evidently we need only to make the following substitutions in the above equations: $\mathbf{E} \to \mathbf{D}, \mathbf{B} \to \mathbf{H}$. The equations then describe *macroscopic averages of the atomic properties*. In this reformulation, the polarisation $\mathbf{P}$ and magnetisation $-\mathbf{M}$ can be combined like $(\mathbf{E}, \mathbf{B})$ and $(\mathbf{D}, \mathbf{H})$ to form an antisymmetric tensor of the second rank, and acquire their physical significance as macroscopic averages of atomic properties in the rest frame of the medium.

Comment

In the force-free case

$$\frac{du^\alpha}{d\tau} = 0, \quad \text{i.e.} \quad \frac{d^2 x^\alpha}{d\tau^2} = 0$$

in agreement with $d^2 x^\alpha/d\tau^2 = 0$ for constant velocity of a particle in an inertial frame. In curvilinear coordinate systems we have (also for $K = 0$) $d^2 u^\alpha/d\tau^2 \neq 0$, and one says, the coefficients of the metric act as potentials of fictitious forces.

17.9   Examples

Example 17.1: Gauge invariance
Is the theory defined by the Lagrangian

$$L = \frac{1}{2}mu^\alpha u_\alpha + eu_\alpha A^\alpha$$

gauge invariant?

Solution: The gauge transformation is

$$A^\alpha \to A^\alpha - \partial^\alpha\chi, \quad \chi = \chi(x^\beta) \text{ arbitrary.}$$

Hence the variation of the Lagrangian is with (see above) $u^\alpha = dx^\alpha/d\tau$

$$\delta L = -eu_\alpha \partial^\alpha \chi = -e\frac{\partial x^\beta}{\partial \tau}\frac{\partial \chi}{\partial x^\beta} = -e\frac{\delta \chi}{\delta \tau}.$$

The answer is therefore : Yes, because (with $\chi = 0$ at τ_1, τ_2)

$$\text{action} \quad S = \int_{\tau_1}^{\tau_2} L d\tau \quad \text{and} \quad \delta S = -e\int_{\tau_1}^{\tau_2} d\tau \frac{\delta \chi}{\delta \tau} \to 0.$$

Example 17.2: Addition of velocities

Let K, K', K'' denote reference frames with parallel axes. K' moves with velocity v away from K, and K'' moves parallel to these with velocity u away from K'. What is the relativistic law of addition of velocities (i.e. obtain the velocity w of K'' with respect to that of K)?

Solution: We use the Lorentz transformation in the form $X' = AX$, where $X = (x, y, z, ict)$ and (cf. Eq. (17.21))

$$A \equiv A_v = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \cosh\phi & i\sinh\phi \\ 0 & 0 & -i\sinh\phi & \cosh\phi \end{pmatrix}, \tag{17.81}$$

with

$$\cosh\phi = 1/\sqrt{1-\beta^2}, \quad \sinh\phi = \beta/\sqrt{1-\beta^2}, \quad \beta = v/c. \tag{17.82}$$

Performing two consecutive Lorentz transformations we obtain

$$A_w = A_u A_v = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \cosh(\phi+\phi') & i\sinh(\phi+\phi') \\ 0 & 0 & -i\sinh(\phi+\phi') & \cosh(\phi+\phi') \end{pmatrix} \equiv \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \cosh\overline{\phi} & i\sinh\overline{\phi} \\ 0 & 0 & -i\sinh\overline{\phi} & \cosh\overline{\phi} \end{pmatrix},$$

where $\phi, \phi', \overline{\phi}$ correspond to the velocities v, u, w. We obtain therefore $\overline{\phi} = \phi + \phi'$, where

$$\cosh\overline{\phi} = \frac{1}{\sqrt{1-\frac{w^2}{c^2}}}, \quad \cosh\phi = \frac{1}{\sqrt{1-\frac{v^2}{c^2}}}, \quad \cosh\phi' = \frac{1}{\sqrt{1-\frac{u^2}{c^2}}}.$$

With the relation

$$\cosh\overline{\phi} \equiv \cosh(\phi+\phi') = \cosh\phi\cosh\phi' + \sinh\phi\sinh\phi'$$

and the relations (17.82) it follows that

$$\frac{1}{\sqrt{1-\frac{w^2}{c^2}}} = \frac{1}{\sqrt{1-\frac{v^2}{c^2}}}\frac{1}{\sqrt{1-\frac{u^2}{c^2}}} + \frac{\frac{v}{c}\frac{u}{c}}{\sqrt{1-\frac{v^2}{c^2}}\sqrt{1-\frac{u^2}{c^2}}}$$

$$= \frac{1+\frac{uv}{c^2}}{\sqrt{1-\frac{u^2}{c^2}}\sqrt{1-\frac{v^2}{c^2}}}.$$

Squaring and taking the reciprocal and later the square root, we obtain

$$\frac{w^2}{c^2} = 1 - \frac{(1-\frac{u^2}{c^2})(1-\frac{v^2}{c^2})}{(1+\frac{uv}{c^2})^2}$$

$$= \frac{(\frac{u}{c}+\frac{v}{c})^2}{(1+\frac{uv}{c^2})^2},$$

and hence

$$w = \frac{u + v}{1 + \frac{uv}{c^2}}. \tag{17.83}$$

We see: For $uv \ll c^2$ the result is simply the sum as in the familiar nonrelativistic considerations.

Example 17.3: The twin problem, or clock paradox

A and B are twin brothers. On their 21st birthday A leaves his twin brother B on Earth and flies for 7 years in the direction of z (7 years measured on his own watch). The velocity v of A relative to that of the Earth is $\tilde{\gamma}c$, the fraction $\tilde{\gamma}$ of the velocity of light c. After 7 years A reverses his direction and returns to Earth with the same speed. How old are A and B then?

Solution: A is evidently $21 + 7 + 7 = 35$ years old according to his watch. We calculate the age of B (as measured by B on his watch) as follows. Reversing the Lorentz transformation of Eq. (17.21), i.e.

$$x' = x, \quad y' = y, \quad z' = z\cosh\phi - ct\sinh\phi, \quad t' = -\frac{z}{c}\sinh\phi + t\cosh\phi, \tag{17.84}$$

we obtain

$$x = x', \quad y = y', \quad z = z'\cosh\phi - ct'\sinh\phi, \quad t = \frac{z'}{c}\sinh\phi + t'\cosh\phi, \tag{17.85}$$

where with $\beta = v/c$ we have $\cosh\phi = 1/\sqrt{1 - \beta^2} = 1/\sqrt{1 - \tilde{\gamma}^2} > 1$. We now use the last of Eqs. (17.85): $t = \frac{z'}{c}\sinh\phi + t'\cosh\phi$. Since A is at rest in his frame K' with primed coordinates, we have $z' = 0$. Then t, the time measured by B on Earth in frame K, is for $t' = 7$ years: $t = 7\cosh\phi$ years, i.e.

$$t = \frac{7}{\sqrt{1 - \tilde{\gamma}^2}} \quad \text{years.} \tag{17.86}$$

The age of B is therefore at the time of return of A

$$= 21 + \frac{2 \times 7}{\sqrt{1 - \tilde{\gamma}^2}} \quad \text{years.}$$

If $\tilde{\gamma} = 4/5$, the age of B follows as:

$$21 + \frac{2 \times 7}{\sqrt{1 - \left(\frac{4}{5}\right)^2}} = 21 + \frac{2 \times 7 \times 5}{3} = 21 + \frac{70}{3} \simeq 44 \quad \text{years.}$$

Thus 35 years and 44 years are the time intervals which A and B measure on their respective watches. Hence in this way, as remarked above, the stationary clock appears to run faster than a moving clock.

Example 17.4: Lifetimes of particles

Establish the lifetime of a particle in the laboratory frame for the cases:
(a) a particle with constant velocity v, and
(b) a particle with arbitrary velocity v.
(c) Is the photon a stable particle?

Solution:
(a) and (b) The spacetime distance ds between the spacetime point 1 of creation of the particle and the spacetime point 2 of its decay is given by

$$ds^2 = c^2 dt^2 - d\mathbf{x}^2 = c^2 dt^2 \left\{ 1 - \left(\frac{d\mathbf{x}}{c\,dt}\right)^2 \right\}.$$

The invariant eigentime or proper time of the particle is the time τ given by $ds^2 = c^2 d\tau^2$. Thus in case (a) the lifetime of the particle measured in the laboratory frame is $(t_2 - t_1)$ given by

$$\tau_2 - \tau_1 = (t_2 - t_1)\sqrt{1 - \beta^2}, \quad v^2 = \beta^2 c^2,$$

and in case (b)

$$\tau_2 - \tau_1 = \int_{t_1}^{t_2} dt\sqrt{1 - \beta(t)^2}.$$

(c) Suppose the photon were unstable. Then the only invariant way to define its lifetime would be as that in its rest frame. However, since the velocity of the photon is constant and unequal to zero, it has no rest frame. This can also be seen as follows:

"Lifetime in rest frame" $= \tau_2 - \tau_1 = \underbrace{(t_2 - t_1)}_{\infty \text{ observed in lab. frame}} \underbrace{\sqrt{1 - \beta^2}}_{0} = 0(\infty) = \text{indeterminate}.$

Thus the question is meaningless. This applies to all particles with mass zero. (See also the remarks at the beginning of this chapter).

Example 17.5: The dipole in a magnetic field*

The charge $+q$ of two charges $+q, -q$, both initially at point **a**, is moved in the (x, y)-plane through the field F_{i3} to the point **b**.
(a) With the gauge choice $\partial_3(A_0, \mathbf{A}) = 0$ determine the mechanical momentum p_3 of the system in the direction 3.
(b) Is this momentum independent of the path from **a** to **b**?
(c) Show that the momentum p_3 is cancelled by a corresponding momentum of the electromagnetic field.

Solution:
(a) Starting from the Lorentz force $\mathbf{F} = \frac{d\mathbf{p}}{dt} = q\frac{d\mathbf{x}}{dt} \times \mathbf{B}, d\mathbf{p} = q d\mathbf{x} \times \mathbf{B}$, we have

$$p_3 = +q\int_{\mathbf{a}}^{\mathbf{b}} F_{3i} dx^i = -q\int_{\mathbf{a}}^{\mathbf{b}} F_{i3} dx^i.$$

Now,

$$\int_{\mathbf{a}}^{\mathbf{b}} F_{i3} dx^i = \int_{\mathbf{a}}^{\mathbf{b}} (F_{13} dx^1 + F_{23} dx^2)$$

$$= \int_{\mathbf{a}}^{\mathbf{b}} (\partial_1 A_3 - \partial_3 A_1) dx^1 + \int_{\mathbf{a}}^{\mathbf{b}} (\partial_2 A_3 - \partial_3 A_2) dx^2.$$

With the condition to be assumed, $\partial_3(A_0, \mathbf{A}) = 0$, we obtain

$$p_3 = -q\int_{\mathbf{a}}^{\mathbf{b}} \mathbf{\nabla} A_3 \cdot d\mathbf{x} = -q\int_{\mathbf{a}}^{\mathbf{b}} dA_3 = -q(A_3(\mathbf{b}) - A_3(\mathbf{a}))$$

and hence
$$p_3 = q(A_3(\mathbf{a}) - A_3(\mathbf{b})),$$

depending only on the endpoints **a** and **b**.
(b) We consider the curl of $d\mathbf{p}$:

$$\mathbf{\nabla} \times d\mathbf{p} = \mathbf{\nabla} \times \{q d\mathbf{x} \times \mathbf{B}\} = q[\underbrace{\mathbf{\nabla} \cdot \mathbf{B}}_{0}(d\mathbf{x}) - (\underbrace{\mathbf{\nabla} \cdot d\mathbf{x}}_{d(\mathbf{\nabla} \cdot \mathbf{x}) = 0})\mathbf{B}] = 0,$$

---

*See also W. Taylor, JHEP **0007** (2000) 039.

since $dF = \partial_1 F_{23} + \partial_2 F_{31} + \partial_3 F_{12} = 0$, i.e. here $\boldsymbol{\nabla} \cdot \mathbf{B} = 0$. As a result of this condition p_3 is independent of the path from $\mathbf{a}$ to $\mathbf{b}$ (compare with the definition of a conservative force in mechanics).

(c) The field momentum of the electromagnetic field has been defined in Sec. 7.7 as

$$\mathbf{P}^{\text{field}} = \int dV (\mathbf{D} \times \mathbf{B}).$$

Thus we have in the present case for this

$$P_3^{\text{field}} = \epsilon_0 c \int dV F_{0i} F_{3i} = -\epsilon_0 c \int dV F_{0i} \partial_i A_3 = \epsilon_0 c \int dV A_3 \partial_i F_{0i}.$$

In the last step the integral over a total divergence is zero. In the next step we use Eq. (17.59), i.e. the Maxwell equation $\boldsymbol{\nabla} \cdot \mathbf{D} = \rho$ in the form $\partial_i F_{0i} = q\mu_0 c\{\delta(\mathbf{x} - \mathbf{b}) - \delta(\mathbf{x} - \mathbf{a})\}$ together with $\epsilon_0 \mu_0 c^2 = 1$, so that

$$P_3^{\text{field}} = \epsilon_0 c \mu_0 c q \int dV A_3\{\delta(\mathbf{x} - \mathbf{b}) - \delta(\mathbf{x} - \mathbf{a})\} = q(A_3(\mathbf{b}) - A_3(\mathbf{a})),$$

which had to be shown. Thus, the mechanical momentum determined by the Lorentz force, is cancelled by a corresponding negative momentum of the field.

Example 17.6: Orthogonality expressed invariantly
Show that when $\mathbf{E} \cdot \mathbf{B} = 0$, also $F_{\mu\nu}\mathcal{F}^{\mu\nu} = 0$.

Solution: Using Eq. (17.36) for $F_{\mu\nu}$ and Eq. (17.37) for $\mathcal{F}^{\mu\nu}$ and multiplying one finds

$$F_{\mu\nu}\mathcal{F}^{\mu\nu} = \frac{\mathbf{E} \cdot \mathbf{B}}{c}\mathbb{1}.$$

This page intentionally left blank

Chapter 18

The Lagrange Formalism for the Electromagnetic Field

18.1 Introductory Remarks

In this chapter we derive Maxwell's equations as the equations of motion of the electromagnetic field in spacetime in the customary way as the Euler–Lagrange equations derived from an action integral. We discuss gauge invariance, transversality and masslessness of the electromagnetic field and touch briefly (restricted by our predominantly classical topic here) its spin. Finally examples are given to illustrate the diversity of some recent explorations motivated by the Lagrangian of Maxwell electrodynamics.

18.2 Euler–Lagrange Equation

We define as *Lagrangian density* the functional

$$\mathcal{L}(A_\mu(x), \partial_\mu A_\nu(x)).$$

We note that $\mathcal{L}$ is *not* $\mathcal{L}(A_\mu(x), \partial_\mu A_\nu(x), t)$. Here $A_\mu(x)$ is the electromagnetic 4-potential describing the local electromagnetic field. We define as *Lagrangian* the volume integral

$$L(t) = \int d^3x \mathcal{L}(A_\mu(x), \partial_\mu A_\nu(x))$$

and as *action* or *action integral*

$$S = \int_{t_1}^{t_2} dt L(t).$$

We demand that S or rather the Lagrangian density be a Lorentz invariant (recall we had verified previously the invariance of the spacetime (Minkowski space) volume element). The time dependence of L is implicitly contained in the fields. Our aim is, to derive the Maxwell equations in analogy to equations of motion in classical mechanics. Hence we construct an action to which we apply a *Hamilton's principle*. This means we demand that in varying the action by varying the fields $A_\mu(x)$ and their derivatives $\partial_\mu A_\nu(x)$ at fixed endpoints, *this action is to remain stationary*. Thus we demand

$$0 = \delta S = \delta \int_{t_1}^{t_2} dt d^3 x \mathcal{L}(A_\mu(x), \partial_\mu A_\nu(x))$$

with

$$\delta A_\mu = 0 \quad \text{at} \quad t = t_1, t_2 \quad \text{and} \quad A_\mu(x) \to 0 \quad \text{for} \quad |\mathbf{x}| \to \infty.$$

Hence

$$0 = \int_{t_1}^{t_2} dt d^3 x \left[\frac{\partial \mathcal{L}}{\partial A_\mu} \delta A_\mu + \frac{\partial \mathcal{L}}{\partial(\partial_\mu A_\nu)} \delta(\partial_\mu A_\nu) \right]$$

$$= \int_{t_1}^{t_2} dt d^3 x \left[\frac{\partial \mathcal{L}}{\partial A_\mu} \delta A_\mu + \frac{\partial \mathcal{L}}{\partial(\partial_\mu A_\nu)} \partial_\mu(\delta A_\nu) \right].$$

As in classical mechanics one varies with respect to a parameter α, whose values label the various paths in configuration space, so that

$$\delta(\partial_\mu A_\nu) = \frac{\partial}{\partial \alpha}(\partial_\mu A_\nu(x, \alpha)) \delta \alpha = \partial_\mu(\delta A_\nu).$$

After a partial integration the variation of S is seen to be

$$0 = \int_{t_1}^{t_2} dt d^3 x \left[\frac{\partial \mathcal{L}}{\partial A_\nu} - \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu A_\nu)} \right) \right] \delta A_\nu + \int dS_\mu \left[\frac{\partial \mathcal{L}}{\partial(\partial_\mu A_\nu)} \delta A_\nu \right].$$

S_μ is a 3-dimensional pseudo-plane orthogonal to x_μ. If the surface S_μ is time-like, i.e. an area orthogonal to the time-axis, the second integral vanishes, because δA_ν is zero at the integration limits $t = t_1, t_2$. But also the non-timelike surface-area contributions vanish by extending the volume of integration far enough so that all field components and their derivatives vanish on the boundary. (This is a plausible and in field theory frequently made assumption on the basis of the localisability of the fields, i.e. finiteness of the velocity of light, c). Since the components A_μ, or rather their variations, are independent of one another, we obtain the *Euler–Lagrange equations*

$$\frac{\partial \mathcal{L}}{\partial A_\mu} = \partial_\nu \left(\frac{\partial \mathcal{L}}{\partial(\partial_\nu A_\mu)} \right). \qquad (18.1)$$

These equations are *Lorentz covariant*, provided $\mathcal{L}$ is a Lorentz invariant.

18.3 Symmetries and Energy-Momentum Tensor

In classical mechanics a very generally valid theorem known as *Noether's theorem* is well known. This theorem says: Every continuous symmetry (group) of $\mathcal{L}$ (i.e. invariance of $\mathcal{L}$ under the transformations of the group) has as its consequence a conservation law. With the help of the Euler–Lagrange equations we can immediately derive one conservation law or conserved current. We have

$$\delta\mathcal{L} = \frac{\partial\mathcal{L}}{\partial A_\mu}\delta A_\mu + \frac{\partial\mathcal{L}}{\partial(\partial_\mu A_\nu)}\delta(\partial_\mu A_\nu).$$

Using the above Euler–Lagrange equations, we have

$$\delta\mathcal{L} = \partial_\nu\left(\frac{\partial\mathcal{L}}{\partial(\partial_\nu A_\mu)}\right)\delta A_\mu + \frac{\partial\mathcal{L}}{\partial(\partial_\mu A_\nu)}\delta(\partial_\mu A_\nu) = \partial_\mu\left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu A_\nu)}\delta A_\nu\right).$$

Consider now a transformation depending on some parameters, with an infinitesimal *"form transformation"* δ_P of the field, i.e. in the present case of A_μ, defined by

$$\delta_P A_\mu := A'_\mu(x) - A_\mu(x)$$

(note the argument x is the same in both quantities on the right). Then, if in the corresponding variation of $\mathcal{L}$ without the use of the equations of motion one has

$$\delta_P\mathcal{L} = \partial_\mu V^\mu, \qquad V^\mu \text{ some 4–vector},$$

one says the transformation represents a continuous symmetry of the theory. Setting $\delta A_\mu = \delta_P A_\mu$ and $\delta\mathcal{L} = \delta_P\mathcal{L}$, one has

$$\partial_\mu j^\mu = 0 \tag{18.2}$$

with

$$j^\mu = \frac{\partial\mathcal{L}}{\partial(\partial_\mu A_\nu)}\delta_P A_\nu - V^\mu. \tag{18.3}$$

Then j^μ is called a *conserved current* or *Noether current*.[*] There are cases for which $V^\mu \neq 0$ (e.g. in the case of translations, rotations) and cases for which $V^\mu = 0$ (in the case of gauge transformations). A simple example for the latter is $\mathcal{L} = \phi^*\phi$, the star meaning complex conjugation. The infinitesimal transformation $\delta\phi = i\epsilon\phi$ (ϵ real) implies $\delta_P\mathcal{L} = 0$, i.e. $V^\mu = 0$. One says, the transformation is an *"internal symmetry"*. Equation (18.2) expresses the conservation of the current j^μ. In Example 18.1 the derivation of the conservation law associated with translations is demonstrated in the case of

---

[*]This procedure, attributed to R. Jackiw, was emphasised in H. Fleming, *Noether theorems*, Univ. of Sao Paulo Report IFUSP/P-517(1985), unpublished.

the simpler theory of a scalar field $\phi(x)$. Instead of deriving the currents we are interested in with lengthy calculations from the above expression,[†] we simply define certain quantities here and then verify that they are currents. We define as the so-called *energy-momentum tensor* the quantity

$$T^{\mu\nu} := \frac{1}{c\mu_0}\left[-\eta^{\mu\nu}\mathcal{L} + \frac{\partial\mathcal{L}}{\partial(\partial_\mu A_\rho)}\partial^\nu A_\rho\right]. \tag{18.4}$$

The constant in front is introduced to ensure agreement of the field momentum with our earlier definition in Sec. 7.7. It follows that

$$\begin{aligned}
c\mu_0\partial_\mu T^{\mu\nu} &= -\partial^\nu\mathcal{L} + \partial_\mu\left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu A_\rho)}\partial^\nu A_\rho\right) \\
&= -\partial^\nu\mathcal{L} + \partial_\mu\left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu A_\rho)}\right)\partial^\nu A_\rho + \left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu A_\rho)}\right)\partial_\mu\partial^\nu A_\rho \\
&= -\partial^\nu\mathcal{L} + \frac{\partial\mathcal{L}}{\partial A_\rho}\partial^\nu A_\rho + \frac{\partial\mathcal{L}}{\partial(\partial_\mu A_\rho)}\partial^\nu(\partial_\mu A_\rho) \\
&= -\partial^\nu\mathcal{L} + \partial^\nu\mathcal{L} \\
&= 0, \tag{18.5}
\end{aligned}$$

where we used in an intermediate step the Euler–Lagrange equations. The result represents a conservation law or equation of continuity. If we integrate the equation with respect to space coordinates at a fixed time, we obtain

$$\int d^3x\,\partial_\mu T^{\mu\nu} = 0,$$

and hence

$$\partial_0\int d^3x\,T^{0\nu} + \int d^3x\,\partial_i T^{i\nu} = 0.$$

The second contribution is the volume integral of a total divergence, which with Gauss' divergence theorem can be converted into a surface integral, and this vanishes on account of a localisation of the fields in a finite part of space (i.e. the fields decrease to zero on the boundary). Thus we obtain

$$0 = \frac{d}{dt}\int d^3x\,T^{0\nu}, \quad \nu = 0, 1, 2, 3. \tag{18.6}$$

For a further evaluation of this expression in electrodynamics, we need to know the components $T^{0\nu}$, which therefore have to be determined.

[†]See e.g. Schweber [19], pp. 207–211.

We define as *4-momentum density* p^μ the 4-vector

$$p^\mu = T^{0\mu} = \frac{1}{c\mu_0}\left[-\eta^{0\mu}\mathcal{L} + \frac{\partial\mathcal{L}}{\partial(\partial_0 A_\rho)}\partial^\mu A_\rho\right],$$

and as *4-momentum P^μ* the spatial integral of the density

$$P^\mu = \int d^3x\, p^\mu = \frac{1}{c\mu_0}\int d^3x[-\eta^{0\mu}\mathcal{L} + \pi^\rho\partial^\mu A_\rho]. \tag{18.7}$$

In this expression, π^ρ is the density of the *conjugate momentum*

$$\pi^\rho = \frac{\partial\mathcal{L}}{\partial(\partial_0 A_\rho)} \tag{18.8}$$

(defined in analogy to the conjugate momentum in classical mechanics i.e. to $p = \partial L/\partial\dot{q}$). The time component of p^μ is

$$p^0 = \frac{1}{c\mu_0}[-\mathcal{L} + \pi^\rho\partial^0 A_\rho]. \tag{18.9}$$

This expression represents a *Legendre transform* and we set

$$\mathcal{H} = \mathcal{H}(A_\mu, \pi_\mu) = cp^0 = \frac{1}{\mu_0}[\pi^\rho\partial^0 A_\rho - \mathcal{L}], \tag{18.10}$$

and we shall see that this is the energy density or *Hamilton density* in agreement with our earlier expression. We observe that contrary to $\mathcal{L}$, $\mathcal{H}$ is *not a Lorentz invariant*, but the 0-component of a 4-vector! For the further evaluation of these expressions we require an explicit expression for the *Lagrangian density*. We then expect that we obtain for the momentum density the field density and for the Hamilton density the energy density of the electromagnetic field — both of which we encountered in earlier chapters.

Example 18.1: Translations as continuous symmetry of a Lagrangian
Show that the translations $x'^\mu = x^\mu + \epsilon^\mu$ (ϵ^μ a constant 4-vector) represent a continuous symmetry of the Lagrangian of a scalar field $\phi(x)$ with density

$$\mathcal{L}(\phi, \partial_\mu\phi) = \frac{1}{2}\partial^\mu\phi\partial_\mu\phi - \frac{1}{2}m^2\phi^2.$$

Solution: According to the above we have to show that

$$\delta_P\mathcal{L} = \partial_\mu V^\mu$$

and have to determine V^μ. Since $\phi(x)$ is a scalar, we have $\phi'(x') = \phi(x)$ (recall that scalars, vectors, tensors in some space are defined by the way they transform under the transformations of this space). But also for infinitesimal ϵ^μ

$$\phi'(x') = \phi'(x) + (x'^\mu - x^\mu)\partial_\mu\phi(x)$$
$$= \phi'(x) + \epsilon^\mu\partial_\mu\phi(x).$$

(It suffices to consider the infinitesimal transformation, since the finite transformation can be generated from a sequence of infinitesimal transformations). Hence the *form variation* $\delta_P \phi \equiv \phi'(x) - \phi(x)$ is with $\phi'(x') = \phi(x)$ in the case of translations given by

$$\delta_P \phi(x) = \phi'(x) - \phi(x) = -\epsilon^\mu \partial_\mu \phi(x).$$

With this we obtain

$$\delta_P \mathcal{L} = (\partial^\mu \phi \partial_\mu - m^2 \phi) \delta_P \phi = (\partial^\mu \phi \partial_\mu - m^2 \phi)(-\epsilon^\rho \partial_\rho \phi)$$

or

$$\delta_P \mathcal{L} = -\epsilon^\rho \partial_\rho \left[\frac{1}{2} \partial^\mu \phi \partial_\mu \phi - \frac{1}{2} m^2 \phi^2 \right].$$

Hence in the present case of translations

$$V^\mu = -\epsilon^\mu \mathcal{L}.$$

This was obtained without recourse to an equation of motion. Applying Eq. (18.3) in the present case, we have

$$j^\mu = \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)} \delta_P \phi - V^\mu = -\epsilon^\nu \left\{ \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)} \partial_\nu \phi - \eta^\mu_\nu \mathcal{L} \right\}$$

and so $\partial_\mu j^\mu = 0$ can be written $\partial_\mu T^{\mu\nu} = 0$, where

$$T^{\mu\nu} = \text{const.} \left\{ \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)} \partial^\nu \phi - \eta^{\mu\nu} \mathcal{L} \right\},$$

which is the corresponding energy-momentum tensor. For the transformations of the full Poincaré group the corresponding considerations are somewhat more involved.

18.4   The Lagrangian of the Electromagnetic Field

We have seen previously that two of the four Maxwell equations can be combined in the following covariant form

$$\partial_\mu F^{\mu\nu} = \mu_o J^\nu. \tag{18.11}$$

We saw that the other two Maxwell equations follow from the Jacobi identity or are trivial consequences of the antisymmetry of the field strength tensor $F^{\mu\nu}$. We now search for an expression of $\mathcal{L}$, from which the upper equation can be derived as Euler–Lagrange equation. We consider the following expression and we shall see that this yields the desired equations:

$$\mathcal{L}(A_\nu, \partial_\mu A_\nu) = -\frac{1}{4} F^{\mu\nu} F_{\mu\nu} - \mu_0 J^\nu A_\nu. \tag{18.12}$$

We see immediately that the expression is Lorentz invariant. With some thought one will realise that (apart from a possible additional term) there

is hardly any possibility to construct some other expression under the conditions it has to satisfy (only first derivatives, Lorentz invariance, gauge invariance). Consider

$$
\frac{\partial}{\partial(\partial_\nu A_\mu)}\left\{\frac{1}{4}(\partial^\rho A^\kappa - \partial^\kappa A^\rho)(\partial_\rho A_\kappa - \partial_\kappa A_\rho)\right\}
$$

$$
= \frac{1}{4}F^{\rho\kappa}(\delta^\nu_\rho\delta^\mu_\kappa - \delta^\nu_\kappa\delta^\mu_\rho) + \frac{1}{4}(\delta^{\rho\nu}\delta^{\mu\kappa} - \delta^{\nu\kappa}\delta^{\mu\rho})F_{\rho\kappa}
$$

$$
= \frac{1}{4}[F^{\nu\mu} - F^{\mu\nu} + F^{\nu\mu} - F^{\mu\nu}]
$$

$$
= -\frac{4F^{\mu\nu}}{4}
$$

$$
= +F^{\nu\mu}. \tag{18.13}
$$

Hence the *equation of motion*

$$
\frac{\partial\mathcal{L}}{\partial A_\mu} = \partial_\nu\left(\frac{\partial\mathcal{L}}{\partial(\partial_\nu A_\mu)}\right)
$$

implies the relation

$$
-\mu_0 J^\mu = -\partial_\nu(F^{\nu\mu}),
$$

i.e.

$$
\partial_\nu F^{\nu\mu} = \mu_0 J^\mu \tag{18.14}
$$

in agreement with the desired equation.

Next we calculate the *density of the canonical momentum*, i.e.

$$
\pi^\mu = \frac{\partial\mathcal{L}}{\partial(\partial_0 A_\mu)} = -F^{0\mu} = F^{\mu 0}, \quad \pi^0 = 0, \quad \pi^i = F^{i0}. \tag{18.15}
$$

According to the expression for $F^{\mu\nu}$ in Eq. (17.35),

$$
F^{i0} = \frac{E_i}{c}. \tag{18.16}
$$

The product $F^{\mu\nu}F_{\mu\nu}$ in the Lagrangian density can also be reexpressed in terms of the field strengths with the help of the matrices (17.35) and (17.36). Thus ("Tr" meaning "*trace*")

$$
-\sum_{\mu,\nu} F^{\mu\nu}F_{\mu\nu} = \mathrm{Tr}(F^{\mu\nu}F_{\nu\mu})
$$

$$
= \mathrm{Tr}\left[\begin{pmatrix} 0 & -\frac{E_1}{c} & -\frac{E_2}{c} & -\frac{E_3}{c} \\ \frac{E_1}{c} & 0 & -B_3 & B_2 \\ \frac{E_2}{c} & B_3 & 0 & -B_1 \\ \frac{E_3}{c} & -B_2 & B_1 & 0 \end{pmatrix}\begin{pmatrix} 0 & \frac{E_1}{c} & \frac{E_2}{c} & \frac{E_3}{c} \\ -\frac{E_1}{c} & 0 & -B_3 & B_2 \\ -\frac{E_2}{c} & B_3 & 0 & -B_1 \\ -\frac{E_3}{c} & -B_2 & B_1 & 0 \end{pmatrix}\right]
$$

$$= 2 \left(\frac{\mathbf{E}^2}{c^2} - \mathbf{B}^2 \right).$$

Hence the Lagrangian density is also

$$\mathcal{L} = \mathcal{L}(E_i, B_j) = \frac{1}{2} \left(\frac{\mathbf{E}^2}{c^2} - \mathbf{B}^2 \right). \tag{18.17}$$

We note the analogy with the simple Lagrangian in classical mechanics: $L = $ kinetic energy $-$ potential energy, and compare also with the energy density (see below).[‡]

We consider now the case of the *free electromagnetic field*, i.e. $J^\nu = 0$. Then

$$\pi^\rho \partial^0 A_\rho = \pi^0 \partial^0 A_0 + \pi^i \partial^0 A_i = 0 - \pi^i \partial^0 A^i = -F^{i0} \partial^0 A^i$$

$$\overset{(18.16)}{=} -\frac{1}{c} E_i \partial^0 A^i = \frac{1}{c} E_i \left(-\frac{1}{c} \frac{\partial A^i}{\partial t} \right) = \frac{1}{c^2} E_i (E_i - \nabla^i \phi),$$

where we used Eq. (17.33) in the last step. Thus in vector form

$$\pi^\rho \partial^0 A_\rho = \frac{\mathbf{E}^2}{c^2} - \frac{1}{c^2} \mathbf{E} \cdot \nabla \phi$$

$$= \frac{\mathbf{E}^2}{c^2} - \frac{1}{c^2} [\nabla \cdot (\mathbf{E}\phi) - \phi \nabla \cdot \mathbf{E}],$$

where the last term vanishes on account of $J^\nu = 0$. We obtain therefore for the energy density $\mathcal{H}$

$$\mathcal{H} = c p^0 = c T^{00} = \frac{1}{\mu_0} [\pi^\rho \partial^0 A_\rho - \mathcal{L}]$$

$$= \frac{1}{\mu_0} \left[\frac{\mathbf{E}^2}{c^2} - \frac{1}{c^2} \nabla \cdot (\mathbf{E}\phi) - \frac{1}{2} \left(\frac{\mathbf{E}^2}{c^2} - \mathbf{B}^2 \right) \right]$$

$$= \frac{1}{\mu_0} \left[\frac{1}{2} \left(\frac{\mathbf{E}^2}{c^2} + \mathbf{B}^2 \right) - \nabla \cdot \left(\frac{\mathbf{E}\phi}{c^2} \right) \right]. \tag{18.18}$$

(Compare the first contribution with $\mathcal{L}$ above). For the energy of the

---

[‡]Observe also that, with $c^2 \epsilon_0 \mu_0 = 1$, one can define $\mathbf{D}$ as

$$\mathbf{D} = \frac{1}{\mu_0} \frac{\partial \mathcal{L}}{\partial \mathbf{E}} \quad \text{and similarly} \quad \mathbf{H} = -\frac{1}{\mu_0} \frac{\partial \mathcal{L}}{\partial \mathbf{B}}.$$

electromagnetic field we obtain by integrating over all space

$$
\begin{aligned}
E_{\text{field}} = \int d^3x \mathcal{H} &= \frac{1}{2\mu_0} \int d^3x \left(\frac{\mathbf{E}^2}{c^2} + \mathbf{B}^2 \right) \\
&= \frac{1}{2\mu_0} \int d^3x (\epsilon_0 \mu_0 \mathbf{E}^2 + \mathbf{B}^2) \\
&= \frac{1}{2} \int d^3x (\mathbf{E} \cdot \mathbf{D} + \mathbf{H} \cdot \mathbf{B}), \qquad (18.19)
\end{aligned}
$$

where with the help of Gauss' divergence theorem the contribution of the divergence again vanishes for localised fields (as here assumed).[§]

In an analogous way we can calculate the *field momentum* P^i, i.e.

$$
\begin{aligned}
c\mu_0 P^i = c\mu_0 \int d^3x p^i &= \int d^3x [-\eta^{0i}\mathcal{L} + \pi^\rho \partial^i A_\rho] = \int d^3x \pi^\rho \partial^i A_\rho \\
&= \int d^3x (F^{\rho 0}) \partial^i A_\rho = \int d^3x F^{j0} \partial^i A_j \\
&= \int d^3x F^{j0} (\partial^i A_j - \partial_j A^i) + \int d^3x F^{j0} \partial_j A^i \\
&= -\int d^3x F^{j0} (\partial^i A^j - \partial^j A^i) - \int d^3x F^{j0} \partial^j A^i \\
&= -\int d^3x \frac{E_j}{c} (\partial^i A^j - \partial^j A^i) - \int d^3x \partial^j (F^{j0} A^i),
\end{aligned}
$$

where, as above, we have

$$
\partial^j (F^{j0} A^i) = (\partial^j A^i) F^{j0} + A^i \partial^j F^{j0}
$$

with $\partial^j F^{j0} = \boldsymbol{\nabla} \cdot \mathbf{E}/c = \rho/(\epsilon_0 c) = 0$. But $E_j(\partial^i A^j - \partial^j A^i)$ is element "i" of

$$
\begin{pmatrix} E_1 & E_2 & E_3 \end{pmatrix}
\begin{pmatrix}
0 & -B_3 & B_2 \\
B_3 & 0 & -B_1 \\
-B_2 & B_1 & 0
\end{pmatrix}
$$

$$
= (E_2 B_3 - E_3 B_2, \; E_3 B_1 - E_1 B_3, \; E_1 B_2 - E_2 B_1) = \mathbf{E} \times \mathbf{B}.
$$

We obtain therefore

$$
P^i = -\frac{1}{c^2 \mu_0} \int d^3x (\mathbf{E} \times \mathbf{B})^i - \frac{1}{c^2 \mu_0} \int d^3x \boldsymbol{\nabla} \cdot (\mathbf{E} A^i), \qquad (18.20)
$$

or with arguments as before

$$
\mathbf{P} \equiv (P_1, P_2, P_3) = \frac{1}{c^2 \mu_0} \int d^3x (\mathbf{E} \times \mathbf{B}) = \int d^3x (\mathbf{D} \times \mathbf{B}). \qquad (18.21)
$$

For P_0, P_i we obtain therefore the expected expressions, although those for the densities differ from the earlier expressions by divergence contributions.

[§]We can always add a divergence to the Lagrangian so that the Lagrangian density is not uniquely defined.

18.5   Gauge Invariance and Charge Conservation

We know from Noether's theorem that every continuous symmetry of the Lagrangian implies a conservation law. Thus invariance under spatial translations implies conservation of momentum and so on, as is familiar from classical mechanics. An additional aspect of electrodynamics is its invariance under gauge transformations. The associated conserved quantity is the *electric charge.* This is the aspect we investigate next.

In general one distinguishes between *global gauge invariance* – corresponding to x-independent transformations — and *local gauge invariance* — corresponding to x-dependent transformations. In the context of our considerations here only the electromagnetic potential A_μ and the current $J_\mu = (c\rho, -\mathbf{j})$ appear, not however single electrons and positrons other than classical pointlike charges. In order to take electrons and positrons properly into account we would have to construct the complete Lagrangian for charged objects, or rather fields, in interaction with the electromagnetic fields, and that would mean the Lagrangian of *quantum electrodynamics.* The corresponding Lagrangian density has the form

$$\mathcal{L}(A_\mu, \partial_\nu A_\mu, \psi, \partial_\mu \psi) = \overline{\psi}(i\gamma^\mu \partial_\mu - e'\gamma^\mu A_\mu - m)\psi - \frac{1}{4}F^{\mu\nu}F_{\mu\nu}, \qquad (18.22)$$

where ψ is the field of the electron, the so-called *Dirac field*, a 4-component column matrix called *spinor*, and $\gamma^\mu, \mu = 0, 1, 2, 3$, are corresponding 4×4 matrices called *Dirac matrices*, which obey a Clifford algebra. We do not enter into a discussion of specific properties of the electron field here, as this belongs into the realm of quantum electrodynamics. However, we want to consider the gauge invariance of the Lagrangian. To this end we observe that $\mathcal{L}$ is invariant under the following local gauge transformation

$$\psi \to \psi' = e^{i\theta(x)e'}\psi, \quad \overline{\psi} \to \overline{\psi}' = e^{-i\theta(x)e'}\overline{\psi},$$
$$A_\mu \to A'_\mu = A_\mu - \partial_\mu \theta(x). \qquad (18.23)$$

The invariance depends crucially on the fact, that $\partial_\mu \theta(x)$-contributions of the free Dirac field $(A_\mu = 0)$ are cancelled by contributions derived from the interaction with the electromagnetic field (physically this implies that it is not possible to separate the electromagnetic field from the electron). This means

$$\mathcal{L} = -\frac{1}{4}F^{\mu\nu}F_{\mu\nu} - \mu_0 J^\nu A_\nu$$

is gauge invariant only if we add to $\mathcal{L}$ the free Lagrangian of those particles that give rise to the current J^ν or if J^ν is a conserved current, $\partial_\nu J^\nu = 0$, and we ignore a divergence term which, integrated over the spatial volume,

gives zero: $\partial_\mu(J^\mu\theta) = J^\mu\partial_\mu\theta + (\partial_\mu J^\mu)\theta$. The *invariance of $F^{\mu\nu}F_{\mu\nu}$* is easily verified:

$$F^{\mu\nu} \to F^{\mu\nu\prime} = \partial^\mu A^{\nu\prime} - \partial^\nu A^{\mu\prime}$$
$$= \partial^\mu(A^\nu - \partial^\nu\theta(x)) - \partial^\nu(A^\mu - \partial^\mu\theta(x)) = F^{\mu\nu}.$$

The invariance of the Lorentz gauge condition $\partial_\mu A^\mu(x) = 0$ requires

$$\partial^\mu\partial_\mu\theta(x) = 0.$$

We assume therefore that the Lagrangian density is gauge invariant, and return to Noether's theorem. The *simplest way* to obtain the corresponding conservation law is by differentiating the equation of motion (here a kind of trick), i.e.

$$\partial_\nu(\partial_\mu F^{\mu\nu} - \mu_0 J^\nu) = 0.$$

Since (with $\partial^\mu\partial_\mu = -\Box$)

$$\partial_\nu\partial_\mu F^{\mu\nu} = \partial_\nu(\partial_\mu\partial^\mu A^\nu - \partial_\mu\partial^\nu A^\mu)$$
$$= -\Box\partial_\nu A^\nu + \Box\partial_\mu A^\mu = 0,$$

it follows that

$$\partial_\nu J^\nu = 0 \quad \text{(current conservation)}. \tag{18.24}$$

The charge q is

$$q = \int d^3x\rho = \frac{1}{c}\int d^3x J^0, \tag{18.25}$$

i.e. (recall Eq. (17.29), where $J^\mu = (c\rho, \mathbf{j})$)

$$\partial_0 q = \frac{1}{c}\int d^3x\partial_0 J^0 = -\frac{1}{c}\int d^3x\partial_j J^j = -\frac{1}{c}\int d^3x\nabla\cdot\mathbf{j}$$
$$= -\frac{1}{c}\int d\mathbf{F}\cdot\mathbf{j} = 0,$$

if we assume (as frequently explained above) that the fields are localised (having a finite velocity of propagation). This result, $\partial_0 q = 0$, is described as *charge conservation*, since $q = \text{const}$.

18.6 Lorentz Transformations and Associated Conservation Laws

The conservation laws which result from the invariance of the Lagrangian under translations and spatial rotations (i.e. conservation of energy, momentum and angular momentum) are familiar from mechanics (with the

difference that here (e.g.) the mechanical momentum has its counterpart in the field momentum, and that one has to distinguish between momentum and momentum density). A natural question is therefore now: What are the conserved quantities resulting from invariance with respect to the special Lorentz transformations? We do not enter into detailed field calculations here, and instead resort to a *plausibility argument*.

The x-component of the conserved angular momentum $\mathbf{L}$ is

$$L_x = yp_z - zp_y.$$

In the case of a special Lorentz transformation along the x-axis the correspondingly conserved quantity is evidently

$$l_x = xp_0 - x_0p_x, \qquad x_0 = ct, \quad p_0 = \mathcal{H}/c,$$

or more generally

$$\mathbf{l} = \mathbf{r}p_0 - x_0\mathbf{p}.$$

Since we have here field densities, the corresponding conservation law is

$$0 = \frac{d}{dx_0} \int d^3x[\mathbf{r}p_0 - x_0\mathbf{p}], \quad \int d^3x\mathbf{p} = \frac{d}{dx_0} \int d^3x\mathbf{r}p_0$$

with e.g. $\mathbf{p}$ the field momentum density. With this we construct

$$\frac{\int \mathbf{p}d^3x}{\int p_0d^3x} = \frac{1}{\int p_0d^3x}\frac{d}{dx_0} \int \mathbf{r}p_0d^3x = \frac{d}{dx_0}\left(\frac{\int \mathbf{r}\mathcal{H}d^3x}{\int \mathcal{H}d^3x}\right) \equiv \frac{d}{dx_0}\mathbf{R}.$$

The vector $\mathbf{R}$ is (for obvious reasons) called *"centre of mass of the field"*. Thus if we visualise the localised field like a cloud travelling in space, this vector would be that of its centre of mass; this is usually not of much interest and explains why this is rarely mentioned.

18.7   Masslessness of the Electromagnetic Field

Our next aim is to demonstrate that the electromagnetic field (and so the photon) is massless. In principle this is already evident by looking at the Lagrangian and observing that this does not contain a mass term, i.e. $m^2A_\mu A^\mu$. A term of this type would violate the gauge invariance of the Lagrangian. We can also argue as follows without reference to the Lagrangian. We consider the Maxwell equations rewritten in the form of equations of motion of the free electromagnetic field ($J_\mu = 0$), i.e.

$$\partial^\mu F_{\mu\nu} = 0, \tag{18.26}$$

or (important: we do not yet choose a gauge fixing condition!)

$$(\partial_\mu \partial^\mu)A_\nu - \partial_\nu(\partial^\mu A_\mu) = 0. \tag{18.27}$$

For the solution we make the ansatz

$$A_\mu(x) = \epsilon_\mu e^{ip_\nu x^\nu}. \tag{18.28}$$

The vector ϵ_μ is the *polarisation vector*. Inserting this into Eq. (18.27), we obtain

$$0 = [-p_\mu p^\mu \epsilon_\nu + p_\nu(p^\mu \epsilon_\mu)]e^{ip_\nu x^\nu}.$$

Since the solution A_μ is not to be zero, we have

$$(p_\mu p^\mu)\epsilon_\nu = (p^\mu \epsilon_\mu)p_\nu. \tag{18.29}$$

Either

$$(a) \quad p_\mu p^\mu \neq 0, \quad \text{or} \quad (b) \quad p_\mu p^\mu = 0. \tag{18.30}$$

In case (a) we have

$$\epsilon_\nu = \left(\frac{p^\mu \epsilon_\mu}{p^\mu p_\mu}\right)p_\nu, \quad \text{i.e.} \quad \epsilon_\nu \propto p_\nu. \tag{18.31}$$

Hence the solution is

$$A_\nu = \alpha p_\nu e^{ip_\mu x^\mu}, \quad \alpha = \text{const.} \tag{18.32}$$

Such solutions are considered to be *trivial*: They can be *"gauged away"* (see below) and are therefore not observable.

"Gauging away": The gauge transformation is

$$A_\nu \to A'_\nu = A_\nu + \partial_\nu \chi.$$

Since χ is arbitrary, we can set

$$\chi = i\alpha e^{ip_\mu x^\mu}.$$

Then

$$\partial_\nu \chi = -\alpha p_\nu e^{ip_\mu x^\mu}$$

and so, since from Eq. (18.32) $A_\nu = \alpha p_\nu e^{ip_\mu x^\mu}$,

$$A'_\nu = \alpha p_\nu e^{ip_\mu x^\mu} - \alpha p_\nu e^{ip_\mu x^\mu} = 0.$$

Thus, these solutions do not yield anything. *Hence we have case* (b), i.e. $p_\mu p^\mu = 0$ and so

$$0 = (p^\mu \epsilon_\mu)p_\nu, \quad \text{i.e.} \quad p^\mu \epsilon_\mu = 0, \tag{18.33}$$

since in general $p_\nu \neq 0$.

We conclude therefore: The solutions are

$$A_\nu = \epsilon_\nu e^{ip_\mu x^\mu} \quad \text{with} \quad p_\mu p^\mu = 0 \quad \text{and} \quad p^\mu \epsilon_\mu = 0. \qquad (18.34)$$

The condition $p_\mu p^\mu = 0$ says: The mass of the field is zero (compare with (17.78): $(T/c)^2 = \mathbf{p}^2 + m^2 c^2$). The condition is also described as *mass shell condition*, implying that the momentum is physical (in field theory intermediate states may have unphysical momenta).

We recognise in the condition $p^\mu \epsilon_\mu = 0$ our earlier condition $\partial_\mu A^\mu = 0$, which we called Lorentz gauge (fixing condition). We shall see below, that with the help of these two conditions we can eliminate the unphysical components of the vector potential.

18.8 Transversality of the Electromagnetic Field

We now demonstrate explicitly that two of the four components of the 4-potential A_μ can be gauged away, i.e. that the free electromagnetic field is transverse (to the direction of propagation). We choose again

$$\chi = i\alpha e^{ip_\mu x^\mu}$$

(α will be chosen suitably later). Then

$$\begin{aligned}
A'_\nu &= A_\nu + \partial_\nu \chi \\
&= \epsilon_\nu e^{ip_\mu x^\mu} - \alpha p_\nu e^{ip_\mu x^\mu} \\
&= (\epsilon_\nu - \alpha p_\nu) e^{ip_\mu x^\mu} \\
&\equiv \epsilon'_\nu e^{ip_\mu x^\mu}
\end{aligned} \qquad (18.35)$$

with

$$\epsilon_\nu \to \epsilon'_\nu = \epsilon_\nu - \alpha p_\nu.$$

This means, with the gauge transformation we also change the polarisation vector. We select a wave travelling in the direction of z:

$$p_\mu = (\omega/c, 0, 0, k), \qquad (\mathbf{p} \equiv (p^i), \ \mathbf{x} = (x^i)),$$

where the above mass shell condition implies

$$p_\mu p^\mu = 0 \quad \Rightarrow \quad \omega^2 - c^2 k^2 = 0, \quad \omega = ck.$$

Since $p^\mu \epsilon_\mu = 0$, we have

$$\frac{\omega}{c} \epsilon_0 - k\epsilon_3 = 0, \quad \text{so that} \quad \epsilon_0 = \epsilon_3.$$

It follows that

$$A_\nu = \epsilon_\nu e^{ip_\mu x^\mu} = (\epsilon_3, \epsilon_1, \epsilon_2, \epsilon_3)e^{i\frac{\omega}{c}(ct-z)}. \tag{18.36}$$

However, we can still perform a gauge transformation, i.e. we can replace ϵ_ν by ϵ'_ν, $\epsilon'_\nu = \epsilon_\nu - \alpha p_\nu$, i.e.

$$\epsilon'_\nu = \left(\epsilon_3 - \alpha\frac{\omega}{c}, \epsilon_1, \epsilon_2, \epsilon_3 - \alpha k\right).$$

Choosing now

$$\alpha = \frac{\epsilon_3}{k},$$

we have

$$\epsilon'_\nu = (0, \epsilon_1, \epsilon_2, 0),$$

i.e. the components ϵ_0, ϵ_3 have completely disappeared and have no physical significance. The part ϵ_0 of A_μ is described as the *scalar part* ("scalar photon"), the other ϵ_3 part as the *longitudinal part*. We see that only the two transverse parts $\propto \epsilon_1, \epsilon_2$ have physical significance. This result is a consequence of the mass shell condition and the gauge fixing condition.

18.9 The Spin of the Photon

We now have a closer look at the polarisation vector. We put

$$(\epsilon_0, \epsilon_1, \epsilon_2, \epsilon_3) = \tilde{\epsilon}_0(1,0,0,0) + \tilde{\epsilon}_3(0,0,0,1) + \underbrace{\tilde{\epsilon}_+(0,1,-i,0)}_{(\alpha)} + \tilde{\epsilon}_-(0,1,i,0),$$

$$\tag{18.37}$$

where

$$\tilde{\epsilon}_\pm = \frac{1}{2}(\tilde{\epsilon}_1 \pm i\tilde{\epsilon}_2).$$

(Since $\mathbf{E} = -\partial\mathbf{A}/\partial t - \boldsymbol{\nabla}\phi$, i.e. $\mathbf{E} \parallel -\mathbf{A}$, the waves $\epsilon_\pm \exp[i(ct-z)/c]$ are circularly polarised). Actually $\tilde{\epsilon}_\mu \equiv \epsilon_\mu$, but we want to distinguish between the two more clearly. We want to examine the solution of a wave travelling in the direction of z:

$$A_\mu = \epsilon_\mu e^{i\frac{\omega}{c}(ct-z)} \quad \text{or} \quad A^*_\mu = \epsilon^*_\mu e^{-i\frac{\omega}{c}(ct-z)}.$$

Naturally this can be decomposed exactly like ϵ_μ. Consider an ordinary rotation about the z-axis, i.e. a special case of a homogeneous Lorentz transformation,

$$x'^\mu = l^\mu{}_\nu x^\nu$$

with rotation operator R_θ for rotation θ about the z-axis:

$$R_\theta: \quad l^\mu{}_\nu = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\theta & \sin\theta & 0 \\ 0 & -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Here θ is the angle of the rotation in the (x, y)-plane. Then

$$\begin{aligned} A'_\nu = l_\nu{}^\mu A_\mu &= A_\mu l^{T\mu}{}_\nu \\ &= \epsilon_\mu \exp(ip_\rho x^\rho) l^{T\mu}{}_\nu \\ &= \epsilon'_\nu \exp(ip_\rho x^\rho), \end{aligned} \qquad (18.38)$$

where

$$\epsilon'_\nu = \epsilon_\mu l^{T\mu}{}_\nu.$$

Thus the rotation affects only the polarisation vector:

$$\tilde{\epsilon}_0(1,0,0,0) \rightarrow \tilde{\epsilon}'_0 = \tilde{\epsilon}_\mu l^{T\mu}{}_0$$

$$= \tilde{\epsilon}_0 \left((1,0,0,0) \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta & 0 \\ 0 & \sin\theta & \cos\theta & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \right)_0 = \tilde{\epsilon}_0(1,0,0,0).$$

Similarly

$$\tilde{\epsilon}_3(0,0,0,1) \rightarrow \tilde{\epsilon}_3(0,0,0,1).$$

However

$$\tilde{\epsilon}_\pm(0,1,\mp i,0) \rightarrow \tilde{\epsilon}_\pm(0,1,\mp i,0) \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta & 0 \\ 0 & \sin\theta & \cos\theta & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

$$= \tilde{\epsilon}_\pm(0, \cos\theta \mp i\sin\theta, -\sin\theta \mp i\cos\theta, 0)$$

$$= \tilde{\epsilon}_\pm e^{\mp i\theta}(0,1,\mp i,0).$$

If we consider R_θ as rotation operator, we have the "*eigenvalue equation*":

$$R_\theta A_\mu = \text{eigenvalue} \times A_\mu.$$

The rotation operator is given by

$$R_\theta = e^{i\theta I_z}$$

where I_z is the z-component of the (eigen)-angular momentum operator ("spin operator"). Then in a rotation about the z-axis

$$R_\theta A_0 = 1 \times A_0$$
$$\equiv e^{i\theta I_z} A_0 \text{ with eigenvalue of } I_z = 0,$$
$$R_\theta A_3 = 1 \times A_3$$
$$\equiv e^{i\theta I_z} A_3 \text{ with eigenvalue of } I_z = 0,$$
$$R_\theta A_\pm = e^{\mp i\theta} A_\pm$$
$$\equiv e^{i\theta I_z} A_\pm \text{ with eigenvalue of } I_z = \mp 1. \qquad (18.39)$$

We go through the last step in detail. We have for instance

$$
\begin{aligned}
R_\theta A_+ &= R_\theta \epsilon_+ \exp\left[i\frac{\omega}{c}(ct - z) \right] \\
&\overset{(18.37)}{=} \tilde{\epsilon}_+ \left. \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\theta & \sin\theta & 0 \\ 0 & -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \\ -i \\ 0 \end{pmatrix} \right|_{(\alpha)} \exp\left[i\frac{\omega}{c}(ct - z) \right] \\
&= \tilde{\epsilon}_+ \begin{pmatrix} 0 \\ e^{-i\theta} \\ -ie^{-i\theta} \\ 0 \end{pmatrix} \exp\left[i\frac{\omega}{c}(ct - z) \right]
\end{aligned}
$$

and hence

$$R_\theta A_+ = \tilde{\epsilon}_+ \begin{pmatrix} 0 \\ 1 \\ -i \\ 0 \end{pmatrix} e^{-i\theta} \exp\left[i\frac{\omega}{c}(ct - z) \right] = e^{-i\theta} A_+,$$

as in the last line of Eq. (18.39). It follows therefore that for a scalar or longitudinal field or photon the spin projection on the z-axis (i.e. the direction of propagation) is zero, whereas the field of the transverse photon has the spin projections $S_z = \pm 1\hbar$. Since only the transverse components are physical, i.e. observable, it is evident that the photon has spin $1(\hbar), s = 1\hbar$ ($\mathbf{S}^2$ has eigenvalue $s(s+1), s = maximal\ s_z, S_z \equiv I_z$). The photon never has spin projection 0 in the direction of the momentum vector, i.e. in the direction of propagation. One says therefore, the photon has *helicity* ± 1.

18.10   Examples

The following examples are motivated by the Lagrangian of electrodynamics, but they also lead somewhat beyond the considerations of primary interest here and, maybe, illustrate the fascinating use of electrodynamics for the exploration of a plethora of other issues.

Example 18.2: The Chern–Simons term and Rydberg atoms

Verify the gauge invariance of the following term, called Chern–Simons term, which one could add to a Maxwell Lagrangian density:

$$T_{CS} := \epsilon^{\lambda\mu\rho} A_\lambda \partial_\mu A_\rho$$

(up to a divergence). Explore the simple choice $A_i = x_i$ with $i = 1, 2$ and $A_0 = 0$.

Solution: With $A_\mu \to A_\mu + \partial_\mu\chi$ and using the antisymmetry of the Levi–Civita tensor, the term T_{CS} becomes

$$T'_{CS} = \epsilon^{\lambda\mu\rho}(A_\lambda + \partial_\lambda\chi)\partial_\mu(A_\rho + \partial_\rho\chi) = T_{CS} + \triangle,$$

where

$$
\begin{aligned}
\triangle &= \epsilon^{\lambda\mu\rho} A_\lambda \partial_\mu \partial_\rho\chi + \epsilon^{\lambda\mu\rho} \partial_\lambda\chi \partial_\mu(A_\rho + \partial_\rho\chi) \\
&= 0 + \epsilon^{\lambda\mu\rho} \partial_\lambda\chi \partial_\mu A_\rho + \underbrace{\epsilon^{\lambda\mu\rho} \partial_\lambda\chi \partial_\mu \partial_\rho\chi}_{0} \\
&= \epsilon^{\lambda\mu\rho}(\partial_\lambda(\chi\partial_\mu A_\rho) - \chi\partial_\lambda\partial_\mu A_\rho) \\
&= \epsilon^{\lambda\mu\rho}\partial_\lambda(\chi\partial_\mu A_\rho) = \partial_\lambda(\chi\epsilon^{\lambda\mu\rho}\partial_\mu A_\rho) \\
&= \partial_\lambda(\chi V^\lambda), \quad V^\lambda = \epsilon^{\lambda\mu\rho}\partial_\mu A_\rho,
\end{aligned}
$$

which had to be shown.

In the special case of $\lambda, \mu, \rho = 0, 1, 2$ with gauge choice $A_0 = 0$ and $i, j = 1, 2$ and m the mass of the atom, we have

$$T_{CS} = -\epsilon_{ij} A_i \partial_0 A_j.$$

Selecting the simple case of $A_i = \alpha x_i$, $\alpha = $ const., so that $\partial_0 A_j = \alpha\dot{x}_j/c = \alpha p_j/mc$, which relates the electric field to momentum, we have

$$T_{CS} = \frac{\alpha^2}{mc}\epsilon_{ij} p_i x_j.$$

This type of term appears in studies of atomic physics where the Hamiltonian of a Rydberg atom of mass m in the presence of electric and magnetic fields (the latter along x_3) is written as

$$
\begin{aligned}
H &= \frac{1}{2m}\left(p_i + \frac{1}{2}g\epsilon_{ij}x_j\right)^2 + \frac{1}{2}\kappa x_i^2 \\
&= \frac{1}{2m}p_i^2 + \frac{1}{2m}g\epsilon_{ij}p_i x_j + \frac{1}{2}\left(\kappa + \frac{g^2}{4m}\right)x_i^2,
\end{aligned}
$$

where g and κ are constants. The central term in the last expression is seen to be of Chern–Simons type. (Observe the difference between the Hamiltonian here and that of Eq. (19.7) to be discussed later. The minimal coupling to A_i discussed there (cf. the end of Sec. 19.2.1) corresponds here effectively to the coupling to the dual of A_i). The so-called *Rydberg atom* is an atom regarded as one with a permanent dipole moment resulting from a single electron in a shell far away from the filled electron shells enclosing the nucleus, so that the structure of the latter can be ignored.

Example 18.3: 6-dimensional gauge field theory

In a $(1+5)$-dimensional gauge field theory the gauge field is the tensor field $A_{\mu\nu} = -A_{\nu\mu}, \mu, \nu = 0, 1, 2, \ldots, 5$, and the field tensor is

$$F_{\mu\nu\lambda} = \partial_\mu A_{\nu\lambda} + \partial_\nu A_{\lambda\mu} + \partial_\lambda A_{\mu\nu}.$$

(a) Verify the invariance of the field strength $F_{\mu\nu\lambda}$ with respect to gauge transformations

$$\delta A_{\mu\nu} = A'_{\mu\nu} - A_{\mu\nu} = \partial_\mu \chi_\nu - \partial_\nu \chi_\mu.$$

(b) What are the Maxwell equations derived from the action integral

$$S = -\frac{1}{3!} \int d^6 x F_{\mu\nu\lambda} F^{\mu\nu\lambda}.$$

Solution: (a) Trivial. (b) Set

$$E_{ij} := F_{0ij}, \quad B_{ij} := \frac{1}{3!} \epsilon_{ijklm} F^{klm}.$$

The equations of motion are

$$(a) \ \partial^\mu F_{\mu\nu\lambda} = 0, \quad (b) \ \epsilon^{\mu\nu\lambda\rho\sigma\delta} \partial_\lambda F_{\rho\sigma\delta} = 0.$$

Consider (a):

$$\partial^\mu F_{\mu i 0} = 0, \quad \partial^0 F_{0i0} + \partial^j F_{ji0} = 0, \quad F_{00i} = \partial_0 A_{0i} + \partial_i A_{00} + \partial_0 A_{i0} = 0,$$

so that

$$\partial_i E_{ij} = 0 \quad \text{and} \quad \partial^0 F_{0jk} + \partial^i F_{ijk} = 0,$$

i.e.

$$\partial_0 E_{jk} = -\partial^i F_{ijk} \overset{\text{see below}}{=} -\frac{1}{2} \partial^i \epsilon_{ijklm} B^{lm}.$$

Verification:

$$\partial^i \epsilon_{ijklm} B^{lm} = \partial^i \epsilon_{ijklm} \frac{1}{3!} \epsilon^{lmabc} F_{abc} = \partial^i \frac{1}{3!} \left(\underbrace{\epsilon_{ijklm} \epsilon^{lmabc}}_{2! \text{Sgn} \begin{bmatrix} i & j & k \\ a & b & c \end{bmatrix}} \right) F_{abc}$$

$$= \partial^i \frac{1}{3!} 2! 3! F_{ijk} = 2! \partial^i F_{ijk}.$$

The other equations follow similarly from (b). (For the ϵ contractions and related topics, see e.g. Felsager [20], p. 354).

Example 18.4: Born–Infeld theory

The static, nonlinear BI Lagrangian, here simplified to that involving only the electric field $\mathbf{E} = -\nabla\phi$ in a flat space, and with $c = 1$, hence $\epsilon_0 \mu_0 = 1$, is given by

$$L_{BI} = \int d\mathbf{r} \mathcal{L}_{BI}, \quad \mathcal{L}_{BI}(\phi, \nabla\phi) = 1 - \sqrt{1 - (\nabla\phi)^2} - 4\pi e \phi \delta(\mathbf{r}).$$

Show that the energy of the charge e at the origin is given by

$$H_{BI} = 4\pi\epsilon_0 (3.09112) e^{3/2},$$

i.e. is finite, and not infinite like the energy of a point-charge, i.e.

$$\frac{1}{2} \int d\mathbf{r} (\nabla \phi)^2,$$

in Maxwell electrodynamics.

Solution: In an attempt to avoid the infinite self-energy of a point charge in Maxwell theory, Born and Infeld[¶] arrived at the above Lagrangian which is now named after them and plays an important role in String Theory. In fact, $\mathcal{L}_{BI}|_{e=0} = \mathcal{L}_{b=1}$, where

$$\mathcal{L}_b = b^2 \left[1 - \sqrt{1 - \frac{2}{b^2} \mathcal{L}_M} \right] \overset{b^2 \to \infty}{=} \mathcal{L}_M,$$

and $\mathcal{L}_M$ is the Maxwell Lagrangian density, in the present case for $\mathbf{B} = 0$ and $c = 1$ simply $\mathcal{L}_M = \mathbf{E}^2/2$.

The Euler–Lagrange equation (18.1) gives the equation corresponding to the Gauss equation:

$$\partial_i \left(\frac{\partial_i \phi}{\sqrt{1 - (\partial_j \phi)^2}} \right) = -4\pi e \delta(\mathbf{r}).$$

Integrating over a sphere of radius r and using the Gauss divergence theorem, this becomes

$$\int \frac{\nabla \phi}{\sqrt{1 - (\phi'(r))^2}} \cdot d\mathbf{F} = -4\pi e, \quad \text{or} \quad \frac{\phi'(r)}{\sqrt{1 - (\phi'(r))^2}} 4\pi r^2 = -4\pi e,$$

i.e.

$$\frac{\phi'(r)}{\sqrt{1 - (\phi'(r))^2}} = -\frac{e}{r^2}.$$

The solution is then seen to be

$$\phi_{BI}(r) = \int_r^\infty \frac{dx}{\sqrt{1 + \frac{x^4}{e^2}}}.$$

With the help of Tables of Integrals one finds that $\phi_{BI}(0) = 1.85407 e^{1/2}$ (this was the intention of Born and Infeld, to obtain this finite in a covariant theory).

Defining

$$\mathbf{D} = \frac{1}{\mu_0} \frac{\partial \mathcal{L}_{BI}}{\partial \mathbf{E}} = \frac{1}{\mu_0} \frac{\mathbf{E}}{\sqrt{1 - \mathbf{E}^2}},$$

we have (cf. Eq.(18.4), here with $c = 1$)

$$\mu_0 T^{00} = F^{0i} \frac{\partial \mathcal{L}_{BI}}{\partial F^{0i}} - \mathcal{L}_{BI} = \mu_0 \mathbf{E} \cdot \mathbf{D} - \mathcal{L}_{BI}$$

$$= \frac{1}{\sqrt{1 - \mathbf{E}^2}} - 1 + 4\pi e \phi \delta(\mathbf{r}).$$

The energy of the charge is then found to be (after some nontrivial integration)

$$H_{BI} = \int d\mathbf{r} T^{00} = 4\pi \epsilon_0 (3.09112) e^{3/2}.$$

¶M. Born and L. Infeld, *Proc. Roy. Soc.* **A144** (1934) 425; see also Jackson [3], p. 10 and G. W. Gibbons and D. A. Rasheed, *Nucl. Phys.* **B454** (1995) 185.

Example 18.5: Non-gauge invariant solutions of Maxwell's equations[||]

A type of nonlinear, non-gauge invariant solutions $G_{\mu\nu}$ of Maxwell's equations has been found and investigated by Schiff. These are given by (cf. Eq. (17.31))

$$G_{\mu\nu} := F_{\mu\nu} + \sqrt{2}g^{-1}A_\mu A_\nu, \quad (g = \text{const.}).$$

Show that $G_{\mu\nu}$ satisfies everywhere the conditions

$$^*G_{\mu\nu}G^{\mu\nu} = 0, \quad G_{\mu\nu}G^{\mu\nu} = 0,$$

which define it as a *"null field"*. Show that static, radially symmetric solutions of the homogeneous part of the non-gauge invariant equation $\partial_\mu G^{\mu\nu} = \mu_0 j^\nu$ are given by

$$A_r = \frac{k_0}{r}, \quad k_0 = \text{const.},$$

and

$$\phi_1 = \frac{b}{r} \quad \text{and} \quad \phi_2 = \frac{c_2}{r^\gamma}, \quad \text{where} \quad \gamma = \sqrt{2}g^{-1}k_0, \quad b, c_2 = \text{const.}$$

What is the total charge $Q = \int dr\rho$? Finally solve the null condition with $\mathbf{B} = 0$ for the Coulombic solution ϕ_1 and determine the charges b.

Solution: The original starting argument of Schiff is simple. Since for a point charge q at rest the Coulomb potential ϕ and electric field strength E are given by

$$\phi = k\frac{q}{r}, \quad E = k\frac{q}{r^2},$$

one has $E^2 = \phi^4/k^2q^2$, or (cf. Eqs. (17.31), (17.35))

$$F_{i0}{}^2 = \frac{E^2}{c^2} = \frac{A_0{}^4}{g^2}, \quad A_0 = \frac{\phi}{c}, \quad g^2 = \frac{k^2q^2}{c^2}.$$

The covariant generalisation is seen to be

$$F_{\mu\nu}F^{\mu\nu} = -\frac{2}{g^2}(A^\mu A_\mu)^2, \quad A^\mu A_\mu = \frac{\phi^2}{c^2} - \mathbf{A}^2,$$

as one can verify by reduction to the case above. For $\mathbf{B} = 0$ (i.e. $F_{ij} = 0$ in the quantity on the left) this null field condition is

$$\frac{\mathbf{E}^2}{c^2} = \frac{1}{g^2}\left(\mathbf{A}^2 - \frac{\phi^2}{c^2}\right)^2.$$

Writing the relation as

$$F_{\mu\nu}F^{\mu\nu} + \frac{2}{g^2}(A^\mu A_\mu)^2 = 0,$$

we see that — due to opposite symmetries of $F_{\mu\nu}$ and $A_\mu A_\nu$ — we can write it as $G_{\mu\nu}G^{\mu\nu} = 0$ with $G_{\mu\nu}$ as defined above. Multiplying $G_{\mu\nu}$ by the four-dimensional Levi–Civita tensor, we see that its dual satisfies $^*G_{\mu\nu} = {}^*F_{\mu\nu}$, so that with $^*F_{\mu\nu}F^{\mu\nu} = 0$ ($\mathbf{E} \perp \mathbf{B}$), we obtain $^*G^{\mu\nu}G_{\mu\nu} = 0$. Thus $G_{\mu\nu}$ satisfies the conditions of a null field.

The static, radially symmetric part of the homogeneous equation $\partial_\mu G^{\mu\nu} = 0$ with $A = A_r(r)$ and so $\mathbf{B} = 0$ is

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

[||]H. Schiff, *Can. J. Phys.* **47** (1969) 2387 and H. Schiff, *Quark-like potentials in an extended Maxwell theory*, hep-th/0308091.

which is solved by $A = k_0/r$, since

$$\nabla \cdot (A\mathbf{A}) = \nabla A \cdot \mathbf{A} + A\nabla \cdot \mathbf{A} = -\frac{k_0\mathbf{r}}{r^3} \cdot \frac{k_0\mathbf{r}}{r^2} + \frac{k_0}{r}\frac{1}{r^2}\frac{\partial}{\partial r}(r^2\frac{k_0}{r}) = 0.$$

The zero component of $\partial_\mu G^{\mu\nu} = \mu_0 j^\nu$ is (with $G_{0i} = F_{0i} - \sqrt{2}(\mathbf{A})_i A_0/g$) the Gauss law

$$\nabla \cdot \left(\mathbf{E} - \frac{\sqrt{2}}{g}\mathbf{A}\phi\right) = \mu_0 c^2\rho(r), \quad \text{or} \quad -\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial\phi}{\partial r}\right) - \frac{\sqrt{2}k_0}{g}\nabla \cdot \left(\frac{\mathbf{r}}{r^2}\phi\right) = \mu_0 c^2\rho(r),$$

which with

$$\nabla \cdot \left(\frac{\mathbf{r}}{r^2}\phi\right) = \frac{\phi}{r^2} + \frac{1}{r}\frac{\partial\phi}{\partial r}$$

can be written (as shown by Schiff)

$$-\frac{1}{r}\frac{d^2(r\phi)}{dr^2} - \gamma\frac{1}{r^2}\frac{d(r\phi)}{dr} = \mu_0 c^2\rho(r),$$

where $\gamma = \sqrt{2}k_0/g$. Irrespective of the value of γ the homogeneous part has solutions for $r\phi = $ const., i.e. $\phi_1 = b/r$, but for $\gamma \neq 1$ also solutions $\phi_2 = c_2 r^{-\gamma}$, which are *"confinement potentials"* for $\gamma \leq -1$.

Integrating the Gauss law, one obtains in the usual way the total charge

$$Q = \int d\mathbf{r}\rho(r) = \epsilon_0 \int d\mathbf{F} \cdot \left(\mathbf{E}_1 - \frac{\sqrt{2}}{g}\mathbf{A}\phi_1\right)$$

(the index referring to solution ϕ_1), i.e.

$$Q = \epsilon_0 \int r^2 d\Omega\frac{1}{r^2}\left(b - \frac{\sqrt{2}}{g}\frac{\gamma g}{\sqrt{2}}b\right) = \epsilon_0 4\pi b(1 - \gamma).$$

Inserting the Coulombic solution ϕ_1 into the null field condition, i.e.

$$\frac{E}{c} = \pm\frac{1}{g}\left(A^2 - \frac{\phi^2}{c^2}\right),$$

one obtains

$$\frac{b^2}{c^2} \pm g\frac{b}{c} - \frac{\gamma^2 g^2}{2} = 0$$

with solutions

$$\frac{b}{c} = \pm\frac{g}{2}[1 \pm \sqrt{1 + 2\gamma^2}].$$

This determines the charge b in the potential ϕ_1. In the paper of Schiff (cited above) these conditions are further extended to the confinement potentials ϕ_2, and it is shown that fractional charges such as those attributed to quarks can be obtained. This is an interesting observation, although, of course, this is not the acceptable overall theory.

Example 18.6: Quantum Hall effect

Starting from the Lagrangian density of the electromagnetic field in a dielectric with polarisation energy $U(\mathbf{E}, \mathbf{B})$,

$$\mathcal{L}(\mathbf{E}, \mathbf{B}) = \frac{1}{2}\left(\frac{\mathbf{E}^2}{c^2} - \mathbf{B}^2\right) - U(\mathbf{E}, \mathbf{B}), \quad \text{with} \quad U := -g(\mathbf{E} \cdot \mathbf{B}),$$

show that if the induced surface charge is ne, i.e. an integral multiple n of the elementary charge e, and the magnetic flux is $\Phi = m\Phi_0$, i.e. an integral multiple of the magnetic flux quantum $\Phi_0 = h/e$ (in MKSA units), that then the surface *Hall conductivity*,[**] σ_H, defined by

$$\mathbf{j}_M = \sigma_H \mathbf{E},$$

is given by

$$\frac{e}{\Phi_0}\frac{n}{m}(\text{length})^{-1} = \frac{e^2}{h}\frac{n}{m}(\text{length})^{-1}.$$

The Hall conductivity plays an important role in the *quantum Hall effect*[††] (this Example does not claim to explain this, instead is only an exercise in considerations around it).

Solution: We had the equations (cf. Eqs. (4.8), (6.15))

$$\mathbf{D} = \epsilon_0 \mathbf{E} + \mathbf{P}, \quad \mathbf{H} = \frac{\mathbf{B}}{\mu_0} - \mathbf{M}.$$

$$\therefore \mathbf{D} = \frac{1}{\mu_0}\frac{\partial \mathcal{L}}{\partial \mathbf{E}} \quad \text{implies} \quad \mathbf{P} = -\frac{1}{\mu_0}\frac{\partial U}{\partial \mathbf{E}},$$

and

$$\mathbf{H} = -\frac{1}{\mu_0}\frac{\partial \mathcal{L}}{\partial \mathbf{B}} \quad \text{implies} \quad \mathbf{M} = -\frac{1}{\mu_0}\frac{\partial U}{\partial \mathbf{B}}.$$

For the given expression of U these imply

$$\mathbf{P} = \frac{g}{\mu_0}\mathbf{B}, \quad \mathbf{M} = \frac{g}{\mu_0}\mathbf{E}.$$

Since we had for induced charge and current densities

$$\rho_P = -\boldsymbol{\nabla} \cdot \mathbf{P}, \quad \mathbf{j}_M = \boldsymbol{\nabla} \times \mathbf{M},$$

we have $\rho_P = 0$ (since $\boldsymbol{\nabla} \cdot \mathbf{B} = 0$), but (cf. Eq. (4.13)) the surface charge density and the magnetic current density resulting from polarisation are

$$\sigma_P = \mathbf{P} \cdot \mathbf{n} = \frac{g}{\mu_0}\mathbf{B} \cdot \mathbf{n}, \quad \mathbf{j}_M = \boldsymbol{\nabla} \times \mathbf{M} = \frac{g}{\mu_0}\boldsymbol{\nabla} \times \mathbf{E},$$

where $\mathbf{n}$ is a unit vector pointing out of the surface or interface area $\mathbf{F}$ with, say, $F = L^2$ and circumference L_0. Then, using the theorem of Stokes,

$$j_M F = \frac{g}{\mu_0}\int_F \boldsymbol{\nabla} \times \mathbf{E} \cdot d\mathbf{F} = \frac{g}{\mu_0}\int_0^L \mathbf{E} \cdot d\mathbf{l} = \frac{g}{\mu_0}EL_0.$$

Hence

$$j_M = \frac{g}{\mu_0}\frac{EL_0}{F} \equiv \sigma_H E, \quad \therefore \sigma_H = \frac{g}{\mu_0}\frac{L_0}{F}.$$

[**]For a recent discussion with references to earlier literature see J. E. Avron, D. Osadchy and R. Seiler, "A Topological Look at the Quantum Hall Effect", *Physics Today*, August 2003, p. 38. The standard definition of the magnetic flux quantum in SI units is $\Phi_0 = h/2e$ (as in Appendix B). In c.g.s. units it is $hc/2e$. This is for superconductors where the charge carriers are Cooper pairs of charge $2e$ (ignoring the sign of the charge). The above article considers single electrons, thus the elementary quantum of magnetic flux there is defined as $\Phi_0 = hc/e$ in c.g.s. units. The author is indebted to Dr. B. Taylor, National Institute of Standards and Technology, USA, for clarifying correspondence.

[††]See K. von Klitzing, G. Dorda and M. Pepper, *Phys. Rev. Lett.* **45** (1980) 494.

The quantity σ_H is the interface *Hall conductivity*. Considering the surface charge density σ_P, we have

$$\text{surface charge} \; \equiv Q = \int \sigma_P dF = \frac{g}{\mu_0} \int \mathbf{B} \cdot d\mathbf{F} = \frac{g}{\mu_0}\Phi,$$

where Φ is the magnetic flux through the surface. Setting $Q = ne$ and $\Phi = m\Phi_0$, we have

$$g = \frac{\mu_0 e}{\Phi_0}\frac{n}{m} \quad \text{and so} \quad \sigma_H = \frac{e}{\Phi_0}\frac{n}{m}\frac{L_0}{F}.$$

In strong magnetic fields $\mathbf{B}$ the Hall conductivity is actually observed to be quantised.

Chapter 19

The Gauge Covariant Schrödinger Equation and the Aharonov–Bohm Effect

19.1 Introductory Remarks

In this chapter we first establish the Schrödinger equation for a charged particle moving in an unspecified electromagnetic field. We demonstrate the covariance of this equation under gauge transformations which ensures that the probability density is independent of the choice of a particular vector potential. We then consider the special case of a homogeneous solenoidal magnetic field restricted to a small domain behind a diaphragm with a double slit and with an electron source on the other side. It is then explained why the resulting interference pattern observed on a screen (on the solenoid side of the slits and some distance away from it) is displaced when the current in the solenoid is switched on or off. This Aharonov–Bohm effect is a quantum mechanical effect resulting from the phase of the Schrödinger wave function, and may be looked at as experimental evidence of the vector potential. The considerations here supplement our considerations of solenoidal fields in Chapters 5 and 8, where, in particular, we investigated the electromagnetic vector potential $\mathbf{A}$ in the neighbourhood of a long solenoid. The Aharonov–Bohm effect is clearly a fundamentally significant phenomenon.

19.2   Schrödinger Equation of a Charged Particle in an Electromagnetic Field

19.2.1   Hamiltonian of a Charge in an Electromagnetic Field

We return once again to the nonrelativistic equation of motion of a particle with mass m and charge e in an electromagnetic field — this time, however, with a view to the formulation of the corresponding Schrödinger equation. Previously we wrote the former equation as a Newton equation with the Lorentz force as external force, i.e.

$$\frac{d}{dt}(m\dot{\mathbf{r}}) = e(\mathbf{E} + \dot{\mathbf{r}} \times \mathbf{B}), \quad |\dot{\mathbf{r}}| \equiv |\mathbf{v}| \ll c. \tag{19.1}$$

This equation can also be derived as in ordinary classical mechanics from the following Lagrangian (with $\mathbf{r} \equiv (q_1, q_2, q_3)$):

$$L(q, \dot{q}) = \frac{1}{2}m\dot{q}_i^2 + e(\dot{q}_i A_i(q, t) - \phi(q)), \tag{19.2}$$

as we verify now. With the Lagrangian we can then (via a Legendre transform) construct the Hamiltonian as a prerequisite for the transition to the quantum mechanical Schrödinger equation. From L we obtain

$$\frac{\partial L}{\partial \dot{q}_i} = m\dot{q}_i + eA_i, \quad \frac{\partial L}{\partial q_j} = e\dot{q}_i\frac{\partial A_i}{\partial q_j} - e\frac{\partial \phi}{\partial q_j}, \tag{19.3}$$

and the Euler–Lagrange equation of motion

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_j}\right) - \frac{\partial L}{\partial q_j} = 0$$

yields

$$m\ddot{q}_j + \frac{d}{dt}(eA_j) = e\dot{q}_i\frac{\partial A_i}{\partial q_j} - e\frac{\partial \phi}{\partial q_j}, \tag{19.4}$$

where d/dt is the total time derivative, i.e.

$$\frac{d}{dt} = \frac{\partial}{\partial t} + \dot{q}_i\frac{\partial}{\partial q_i},$$

so that

$$m\ddot{q}_j + e\frac{\partial A_j}{\partial t} + e\frac{\partial A_j}{\partial q_i}\dot{q}_i = e\dot{q}_i\frac{\partial A_i}{\partial q_j} - e\frac{\partial \phi}{\partial q_j}$$

and hence

$$m\ddot{q}_j = e\left(-\frac{\partial A_j}{\partial t} - \frac{\partial \phi}{\partial q_j}\right) + e\left(\frac{\partial A_i}{\partial q_j}\dot{q}_i - \frac{\partial A_j}{\partial q_i}\dot{q}_i\right).$$

However, we know that

$$\mathbf{E} = -\frac{\partial \mathbf{A}}{\partial t} - \nabla \phi \quad \text{and} \quad \mathbf{B} = \nabla \times \mathbf{A},$$

and thus

$$\mathbf{v} \times \mathbf{B} = \mathbf{v} \times (\nabla \times \mathbf{A}) = \nabla(\mathbf{v} \cdot \mathbf{A}) - (\mathbf{v} \cdot \nabla)\mathbf{A}$$

$$= \frac{\partial}{\partial \mathbf{r}}(\mathbf{v} \cdot \mathbf{A}) - \left(\mathbf{v} \cdot \frac{\partial}{\partial \mathbf{r}}\right)\mathbf{A},$$

i.e.

$$(\mathbf{v} \times \mathbf{B})_j = \frac{\partial}{\partial x_j}(v_i A_i) - v_i \frac{\partial}{\partial x_i} A_j,$$

where

$$\frac{\partial}{\partial x_j}(v_i A_i) = v_i \frac{\partial A_i}{\partial x_j} + A_i \frac{\partial v_i}{\partial x_j} = v_i \frac{\partial A_i}{\partial x_j}$$

and hence

$$m\ddot{q}_j = eE_j + e(\mathbf{v} \times \mathbf{B})_j,$$

as had to be shown, i.e. that the Euler–Lagrange equation is identical with Newton's equation.

One should note: The *conjugate momentum* is $p_i = \partial L/\partial \dot{q}_i$; this is also the *canonical momentum* in the Hamilton formalism. We have (cf. Eq. (19.3))

$$\mathbf{p} = m\mathbf{v} + e\mathbf{A}. \tag{19.5}$$

However, the so-called *mechanical momentum* is $m\mathbf{v}$ or

$$m\mathbf{v} = \mathbf{p} - e\mathbf{A} \equiv m\dot{\mathbf{q}}. \tag{19.6}$$

This distinction is very important.

We obtain the *Hamiltonian $H(q,p)$* as in mechanics with a Legendre transform, i.e. by defining

$$H(q,p) := p_i \dot{q}_i - L(q, \dot{q})$$

$$= p_i \dot{q}_i - \frac{1}{2}m\dot{q}_i^2 - e(\dot{q}_i A_i - \phi)$$

$$= \frac{p_i}{m}(p_i - eA_i) - \frac{1}{2}\frac{m}{m^2}(p_i - eA_i)^2 + e\phi - \frac{eA_i}{m}(p_i - eA_i),$$

or

$$H(q,p) = \frac{1}{2m}(p_i - eA_i)^2 + e\phi = \frac{1}{2m}(\mathbf{p} - e\mathbf{A})^2 + e\phi$$

$$= \left[\frac{1}{2m}(m\mathbf{v})^2 + e\phi\right]_{m\mathbf{v} \to \mathbf{p} - e\mathbf{A}} \tag{19.7}$$

The contribution $-e\mathbf{A}$ is seen to be the effect of the Lorentz force expressed in terms of the vector potential. The substitution

$$\mathbf{p}_{\mathrm{mech}} = m\mathbf{v} \to \mathbf{p} - e\mathbf{A}$$

for the ordinary momentum in mechanics is frequently described as *minimal electromagnetic coupling.*

19.2.2   The Gauge Covariant Schrödinger Equation

We now construct the time-independent Schrödinger equation and explain, what is meant by *"gauge covariance"*. We saw above: For a particle with charge e the relation between mechanical momentum and conjugate momentum is changed by the effect of the Lorentz force, i.e. the momentum $\mathbf{p}$ becomes $\mathbf{p} - e\mathbf{A}$. Correspondingly we have in nonrelativistic quantum mechanics the substitution

$$\mathbf{p}_{\mathrm{mech}} \to -i\hbar\boldsymbol{\nabla} - e\mathbf{A} = -i\hbar\left(\boldsymbol{\nabla} - \frac{ie}{\hbar}\mathbf{A}\right)$$

and $(\boldsymbol{\nabla} - \frac{ie}{\hbar}\mathbf{A})$ is called *gauge covariant derivative.*

The *canonical variables* required for quantisation are the spatial Cartesian coordinates x_i and the appropriate *canonical momenta* p_i. In the present case *quantisation* requires the replacement of $\mathbf{p}$ by the operators $\boldsymbol{\nabla}$, and to consider all quantities as operators in the space of states ψ. The corresponding Schrödinger equation in position space and with eigenvalue E is then (cf. Hamiltonian (19.7))

$$-\frac{\hbar^2}{2m}\left(\boldsymbol{\nabla} - \frac{ie}{\hbar}\mathbf{A}\right)^2\psi = (E - e\phi)\psi. \tag{19.8}$$

One expects a sensible formulation of the theory to be such that Schrödinger equations with different potentials $\mathbf{A}$ are "equivalent" in the sense that they yield the same observable results, which depend only on the probability density $|\psi|^2$. To achieve this, the wave function ψ has to be transformed along with the potential in the gauge transformation from one potential $\mathbf{A}$ to a different one. Thus, with the gauge transformation of the vector potential,

$$\mathbf{A} \to \mathbf{A}' = \mathbf{A} + \boldsymbol{\nabla}\chi$$

(where our time-independent considerations leave ϕ unchanged), we also transform the phase of the wave function by writing

$$\psi \to \psi' = \exp\left[\frac{ie\chi}{\hbar}\right]\psi.$$

This equivalence (to be verified below) is described as *gauge covariance.* We now verify that these transformations convert a Schrödinger equation in ψ' and $\mathbf{A}'$ into one in ψ and $\mathbf{A}$ or vice versa. We start from the Schrödinger equation in primed quantities:

$$-\frac{\hbar^2}{2m}\left(\nabla - \frac{ie}{\hbar}\mathbf{A}'\right)^2\psi' = (E - e\phi)\psi'. \tag{19.9}$$

Now

$$\left(\nabla - \frac{ie}{\hbar}\mathbf{A}'\right)\psi' = \left(\nabla - \frac{ie}{\hbar}(\mathbf{A} + \nabla\chi)\right)e^{\frac{ie\chi}{\hbar}}\psi$$

$$= e^{\frac{ie\chi}{\hbar}}\left(\nabla + \frac{ie}{\hbar}\nabla\chi - \frac{ie}{\hbar}(\mathbf{A} + \nabla\chi)\right)\psi$$

$$= e^{\frac{ie\chi}{\hbar}}\left(\nabla - \frac{ie}{\hbar}\mathbf{A}\right)\psi,$$

i.e.

$$\left(\nabla - \frac{ie}{\hbar}\mathbf{A}'\right)^2\psi' = e^{\frac{ie\chi}{\hbar}}\left(\nabla - \frac{ie}{\hbar}\mathbf{A}\right)^2\psi.$$

This implies the equation in $\mathbf{A}, \psi$:

$$-\frac{\hbar^2}{2m}\left(\nabla - \frac{ie}{\hbar}\mathbf{A}\right)^2\psi = (E - e\phi)\psi. \tag{19.10}$$

Since the only physically observable quantity in ψ is the probability density $|\psi|^2$, and this remains unchanged under the gauge transformation, we see that all gauge potentials, which are related by gauge transformations, describe the same physical state. (A wave function is a probability amplitude; the phases of wave functions have a *relative* significance, not an absolute significance, as quantum mechanical interference experiments show). In quantum mechanics the wave function of a free particle of mass m, energy $E = p^2/2m$ and momentum p is given by

$$\psi(x,t) = \exp\left[\frac{i}{\hbar}(px - Et)\right].$$

Recall briefly how quantum mechanics enters here. *Canonical quantisation* implies, that the Cartesian quantities x_i and p_i (as operators in the space of states) obey the following commutator algebra:

$$[x_i, x_j] = 0, \quad [p_j, x_k] = -i\hbar\delta_{jk}, \quad [p_j, p_k] = 0.$$

These operator relations are satisfied by the position space representation $p_j = -i\hbar\partial/\partial x_j$ of the momentum operator. Now assume $E = p^2/2m + V(x)$. Then, classically in the non-free case (i.e. potential $V(x) \neq 0$) $p^2/2m = E - V(x)$ and quantum mechanically

$$-\frac{\hbar^2\boldsymbol{\nabla}^2}{2m}\psi = (E - V(x))\psi,$$

where $\psi(x)$ is a wave function. In general one can separate the angles and one obtains a one-dimensional differential equation analogous to the Schrödinger equation in one dimension (with centrifugal term). An approximation ψ_0 (for $V \neq 0$, exact for $V = 0$) of ψ is obtained from the classical $p = \sqrt{2m(E - V(x))}$ as solution of

$$-i\hbar\frac{d}{dx}\psi_0 = \sqrt{2m(E - V(x))}\psi_0,$$

i.e.

$$\int\frac{d\psi_0}{\psi_0} = \int\frac{i}{\hbar}\sqrt{2m(E - V(x))}dx$$

or

$$\psi_0 = \exp\left[\frac{i}{\hbar}\int\sqrt{2m(E - V(x))}dx\right] = \exp\left[\frac{i}{\hbar}\int p(x)dx\right].$$
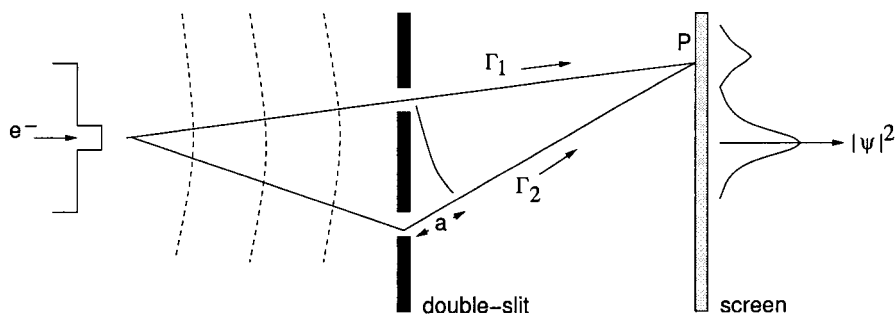
Fig. 19.1 The two paths Γ_1, Γ_2 of the electron wave
and their interference spectrum.

Considering now the case of a particle travelling in the Maxwell field $\mathbf{A}$, then according to the above prescription of minimal coupling, we have to replace the free particle momentum $\mathbf{p}$ by $\mathbf{p} - e\mathbf{A}$, so that the wave function has the form

$$\psi = \exp\left[\frac{i}{\hbar}\int\mathbf{p}\cdot d\mathbf{r} - \frac{ie}{\hbar}\int\mathbf{A}\cdot d\mathbf{r}\right].$$

The contribution

$$-\frac{ie}{\hbar} \int \mathbf{A} \cdot d\mathbf{r}$$

is the effect of the Lorentz force in the quantum mechanical case. We can consider it as the corresponding quantum mechanical law. It plays an important role in the interpretation of the Aharonov–Bohm effect. One should note that the expression $\oint \mathbf{A} \cdot d\mathbf{r}$ and hence the wave function ψ (for integration around a closed path) is invariant under nonsingular gauge transformations, since

$$\oint \mathbf{A}' \cdot d\mathbf{r} = \oint \mathbf{A} \cdot d\mathbf{r} + \oint \boldsymbol{\nabla}\chi \cdot d\mathbf{r} = \oint \mathbf{A} \cdot d\mathbf{r},$$

because

$$\oint \boldsymbol{\nabla}\chi \cdot d\mathbf{r} = \oint d\chi = 0.$$

19.3 The Aharonov–Bohm Effect

We restrict ourselves here largely to a qualitative treatment.[*] Electrons with momentum $p = h/\lambda$ are sent through a diaphragm or partition with two slits. If the difference in the lengths of the two possible paths, a, is of the order of the wavelength λ of the waves one expects and observes an interference pattern on the screen behind the slits (due to constructive or destructive interference). We write the difference between the phases of the wave function of the electrons due to the different paths

$$\triangle \left(\frac{1}{\hbar} \int \mathbf{p} \cdot d\mathbf{r} \right)_{p=\hbar k} = ka \equiv \delta. \tag{19.11}$$

At a maximum P as indicated in Fig. 19.1 we have the effect of the superposition of two waves which have a phase difference ka. If we arrange the experiment such that electrons can only pass through one slit, then $\delta = 0$, because $a = 0$ (as well as $\psi = 0$). In this case we observe no interference and the conditions of quantum mechanics are not provided.

The present problem with (as we shall see) a homogeneous magnetic field vertically out of the plane of the diagram of Fig. 19.1 in a small domain behind the slits is seen to have *cylindrical symmetry*. This suggests one to use cylindrical coordinates r, θ, z . In these coordinates, with field components

$$\mathbf{A} = \mathbf{e}_\theta A, \quad \mathbf{B} = \boldsymbol{\nabla} \times \mathbf{A} \parallel \mathbf{e}_z,$$

[*]For detailed literature see e.g. M. Peshkin and A. Tonomura, *The Aharonov–Bohm Effect*, Lecture Notes in Physics, Vol. 340 (Springer, 1989).

the gradient is given by

$$\boldsymbol{\nabla} \Leftrightarrow \mathbf{e}_r \left\{ \frac{\partial}{\partial r} \right\} + \mathbf{e}_\theta \left\{ \frac{1}{r} \frac{\partial}{\partial \theta} \right\} + \mathbf{e}_z \left\{ \frac{\partial}{\partial z} \right\},$$

(19.12)

and the divergence by

$$\boldsymbol{\nabla} \cdot \mathbf{C} = \frac{1}{r} \frac{\partial}{\partial r} (r C_r) + \frac{1}{r} \frac{\partial C_\theta}{\partial \theta} + \frac{\partial C_z}{\partial z}.$$

(19.13)

Hence

$$\boldsymbol{\nabla} \cdot \boldsymbol{\nabla} = \frac{1}{r} \frac{\partial}{\partial r} \left\{ r \frac{\partial}{\partial r} \right\} + \frac{1}{r} \frac{\partial}{\partial \theta} \left\{ \frac{1}{r} \frac{\partial}{\partial \theta} \right\} + \frac{\partial}{\partial z} \left\{ \frac{\partial}{\partial z} \right\}.$$

(19.14)

For the chosen vector potential $\mathbf{A}$ in the direction of $\mathbf{e}_\theta$ (in the outside region around the solenoid as we saw in Chapter 8) we have

$$
\begin{aligned}
\left(\boldsymbol{\nabla} - \frac{ie}{\hbar} \mathbf{A} \right)^2 &= \left(\boldsymbol{\nabla} - \frac{ie}{\hbar} A \mathbf{e}_\theta \right) \cdot \left[\frac{\partial}{\partial r}, \frac{1}{r} \frac{\partial}{\partial \theta} - \frac{ie}{\hbar} A, \frac{\partial}{\partial z} \right] \\
&= \frac{1}{r} \frac{\partial}{\partial r} r \frac{\partial}{\partial r} + \frac{1}{r} \frac{\partial}{\partial \theta} \left(\frac{1}{r} \frac{\partial}{\partial \theta} - \frac{ie}{\hbar} A \right) \\
&\quad + \frac{\partial}{\partial z} \left(\frac{\partial}{\partial z} \right) - \frac{ie}{\hbar} A \left(\frac{1}{r} \frac{\partial}{\partial \theta} - \frac{ie}{\hbar} A \right) \\
&= \frac{\partial^2}{\partial r^2} + \frac{1}{r} \frac{\partial}{\partial r} + \left(\frac{1}{r} \frac{\partial}{\partial \theta} - \frac{ie}{\hbar} A \right)^2 + \frac{\partial^2}{\partial z^2} \\
&= \frac{\partial^2}{\partial r^2} + \frac{1}{r} \frac{\partial}{\partial r} + \frac{1}{r^2} \left(\frac{\partial}{\partial \theta} - \frac{ie2\pi r}{h} A \right)^2 + \frac{\partial^2}{\partial z^2}.
\end{aligned}
$$

(19.15)

But now

$$\oint_{C(F)} \mathbf{A} \cdot d\mathbf{l} = A 2\pi r = \int (\boldsymbol{\nabla} \times \mathbf{A}) \cdot d\mathbf{F} = \int_{F(C)} \mathbf{B} \cdot d\mathbf{F} \equiv \Phi,$$

(19.16)

where Φ is the magnetic flux through the surface F. Hence with

$$\alpha := -\frac{e 2\pi r A}{h} = -\frac{e\Phi}{h} = -\frac{e\Phi}{2\pi\hbar}$$

the expression (19.15) becomes

$$\left(\boldsymbol{\nabla} - \frac{ie}{\hbar} \mathbf{A} \right)^2 = \frac{\partial^2}{\partial r^2} + \frac{1}{r} \frac{\partial}{\partial r} + \frac{1}{r^2} \left(\frac{\partial}{\partial \theta} + i\alpha \right)^2 + \frac{\partial^2}{\partial z^2}.$$

The corresponding Schrödinger equation for scattering in 2 dimensions is

$$\left(\nabla - \frac{ie}{\hbar}\mathbf{A}\right)^2 \psi = -\frac{2mE}{\hbar^2}\psi \equiv -k^2\psi, \qquad (19.17)$$

or with $\psi(r, \theta, z) \rightarrow \psi(r, \theta)$

$$\left[\frac{\partial^2}{\partial r^2} + \frac{1}{r}\frac{\partial}{\partial r} + \frac{1}{r^2}\left(\frac{\partial}{\partial \theta} + i\alpha\right)^2 + k^2\right]\psi(r, \theta) = 0. \qquad (19.18)$$

Setting $\psi(r, \theta) \propto e^{im\theta} = e^{im(\theta + 2\pi)}$, $\quad m = 0, \pm 1, \pm 2, \ldots$, we obtain

$$\left[\frac{\partial^2}{\partial r^2} + \frac{1}{r}\frac{\partial}{\partial r} - \frac{(\alpha + m)^2}{r^2} + k^2\right]\psi(r, \theta) = 0. \qquad (19.19)$$

Comparison with the Bessel equation

$$\frac{d^2 Z_\nu}{dz^2} + \frac{1}{z}\frac{dZ_\nu}{dz} + \left(1 - \frac{\nu^2}{z^2}\right)Z_\nu = 0 \qquad (19.20)$$

shows that the general solution of the Schrödinger equation has the form

$$\psi = \sum_{m=-\infty}^{\infty} e^{im\theta}\left[a_m J_{|m+\alpha|}(kr) + b_m J_{-|m+\alpha|}(kr)\right], \qquad (19.21)$$

where a_m, b_m are arbitrary constants and $J_\nu, J_{-\nu}$ are Bessel functions. Since the wave function has to be regular at $r = 0$, all coefficients b_m must be zero, i.e.

$$\psi(r, \theta) = \sum_{m=-\infty}^{\infty} a_m e^{im\theta} J_{|m+\alpha|}(kr). \qquad (19.22)$$

This is the small-r expansion. The Bessel functions can also be expressed as the sum of two Hankel functions valid at large r, i.e.

$$J_\nu(z) = \frac{1}{2}(H_\nu^{(1)}(z) + H_\nu^{(2)}(z)),$$

where for $|z| \rightarrow \infty$ (see books on Special Functions)

$$H_\nu^{(1)}(z) = \sqrt{\frac{2}{\pi z}}e^{i(z - \frac{\nu\pi}{2} - \frac{\pi}{4})}(1 + O(z^{-1})),$$

$$H_\nu^{(2)}(z) = \sqrt{\frac{2}{\pi z}}e^{-i(z - \frac{\nu\pi}{2} - \frac{\pi}{4})}(1 + O(z^{-1})).$$
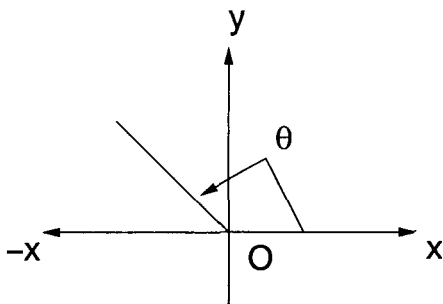
Fig. 19.2 The angle θ of cylindrical coordinates.

Hence for $r \to \infty$ the wave function can be written

$$\psi(r,\theta) = \frac{1}{2} \sum_{m=-\infty}^{\infty} a_m e^{im\theta} \sqrt{\frac{2}{\pi k r}} \left\{ e^{i(kr-|m+\alpha|\frac{\pi}{2}-\frac{\pi}{4})} \right.$$

$$\left. + e^{-i(kr-|m+\alpha|\frac{\pi}{2}-\frac{\pi}{4})} \right\} (1 + O(1/r)). \qquad (19.23)$$

The wave function $\psi(r,\theta)$ can now be subdivided for large values of r into an incident or ingoing wave (let's say from the right) and a scattered or outgoing wave whose asymptotic behaviour defines the scattering amplitude $f(\theta)$, i.e.

$$\psi(r,\theta) = \langle \underbrace{e^{-ikr\cos\theta}}_{e^{-ikx}} \rangle + \psi_{\text{out}} \qquad (19.24)$$

(the braces $\langle \dots \rangle$ are meant to indicate that the enclosed expression is not the strictly correct free wave, as is explained below) with (for $r \to \infty$ we deduce from the differential equation or the above Hankel functions that $\psi \sim e^{ikr}/\sqrt{r}$)

$$\psi_{\text{out}} = \frac{1}{r^{1/2}} e^{ikr} f(\theta)(1 + O(1/r)). \qquad (19.25)$$

We expand the ingoing wave in terms of Bessel functions. We take this expansion from the literature,[†] i.e.

$$e^{-ikr\cos\theta} = \sum_{n=-\infty}^{\infty} (-i)^n J_n(kr) e^{in\theta}$$

$$\simeq \sum_{n=-\infty}^{\infty} (-i)^n e^{in\theta} \left(\frac{2}{\pi k r} \right)^{\frac{1}{2}} \cos\left(kr - n\frac{\pi}{2} - \frac{\pi}{4} \right). \qquad (19.26)$$

[†]I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series and Products* (Academic, 1965), p. 973.
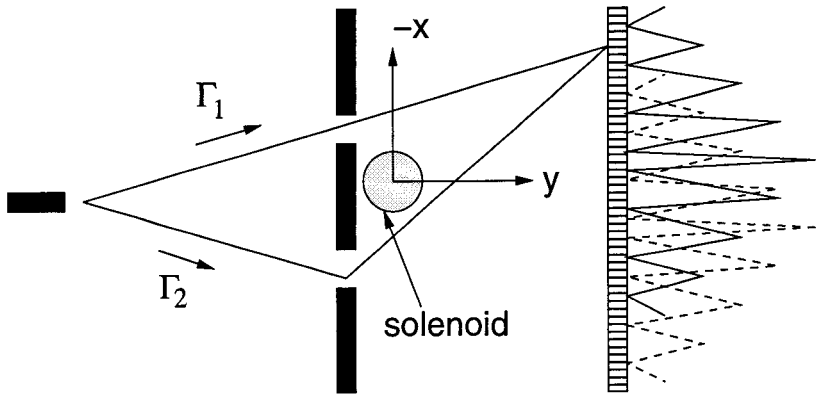
Fig. 19.3 Displacement of the interference spectrum
when a current flows in the solenoid.

We substitute (19.26) and (19.25) into (19.24). Comparison of the coefficients of e^{-ikr} in this expression and in (19.23) then yields the expression for a_m, i.e.

$$a_m = e^{-i\frac{\pi}{2}|m+\alpha|}. \tag{19.27}$$

Comparison of the coefficients of e^{+ikr} yields the scattering amplitude, scattering amplitude,

$$f(\theta) = \frac{1}{\sqrt{2\pi ik}} \sum_{m=-\infty}^{\infty} e^{im(\theta-\pi)}(e^{2i\delta_m} - 1) \tag{19.28}$$

with the *scattering phase*

$$\delta_m = -\frac{\pi}{2}|m+\alpha| + \frac{\pi}{2}|m|. \tag{19.29}$$

We see therefore, that for $\alpha \neq 0$ the scattering phase $\delta_m \neq 0$. Actually the ingoing wave in (19.24) is in the present case wrong. In the usual case of the scattering off a potential like the Coulomb potential (as in the familiar case of 3-dimensional scattering theory), one assumes that the potential vanishes for $r \to \infty$. Here we consider the scattering of electrons (or the electron wave) in the field of an infinitely long solenoid, i.e. off magnetic flux lines which do not return. Aharonov and Bohm showed in their work,[‡] that the ingoing wave therefore carries a phase factor, i.e. that Eq. (19.24) must be written

$$\psi(r,\theta) = e^{-i\alpha\theta(x,y)}e^{-ikr\cos\theta} + \psi_{\text{out}}. \tag{19.30}$$

[‡]The original papers are: D. Bohm and Y. Aharonov: "Significance of electromagnetic potentials in the quantum theory", *Phys. Rev.* **115** (1959) 485 and "Further considerations on electromagnetic potentials in the quantum theory", *Phys. Rev.* **123** (1961) 1511.

This change is by no means self-evident. Indeed the calculation of the scattering amplitude $f(\theta)$ here depends on the order in which the sum $\sum_m$ and the limit $\lim_{r\to\infty}$ are taken. Aharonov and Bohm used in their complicated calculation the order $\lim_{r\to\infty}\sum_m$, whereas in the above partial wave expansion these steps are taken in the opposite order. It is found that the summed expressions for $f(\theta)$ differ in forward direction by a contribution proportional to $\delta(\theta - \pi)$. We do not enter into a deeper analysis of these mathematical details here.[§]

We now consider the experimental setup shown schematically in Fig. 19.3 and proposed by Aharonov and Bohm with a solenoid (infinitely long, so that the field outside is zero) immediately behind the double slit and perpendicular to the diagram. In practice, in real experiments, the solenoid is replaced by microscopically thin magnetised iron crystals. One can see that with the current in the solenoid switched on, i.e. the magnetic flux, the phase of the electron wave function changes by an amount given by[¶]

$$\frac{\pi}{2}\alpha = \frac{\pi}{2}\left(- \frac{e\Phi}{2\pi\hbar}\right) = -\frac{\pi}{2}\frac{e}{2\pi\hbar}\oint \mathbf{A}\cdot dl,$$

which may be rewritten

$$\frac{\pi}{2}\alpha = -\frac{\pi}{2}\frac{e}{2\pi\hbar}\left[\int_{\Gamma_1} \mathbf{A}\cdot dl - \int_{\Gamma_2} \mathbf{A}\cdot dl\right]. \tag{19.31}$$

Here Γ_1 and Γ_2 are the two paths from the electron source through the two slits to the interference point on the screen. The interference spectrum suffers a corresponding shift as indicated in Fig. 19.3. This effect is known as *Aharonov–Bohm effect*.[‖] The effect has been confirmed experimentally[**] and plays an important role in numerous theoretical considerations.

What makes the Aharonov–Bohm effect so remarkable is:

[§]The partial wave treatment can be found in particular in S. N. M. Ruijsenhaars, *Ann. Phys.* (N.Y.) **146** (1983) 1. The discussion on the order of $\sum_m$ and $\lim_{r\to\infty}$ can be found in C.R. Hagen, *The Aharonov–Bohm Scattering Amplitude*, Univ. of Rochester Report UR-1103 (1989). Further discussions are given in the book cited above, of Peshkin and Tonomura.

[¶]In some considerations the wave function with phase factor $\exp\{ie\int_\Gamma \mathbf{A}\cdot dl/\hbar\}$ is not unique, since several revolutions around the field **B** are possible. This problem can be circumvented by summing over arbitrarily many revolutions or by using only uniquely defined wave functions (see e.g. the discussion in the last paragraph of M. V. Berry, *Proc. Roy. Soc. Lond.* **A392** (1984) 45).

[‖]The original papers have been cited above. A readable account of the Aharonov–Bohm effect can be found in Felsager [20], pp. 49–55.

[**]An early experimental verification can be found in G. Möllenstaedt and W. Bayh: "The continuous variation of the phase of electron waves in field free space by means of the magnetic vector potential of a solenoid", *Phys. Blätter* **18** (1962) 299. A more recent confirmation has been reported in A. Tonomura, O Noboyuki, T. Matsuda, T. Kawasaki, J. Endo, S. Yano and H. Yamada: "Evidence for Aharonov–Bohm effect with magnetic field completely shielded from electron wave", *Phys. Rev. Lett.* **56** (1986) 729.
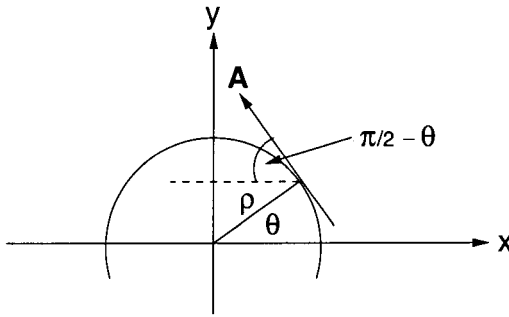
Fig. 19.4 The direction of $\mathbf{A}$ with respect to the angle θ.

(1) that it is a purely quantum mechanical (wave function) effect,
(2) that the electrons travel only in such domains in which $\mathbf{E}$ and $\mathbf{B}$ are zero, and
(3) (as we saw in Chapter 8) that also the vector potential $\mathbf{A}$ can be almost everywhere zero, i.e. where the singular gauge transformation is defined. However, there is no need to perform a singular gauge transformation. Thus, in quantum mechanics the vector potential has physical significance and becomes effectively observable in the shift of the interference spectrum, i.e. in an indirect way.

Example 19.1: Eigenvalues of $m\mathbf{v}$

For the case that α is proportional to the magnetic flux in the direction of z as in the above, and for propagation of electrons in the direction of x, so that the wave function ψ of the electrons is

$$\psi = e^{-i\alpha\theta(x,y)}e^{-ikx}, \quad \cot\theta = \frac{x}{y},$$

verify that

$$m v_x \psi = -\hbar k \psi, \quad m v_y \psi = m v_z \psi = 0.$$

Solution: We have the operator relation

$$m\mathbf{v} = -i\hbar\boldsymbol{\nabla} \pm e\mathbf{A},$$

and

$$\Phi = \int_F \mathbf{B}\cdot d\mathbf{F} = \int_F \boldsymbol{\nabla}\times\mathbf{A}\cdot d\mathbf{F} = \oint_C \mathbf{A}\cdot d\mathbf{l} = A(\rho)2\pi\rho,$$

so that

$$\alpha = \pm\frac{e\Phi}{2\pi\hbar} = \pm\frac{eA(\rho)2\pi\rho}{2\pi\hbar} = \pm\frac{eA(\rho)\rho}{\hbar}.$$

Hence

$$m v_x \psi = \left(-i\hbar\frac{\partial}{\partial x} \pm eA_x\right)e^{-i\alpha\theta(x,y)}e^{-ikx}$$

with, cf. Fig. 19.4,

$$\cot\theta = \frac{x}{y}, \quad \theta = \cot^{-1}\frac{x}{y}, \quad \frac{d}{dx}\theta = \frac{d}{dx}\cot^{-1}\frac{x}{y} = -\frac{y}{x^2+y^2} = -\frac{y}{\rho^2} = -\frac{\sin\theta}{\rho}.$$

It follows that

$$mv_x\psi = \left(- i\hbar(-i\alpha)\frac{\partial\theta(x,y)}{\partial x} - i\hbar(-ik) \pm eA_x \right)e^{-i\alpha\theta}e^{-ikx}.$$

This equation, i.e.

$$mv_x\psi = \left(- \alpha\hbar\frac{\partial\theta}{\partial x} - \hbar k \pm eA_x \right)e^{-i\alpha\theta}e^{-ikx},$$

may now be rewritten as

$$
\begin{aligned}
mv_x\psi &= \left(- k\hbar + \alpha\hbar\frac{\sin\theta}{\rho} \pm eA_x \right)e^{-i\alpha\theta}e^{-ikx} \\
&= \left(- k\hbar \pm eA(\rho)\rho\frac{\sin\theta}{\rho} \pm eA_x \right)e^{-i\alpha\theta}e^{-ikx} \\
&= (-k\hbar \pm eA(\rho)\sin\theta \pm eA_x)e^{-i\alpha\theta}e^{-ikx} \\
&= -k\hbar e^{-i\alpha\theta}e^{-ikx} = -k\hbar\psi,
\end{aligned}
$$

where in an intermediate step we replaced A_x by $-A(\rho)\cos(\pi/2-\theta) = -A(\rho)\sin\theta$. In an analogous way one can verify that

$$mv_y = 0, \quad \text{and} \quad mv_z = 0.$$

Chapter 20

Quantisation of the Electromagnetic Field and the Casimir Effect

20.1 Introductory Remarks

The crux of the *Casimir effect*[*] can already be explained with the help of the simple harmonic oscillator. In the case of the one-dimensional oscillator defined on $x \in [-\infty, \infty]$ one obtains the ground state energy $E_0 = \frac{1}{2}\hbar\omega$, also called zero point energy, where ω is the frequency of the oscillator. If the oscillator is restricted to a domain $x \in [-a, a]$, a large but finite, i.e. if the wave function ψ is subjected to the boundary condition $\psi(\pm a) = 0$, the eigenvalues naturally change and become

$$E_0 = \frac{\hbar\omega}{2}(1 + O(1/a)),$$

so that for $a \to \infty$ the zero point energy $E_0 = \frac{1}{2}\hbar\omega$ is regained. Thus with the boundary condition one obtains a contribution $\propto 1/a$ (or similar) to the energy, and hence a

$$\text{force} \quad = -\frac{\partial}{\partial a}(\text{energy}) \propto \frac{1}{a^2}.$$

Thus as a consequence of the boundary conditions, quantum mechanics implies a force proportional to $1/a^2$. The additional force derived from such a boundary-dependence in the case of quantised electrodynamics (in which

---

[*]The first book on this topic is that of K. A. Milton, *The Casimir Effect: Physical Manifestation of Zero-Point Energy* (World Scientific, 2002).

461

conductors play the role of the walls at $x = \pm a$ in the above) is referred to as Casimir-effect.

In the following we introduce first the simple canonical quantisation of the n-dimensional harmonic oscillator. This method is then used to perform an analogous quantisation of the electromagnetic field which naturally implies divergences if the field is visualised as providing harmonic oscillators at every point in space. Thus the calculation requires the definition of an artificial but sensible regularisation procedure which leaves a physically sensible result. The method is first applied to the academic example of electrodynamics in one space dimension but with boundary conditions. Thereafter the three-dimensional case is treated along similar lines, and it is shown how the Casimir force arises. Our considerations here can be regarded as a natural supplement to those of Chapter 14 on wave guides and resonators.

20.2   Quantisation of the n-Dimensional Harmonic Oscillator

In the canonical quantisation of the n-dimensional harmonic oscillator one usually proceeds as follows. The Hamiltonian of the problem to be quantised can be taken as (with a suitable choice of coefficients which are irrelevant here)

$$H(q, p) = \sum_{i=1}^{n} \frac{1}{2}(p_i^2 + q_i^2) \equiv \sum_{i=1}^{n} \frac{1}{2}(p_i^2 + \omega_i^2 q_i^2)|_{\omega_i = 1}.$$

Canonical quantisation implies, that we associate with every dynamical variable q_i, p_i (hermitian) operators $\hat{q}_i, \hat{p}_i$, which obey the following algebra:

$$[\hat{q}_i, \hat{q}_j] = 0, \quad [\hat{p}_i, \hat{p}_j] = 0, \quad [\hat{p}_i, \hat{q}_j] = -i\hbar\delta_{ij} \tag{20.1}$$

(in the following we usually set $\hbar = 1$), where the generalised coordinates q_i have to be Cartesian coordinates (i.e. $q_1 = x, q_2 = y, q_3 = z$). The momentum operator $\hat{p}_i$ has in position space the differential operator representation $-i\hbar\partial/\partial q_i$. Next one defines the "annihilation operators" $\hat{a}_i$ and "creation operators" $\hat{a}_i^\dagger$ by the relations

$$\hat{a}_i := \frac{1}{\sqrt{2}}(\hat{q}_i + i\hat{p}_i), \quad \hat{a}_i^\dagger := \frac{1}{\sqrt{2}}(\hat{q}_i - i\hat{p}_i), \tag{20.2}$$

and it is readily verified that in view of the relations (20.1) these satisfy the following commutator relations

$$[\hat{a}_i, \hat{a}_j] = 0, \quad [\hat{a}_i^\dagger, \hat{a}_j^\dagger] = 0, \quad [\hat{a}_i, \hat{a}_j^\dagger] = \delta_{ij}. \tag{20.3}$$

With a little algebra one can verify that $H(q, p)$ becomes the Hamilton operator given by

$$\hat{H}(\hat{q}, \hat{p}) = \sum_{i=1}^{n} \left(\hat{a}_i^\dagger \hat{a}_i + \frac{1}{2} \right) \equiv \sum_{i=1}^{n} \omega_i \left(N_i + \frac{1}{2} \right) \bigg|_{\omega_i=1}. \qquad (20.4)$$

The operator

$$\hat{N}_i = \hat{a}_i^\dagger \hat{a}_i$$

is called *particle number operator*. In the case of the harmonic oscillator we do not have real particles and therefore one uses instead the terms "quasi-particles" and "quasi-particle operators". The *"vacuum state"* or *"ground state"* $|0\rangle$ with wave function

$$\psi_0 \equiv \langle q_i | 0 \rangle$$

in position space representation is defined by the condition of "all annihilation operators applied to the vacuum give zero", i.e.

$$\hat{a}_i | 0 \rangle = 0 \quad \text{for all} \quad i = 1, 2, \ldots, n.$$

In this state the system has an energy given by

$$E_0 = \langle 0 | \hat{H}(\hat{q}, \hat{p}) | 0 \rangle = \frac{1}{2} \sum_{i=1}^{n} \omega_i, \qquad (20.5)$$

called *"zero point energy"*. It is an amusing exercise to find a way to Maxwell's equations with the help of the commutator relations (20.1), as in the following example.

Example 20.1: Feynman's "proof" of Maxwell's equations

Assume a particle has mass m and coordinates $x_j (j = 1, 2, 3)$ and velocity $\dot{x}_j$. Also assume the velocity satisfies Newton's equation $m\ddot{x}_j = F_j(x, \dot{x}, t)$, and the coordinates and velocity components satisfy the commutator relations

$$[x_j, x_k] = 0, \quad m[x_j, \dot{x}_k] = i\mu_0 \hbar \delta_{jk}.$$

Then show that there are fields $\mathbf{E}_i(x, t)$ and $\mathbf{H}_j(x, t)$, which satisfy the Lorentz-force equation $F_j = E_j + \epsilon_{jkl} \dot{x}_k H_l$ and the Maxwell equations $\nabla \cdot \mathbf{H} = 0$, $\mu_0 \frac{\partial \mathbf{H}}{\partial t} + \nabla \times \mathbf{E} = 0$. The remaining two Maxwell equations (with $\epsilon_0 \mu_0 c^2 = 1$) $\nabla \cdot \mathbf{E} = \rho/\epsilon_0$, $-\epsilon_0 \frac{\partial \mathbf{E}}{\partial t} + \nabla \times \mathbf{H} = \mathbf{j}$, then simply define the external charge and current densities ρ and $\mathbf{j}$.

Solution: The solution can be found in the paper of Dyson[†] and is therefore not reproduced here. This publication gave rise to further discussion in the literature.[‡]

[†]F. J. Dyson, *Am. J. Phys.* **58** (1990) 209.

[‡]See e.g. I.E. Farquhar, *Phys. Lett.* **A151** (1990) 203 and A. Vaidya and C. Farina, *Phys. Lett.* **A153** (1991) 265.

20.3   Hamiltonian of the Gauge Field

Our next objective is to derive the Hamiltonian density of the free electromagnetic field in analogy to the Hamiltonian of the n-dimensional harmonic oscillator as a necessary step towards its quantisation. For reasons of simplicity we use from now on again the same symbols q, p for c-number quantities as well as for operators $\hat{q}, \hat{p}$, since it is clear from the context which case is meant. Indeed, in the case of the free electromagnetic field we can proceed parallel to the case of the harmonic oscillator. The Lagrangian density $\mathcal{L}$ and the Lagrangian L are, as we saw in Chapter 18, given by

$$\mathcal{L} = -\frac{1}{4}F^{\mu\nu}F_{\mu\nu} \quad \text{and} \quad L = \int d^3x \mathcal{L}. \tag{20.6}$$

We have four dynamical (i.e. time-dependent) variables $A_\mu, \mu = 0, 1, 2, 3$. The Euler–Lagrange equations are, as we saw,

$$\partial_\mu \left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu A_\nu)} \right) - \frac{\partial \mathcal{L}}{\partial A_\nu} = 0,$$

i.e.

$$\partial_\mu F^{\mu\nu} = 0 \quad \left(\frac{\partial \mathcal{L}}{\partial A_\nu} = 0 \right). \tag{20.7}$$

These equations separate into (for $\nu = i, 0$):

(1) $\qquad\qquad \partial_0 F^{0i} = -\partial_j F^{ji} \quad \left(\frac{\partial \mathcal{L}}{\partial A_i} = 0 \right)$

(because of the time derivative this is an *equation of motion*), and

(2) $\qquad\qquad \partial_k F^{k0} = 0 \qquad \left(F^{00} = 0, \ \frac{\partial \mathcal{L}}{\partial A_0} = 0 \right)$

(this is a *constraint*).

We note that (2) results from the variation with respect to A^0. If one would impose from the very beginning the gauge fixing condition $A^0 = 0$, then (2), and hence (see below) the Gauss law would be lost!

The momenta canonically conjugate to the components A_μ are given by (again as we saw earlier)

$$\pi_k := \frac{\partial \mathcal{L}}{\partial(\partial_0 A^k)} = F_{k0} = \partial_k A_0 - \partial_0 A_k. \tag{20.8}$$

The result

$$\pi_0 := F_{00} = 0$$

is an additional constraint. From (2), and the definition of the momentum density components given by Eq. (20.8), we obtain

$$\partial_k \pi_k = 0.$$

This is the condition $\nabla \cdot \mathbf{E} = 0$, which is the Gauss law. We observed earlier (see Eq. (18.17)), that the Lagrangian density can be expressed as

$$\mathcal{L} = \frac{1}{2}\left(\frac{\mathbf{E}^2}{c^2} - \mathbf{B}^2\right).$$

The quantities entering here do not involve A_0. This suggests that we proceed as follows. We add the Gauss law with a Lagrangian multiplier A^0 to $\mathcal{L}$, so that (with — as we saw — $\partial\mathcal{L}/\partial A^0 = 0 = \nabla \cdot \mathbf{E}/c$)

$$\mathcal{L} = \frac{1}{2}\left(\frac{\mathbf{E}^2}{c^2} - \mathbf{B}^2\right) + A^0 \nabla \cdot \frac{\mathbf{E}}{c}. \tag{20.9}$$

The canonical Hamiltonian density $\mathcal{H}$ is as usual given by a Legendre transform for which (cf. Eq. (18.10))

$$
\begin{aligned}
\mu_0 \mathcal{H} &= \pi^\rho \partial^0 A_\rho - \mathcal{L} \\
&= -\pi_k(\partial_0 A_k) - \mathcal{L} \\
&\stackrel{(20.8)}{=} -\pi_k(-\pi_k + \partial_k A_0) - \frac{1}{2}(\pi_k^2 - B_k^2) - A_0(\partial_k \pi_k) \\
&= \frac{1}{2}(\pi_k^2 + B_k^2) - \pi_k(\partial_k A_0) - A_0(\partial_k \pi_k) \\
&= \frac{1}{2}(\pi_k^2 + B_k^2) - \partial_k(\pi_k A_0).
\end{aligned} \tag{20.10}
$$

Hence if we integrate over all space

$$H = \frac{1}{\mu_0}\int d^3x \mathcal{H} = \frac{1}{\mu_0}\int d^3x \frac{1}{2}(\pi_k^2 + B_k^2). \tag{20.11}$$

We now choose the gauge $A_0 = 0$ and $\nabla \cdot \mathbf{A} = 0$, so that (see Eq. (10.11))

$$\Box \mathbf{A} = 0. \tag{20.12}$$

The condition $A_0 = 0$ suggests itself, since $\pi_0 = 0$; i.e. the 0-degree of freedom is completely eliminated. The other constraint $\nabla \cdot \mathbf{A} = 0$ together with Eq. (20.12) then ensures that only two components of $\mathbf{A}$ are independent.

20.4 Quantisation of the Electromagnetic Field

Equation (20.12) allows us to write the solutions

$$\mathbf{A} \propto \frac{1}{\sqrt{V}}\epsilon_{k\lambda}e^{i(\mathbf{k}\cdot\mathbf{x}-\omega t)} \quad (\text{or sin, cos}). \tag{20.13}$$

For these we have the dispersion relation

$$\omega = kc \equiv \omega_k.$$

The solutions of $\Box \mathbf{A} = 0$ are therefore periodic and we can write them as Fourier expansions (with $q_{k,\lambda}(t) = q_{k,\lambda}(0)e^{-i\omega t}$)

$$\mathbf{A}(\mathbf{x}, t) = \frac{c}{\sqrt{V}} \sum \int \epsilon_{k,\lambda} e^{i\mathbf{k}\cdot\mathbf{x}} q_{k,\lambda}(t),$$

$$\boldsymbol{\pi}(\mathbf{x}, t) = \frac{1}{\sqrt{V}} \sum \int \epsilon_{k,\lambda} e^{i\mathbf{k}\cdot\mathbf{x}} p_{k,\lambda}(t), \tag{20.14}$$

where V is the space volume and $\lambda = 1, 2$ label the two possible, mutually orthogonal directions of polarisation with

$$\epsilon_{k,\lambda} \cdot \epsilon^*_{k,\lambda'} = \delta_{\lambda,\lambda'}. \tag{20.15}$$

The symbol $\sum \int$ is meant to indicate summation over discrete parameters (λ) and integration over continuous parameters (the components of $\mathbf{k}$). The condition $\nabla \cdot \mathbf{A} = 0$ implies

$$\epsilon_{k,\lambda} \cdot \mathbf{k} = 0. \tag{20.16}$$

Then with $\frac{1}{V} \int d^3x e^{i\mathbf{k}\cdot\mathbf{x}} = \delta(\mathbf{k})$ in the second step we have

$$\int \pi^2 d^3x = \frac{1}{V} \int d^3x \sum \int_{k,\lambda} \sum \int_{k',\lambda'} \epsilon_{k,\lambda} \cdot \epsilon_{k',\lambda'} e^{i(\mathbf{k}+\mathbf{k}')\cdot\mathbf{x}} p_{k,\lambda}(t) p_{k',\lambda'}(t)$$

$$= \sum \int_{k,\lambda} \sum \int_{k',\lambda'} \epsilon_{k,\lambda} \cdot \epsilon_{k',\lambda'} \delta(\mathbf{k} + \mathbf{k}') p_{k,\lambda}(t) p_{k',\lambda'}(t)$$

$$= \sum \int_{k,\lambda,\lambda'} \epsilon_{k,\lambda} \cdot \epsilon_{-k,\lambda'} p_{k,\lambda}(t) p_{-k,\lambda'}(t). \tag{20.17}$$

We demand that $\mathbf{A}$ and $\boldsymbol{\pi}$ are real and as operators hermitian (since they represent quantum mechanical observables). It follows that for $\boldsymbol{\pi}(\mathbf{x}, t)$ (observe that $\sum_k$ contains the summation over positive as well as negative components of $\mathbf{k}$, which is used in the last step in the following)

$$\boldsymbol{\pi}(\mathbf{x}, t) = \frac{1}{\sqrt{V}} \sum \int_{k,\lambda} e^{i\mathbf{k}\cdot\mathbf{x}} \epsilon_{k,\lambda} p_{k,\lambda}(t)$$

$$= \frac{1}{\sqrt{V}} \sum \int_{k,\lambda} e^{-i\mathbf{k}\cdot\mathbf{x}} \epsilon^*_{k,\lambda} p^*_{k,\lambda}(t) \ \ (= \boldsymbol{\pi}^*(\mathbf{x}, t))$$

$$= \frac{1}{\sqrt{V}} \sum \int_{k,\lambda} e^{+i\mathbf{k}\cdot\mathbf{x}} \epsilon^*_{-k,\lambda} p^*_{-k,\lambda}(t), \tag{20.18}$$

i.e. (comparison of the first line with the third)

$$\epsilon_{k,\lambda} p_{k,\lambda}(t) = \epsilon^*_{-k,\lambda} p^*_{-k,\lambda}(t), \tag{20.19}$$

or

$$\epsilon^*_{k,\lambda} p^*_{k,\lambda}(t) = \epsilon_{-k,\lambda} p_{-k,\lambda}(t), \tag{20.20}$$

so that with Eq. (20.15)

$$
\begin{aligned}
\int \pi^2 d^3x &= \sum_{k,\lambda,\lambda'} \int \epsilon_{k,\lambda} \cdot \epsilon_{-k,\lambda'} p_{k,\lambda}(t) p_{-k,\lambda'}(t) \\
&= \sum_{k,\lambda,\lambda'} \int \epsilon_{k,\lambda} \cdot \epsilon^*_{k,\lambda'} p_{k,\lambda}(t) p^*_{k,\lambda'}(t) \\
&= \sum_{k,\lambda} \int p^*_{k,\lambda}(t) p_{k,\lambda}(t).
\end{aligned} \tag{20.21}
$$

In order to be able to handle the other terms in H in an analogous manner, we first have to rewrite these. It can be shown, as we do below, that

$$\int d^3x (\nabla \times \mathbf{A}) \cdot (\nabla \times \mathbf{A}) = -\int d^3x \mathbf{A} \cdot (\triangle \mathbf{A}). \tag{20.22}$$

We use Gauss' divergence theorem

$$\int d^3x \nabla \cdot \mathbf{D} = \int \mathbf{D} \cdot d\mathbf{F}. \tag{20.23}$$

Here we set

$$\mathbf{D} := \mathbf{A} \times (\nabla \times \mathbf{A}).$$

Then

$$\nabla \cdot \mathbf{D} = \nabla \cdot \{\mathbf{A} \times (\nabla \times \mathbf{A})\}.$$

We use

$$\nabla \cdot (\mathbf{A} \times \mathbf{B}) = \mathbf{B} \cdot (\nabla \times \mathbf{A}) - \mathbf{A} \cdot (\nabla \times \mathbf{B}),$$

so that

$$\nabla \cdot \{\mathbf{A} \times (\nabla \times \mathbf{A})\} = (\nabla \times \mathbf{A}) \cdot (\nabla \times \mathbf{A}) - \mathbf{A} \cdot (\nabla \times \{\nabla \times \mathbf{A}\}).$$

Recalling the relation "*curl curl = grad div − div grad* ", we have

$$\nabla \times (\nabla \times \mathbf{A}) = \nabla(\nabla \cdot \mathbf{A}) - \triangle \mathbf{A}.$$

Hence

$$\nabla \cdot \{\mathbf{A} \times (\nabla \times \mathbf{A})\} = (\nabla \times \mathbf{A}) \cdot (\nabla \times \mathbf{A}) - (\mathbf{A} \cdot \nabla)(\nabla \cdot \mathbf{A}) + \mathbf{A} \cdot (\triangle \mathbf{A}).$$

We insert this in Eq. (20.23) and obtain:

$$\underbrace{\int d^3x(\boldsymbol{\nabla} \times \mathbf{A}) \cdot (\boldsymbol{\nabla} \times \mathbf{A}) - \int d^3x(\mathbf{A} \cdot \boldsymbol{\nabla})(\boldsymbol{\nabla} \cdot \mathbf{A}) + \int d^3x \mathbf{A} \cdot (\triangle \mathbf{A})}_{\int d^3x \boldsymbol{\nabla} \cdot \mathbf{D}}$$

$$= \int \{\underbrace{\mathbf{A} \times (\boldsymbol{\nabla} \times \mathbf{A})}_{\mathbf{D}}\} \cdot d\mathbf{F}$$

$$= \int d\mathbf{F} \cdot \{A_i \boldsymbol{\nabla} A_i - (\mathbf{A} \cdot \boldsymbol{\nabla})\mathbf{A}\}. \tag{20.24}$$

The right side is an integral over the surface area of the volume V on which $\mathbf{A} \to 0$. Hence this integral vanishes. With the gauge fixing condition $\boldsymbol{\nabla} \cdot \mathbf{A} = 0$ another contribution drops out, and we obtain the result (20.22), i.e.

$$\int d^3x(\boldsymbol{\nabla} \times \mathbf{A}) \cdot (\boldsymbol{\nabla} \times \mathbf{A}) = -\int d^3x \mathbf{A} \cdot (\triangle \mathbf{A}).$$

Hence

$$\int d^3x B_k^2 = \int d^3x(\boldsymbol{\nabla} \times \mathbf{A}) \cdot (\boldsymbol{\nabla} \times \mathbf{A}) = -\int d^3x \mathbf{A} \cdot (\triangle \mathbf{A})$$

$$= k^2 \int d^3x \mathbf{A} \cdot \mathbf{A} \tag{20.25}$$

since $\triangle \mathbf{A} = -k^2 \mathbf{A}$. Repeating the procedure of above we obtain evidently

$$\int d^3x B_k^2 = \sum \int_{k,\lambda} \omega_k^2 q_{k,\lambda}^*(t) q_{k,\lambda}(t). \tag{20.26}$$

Thus the expression for the Hamiltonian becomes

$$H = \frac{1}{2\mu_0} \sum \int_{k,\lambda} [p_{k,\lambda}^*(t) p_{k,\lambda}(t) + \omega_k^2 q_{k,\lambda}^*(t) q_{k,\lambda}(t)]. \tag{20.27}$$

The next step is quantisation. We demand as quantisation conditions at equal times t (with transition from c-number valued A_i, π_j to operators, and in analogy with the harmonic oscillator)

$$[A_i(\mathbf{x}, t), A_j(\mathbf{x}', t)] = 0, \quad [\pi_i(\mathbf{x}, t), \pi_j(\mathbf{x}', t)] = 0,$$

$$[\pi_j(\mathbf{x}, t), A_i(\mathbf{x}', t)] = -\frac{i\mu_0 \hbar}{c} \delta_{ij} \delta(\mathbf{x} - \mathbf{x}'), \tag{20.28}$$

or

$$[q_{k,\lambda}(t), q_{k',\lambda'}(t)] = 0, \quad [p_{k,\lambda}(t), p_{k',\lambda'}(t)] = 0,$$

$$[p_{k,\lambda}(t), q_{k',\lambda'}(t)] = -i\mu_0 \hbar \delta_{k,k'} \delta_{\lambda,\lambda'}. \tag{20.29}$$

We now define annihilation and creation operators

$$a_{\lambda,k} = \sqrt{\frac{\omega_k}{2}}\left(q_{k,\lambda} + \frac{i}{\omega_k}p_{k,\lambda}\right), \quad a_{\lambda,k}^\dagger = \sqrt{\frac{\omega_k}{2}}\left(q_{k,\lambda} - \frac{i}{\omega_k}p_{k,\lambda}\right) \quad (20.30)$$

with

$$[a_{\lambda,k}, a_{\lambda',k'}] = 0, \quad [a_{\lambda,k}^\dagger, a_{\lambda',k'}^\dagger] = 0,$$
$$[a_{\lambda,k}, a_{\lambda',k'}^\dagger] = \mu_0\hbar\delta_{\lambda,\lambda'}\delta_{k,k'}, \quad (20.31)$$

so that in analogy with the case of the harmonic oscillator

$$H = \sum \int_{k,\lambda} \omega_k\left(\frac{a_{k,\lambda}^\dagger a_{k,\lambda}}{\mu_0} + \frac{\hbar}{2}\right). \quad (20.32)$$

The *vacuum state* $|0\rangle$ or ground state of the quantised electromagnetic field is defined by

$$a_{\lambda,k}|0\rangle = 0, \quad \text{all } k, \lambda. \quad (20.33)$$

The energy of the system in this state is the zero point energy

$$E_0 = \langle 0|H|0\rangle = \frac{1}{2}\sum_{k,\lambda}\hbar\omega_k. \quad (20.34)$$

Since we have to sum here over infinitely many values of **k**, the sum is (different from the case of the n-dimensional harmonic oscillator) infinite. This result seems very natural because in field theory we have effectively at every point in space one or more oscillators, each of which contributes to the zero point energy. If the physical field is to satisfy boundary conditions at boundaries, these change the spectrum, i.e. the values of ω_k and hence the zero point energy. For this reason it is plausible to define the finite *physical vacuum energy* as a difference in zero point energies, i.e. as

$$E_{\text{vac}} = E_{0,\text{with boundary}} - E_{0,\text{without boundary}}. \quad (20.35)$$

If $E_{\text{vac}} \neq 0$, this effect or the corresponding force has to be attributed to the presence of boundaries. The first investigation of this effect was published by Casimir in 1948.[§] Casimir considered the zero point energy of the electromagnetic field between two infinitely extended parallel metal plates a distance d apart, and showed that as a result of the quantum nature of the field an attractive force $\propto -1/d^4$ acts between the plates.

[§]H. B. G. Casimir, *Proc. Kon. Ned. Akad. Wet.* **51** (1948) 793. A very readable review paper is that of G. Plunien, B. Müller and W. Greiner, *Physics Reports* **134** (1986) 87.

20.5 One-dimensional Illustration of the Casimir Effect

In his original work (see above) Casimir considered two *uncharged, parallel, perfectly conducting* metal plates, and showed that an attractive force proportional to $-1/d^4$ acts between the plates, where d is the distance between the plates. We now want to consider explicitly the analogous calculation for the academic, though simpler, 1-dimensional case. We shall see that it is not enough to define the vacuum energy as a difference. In order to obtain well defined, i.e. finite expressions, one must also carry out a *regularisation* of the expressions. In order to justify this regularisation physically, we recall that at high frequencies of the radiation, $\omega > \omega_c$ (ω_c a critical or so-called cutoff frequency) the metal behaves like a dielectric and hence is practically transparent to the radiation.* This means that the eigenfrequencies of very high modes ($\omega_n > \omega_c$) remain unaffected by the boundaries and hence do not contribute to the vacuum energy.

We consider a "one-dimensional box" of length a, i.e. a box with cross section practically zero (cf. the rectangular resonator we discuss later), between two conductor planes, and we let the one-dimensional separation of these be between $x = 0$ and $x = a$. Moreover, we imagine a third and parallel conductor plane at a farther distance $L - a$ as indicated in Fig. 20.1. This additional plate serves the purpose of permitting the construction of the difference of energies referred to above. At the metal boundaries the electromagnetic field in the intermediate space must satisfy boundary conditions.
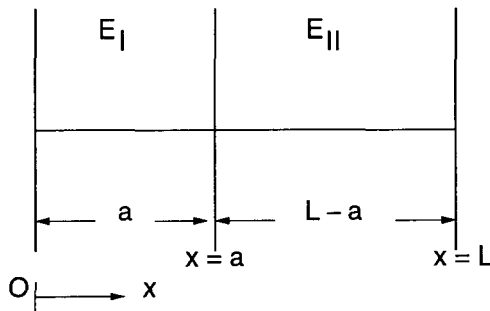
Fig. 20.1 The three conducting plates.

In the present idealised one-dimensional case this means, we have to select conditions that correspond to the boundary conditions we discussed at length

---

*See the discussion at the end of Sec. 9.4.

earlier, i.e.

$$\mathbf{n} \cdot \mathbf{B} = 0, \quad \mathbf{n} \times \mathbf{E} = 0. \tag{20.36}$$

Since the radiation is confined to the one-dimensional space between the "walls" (i.e. endpoints), this implies, that we have to demand periodicity, so that we obtain stationary waves.

This means we have instead of the vector potential $\mathbf{A}(\mathbf{x}, t)$ only a scalar field $A(x, t)$ (i.e. we have the case of only $\mathbf{E} = -\nabla\phi, \phi = A$) with $A(x, t) = 0$ at $x = 0$ and $x = a$. Since

$$\Box A(x, t) = 0,$$

we have

$$A(x, t) \propto e^{i\omega t} \sin kx$$

with

$$ka = n\pi, \quad n = 0, \pm 1, \pm 2, \ldots.$$

Thus the momentum is discretised. This is why we considered immediately discretised momenta above. It follows that $\omega = kc$ leads to the eigenfrequencies

$$\omega_n = \frac{n\pi}{a} c. \tag{20.37}$$

The vacuum energy in the domain I in Fig. 20.1 is therefore (since the Hamilton operator is positive, semidefinite, only values $n \geq 0$ make sense)

$$E_I(a) = \frac{1}{2} \sum_{k,\lambda} \omega_k = \frac{1}{2} \sum_{n>0} \frac{n\pi}{a} \tag{20.38}$$

with $\hbar = c = 1$. The summation over λ yields in general the factor 2, in view of the two possible directions of polarisation, here in the one-dimensional case, of course, not. This means we obtain $E_I(a) \simeq \frac{1}{2} \sum_{n>0} n\pi/a$, an absolutely divergent series! We assume, as argued at the beginning, that high eigenfrequencies, i.e. modes, do not contribute significantly. This implies that the sum does not have to be extended up to $n = \infty$. For this reason we introduce a parameter α as a cutoff factor, and write

$$E_I(a) \rightarrow E_I(a, \alpha) = \frac{1}{2} \sum_{n=1}^{\infty} \left(\frac{n\pi}{a}\right) e^{-\alpha(\frac{n\pi}{a})}$$

$$= -\frac{\partial}{\partial\alpha} \frac{1}{2} \sum_{n=1}^{\infty} e^{-\alpha(\frac{n\pi}{a})}.$$

The sum is now a geometric progression which is found to give

$$E_I(a, \alpha) = -\frac{\partial}{\partial \alpha} \frac{1}{2} \frac{1}{1 - e^{-\alpha(\frac{\pi}{a})}}$$

$$= \frac{\pi}{2a} e^{-\alpha\frac{\pi}{a}} [1 - e^{-\alpha\frac{\pi}{a}}]^{-2}.$$

Expanding the exponentials we have

$$E_I(a, \alpha) = \frac{\pi}{a} \left[1 - \alpha\left(\frac{\pi}{a}\right) + \frac{1}{2}\alpha^2\left(\frac{\pi}{a}\right)^2 - \cdots \right]$$

$$\times \frac{1/2}{\alpha^2(\frac{\pi}{a})^2} \left[1 + \alpha\left(\frac{\pi}{a}\right) + \alpha^2\left(-\frac{2}{6} + \frac{3}{2^2}\right)\left(\frac{\pi}{a}\right)^2 - \cdots \right],$$

and hence

$$E_I(a, \alpha) = \frac{1/2}{\alpha^2(\frac{\pi}{a})} \left[1 + \alpha^2\left(\frac{\pi}{a}\right)^2 \left(\frac{1}{2} + \frac{5}{12} - 1\right) + \cdots \right]$$

$$= \frac{1/2}{\alpha^2(\frac{\pi}{a})} \left[1 - \frac{\alpha^2}{12}\left(\frac{\pi}{a}\right)^2 + \cdots \right].$$

Thus the result becomes

$$E_I(a, \alpha) = \frac{1}{2}\left[\left(\frac{a}{\pi}\right)\frac{1}{\alpha^2} - \frac{1}{12}\left(\frac{\pi}{a}\right) + O(\alpha^2) \right]. \tag{20.39}$$

For the domain II in Fig. 20.1, the corresponding expression is (replacing in Eq. (20.39) a by $L - a$)

$$E_{II}(L - a, \alpha) = \frac{1}{2}\left[\left(\frac{L-a}{\pi}\right)\frac{1}{\alpha^2} - \frac{1}{12}\left(\frac{\pi}{L-a}\right) + O(\alpha^2) \right]. \tag{20.40}$$

Hence for $L \gg a$

$$E_I(a, \alpha) + E_{II}(L-a, \alpha) = \frac{1}{2}\left[\left(\frac{L}{\pi}\right)\frac{1}{\alpha^2} - \frac{1}{12}\left(\frac{\pi}{a}\right) - \frac{1}{12}\left(\frac{\pi}{L}\right) + O(\alpha^2) \right]. \tag{20.41}$$

We see that this sum removed the contribution $a/2\pi\alpha^2$ in Eq. (20.39), but produced the contribution $L/2\pi\alpha^2$. We now construct the procedure which allows the removal of the wall at $x = L$ to infinity without, however, producing a new contribution $\propto 1/\alpha^2$. This is achieved by another pair of walls as depicted in Fig. 20.2 with L thereafter taken to infinity. In other words, we single out the finite boundary effect by subtracting the corresponding expression with a replaced by $L/\eta, \eta > 1$ as indicated in Fig. 20.2 (we demand

$\eta > 1$, so that $L - L/\eta \neq 0$). The additional "walls" (endpoints) are then removed by taking the limit $L \to \infty$. This means we consider the difference

$$[E_I(a, \alpha) + E_{II}(L - a, \alpha)] - [E_{III}(L/\eta, \alpha) + E_{IV}(L - L/\eta, \alpha)]. \quad (20.42)$$

Here, according to Eq. (20.39),

$$E_{III}(L/\eta, \alpha) = \frac{1}{2}\left[\left(\frac{L/\eta}{\pi}\right)\frac{1}{\alpha^2} - \frac{1}{12}\left(\frac{\pi}{L/\eta}\right) + O(\alpha^2)\right] \quad (20.43)$$

and

$$E_{IV}(L - L/\eta, \alpha) = \frac{1}{2}\left[\left(\frac{L - L/\eta}{\pi}\right)\frac{1}{\alpha^2} - \frac{1}{12}\left(\frac{\pi}{L - L/\eta}\right) + O(\alpha^2)\right]. \quad (20.44)$$

Hence for $L \to \infty$
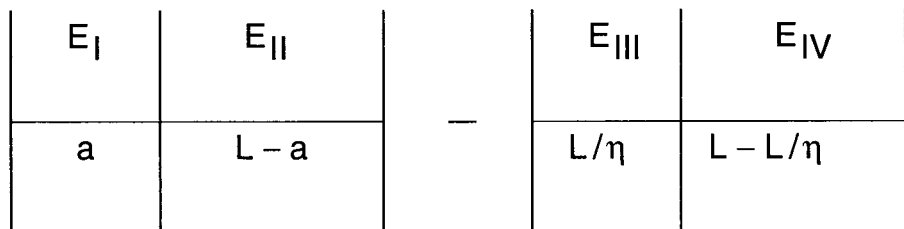
Fig. 20.2 The subtraction procedure.

$$E_{III} + E_{IV} = \frac{1}{2}\left(\frac{L}{\pi}\right)\frac{1}{\alpha^2} + O\left(\frac{1}{L}\right) + O(\alpha^2).$$

Thus finally

$$\lim_{\alpha \to 0, L \to \infty}[(E_I + E_{II}) - (E_{III} + E_{IV})] = -\frac{\pi}{24a}. \quad (20.45)$$

Hence there remains a finite energy or potential

$$V(a) = -\frac{\pi}{24a} \quad (20.46)$$

with the repulsive force

$$F = -\frac{\partial V}{\partial a} = +\frac{\pi}{24a^2}. \quad (20.47)$$

This quantum mechanical effect which yields a force between the walls, is called *Casimir-effect*.

20.6   The Three-Dimensional Case

In the original 3-dimensional problem considered by Casimir the geometry is analogous to that of the previous case, but now areas in the (x, y)-plane represent metal plates. This geometry reminds us immediately of that of resonators, which we discussed in Chapter 14, where boundary conditions resulted from the plates at both ends of an otherwise open, cylindrical wave guide.
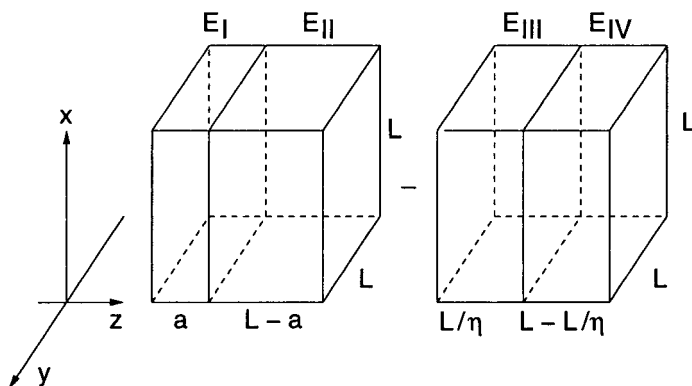
Fig. 20.3 The subtraction procedure in the 3-dimensional case.

From there we obtain (see Eq. (14.152)) for the case when the interior is the vacuum, and for $-L/2 \leq x, y \leq L/2$ as in Figs. 20.3 and 20.4 (instead of $0 \leq x, y < L$), so that with $L \to \infty$ these walls can be removed,

$$\frac{\omega^2}{c^2} = \left(\frac{n\pi}{a}\right)^2 + k_\perp^2, \quad n = 0, \pm 1, \pm 2, \ldots, \tag{20.48}$$

or (see Sec. 14.10)

$$\frac{\omega^2}{c^2} = \left(\frac{n\pi}{a}\right)^2 + \left(\frac{n_x\pi}{L/2}\right)^2 + \left(\frac{n_y\pi}{L/2}\right)^2, \tag{20.49}$$

considering a resonator of volume $L^2 a$. In the present case n_x, n_y are not fixed as integers, since we allow the corresponding areas to extend to infinity (with L). The summations over n_x, n_y must therefore be replaced by integrations. Because

$$k_x = \frac{n_x\pi}{L/2},$$

we have

$$dn_x = \frac{L}{2\pi}dk_x, \quad dn_x dn_y = \left(\frac{L}{2\pi}\right)^2 dk_\perp^2, \tag{20.50}$$
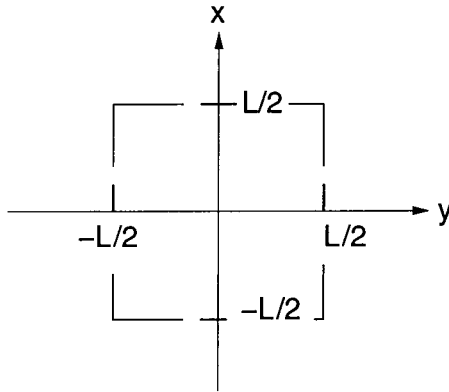
Fig. 20.4 The cross section.

and the sum

$$E_0 = \frac{1}{2} \sum_{\lambda, k} \hbar \omega_k$$

yields

$$E_I(a) = \frac{\hbar}{2} \sum_{\lambda, n} c \int dn_x dn_y \sqrt{\left(\frac{n\pi}{a}\right)^2 + \left(\frac{n_x \pi}{L/2}\right)^2 + \left(\frac{n_y \pi}{L/2}\right)^2}.$$

We can write this

$$E_I(a) = \frac{1}{2} \sum_{\lambda, n} \hbar c \left(\frac{L}{2\pi}\right)^2 \int dk_\perp^2 \sqrt{\left(\frac{n\pi}{a}\right)^2 + k_x^2 + k_y^2}.$$

Separating the $n = 0$ part and introducing the cutoff exponential, this becomes

$$E_I(a) = \frac{\hbar c}{2} \sum_{\lambda} \left(\frac{L}{2\pi}\right)^2 \int d^2 k_\perp \left[\underbrace{\sqrt{k_x^2 + k_y^2} e^{-\alpha k_\perp^2}}_{n=0 \text{ part}} \right.$$

$$\left. + 2 \sum_{n=1}^{\infty} e^{-\alpha k_\perp^2} \sqrt{\left(\frac{n\pi}{a}\right)^2 + k_x^2 + k_y^2} \right]. \qquad (20.51)$$

Here $\sum_\lambda$ is the sum over the two possible directions of polarisation. Since the integral is in any case divergent, we have again introduced a frequency cutoff

α, which will be sent to zero at the end of the calculation. The arguments now run parallel to those of the previous case. Finally, one obtains the energy

$$E(a) = \lim_{L \to \infty} [\{E_I(a) + E_{II}(L - a)\} - \{E_{III}(L/\eta) + E_{IV}(L - L/\eta)\}]$$

$$= -\frac{L^2 \pi^2}{720} \frac{\hbar c}{a^3}, \tag{20.52}$$

with the attractive force per unit area (vacuum pressure) (i.e. divided by the area L^2)[†]

$$F = -\frac{\partial E(a)}{\partial a} = -\frac{\pi^2}{240} \frac{\hbar c}{a^4_{\text{meter}}} \text{ newton/meter}^2$$

$$= -\frac{1.30 \times 10^{-26}}{a^4_{\text{cm}}} \text{ dyne/cm}^2. \tag{20.53}$$

The earliest experimental investigation of the Casimir effect was carried out by M. J. Sparnaay.[‡] Sparnaay investigated the effect in the domain

$$0.5\mu \lesssim a \lesssim 2\mu \quad (1\mu = 10^{-6}\text{cm})$$

and observed a force, which did not contradict Casimir's original theoretical findings. By now the Casimir effect has been investigated in a large number of field theoretical models and has been verified experimentally. The latest experiment on the verification of the Casimir effect is that of Lamoreux,[§] who achieved agreement with theoretical prediction to an error of at most 5%. In recent years the Casimir effect has become a very active area of highly specialised research and therefore will not be treated in more detail in the present context.[¶] The case in which the metal plates of the above example are replaced by concentric spherical shells has also been treated in the literature.[‖] We add, that external boundary conditions like those we considered here, play a role analogous to that of external fields, which supply observable effects via quantum fluctuations.

[†]See e.g. Plunien, B. Müller and W. Greiner, *Physics Reports* **134** (1986) 87, or M. Fierz, *Helv. Phys. Acta* **33** (1960) 855.

[‡]M. J. Sparnaay, *Physica* **XXIV** (1958) 751.

[§]S. K. Lamoreux, *Phys. Rev. Lett.* **78** (1997) 5.

[¶]Some references from which also earlier literature can be traced back are G. Barton and C. Eberlein, *Ann. Phys.* (N.Y.) **227** (1993) 222, and J. B. Pendry, *J. Phys. Cond. Matt.* **9** (1997) 10301. These also treat the case with motion, which is the reason that the second of these papers also discusses the concept of "quantum friction". A critical evaluation of various aspects of calculations of the Casimir effect can be found in an article by C. R. Hagen, *Casimir energy for spherical boundaries*, hep-th/9902057.

[‖]T. H. Boyer, *Phys. Rev.* **174** (1968) 1764.

Example 20.2: The fine-structure constant

For a perfectly conducting sphere of radius a the attractive Casimir energy has been calculated[*] to be

$$E_C(a) = -\frac{1}{8\pi a}.$$

Assuming that the charge of the electron were uniformly distributed over the surface of a sphere of radius a, calculate the energy of this electron's repulsive Coulomb force as that of a spherical shell. Assuming that these energies cancel each other, calculate the value of the fine-structure constant $\alpha = e^2/4\pi\epsilon_0\hbar c$.

Solution: We obtain the Coulomb energy from the formula for the energy of charge distributions, i.e.

$$W = \frac{1}{2}\int dV \rho(\mathbf{r})\phi(\mathbf{r}), \quad \text{with} \quad \rho(\mathbf{r}) = e\delta(\mathbf{r}-\mathbf{a}), \quad \phi(\mathbf{r}) = \frac{e}{4\pi\epsilon_0 r},$$

so that

$$W = \frac{1}{2}\int dr e\delta(\mathbf{r}-\mathbf{a})\frac{e}{4\pi\epsilon_0 a}, \quad \text{i.e.} \quad W = \frac{e^2}{8\pi\epsilon_0 a}.$$

The corresponding force is

$$-\frac{\partial W}{\partial a} = +\frac{e^2}{8\pi\epsilon_0 a^2}, \quad \text{repulsive.}$$

We set

$$E_C(a) + W = 0, \quad \frac{1}{8\pi a} = \frac{e^2}{8\pi a\epsilon_0}, \quad \text{i.e.} \quad \frac{e^2}{\epsilon_0} = 1.$$

Then (with $\hbar = c = 1$)

$$\alpha = \frac{e^2}{4\pi\epsilon_0} = \frac{1}{4\pi} = \frac{1}{12.56}.$$

Hence the value obtained here for the fine-structure constant is too large by a factor of 10 (the correct value being $\sim 1/137$).

---

[*]K. A. Milton, *Ann. Phys.* (N.Y.) **127** (1980) 49.

This page intentionally left blank

Chapter 21

Duality and Magnetic Monopoles

21.1 Symmetrisation of the Maxwell Equations

A conspicuous and eye-catching aspect of the four Maxwell equations is a certain symmetry they exhibit between electric and magnetic fields. But it is also obvious that this symmetry is violated if one looks at charge and current densities. This symmetry is described as *duality*. In this chapter we consider attempts to gain some understanding of these observations.

In our treatment of the Aharonov–Bohm effect we also encountered singular fields. A deeper investigation into such fields leads to the topic of magnetic monopoles, i.e. to that of single magnetic poles. The full symmetrisation of the Maxwell equations also requires the introduction of these monopoles.

We first recapitulate the Maxwell equations with new notation for charge and current densities to which we affix an index "e" for "electric":

$$\boldsymbol{\nabla} \cdot \mathbf{D} = \rho_e, \quad \boldsymbol{\nabla} \cdot \mathbf{B} = 0,$$
$$\boldsymbol{\nabla} \times \mathbf{H} = \mathbf{j}_e + \frac{\partial \mathbf{D}}{\partial t}, \quad \boldsymbol{\nabla} \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}, \quad (21.1)$$

and the equation of continuity is

$$\frac{\partial \rho_e}{\partial t} + \boldsymbol{\nabla} \cdot \mathbf{j}_e = 0. \quad (21.2)$$

Thus, as far as these Maxwell equations are concerned, there are electric charges as sources of the electric field, but no magnetic charges or poles as sources of the magnetic field, which are here provided by electric currents.

479

The equation $\nabla \cdot \mathbf{B} = 0$ has no source term. This asymmetry of the equations is not appealing. It is therefore a natural step to introduce magnetic poles and to completely symmetrise the Maxwell equations and to explore the consequences which result from this. For symmetry reasons it is helpful in this case to use the formulation in terms of the MKSA-units as we do here, and not Gaussian units (but, of course, this is not essential). Thus we symmetrise the equations by adding magnetic charge and current densities with an index "m". The equations are then

$$\nabla \cdot \mathbf{D} = \rho_e, \quad \nabla \cdot \mathbf{B} = \rho_m \mu_0 c,$$
$$\nabla \times \mathbf{H} = \mathbf{j}_e + \frac{\partial \mathbf{D}}{\partial t}, \quad -\nabla \times \mathbf{E} = \mathbf{j}_m \mu_0 c + \frac{\partial \mathbf{B}}{\partial t}. \tag{21.3}$$

Here μ_0 is the magnetic permeability of the vacuum.

Previously we encountered the following quantities which were already symmetric:

(a) Energy density

$$u = \frac{1}{2}(\mathbf{E} \cdot \mathbf{D} + \mathbf{H} \cdot \mathbf{B}), \tag{21.4}$$

(b) Poynting vector

$$\mathbf{S} = \mathbf{E} \times \mathbf{H}, \tag{21.5}$$

(c) Maxwell stress tensor and field momentum density

$$T_{ij} = E_i D_j + H_i B_j - \frac{1}{2} u \delta_{ij},$$
$$\mathbf{p}^{\text{field}} = \mathbf{D} \times \mathbf{B}. \tag{21.6}$$

However, the Lorentz force $\mathbf{F}_e$ must be supplemented by a force $\mathbf{F}_m$ to yield the pair of symmetrised forces.

(d) Symmetrised forces

$$\mathbf{F}_e = e(\mathbf{E} + \mathbf{v} \times \mathbf{B}), \quad \mathbf{F}_m = g\mu_0 c(\mathbf{H} - \mathbf{v} \times \mathbf{D}) \tag{21.7}$$

with magnetic charge density ρ_m defined by

$$\rho_m = g\delta(\mathbf{r}). \tag{21.8}$$

This system of equations is now completely symmetric. The symmetry can be expressed in terms of a transformation called *duality transformation* of the fields and charge or pole densities, i.e. the relations (with $\epsilon_0 \mu_0 c^2 = 1$)

$$\mathbf{E} \rightarrow \mathbf{E}' = \mathbf{E} \cos \phi + c\mathbf{B} \sin \phi,$$
$$c\mathbf{B} \rightarrow c\mathbf{B}' = -\mathbf{E} \sin \phi + c\mathbf{B} \cos \phi \tag{21.9}$$

and

$$\rho_e \rightarrow \rho'_e = \rho_e \cos \phi + \rho_m \sin \phi,$$
$$\rho_m \rightarrow \rho'_m = -\rho_e \sin \phi + \rho_m \cos \phi. \tag{21.10}$$

Setting $\phi = 0$, one has $\rho'_e = \rho_e$, i.e. charges $q'_e = q_e \equiv e$ and the magnetic charges or pole strengths are correspondingly $q'_m = q_m = 0$. Setting $\phi = \pi/2$, one has $q'_e = q_m \equiv 0$ and $q'_m = -q_e = -e$. The question is: Is a certain angle ϕ for all particles a universal entity? Within the limits of present-day experiments $\phi = 0$ is the correct choice and in agreement with experiment. From this point of view the symmetry of classical electrodynamics is no argument for or against the existence of single magnetic poles. The question is not, do all particles in Nature have the same ratio $\rho_m/\rho_e = \tan \phi$ (in this case $\rho'_m = 0$ and we could choose units suitably, so that the usual Maxwell equations remain valid). The decisive question is really: Are there particles with different ratios q_m/q_e?

Example 21.1: Magnetic pole and electric charge at rest

Show that a system consisting of a magnetic pole g and a charge e, at rest at time $t = 0$, is always at rest.

Solution: We consider the expressions (21.7) for the forces acting on the pole and the charge. In $\mathbf{F}_e$ the vectors $\mathbf{E}$ and $\mathbf{B}$ are fields of the monopole, which act on the charge. Since the monopole is at rest the field $\mathbf{E}$, which would be generated by the monopole's motion (cf. Eq. (21.3)), is zero. Also since $\mathbf{v} = 0$, we have $\mathbf{v} \times \mathbf{B} = 0$. Hence $\mathbf{F}_e = 0$. Similarly we have $\mathbf{F}_m = 0$.

Example 21.2: Separating the charge-pole pair

The magnetic pole g of a dyon, i.e. charge-pole pair, with negligible binding energy is moved away from the charge e to a point $\mathbf{x}'$. Show that the mechanical momentum $\mathbf{p}^{\text{mech}}$ of the pole g is exactly cancelled by the field momentum. (Compare with Examples 7.2 and 17.5).

Solution: The Lorentz force acting on a pole g is given by Eq. (21.7), $\mathbf{F}_m = -g\mu_0 c \mathbf{v} \times \mathbf{D}$. Suppose we give the pole mechanical momentum $\mathbf{p}^{\text{mech}}$ by pulling it away from the charge with a force equal and opposite to that of the Lorentz force. Then Newton's equation of motion for the pole is

$$\frac{dp_i^{\text{mech}}}{dt} = g\mu_0 \epsilon_0 c^2 \, \epsilon_{ijk} \frac{dx_j}{dt} F_{0k}.$$

In moving the pole it generates field momentum which has to be such that the total momentum is conserved, i.e. here equal to that initially, i.e. zero. Thus

$$p_i^{\text{mech}} = g\epsilon_{ijk} \int_0^{x_j'} dx_j \, F_{0k}.$$

Since $\nabla \cdot \mathbf{B} = \mu_0 cg\delta(\mathbf{r})$, we have $\mu_0 cg = \int \mathbf{B} \cdot d\mathbf{a} = \int B_l da_l$, where da_l is an element of area with

normal in the direction l, and $dx_j da_l = \delta_{jl} dV$. Hence

$$
\begin{aligned}
p_i^{\text{mech}} &= \frac{1}{\mu_0 c} \epsilon_{ijk} \int dx_j \, F_{0k} B_l da_l \\
&= \frac{1}{\mu_0 c} \epsilon_{ijk} \int dV \, F_{0k} B_j \\
&= \epsilon_0 c \epsilon_{ijk} \int dV \, F_{0k} B_j \\
&= - \int dV (\mathbf{D} \times \mathbf{B})_i \\
&= -P_i^{\text{field}}.
\end{aligned}
$$

It follows that the total momentum vanishes,

$$
p_i^{\text{mech}} + P_i^{\text{field}} = 0,
$$

as in the original state of the system.

Example 21.3: Angular momentum of separated charge-pole pair at rest[*]

Consider a magnetic pole g at the origin $\mathbf{r} = 0$ and a charge q at $\mathbf{r} = r_0 \mathbf{e}_z$ at rest and the distance r_0 apart along the z-axis. Show that the system has a nonvanishing angular momentum only along the line drawn from the pole to the charge and determine this angular momentum.

Solution: The angular momentum $\mathbf{L}$ is given by

$$
\mathbf{L} = \int d\mathbf{r} \, \mathbf{l}, \quad \mathbf{l} = \mathbf{r} \times \mathbf{p}^{\text{field}}, \quad \mathbf{p}^{\text{field}} = \mathbf{D} \times \mathbf{B}.
$$

In the present case we have explicitly

$$
\mathbf{B} = \frac{\mu_0 c g \mathbf{r}}{4 \pi r^3}, \quad \mathbf{D} = \frac{q(\mathbf{r} - \mathbf{r}_0)}{4 \pi |\mathbf{r} - \mathbf{r}_0|^3}.
$$

The vector $\mathbf{D} \times \mathbf{B}$ is perpendicular to the z-axis which points from the pole to the charge at $z = r_0$, as shown in Fig. 21.1. At fixed angle θ between $\mathbf{r}$ and $\mathbf{r}_0 = r_0 \mathbf{e}_z$ but varying azimuthal angle φ in the (x, y)-plane the point $\mathbf{r}$ traces a circle with $\mathbf{D} \times \mathbf{B} \propto (\mathbf{r} - \mathbf{r}_0) \times \mathbf{r}$ along its tangential direction $\mathbf{e}_\varphi$. Thus the geometry is analogous to that of a circular current with the Biot–Savart field to be determined at the origin. The components perpendicular to the z-axis cancel each other, so that only a z-component remains, in the present case L_z. We obtain this as follows.

We have

$$
l_z = [\mathbf{r} \times (\mathbf{D} \times \mathbf{B})] \cdot \mathbf{e}_z = (\mathbf{r} \cdot \mathbf{B}) D_z - (\mathbf{r} \cdot \mathbf{D}) B_z.
$$

Inserting the explicit expression for $\mathbf{B}$ we obtain

$$
L_z = \int d\mathbf{r} \, l_z = \frac{\mu_0 c g}{4 \pi} \int d\mathbf{r} \frac{1}{r^3} [r^2 (\mathbf{D} \cdot \mathbf{e}_z) - (\mathbf{r} \cdot \mathbf{D})(\mathbf{r} \cdot \mathbf{e}_z)].
$$

With the help of the formula

$$
\nabla(\mathbf{u} \cdot \mathbf{v}) = (\mathbf{v} \cdot \nabla)\mathbf{u} + (\mathbf{u} \cdot \nabla)\mathbf{v} + \mathbf{v} \times (\nabla \times \mathbf{u}) + \mathbf{u} \times (\nabla \times \mathbf{v}),
$$

we have (setting $\mathbf{u} = \mathbf{r}/r$, $\mathbf{v} = \mathbf{e}_z$)

$$
\nabla\left(\frac{\mathbf{r} \cdot \mathbf{e}_z}{r}\right) = (\mathbf{e}_z \cdot \nabla)\left(\frac{\mathbf{r}}{r}\right) + \mathbf{e}_z \times \underbrace{\left(\nabla \times \frac{\mathbf{r}}{r}\right)}_{0}
$$

---

[*]See also Felsager [20], p. 487.

implying

$$\nabla\left(\frac{\mathbf{r}\cdot\mathbf{e}_z}{r}\right) = \frac{1}{r^3}[r^2\mathbf{e}_z - \mathbf{r}(\mathbf{r}\cdot\mathbf{e}_z)].$$
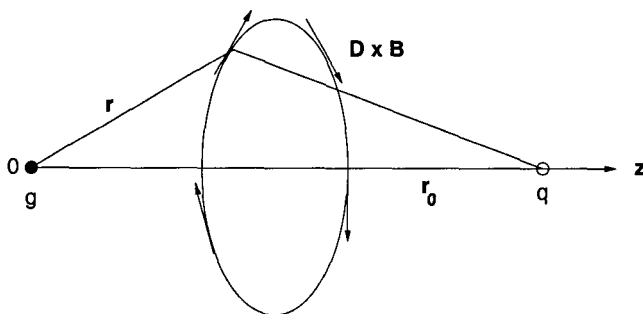
Fig. 21.1 Field momentum density $\mathbf{D}\times\mathbf{B}$ in the (x,y)-plane.

Hence

$$\mathbf{D}\cdot\nabla\left(\frac{\mathbf{r}\cdot\mathbf{e}_z}{r}\right) = \frac{1}{r^3}[r^2(\mathbf{D}\cdot\mathbf{e}_z) - (\mathbf{D}\cdot\mathbf{r})(\mathbf{r}\cdot\mathbf{e}_z)],$$

and so

$$L_z = \frac{\mu_0 cg}{4\pi}\int d\mathbf{r}\mathbf{D}\cdot\nabla\left(\frac{\mathbf{r}\cdot\mathbf{e}_z}{r}\right).$$

But

$$\mathbf{D}\cdot\nabla\left(\frac{\mathbf{r}\cdot\mathbf{e}_z}{r}\right) = \nabla\cdot\left(\frac{\mathbf{r}\cdot\mathbf{e}_z}{r}\mathbf{D}\right) - \frac{\mathbf{r}\cdot\mathbf{e}_z}{r}\nabla\cdot\mathbf{D}.$$

Ignoring again the total derivative with the usual reasoning we obtain therefore

$$L_z = -\frac{\mu_0 cg}{4\pi}\int d\mathbf{r}\frac{\mathbf{r}\cdot\mathbf{e}_z}{r}\nabla\cdot\mathbf{D}.$$

Here $\nabla\cdot\mathbf{D} = q\delta(\mathbf{r}-\mathbf{r}_0)$, so that

$$L_z = -\frac{\mu_0 cg}{4\pi}\int d\mathbf{r}\frac{\mathbf{r}\cdot\mathbf{e}_z}{r}q\delta(\mathbf{r}-\mathbf{r}_0)$$

$$= -\frac{\mu_0 cgq}{4\pi}.$$

This expression is seen to involve already the product $qg\mu_0/2\pi$, which will appear later in Dirac's charge quantisation condition. In fact, the latter follows from the present result with the half-integral quantisation condition $|L_z| = n\hbar/2$, n an integer.

21.2 Quantisation of Electric Charge

The question Dirac asked himself in 1931 was: Is the existence of magnetic "monopoles" compatible with the principles of quantum mechanics? His

answer was: Yes, provided the elementary electric charge e and the magnetic charge g satisfy the condition

$$\frac{eg}{\hbar} = \frac{2\pi n}{\mu_0 c}, \quad (n \text{ an integer}). \qquad (21.11)$$

The basic idea of this connection can be explained in the following elementary way (easiest possibility to make the quantisation condition (21.11) plausible). We let g be the magnetic charge located at the origin of coordinates which, similar to the Coulomb potential, generates at a point $\mathbf{r}$ the magnetic flux density

$$\mathbf{B}(\mathbf{r}) = \frac{g\mu_0 c \mathbf{r}}{4\pi r^3}. \qquad (21.12)$$

We now allow a particle with electric charge e — as illustrated in Fig. 21.2 — to travel along a straight line with shortest distance b from the origin and with velocity $\mathbf{v} = v\mathbf{e}_z$ past the pole charge g. On its flight this particle experiences the Lorentz force (pointing out of the paper in Fig. 21.2)

$$\mathbf{F} = F_y \mathbf{e}_y = evB_x \mathbf{e}_y = \frac{e\mu_0 cg}{4\pi} \frac{vb\mathbf{e}_y}{(b^2 + v^2 t^2)^{3/2}}. \qquad (21.13)$$

Thereby the particle receives an impulse with momentum transfer

$$\underbrace{\triangle P_y = \int_{-\infty}^{\infty} F_y dt}_{\frac{d}{dt}P_y = F_y} = \frac{egvb\mu_0 c}{4\pi} \underbrace{\int_{-\infty}^{\infty} \frac{dt}{(b^2 + v^2 t^2)^{3/2}}}_{2/vb^2} = \frac{eg\mu_0 c}{2\pi b}. \qquad (21.14)$$

This momentum transfer leads to a change in angular momentum L_z along the direction of flight given by

$$\triangle L_z = b\triangle P_y = \frac{eg\mu_0 c}{2\pi}. \qquad (21.15)$$

According to quantum mechanics orbital angular momentum is always quantised, so that

$$L_z = n\hbar, \quad n = 0, \pm 1, \pm 2, \pm 3, \ldots. \qquad (21.16)$$

It follows that

$$\frac{eg\mu_0 c}{2\pi} = n\hbar, \quad g = n\frac{2\pi\hbar}{e\mu_0 c}, \qquad (21.17)$$

i.e.

$$\frac{eg}{\hbar} = \frac{2\pi n}{\mu_0 c}. \qquad (21.18)$$

This implies that

$$\frac{e}{4\pi} = \frac{n}{2}\left(\frac{\hbar}{g\mu_0 c}\right) \qquad (21.19)$$

must be quantised, i.e. the electric charge can appear only in positive or negative integral multiples of the unit $4\pi\hbar/2g\mu_0$. The charge is therefore quantised as a consequence of the quantisation of angular momentum.
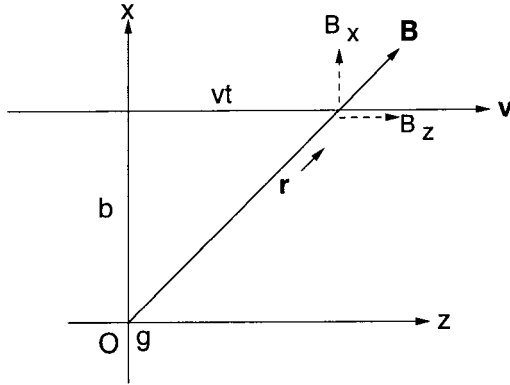
Fig. 21.2 Velocity **v** of the electric charge
in passing the magnetic pole at the origin O.

One should note: For this result to be physical, it suffices that a single magnetic pole exists somewhere in the universe, i.e. one we can use for the above argument.

A presumably more respectable derivation of the result (21.18) is based like Dirac's original derivation on the uniqueness of the Schrödinger wave function ψ, which we encountered already in Chapter 19. To be able to proceed to this, however, we first need to take a closer look at the field of the monopole.

The coupling constant of electromagnetic fields is given by the so-called *fine-structure constant*

$$\frac{e^2}{4\pi\hbar c\epsilon_0} = \frac{1}{137}.$$

With the above result the corresponding "magnetic fine-structure constant" is

$$\frac{\mu_0}{4\pi\hbar c}g^2 = \frac{\mu_0(\mu_0\epsilon_0 c^2)}{4\pi\hbar c}g^2 = \frac{\mu_0(\mu_0\epsilon_0 c^2)}{4\pi\hbar c}\left(\frac{2\pi\hbar n}{e\mu_0}\right)^2$$

$$= \frac{4\pi\hbar c\epsilon_0}{e^2}\left(\frac{n^2}{4}\right) = \frac{137n^2}{4} \overset{n=1}{=} 34.25,$$

for $n = 1$. Thus, instead of

$$\frac{e^2}{4\pi\hbar c\epsilon_0} = \frac{1}{137}$$

one has

$$\frac{g^2 \mu_0}{4\pi \hbar c} = 34.25$$

as coupling of the electromagnetic field to the magnetic pole, which is seen to have a much larger value. One can say, a unit magnetic charge exerts a force equivalent to that of an electrically charged nucleus with $Z = 137/4$ protons.

21.3   The Field of the Monopole

If we want to take magnetic poles in a Lagrangian formalism into account, we immediately face problems. If we insist, that the field strengths $F_{\mu\nu}$ are defined by gauge potentials, we can do this only at the price of admitting singularities in the gauge potentials, i.e. the *"Dirac strings"*. Consider once again the pure monopole field[†] (21.12), i.e.

$$\mathbf{B} = \frac{g\mu_0 c \mathbf{r}}{4\pi r^3}, \quad \mathbf{r} = (x, y, z), \quad r = \sqrt{x^2 + y^2 + z^2}.$$

This field is singular at the origin (like the Coulomb field). The field $\mathbf{B} = \nabla \times \mathbf{A}$ can be obtained from the vector potential (Cartesian components written as column elements)

$$\mathbf{A} = \frac{\mu_0 c g}{4\pi} \begin{pmatrix} y/r(r-z) \\ -x/r(r-z) \\ 0 \end{pmatrix} = -\frac{g\mu_0 c}{4\pi} \frac{(1 + \cos\theta)}{r \sin\theta} (\sin\varphi, \cos\varphi, 0) \quad (21.20)$$

(one expression in Cartesian coordinates, the other in spherical polar coordinates). For instance

$$B_x \propto \frac{\partial}{\partial y}(0) + \frac{\partial}{\partial z}\left(\frac{x}{r(r-z)}\right) = -\frac{x\frac{\partial r}{\partial z}}{r^2(r-z)} - \frac{x(\frac{\partial r}{\partial z} - 1)}{r(r-z)^2}$$

$$= -\frac{x\frac{z}{r}}{r^2(r-z)} - \frac{x(\frac{z}{r} - 1)}{r(r-z)^2} = -\frac{x\frac{z}{r} - x}{r^2(r-z)}$$

$$= \frac{x}{r^3}, \qquad \qquad \text{Q.E.D.}$$

We observe that $\mathbf{A}$ is singular at the origin $r = 0$ (like $\mathbf{B}$), but also along the positive z-axis, $z = r \geq 0$ (different from $\mathbf{B}$), or for $\theta = 0$. This second singularity from $r = 0$ along the positive z-axis is described as *Dirac*

[†]The topic we discuss here has also been treated in H. Fraas, "Magnetische Monopole", *Physik in unserer Zeit* **15** (1984) 173.

string (this is a singularity and is not to be confused with the strings of String Theory!). One may note: At $\theta = \pi$ we have

$$\sin \theta = \sin((\theta - \pi) + \pi) \simeq -(\theta - \pi),$$

$$\cos \theta = \cos((\theta - \pi) + \pi) \simeq -1 + \frac{1}{2}(\theta - \pi)^2,$$

so that at $\theta = \pi$ the factor in Eq. (21.20)

$$\frac{1 + \cos \theta}{\sin \theta} \propto (\theta - \pi) = \text{finite}.$$

Thus the vector potential is singular at the "*north pole*" where $\theta = 0$, but regular at the "*south pole*" where $\theta = \pi$. With a gauge transformation one can find another and different vector potential which is regular at the north pole, but singular at the south pole; in the overlap region around the equator the potentials are proportional to one another. In this way, i.e. by patching together different vector potentials, as in Sec. 21.4, one can avoid the Dirac string (singularity). The price one has to pay is this patching together of different potentials.[‡]

21.4 Uniqueness of the Wave Function

We now use two vector potentials, $\mathbf{A}_1, \mathbf{A}_2$, on a sphere $\mathbf{S}^2$ surrounding the magnetic charge g at the centre. We choose the vector potentials such that one can be continued smoothly into the other in the equatorial region. We choose the potential $\mathbf{A}_1$ to be regular on the northern hemisphere $\mathbf{S}^2_+$ (with Dirac string along the negative z-axis through the south pole), and $\mathbf{A}_2$ correspondingly regular on the southern hemisphere $\mathbf{S}^2_-$. We thus have for one and the same system two Schrödinger equations with different gauge potentials $\mathbf{A}_1, \mathbf{A}_2$ and ϕ_1, ϕ_2, and so with Hamilton operators

$$\hat{H}_1 = -\frac{\hbar^2}{2m}\left(\nabla - \frac{ie}{\hbar}\mathbf{A}_1\right)^2 + e\phi_1, \ \hat{H}_2 = -\frac{\hbar^2}{2m}\left(\nabla - \frac{ie}{\hbar}\mathbf{A}_2\right)^2 + e\phi_2. \quad (21.21)$$

The two gauge potentials are related by the gauge transformation

$$\mathbf{A}_1 \rightarrow \mathbf{A}_2 = \mathbf{A}_1 + \nabla \chi,$$

[‡]The well known reference for this is: T. T. Wu and C. N. Yang, *Phys. Rev.* **D12** (1975) 3845. The Dirac string is also discussed in the book of Felsager [20], p. 476. The same book, p. 486, also discusses the angular momentum as a consequence of the monopole field. See also E. Witten, *Nucl. Phys.* **B223** (1983) 422, pp. 424 to 425.

however, the function χ is then singular along the z-axis (i.e. along the Dirac strings) (otherwise the left side would possess the singularity but not the right side). The necessary associated transformation of the wave function ψ for the gauge covariance of the Schrödinger equation, $\hat{H}\psi = E\psi$, is (as we saw in Chapter 19),

$$\psi \to \psi' = \exp\left(\frac{ie\chi}{\hbar}\right)\psi, \quad \chi = \chi(\mathbf{r}, t),$$

i.e. in the present case

$$\psi_1 \to \psi_2 = \exp\left(\frac{ie\chi}{\hbar}\right)\psi_1.$$

The uniqueness of the wave function in the nonsingular domain requires therefore, that after one closed orbit about the string this wave function has the same value, i.e.

$$\frac{e}{\hbar}\oint_C d\chi = 2\pi n, \quad n = 1, 2, 3, \ldots. \tag{21.22}$$

The left hand side is really the discontinuity across the singularity of the function χ. We compute this as follows. We have

$$\begin{aligned}
\frac{2\pi n\hbar}{e} &= \oint_C d\chi = \oint_C \boldsymbol{\nabla}\chi \cdot d\mathbf{l} = \oint_C (\mathbf{A}_2 - \mathbf{A}_1) \cdot d\mathbf{l} \\
&= \int_{\mathbf{S}_+^2} (\boldsymbol{\nabla}\times\mathbf{A}_2) \cdot d\mathbf{F} - \int_{\mathbf{S}_-^2} (\boldsymbol{\nabla}\times\mathbf{A}_1) \cdot d\mathbf{F} \\
&= \int_{\mathbf{S}_+^2} \mathbf{B}\cdot d\mathbf{F} - \int_{\mathbf{S}_-^2} \mathbf{B}\cdot d\mathbf{F} = \int_{\mathbf{S}^2} \mathbf{B}\cdot d\mathbf{F} \\
&= \int_{\mathbf{S}^2} \frac{g\mu_0 c \, d\mathbf{F}}{4\pi r^2} = g\mu_0 c.
\end{aligned}$$

Hence

$$\frac{2\pi\hbar n}{e} = g\mu_0 c, \quad \frac{eg}{\hbar} = \frac{2\pi n}{\mu_0 c}.$$

We see here that the magnetic charge provides the discontinuity across the Dirac string singularity. The Dirac string itself, however, has only a virtual, not a real significance.

There is another, more complicated argument, which leads to the same result.[§] This argument is closely related to the arguments involved in the

---

[§]A detailed treatment can be found in the book of Felsager [20], p. 501.

explanation of the Aharonov–Bohm effect. There we saw that when the current in the solenoid is switched on, the interference pattern observed on the screen undergoes a displacement depending on the magnetic flux through the solenoid. The interference pattern results from interference of electron waves along the paths Γ_1, Γ_2 passing through the two slits, i.e. from the phase factors

$$\exp\left(\frac{ie}{\hbar}\int_{\Gamma_2} \mathbf{A}\cdot d\mathbf{l}\right)\exp\left(-\frac{ie}{\hbar}\int_{\Gamma_1} \mathbf{A}\cdot d\mathbf{l}\right) = \exp\left(\frac{ie}{\hbar}\oint_{\Gamma} \mathbf{A}\cdot d\mathbf{l}\right). \quad (21.23)$$

Here

$$\oint_{\Gamma} \mathbf{A}\cdot d\mathbf{l} = \int \mathbf{B}\cdot d\mathbf{F}$$

is the magnetic flux through the solenoid. Thus the entire system consists of the charged particle (electron) and the magnetic field. Consider again the case of an electron and a magnetic pole. These generate the electromagnetic field $F_{\mu\nu}$, which, as we saw, has to be singular and cannot be derived from a single global vector potential. However, one can imagine an additional tensor field $S_{\mu\nu}$ and can construct it such that the sum of both, $F + S$ with $\int_{\mathbf{S}^2} F = g\mu_0 c, \int_{\mathbf{S}^2} S = -g\mu_0$ is generated by a still singular $\mathbf{A}$. We take $\int_{\mathbf{S}^2} S$ to be the magnetic flux through $\mathbf{S}^2$ due to the Dirac string. Then we have

$$\oint_{\Gamma} \mathbf{A}\cdot d\mathbf{l} = \int_{\mathbf{S}^2} F + \int_{\mathbf{S}^2} S.$$

This separation is now inserted into the phase factor (21.23). However, the string-part S may not contribute to this, since otherwise the interference would depend on its position in space. It follows therefore that we must have

$$\exp\left(\frac{ie}{\hbar}\int_{\mathbf{S}^2} S\right) = 1,$$

and this implies

$$\frac{eg}{\hbar} = \frac{2\pi n}{\mu_0 c}.$$

Finally we add yet another derivation of the Dirac charge quantisation condition which, in view of its brevity or even the associated geometry of Fig. 21.3, may be described as a *"nutshell proof"*.[¶] We saw in Chapter 19 that for a particle with charge e moving in a magnetic field (here that of a pole g), the Lorentz force appears in the quantum mechanical wave function as the phase factor

$$\exp\left(\frac{ie}{\hbar}\oint_{\Gamma} \mathbf{A}\cdot d\mathbf{l}\right).$$

[¶]J. Polchinski, *String Theory*, Vol. II (Cambridge University Press, 1998), p. 148.
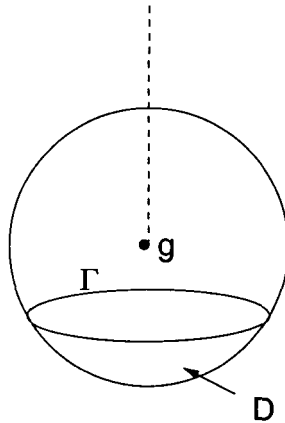
Fig. 21.3 The Dirac string emanating from the magnetic pole g, and
the orbit Γ as boundary of area D.

We also saw that the vector potential cannot be gauged away in view of
its Dirac string singularity extending, as shown in Fig. 21.3, from the pole
through the north pole to infinity. However, we can proceed as follows. The
closed path Γ is the boundary of the part D of the surface of a sphere around
the south pole. Pulling the path Γ across the equator and contracting it to
a small circle around the Dirac string through the north pole, the area of D
becomes that of the sphere and we can write the phase factor

$$\exp\left(\frac{ie}{\hbar}\oint_{\mathbf{S}^2}\mathbf{B}\cdot d\mathbf{F}\right) = \exp\left(\frac{ie}{\hbar}g\mu_0 c\right).$$

Since there may not be any trace of the Dirac string, this factor must be 1
giving again

$$\frac{e}{\hbar}g\mu_0 c = 2\pi n, \quad n \text{ integral.}$$

21.5   Regularisation of the Monopole Field

Finally we have a closer look at the vector potential (21.20) in order to gain
some physical understanding of the Dirac string, i.e. what we would have to
do in order to remove its singular behaviour. We shall see that a thin solenoid
along the position of the string will achieve this. To this end we regularise
the vector potential (21.20), by replacing in it r by $\sqrt{r^2 + \epsilon^2}$, where ϵ is
small. We rename the correspondingly modified fields $\mathbf{A}, \mathbf{B}, \mathbf{E}$ respectively
$\mathbf{A}_\epsilon, \mathbf{B}_\epsilon, \mathbf{E}_\epsilon$. With this modification the field $\mathbf{A}_\epsilon$ is well defined in the whole
of space. Since this regularisation does not affect the time dependence, we

have $\mathbf{E}_\epsilon = 0$. We maintain, of course, $\mathbf{B}_\epsilon = \nabla \times \mathbf{A}_\epsilon$ and see, that changes ensue. First we replace in $\mathbf{A}$ of Eq. (21.20) r by $\sqrt{r^2 + \epsilon^2}$ and expand the expressions up to the leading contributions in powers of ϵ^2. With this we obtain

$$\mathbf{A}_\epsilon = \frac{\mu_0 cg}{4\pi} \begin{pmatrix} \frac{y}{r(r-z)}\left\{1 - \frac{\epsilon^2(2r-z)}{2r^2(r-z)}\right\} \\ -\frac{x}{r(r-z)}\left\{1 - \frac{\epsilon^2(2r-z)}{2r^2(r-z)}\right\} \\ 0 \end{pmatrix}. \tag{21.24}$$

These expressions *seem* to be singular at $r = z$. But consider once again the starting expressions, by taking for instance, the x-component of $\mathbf{A}_\epsilon$:

$$A_{\epsilon x} \propto \frac{y}{\sqrt{r^2 + \epsilon^2}(\sqrt{r^2 + \epsilon^2} - z)}.$$

This expression is regular at $r = z$. We avoid these problems in the following by replacing the factor $(r - z)$ where it seems to be problematic, by $(\sqrt{r^2 + \epsilon^2} - z)$. With this we calculate the components of $\mathbf{B}_\epsilon = \nabla \times \mathbf{A}_\epsilon$ and obtain

$$B_{\epsilon x} = \frac{\mu_0 cg}{4\pi} \frac{x}{r^3}\left(1 - \frac{3\epsilon^2}{2r^2}\right),$$

$$B_{\epsilon y} = \frac{\mu_0 cg}{4\pi} \frac{y}{r^3}\left(1 - \frac{3\epsilon^2}{2r^2}\right),$$

$$B_{\epsilon z} = \frac{\mu_0 cg}{4\pi} \frac{z}{r^3}\left(1 - \frac{3\epsilon^2}{2r^2}\right) - \frac{\mu_0 g}{4\pi} \frac{\epsilon^2(2r-z)}{r^3(r-z)^2}. \tag{21.25}$$

We now see that the z-component displays an additional contribution of the order of ϵ^2, i.e. a B-field along the string. We have a closer look at this expression, i.e. at

$$w_\epsilon := \frac{\epsilon^2(2r-z)}{r^3(\sqrt{r^2 + \epsilon^2} - z)^2},$$

in which we replaced $(r - z)$ as explained above. We choose a point $z_0 > 0$ along the z-axis and explore the behaviour of w_ϵ in its neighbourhood, i.e. around $r \sim z_0$. First we have with $\rho^2 = x^2 + y^2$

$$r - z_0 \rightarrow \sqrt{r^2 + \epsilon^2} - z_0$$
$$= \sqrt{\rho^2 + z_0^2 + \epsilon^2} - z_0$$
$$\simeq z_0\left(1 + \frac{\rho^2 + \epsilon^2}{2z_0^2}\right) - z_0$$
$$= \frac{\rho^2 + \epsilon^2}{2z_0}.$$
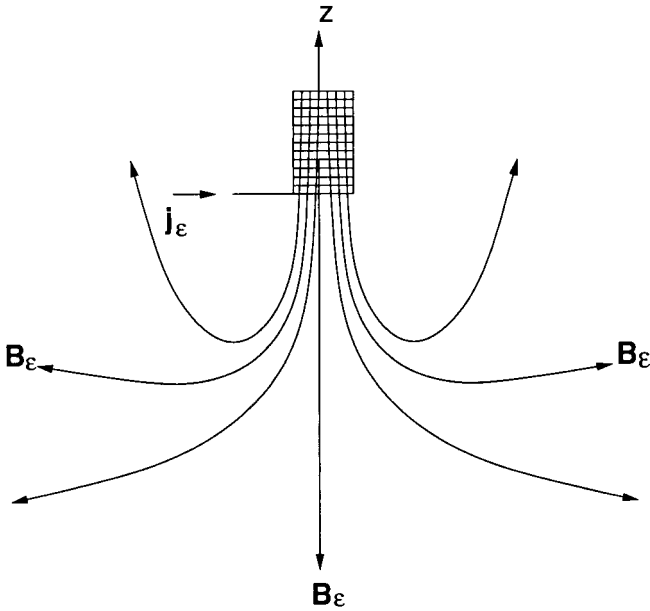
Fig. 21.4 The regularised field **B** of a monopole.

Also

$$(2r - z_0) = (r - z_0) + r \to \frac{\rho^2 + \epsilon^2}{2z_0} + r.$$

With these expressions we obtain

$$w_\epsilon \to \frac{\epsilon^2 \left\{ \frac{\rho^2 + \epsilon^2}{2z_0} + r \right\} (2z_0)^2}{r^3 (\rho^2 + \epsilon^2)^2} \overset{r \sim z_0}{=} \frac{4\epsilon^2 \left\{ 1 + \frac{\rho^2 + \epsilon^2}{2z_0^2} \right\}}{(\rho^2 + \epsilon^2)^2}$$

$$= \frac{4\epsilon^2 (2z_0^2 + \rho^2)}{2z_0^2 (\rho^2 + \epsilon^2)^2} \simeq \frac{4\epsilon^2}{(\rho^2 + \epsilon^2)^2}.$$

and hence with Eq. (21.25)

$$B_{\epsilon z} = \frac{\mu_0 c g}{4\pi} \frac{z}{r^3} - \mu_0 c g \frac{\epsilon^2}{\pi (\rho^2 + \epsilon^2)^2}. \tag{21.26}$$

We now use a 2-dimensional representation of the delta function which we found in Example 2.6, i.e.

$$\lim_{\epsilon \to 0} \frac{\epsilon^2}{\pi (\epsilon^2 + \rho^2)^2} = \delta(x)\delta(y),$$

and obtain

$$\lim_{\epsilon \to 0} B_{\epsilon z} = \frac{\mu_0 c g}{4\pi} \frac{z}{r^3} - \mu_0 c g \delta(x)\delta(y)\theta(z). \tag{21.27}$$

Here we have multiplied the second contribution by the step function since the string runs along the positive z-axis. The magnetic field concentrated in the Dirac string can therefore be written

$$\mathbf{B}_{\text{string}} = -\mu_0 c g \delta(x)\delta(y)\theta(z)\mathbf{e}_z. \tag{21.28}$$

The divergence of this expression is

$$\boldsymbol{\nabla} \cdot \mathbf{B}_{\text{string}} = -\mu_0 c g \delta(x)\delta(y)\delta(z). \tag{21.29}$$

On the other hand, the divergence of the monopole part of the field $\mathbf{B}$, i.e. the first part on the right of (21.27) and the corresponding components contained in Eq. (21.25), is

$$\boldsymbol{\nabla} \cdot \mathbf{B}_{\text{monopole}} = \mu_0 c g \delta(x)\delta(y)\delta(z). \tag{21.30}$$

We see that the sum of the two sources cancels exactly as a consequence of our regularisation of the originally singular vector potential.

Finally we calculate the curl of the string field. We obtain evidently

$$\boldsymbol{\nabla} \times \mathbf{B}_{\text{string}} = -\mu_0 c g \begin{pmatrix} \delta(x)\delta'(y) \\ -\delta'(x)\delta(y) \\ 0 \end{pmatrix} \theta(z) \equiv \mu_0 \mathbf{j}_{\text{string}} \tag{21.31}$$

with a singular current density. We can visualise the current associated with this current density in the (x,y)-plane as a current in a very narrow, semi-infinite solenoid which supplies the magnetic field concentrated along the z-axis. This situation together with the appropriate field $\mathbf{B}$ of (21.25) are illustrated in Fig. 21.4. The regularisation we explored above is therefore equivalent to introducing this solenoidal flux along the z-axis, so that the effect of the original singularity is cancelled.

21.6   Concluding Remarks

In the above we investigated a few important elementary aspects related to the introduction of classical magnetic monopoles, and other consequences of the symmetrisation of Maxwell's equations. The experimental search for evidence of magnetic monopoles received considerable impetus in 1982 with the possibility that an appropriate event had been observed, at least it was not easy to propose some other explanation for the observations of Cabrera.[∥] A review of the experimental situation thereafter, which has hardly changed

---

[∥]B. Cabrera, *Phys. Rev. Lett.* **48** (1982) 1378.

since then, and which also recapitulates the historical development and cites extensive literature can be found in the paper of Fryberger.[**] The principle which is used in experimental attempts to observe a monopole, is similar to that discussed in Sec. 17.6, i.e. that a monopole passing an atom at some distance generates a pulse of magnetic flux which can be observed. A complicated MACRO-detector search for magnetic monopoles was carried out throughout the 1990s. It found none but set stringent limits. The final results have recently been published.[††] Monopoles which play an important role in field theories as topologically nontrivial field configurations are of a completely different nature. However, seen from far away, they assume the structure of the Dirac monopole discussed above.[‡‡] The question of whether magnetic monopoles really exist is therefore the remaining bigg puzzle of electrodynamics.

[**]D. Fryberger, *Magnetic Monopoles*, Stanford Linear Accelerator Center Report, SLAC-PUB-3535 (1984), Invited paper at 1984 Applied Superconductivity Conf., San Diego, Calif., unpublished.

[††]M. Ambrosio *et al.* *Eur. Phys. J.* **C25** (2002) 511 and **C26** (2002) 163; see also CERN *Courier* **43** (2003), No. 4, 23.

[‡‡]See, for instance, P. Goddard and D. Olive, *Rep. Progr. Phys.* **41** (1978) 1357.

Appendix A

The Delta Distribution

The delta-distribution* or delta function, as it was originally called, is a singular function beyond the realm of classical analysis. As we observed in the text, in physical contexts the delta distribution arises in the consideration of point charges, and similarly in that of mass points. The density of a unit charge or unit mass at (say) $x = 0$ which is written $\delta(x)$ is everywhere zero except at $x = 0$, where it is so large that the total charge, i.e. its integral over all space, becomes 1, i.e.

$$\delta(x) = 0 \ \text{ for } \ x \neq 0, \quad \int_{-\infty}^{\infty} \delta(x)dx = 1. \tag{A.1}$$

No function of classical analysis has such properties since for any function which is everywhere zero except at one point, the integral must vanish (irrespective of the concept of the integral). As a further example, which leads to a singular function, we consider the case of two charges of opposite signs with intensities $\pm 1/\epsilon$ and located at the points $x = 0, x = \epsilon$. The density distributions of the charges are $\delta(x)/\epsilon, -\delta(x - \epsilon)/\epsilon$. In the limit $\epsilon \to 0$ the charges approach each other, with the product of intensity $1/\epsilon$ and mutual separation ϵ, i.e. the dipole moment, remaining constant. One thus obtains a dipole with density

$$\delta'(x) = \lim_{\epsilon \to 0} \left(\frac{\delta(x) - \delta(x - \epsilon)}{\epsilon} \right). \tag{A.2}$$

This limiting value is undefined and does not exist in the context of classical analysis. However, in the theory of distributions developed by L. Schwartz, the delta function, its derivative $\delta'(x)$ and similar quantities, find exact definitions which, moreover, permit application of customary operations of classical analysis. Physical considerations point the way to how best to proceed

---

*We follow here W. Güttinger, *Fortschr. Physik* **14** (1966) 483.

in developing such a theory. In order to localise an object with a certain
density distribution, one observes its reaction to some action of a test object.
If there is no reaction, one concludes that the object was not present in the
tested domain. We can characterise the density distribution of the object
by $f(x)$, and that of the test object by $\varphi(x)$. Then the product $f(x)\varphi(x)$
represents the result of the test for the spatial point x, because $f(x)\varphi(x) =$
0 if not simultaneously $f(x) \neq 0, \varphi \neq 0$, i.e. if there is no encounter of
the object with the test object. The integral $\int f(x)\varphi(x)dx$ then describes
the result of the testing or scanning procedure over the entire space. If we
perform this procedure with different test objects, i.e. with different test
functions $\varphi_i(x), i = 1, 2, \ldots$, we obtain as the result over the whole of space
different numbers corresponding to the individual φ_i. These φ-dependent
numbers are written

$$f\langle\varphi\rangle = \int f(x)\varphi(x)dx. \tag{A.3}$$

If $f\langle\varphi\rangle = 0$ for all continuous differentiable functions $\varphi(x)$, then $f(x) = 0$.
In general one expects that a knowledge of the numbers $f\langle\varphi\rangle$ and the test
functions $\varphi(x)$ characterises the function $f(x)$ itself provided the set of test
functions is complete. In this way one arrives at a new concept of functions:
Instead by its values $y = f(x)$, the function f is now determined by its
action on all test functions $\varphi(x)$. One calls f a *functional*: The functional f
assigns every test function φ a number $f\langle\varphi\rangle$. The functional is therefore the
mapping of a space of functions onto a space of numbers. $f\langle\varphi\rangle$ is the value
of the functional at the "point" φ. The concept of a functional permits us
to define objects which are not functions in the sense of classical analysis.
Consider e.g. the delta function. This is now defined as the functional $\delta\langle\varphi\rangle$
which assigns every test function $\varphi(x)$ a number, in this case the value of
the test function at $x = 0$, i.e.

$$\delta\langle\varphi\rangle = \varphi(0), \tag{A.4}$$

where with (A.3)

$$\delta\langle\varphi\rangle = \int \delta(x)\varphi(x)dx. \tag{A.5}$$

The result of the action of the delta function on the test function $\varphi(x)$
is the number $\varphi(0)$. Writing this $\int \delta(x)\varphi(x)dx$ is to be understood only
symbolically. The example of the delta function demonstrates that a function
does not have to be given in order to be able to define a functional.

    To ensure that in the transition from a function $f(x)$ defined in the clas-
sical sense to its corresponding functional $f\langle\varphi\rangle$ no information about f is
lost, i.e. so that $f\langle\varphi\rangle$ is equivalent to $f(x)$, the class of test functions has

to be sufficiently large. This means: If the integral $\int f(x)\varphi(x)dx$ is to exist even for a function $f(x)$, which increases with x beyond all bounds, then the test functions must decrease to zero sufficiently rapidly for large values of x (how, depends of course on the given physical situation). However, it is clear that the space of test functions must contain those functions φ, which vanish beyond a bounded and closed domain, because these correspond to the possibility, to measure mass distributions that are restricted to a finite domain. The functions φ must also have a sufficiently regular behaviour so that the integral of Eq. (A.3) exists. One demands continuous differentiability of any high order. Thus we define D as the space of all infinitely often differentiable functions $\varphi(x)$, which like all their derivatives $\varphi^{(n)}(x), n = 0, 1, 2, \ldots$, vanish at $x = \pm\infty$. The functionals $f\langle\varphi\rangle$ defined on D are called *distributions*.

We expect also that certain continuity properties of the function $f(x)$ are reflected in the associated functional. The reaction of a mass distribution $f(x)$ to a test object $\varphi(x)$ is the weaker, the weaker $\varphi(x)$ is. It makes sense, therefore, that if a sequence of test functions $\varphi_\nu(x)$ as well as the sequences of their derivatives of arbitrary order $\varphi_\nu^{(n)}(x)$ converge uniformly towards zero, then also the sequence of numbers $f\langle\varphi_\nu\rangle$ converges towards zero.

The linearity

$$f\langle\varphi_1 + \varphi_2\rangle = f\langle\varphi_1\rangle + f\langle\varphi_2\rangle \tag{A.6}$$

follows from the superposition principle. The derivative of a distribution $f\langle\varphi\rangle$ is defined by the following equation:

$$f'\langle\varphi\rangle = -f\langle\varphi'\rangle. \tag{A.7}$$

This definition is natural. Because if we associate with the function $f(x)$ the distribution

$$f\langle\varphi\rangle = \int_{-\infty}^{\infty} dx\, f(x)\varphi(x),$$

then the derivative $f'(x)$ is assigned the functional

$$f'\langle\varphi\rangle = \int_{-\infty}^{\infty} dx\, f'(x)\varphi(x).$$

Partial integration of this integral yields, since $\varphi(\pm\infty) = 0$,

$$f'\langle\varphi\rangle = [f(x)\varphi(x)]\Big|_{-\infty}^{\infty} - \int_{-\infty}^{\infty} dx\, f(x)\varphi'(x) = -f\langle\varphi'\rangle$$

as in Eq. (A.7). Equation (A.7) defines the derivative of the functional $f\langle\varphi\rangle$, even if no function $f(x)$ exists, which defines the functional. For instance, we have in the case of the delta function with (A.4)

$$\delta'\langle\varphi\rangle = -\delta\langle\varphi'\rangle = -\varphi'(0). \tag{A.8}$$

Formally one writes

$$\int_{-\infty}^{\infty} dx \delta'(x)\varphi(x) = [\delta(x)\varphi(x)]\Big|_{-\infty}^{\infty} - \int_{-\infty}^{\infty} dx \delta(x)\varphi'(x)$$
$$= -\delta\langle\varphi'\rangle = -\varphi'(0). \tag{A.9}$$

For an infinitely often differentiable function $g(x)$ we have evidently

$$(g \cdot f)\langle\varphi\rangle = f\langle g\varphi\rangle, \tag{A.10}$$

so that

$$(x\delta)\langle\varphi\rangle = \delta\langle x\varphi\rangle = [x\varphi]_{x=0} = 0 \cdot \varphi(0) = 0,$$

or

$$\int x\delta(x)\varphi(x)dx = 0, \quad x\delta(x) = 0. \tag{A.11}$$

Thus formally one can operate with the delta function like with a function of classical analysis. As another example we consider the relation

$$f(x)\delta(x) = f(0)\delta(x). \tag{A.12}$$

According to Eq. (A.10) we have:

$$\int f(x)\delta(x)\varphi(x)dx = \delta\langle f\varphi\rangle = [f(x)\varphi(x)]_{x=0}$$
$$= f(0)\varphi(0) = f(0)\delta\langle\varphi\rangle$$
$$= \int f(0)\delta(x)\varphi(x)dx,$$

as claimed with (A.12). Formal differentiation of the relation (A.12) gives

$$f(x)\delta'(x) = f(0)\delta'(x) - f'(x)\delta(x).$$

It is easy to convince oneself that this formal relation also follows from the defining equation (A.3). In the special case $f(x) = x$ we obtain the useful result

$$x\delta'(x) = -\delta(x). \tag{A.13}$$

A function which is important in applications is the *Heaviside* or *step function*, which is defined as follows:

$$\theta(x) = \begin{cases} 1 & \text{for } x > 0, \\ 0 & \text{for } x < 0. \end{cases} \tag{A.14}$$

From (A.14) we derive the important relation

$$\theta'(x) = \delta(x).\tag{A.15}$$

In order to prove this relation, we associate with the step function the functional

$$\theta(x)\langle\varphi\rangle = \int_{-\infty}^{\infty} \theta(x)\varphi(x)dx = \int_{0}^{\infty} \varphi(x)dx.$$

For the derivative we have according to (A.7):

$$\begin{aligned}
\theta'(x)\langle\varphi\rangle &= -\theta(x)\langle\varphi'\rangle \\
&= -\int_{-\infty}^{\infty} dx\,\theta(x)\varphi'(x) \\
&= -\int_{0}^{\infty} dx\,\varphi'(x) = \varphi(0) - \varphi(\infty) \\
&= \varphi(0) \\
&= \delta(x)\langle\varphi\rangle
\end{aligned}$$

or symbolically

$$\theta'(x) = \delta(x),\tag{A.16}$$

which is Eq. (A.15).

Fourier integrals were formally introduced in Chapter 3. We abstain here from delving into rigorous proofs. We recall that the function represented by the Fourier series is piece-wise smooth. The Fourier integrals of Chapter 3 therefore supply us an integral representation of such a piece-wise smooth function. The Fourier theorem says: If $f(x)$ is a piece-wise continuous function and if the integral $\int_{-\infty}^{\infty} |f(x)|dx$ exists, and if $f(x)$ is defined at a discontinuity x_0 by the relation

$$f(x_0) = \frac{1}{2}[f(x_0 - 0) + f(x_0 + 0)],$$

then the following representation of the function $f(x)$ exists:

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} dk \int dy\, f(y)e^{ik(y-x)}.\tag{A.17}$$

In order to verify (A.17), we consider first the function

$$\delta_\kappa(x) = \frac{1}{2\pi} \int_{-\kappa}^{\kappa} dk\, e^{ikx} = \frac{1}{\pi}\frac{\sin\kappa x}{x}.\tag{A.18}$$

The behaviour of this function for $\kappa > 0$ but not too large is illustrated in Fig. A.1. We observe that as the central maximum κ/π grows with $\kappa \to \infty$, the zeros at $n\pi/\kappa$ converge toward the origin and we have

$$\delta(x) = \lim_{\kappa \to \infty} \delta_\kappa(x), \qquad (A.19)$$
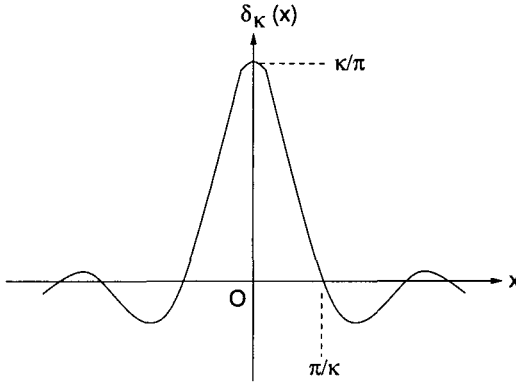
Fig. A.1 The function $\delta_\kappa(x)$.

because $\delta(x)$ is defined by the properties $\delta(x) = 0$ for $x \neq 0$, $\delta(x) = \infty$ for $x = 0$. From (A.18) and (A.19) we obtain the important integral representation of the delta function

$$\delta(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} dk\, e^{ikx}. \qquad (A.20)$$

From (A.18) and (A.19) we also obtain

$$\delta(x) = \delta(-x). \qquad (A.21)$$

The representation (A.20) also proves the relation (A.17) claimed by Fourier's theorem. Because if we perform the integration with respect to k on the right hand side, we obtain

$$\int dy\, f(y)\delta(y - x) = f(x),$$

as claimed.

In the following example we use yet another representation of the delta function which is very similar to the representation (A.19) (effectively the square of the latter). However, in this example our interest is focussed on the explicit use of a test function and the consequent necessity to evaluate the integral defining the distribution.

Example A.1: Explicit application of test functions

The delta function can be represented by the limit

$$\delta(x) = \lim_{t \to \infty} \delta_t(x), \tag{A.22a}$$

where

$$\delta_t(x) := \frac{1}{\pi} \frac{\sin^2 xt}{x^2 t} = \begin{cases} \dfrac{t}{\pi} & \text{for } x \to 0, \\[2mm] \leq \dfrac{1}{\pi x^2 t} & \text{for } x \neq 0. \end{cases} \tag{A.22b}$$

Use the method of test functions to verify the above result in the sense of distribution theory.

Solution: For test functions $\phi(x)$ the delta distribution is defined by the functional

$$\int \delta(x)\phi(x)dx = \phi(0).$$

In the case of the representation given by Eqs. (A.22a) and (A.22b), this is the relation

$$\lim_{t \to \infty} \frac{1}{\pi} \int_{-\infty}^{\infty} dx \frac{\sin^2 xt}{x^2 t} \phi(x) = \phi(0).$$

Thus we have to verify that the right hand side of this equation is obtained with a suitable choice of test functions $\phi(x)$. We choose the test function

$$\phi(x) = e^{-|x|}, \quad \phi(0) = 1.$$

This is a test function $\phi(x) \in S(\mathbb{R})$, which is the set of those test functions which together with all of their derivatives fall off at infinity faster than any inverse power of $|x|$ (the symbol S being derived from the name of Schwartz referred to above). We therefore have to evaluate the integral

$$I = \int_0^{\infty} dx \frac{\sin^2 xt}{x^2} e^{-x}.$$

We do this with the help of Tables of Integrals.[†] We have

$$\int_0^{\infty} e^{-|p|x} \sin ax \sin bx \frac{dx}{x^2} = \frac{p}{4} \ln \frac{p^2 + (a-b)^2}{p^2 + (a+b)^2} + \frac{a}{2} \tan^{-1} \frac{2pb}{p^2 + a^2 - b^2}$$
$$+ \frac{b}{2} \tan^{-1} \frac{2pa}{p^2 + b^2 - a^2},$$

so that for $a = b = t, p = 1$

$$I = -\frac{1}{4} \ln(1 + 4t^2) + t \tan^{-1}(2t).$$

Then

$$\lim_{t \to \infty} \frac{1}{\pi} \int_{-\infty}^{\infty} dx \frac{\sin^2 xt}{x^2 t} e^{-|x|} = 2 \lim_{t \to \infty} \left[-\frac{1}{4t\pi} \ln(1 + 4t^2) + \frac{t}{t\pi} \tan^{-1}(2t) \right]$$

$$= \frac{2}{\pi} \tan^{-1}(\infty) = \frac{2}{\pi} \frac{\pi}{2} = 1.$$

Hence the result is $\phi(0) = 1$ as expected.

[†]See I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series and Products* (Academic Press, 1965), formula 3.947, p. 491.

Finally we consider the following two examples. The first of these deals with the Fourier transform, but contains in addition a further useful representation of the delta function. In the second example we consider the first of the relations (2.9), which remains to be verified.

Example A.2: Determine the Fourier transformation of e^{-ax^2} for $a > 0$

The function required is

$$g(k) = \int_{-\infty}^{\infty} dx\, e^{-ax^2} e^{ikx}. \tag{A.23}$$

Solution: In general one uses for the evaluation of such an integral the method of contour integration. In the present case, however, a simpler method suffices. Differentiation of (A.23) with respect to k yields the same as

$$g'(k) = -\frac{i}{2a} \int_{-\infty}^{\infty} dx\, \frac{d}{dx}(e^{-ax^2}) e^{ikx}.$$

Partial integration of the right hand side implies

$$g'(k) = \frac{i}{2a} \int_{-\infty}^{\infty} ike^{ikx} e^{-ax^2}\, dx = -\frac{k}{2a} g(k).$$

Thus $g(k)$ satisfies the differential equation

$$g'(k) + \frac{k}{2a} g(k) = 0.$$

Simple integration yields $g(k) = Ce^{-k^2/4a}$ with the constant $C = g(0)$. Since

$$g(0) = \int_{-\infty}^{\infty} dx\, e^{-ax^2} = \sqrt{\frac{\pi}{a}},$$

we obtain

$$g(k) = \sqrt{\frac{\pi}{a}} e^{-k^2/4a}. \tag{A.24}$$

From (A.23) and (A.24) we obtain a representation of the delta function:

$$\begin{aligned}
\delta(k) &= \lim_{a \to 0} \frac{1}{2\pi} \int_{-\infty}^{\infty} dx\, e^{-ax^2} e^{ikx} \\
&= \lim_{a \to 0} \frac{1}{2\sqrt{\pi a}} e^{-k^2/4a} \\
&= \lim_{\epsilon \to 0} \frac{e^{-k^2/\epsilon^2}}{\epsilon\sqrt{\pi}}. \tag{A.25}
\end{aligned}$$

Example A.3: Another important property of the delta function

Verify the relation

$$\delta[(x - a)(x - b)] = \frac{1}{|a - b|}[\delta(x - a) + \delta(x - b)] \tag{A.26}$$

for $a \neq b$. The formula can also be written

$$\delta[g(x)] = \sum_n \frac{1}{|g'(x_n)|} \delta(x - x_n),$$

where $g(x_n) = 0$, $g'(x_n) \neq 0$.

Solution: We start with the integral (for $\epsilon > 0$)

$$\frac{1}{2\pi} \int_{-\infty}^{\infty} dk\, e^{-\epsilon|k|} e^{ikx} = \frac{1}{\pi} \int_{0}^{\infty} dk\, e^{-\epsilon k} \cos kx = \frac{1}{\pi} \frac{\epsilon}{\epsilon^2 + x^2}.$$

From this we obtain as representation of the delta function:

$$\delta(x) = \lim_{\epsilon \to 0} \frac{1}{2\pi} \int_{-\infty}^{\infty} dk\, e^{-\epsilon|k|} e^{ikx} = \lim_{\epsilon \to 0} \frac{1}{\pi} \frac{\epsilon}{\epsilon^2 + x^2}. \tag{A.27}$$

With the help of this representation of the delta function one can show by partial fraction decomposition that the above relation (A.26) holds for $a \neq b$. Thus consider the sum

$$\frac{\epsilon}{(x-a)^2 + \epsilon^2} + \frac{\epsilon}{(x-b)^2 + \epsilon^2}$$

$$= \frac{1}{2i} \left[\frac{1}{x-a-i\epsilon} - \frac{1}{x-a+i\epsilon} + \frac{1}{x-b-i\epsilon} - \frac{1}{x-b+i\epsilon} \right]$$

$$= \frac{1}{2i} \left[\frac{(a-b+2i\epsilon)}{[(x-a)-i\epsilon][(x-b)+i\epsilon]} - \frac{(a-b-2i\epsilon)}{[(x-a)+i\epsilon][(x-b)-i\epsilon]} \right].$$

For ϵ small and $a \neq b$ this is

$$\frac{(a-b)}{2i} \left[\frac{1}{(x-a)(x-b) + i\epsilon(b-a)} - \frac{1}{(x-a)(x-b) - i\epsilon(b-a)} \right]$$

$$= \frac{\epsilon(a-b)^2}{(x-a)^2(x-b)^2 + \epsilon^2(a-b)^2}$$

$$= |a-b| \left[\frac{\epsilon'}{(x-a)^2(x-b)^2 + \epsilon'^2} \right],$$

where we set $\epsilon' = |a-b|\epsilon$. Hence

$$\frac{1}{\pi} \left[\frac{\epsilon}{(x-a)^2 + \epsilon^2} + \frac{\epsilon}{(x-b)^2 + \epsilon^2} \right] = \frac{1}{\pi} |a-b| \left[\frac{\epsilon'}{(x-a)^2(x-b)^2 + \epsilon'^2} \right].$$

Again using Eq. (A.27) and taking the limit $\epsilon \to 0$ on both sides, we obtain the result (A.26).

This page intentionally left blank

Appendix B

Units and Physical Constants

The units used in the text are those internationally agreed on and called SI units or SIU, *Systeme International d'Unité*, or often as earlier also the MKSA system of units, depending on meter, kilogramme, second and ampere. The international abbreviation is SI (cf. references below). One distinguishes between two types of units, those of the first type being the *fundamental units*, of which the relevant ones for our purposes here are meter (m), kilogramme (kg), second (s), ampere (A) and kelvin (K), the last for thermodynamical temperature. All other units, like newton, hertz, etc. are *derived SI units*. The most important derived units for our purposes here, and possibly related uses, are summarised in Table 1, especially as a reference of internationally agreed symbols for the various units.[*] According to international agreement, symbols for units which are proper names, as in Table 1, are written with a first capital letter, their special names are written with small letters (e.g. hertz); plurals of unit names are formed according to the usual rules of grammar with irregular exceptions (hertz, siemens). Symbols for physical quantities are set in italic (sloping) type, while symbols for units are set in roman (upright) type.[†] Multiplication of units is indicated by a raised dot (Table 1) or a space between the units (in text) (cf. again R.A. Nelson, above). Note that in the past Ω^{-1} was occasionally described as mho with the MKSA unit of conductivity 1 mho m^{-1}. Also note that capacitance is American usage for capacity.

Table 2 summarises important fundamental constants [‡] (for frequent use-

---

[*]Extracted (with permission gratefully granted by the author and with notification to the Rights Office, AIP) from R. A. Nelson, " Guide for Metric Practice", *Physics Today* **56**, August 2003, BG 15.

[†]For all of these conventions see in particular R. A. Nelson, above.

[‡]Extracted (with permission gratefully granted by the authors and with notification to the Rights Office, AIP) from P. J. Mohr and B. N. Taylor, "The Fundamental Physical Constants",

fulness with inclusion of the standard value of the gravitational acceleration g_n).

Table 1:	Some derived SI units and their symbols		
Quantity		Unit	
	Special name	Symbol	Equivalent
plane angle	radian	rad	$m/m = 1$
solid angle	steradian	sr	$m^2/m^2 = 1$
speed, velocity			m/s
acceleration			m/s^2
angular velocity			rad/s
angular acceleration			rad/s^2
frequency	hertz	Hz	s^{-1}
force	newton	N	$kg \cdot m/s^2$
pressure, stress	pascal	Pa	N/m^2
work, energy, heat	joule	J	$N \cdot m,\ kg \cdot m^2/s^2$
impulse, momentum			$N \cdot s,\ kg \cdot m/s$
power	watt	W	J/s
electric charge	coulomb	C	$A \cdot s$
electric potential, emf	volt	V	$J/C, W/A$
resistance	ohm	Ω	V/A
conductance	siemens	S	$A/V,\ \Omega^{-1}$
magnetic flux	weber	Wb	$V \cdot s$
inductance	henry	H	Wb/A
capacitance	farad	F	C/V
electric field strength			$V/m,\ N/C$
magnetic flux density	tesla	T	$Wb/m^2,\ N/(A \cdot m)$
electric displacement			C/m^2
magnetic field strength			A/m
radioactivity	becquerel	Bq	s^{-1}

In many cases the Gauss units mentioned in Chapter 1 are still in use. In Table 3 the factors for the appropriate conversions to Gaussian units of equations given in the text in MKSA units are summarised (Jackson [3] uses Gaussian units and therefore gives a Table in the reverse direction). Units for force and other non-electromagnetic quantities remain unaffected.

Table 2: Fundamental Physical Constants
(relevant for electrodynamics and related areas)

Quantity	Symbol	Value	Unit
speed of light in vacuum	c	299 792 458	$\mathrm{m\,s^{-1}}$
magnetic constant	μ_0	$4\pi \times 10^{-7}$	$\mathrm{N\,A^{-2}}$
		$= 12.566\ 370\ 614... \times 10^{-7}$	$\mathrm{N\,A^{-2}}$
electric constant $1/\mu_0 c^2$	ϵ_0	$8.854\ 187\ 817... \times 10^{-12}$	$\mathrm{F\,m^{-1}}$
vacuum impedance $\sqrt{\mu_0/\epsilon_0}$	Z_0	$376.730...$	Ω
Newton grav. constant	G	$6.673... \times 10^{-11}$	$\mathrm{m^3\,kg^{-1}\,s^{-2}}$
Planck constant	h	$6.626\ 068... \times 10^{-34}$	$\mathrm{J\,s}$
Planck constant	h	$4.135\ 667... \times 10^{-15}$	$\mathrm{eV\,s}$
$h/2\pi$	$\hbar$	$1.054\ 571... \times 10^{-34}$	$\mathrm{J\,s}$
Planck mass $\sqrt{\hbar c/G}$	m_P	$2.176\ 7 \times 10^{-8}$	kg
Planck length $\hbar/m_P c$	l_P	$1.616\ 0 \times 10^{-35}$	m
Planck time l_P/c	t_P	$5.390\ 6 \times 10^{-44}$	s
elementary charge	e	$1.602\ 176 \times 10^{-19}$	C
magnetic flux quantum $h/2e$	Φ_0	$2.067\ 833 \times 10^{-15}$	Wb
Bohr magneton $e\hbar/2m_e$	μ_B	$927.400\ 899 \times 10^{-26}$	$\mathrm{J\,T^{-1}}$
Bohr magneton $e\hbar/2m_e$	μ_B	$5.788\ 381\ 7 \times 10^{-5}$	$\mathrm{eV\,T^{-1}}$
fine-structure constant $e^2/4\pi\epsilon_0\hbar c$	α	$7.297\ 352 \times 10^{-3}$	
$1/\alpha$	α^{-1}	$137.035\ 999\ 76...$	
electron mass	m_e	$9.109\ 381\ 88 \times 10^{-31}$	kg
m_e equivalent energy	$m_e c^2$	$8.187\ 104 \times 10^{-14}$	J
m_e equivalent energy	$m_e c^2$	$0.510\ 998\ 9$	MeV
electron charge/mass	$-e/m_e$	$-1.758\ 820 \times 10^{11}$	$\mathrm{C\,kg^{-1}}$
grav. acceleration	g_n	$9.806\ 65$	$\mathrm{m\,s^{-2}}$

Examples:

(1). MKSA: $\nabla \cdot \mathbf{D} = \rho$,

$\therefore$ Gauss: $\nabla \cdot \sqrt{\dfrac{\epsilon_0}{4\pi}}\mathbf{D} = \sqrt{4\pi\epsilon_0}\rho$,

i.e.

$$\nabla \cdot \mathbf{D} = 4\pi\rho.$$

(2). MKSA: $\nabla \times \mathbf{E} = -\dfrac{\partial \mathbf{B}}{\partial t}$,

$\therefore$ Gauss: $\nabla \times \dfrac{\mathbf{E}}{\sqrt{4\pi\epsilon_0}} = -\dfrac{\partial}{\partial t}\sqrt{\dfrac{\mu_0}{4\pi}}\mathbf{B}$,

i.e.

$$\nabla \times \mathbf{E} = -\sqrt{\mu_0\epsilon_0}\mathbf{B} = -\frac{1}{c}\frac{\partial \mathbf{B}}{\partial t}.$$

Table 3: Conversion from MKSA to Gaussian units

Quantity in MKSA units	$\rightarrow$	Quantity in Gaussian units
$1/\sqrt{\mu_0\epsilon_0}$	$\rightarrow$	c
$(\mathbf{E}, \phi, V)$	$\rightarrow$	$\frac{1}{\sqrt{4\pi\epsilon_0}}(\mathbf{E}, \phi, V)$
$\mathbf{D}$	$\rightarrow$	$\sqrt{\frac{\epsilon_0}{4\pi}}\mathbf{D}$
$\rho, q, \mathbf{j}, I, \mathbf{P}$	$\rightarrow$	$\sqrt{4\pi\epsilon_0}(\rho, q, \mathbf{j}, I, \mathbf{P})$
$\mathbf{B}$	$\rightarrow$	$\sqrt{\frac{\mu_0}{4\pi}}\mathbf{B}$
$\mathbf{H}$	$\rightarrow$	$\frac{1}{\sqrt{4\pi\mu_0}}\mathbf{H}$
$\mathbf{M}$	$\rightarrow$	$\sqrt{\frac{4\pi}{\mu_0}}\mathbf{M}$
σ	$\rightarrow$	$4\pi\epsilon_0\sigma$
ϵ	$\rightarrow$	$\epsilon_0\epsilon$
μ	$\rightarrow$	$\mu_0\mu$
(R, Z)	$\rightarrow$	$\frac{1}{4\pi\epsilon_0}(R, Z)$
L	$\rightarrow$	$\frac{1}{4\pi\epsilon_0}L$
C	$\rightarrow$	$4\pi\epsilon_0 C$

---

(3). MKSA: $\boldsymbol{\nabla} \times \mathbf{H} = \mathbf{j} + \frac{\partial}{\partial t}\mathbf{D}$,

$\therefore$ Gauss: $\boldsymbol{\nabla} \times \frac{\mathbf{H}}{\sqrt{4\pi\mu_0}} = \sqrt{4\pi\epsilon_0}\mathbf{j} + \frac{\partial}{\partial t}\sqrt{\frac{\epsilon_0}{4\pi}}\mathbf{D}$,

i.e.

$$\boldsymbol{\nabla} \times \mathbf{H} = 4\pi\sqrt{\mu_0\epsilon_0}\mathbf{j} + \frac{\partial}{\partial t}\sqrt{\epsilon_0\mu_0}\mathbf{D} = \frac{4\pi}{c}\mathbf{j} + \frac{1}{c}\frac{\partial\mathbf{D}}{\partial t}.$$

(4). MKSA: $\mathbf{F} = \rho\mathbf{E} + \mathbf{j} \times \mathbf{B}$,

$\therefore$ Gauss: $\mathbf{F} = \sqrt{4\pi\epsilon_0}\rho\frac{\mathbf{E}}{\sqrt{4\pi\epsilon_0}} + \sqrt{4\pi\epsilon_0}\mathbf{j} \times \sqrt{\frac{\mu_0}{4\pi}}\mathbf{B} = \rho\mathbf{E} + \frac{1}{c}\mathbf{j} \times \mathbf{B}$,

i.e.

$$\mathbf{F} = \rho\left(\mathbf{E} + \frac{\mathbf{v} \times \mathbf{B}}{c}\right).$$

These equations show that the factor c always appears whenever a magnetic (or electric) quantity is related to an electric (magnetic) quantity respectively, as we observe this also in the field tensor $F_{\mu\nu}$ (see e.g. Eq. (17.35)). Thus the factor $c = 1/\sqrt{\epsilon_0\mu_0}$ appears in the transition from purely electric units

(long ago called e.s.u.), to purely magnetic units (long ago called e.m.u.). In the Gaussian (c.g.s. = cm gm sec) system of units (see Appendix in Jackson [3]), as in the original e.s.u. ('s' for 'static') the quantities charge, current, potential (voltage) and capacity are given in *statcoulomb, statampere, statvolt* and *statfarad*. With $c = 3 \times 10^{10}$ cm sec^{-1} ('3' as good approximation for 2.997 924 58...) one defines

$$1 \text{ C} = 3 \times 10^9 \text{ e.s.u.} \equiv 3 \times 10^9 \text{ statcoulombs}, \quad 1 \text{ A} = 3 \times 10^9 \text{ statamperes}.$$

Then, using the equivalents of Table 1,

$$1 \text{ V(volt)} = \frac{1 \text{ J(joule)}}{1 \text{ C(coulomb)}} = \frac{10^7 \text{ erg}}{3 \times 10^9 \text{ e.s.u.}}$$

$$= \frac{1}{300} \frac{\text{erg}}{\text{e.s.u.}} = \frac{1}{300} \text{ statvolt},$$

$$1 \text{ F(farad)} = \frac{1 \text{ C(coulomb)}}{1 \text{ V(volt)}} = \frac{3 \times 10^9 \text{ statcoulombs}}{1V}$$

$$= \frac{3 \times 10^9 \text{ statcoulombs}}{\frac{1}{300} \text{ statvolt}} = 9 \times 10^{11} \text{ statfarads}$$

with
$$1 \text{ statfarad} = 1 \text{ statcoulomb}/1 \text{ statvolt}.$$

For the electric field strength E one has therefore

$$E: \qquad 1 \text{ V m}^{-1} = \frac{1}{3} \times 10^{-4} \text{ statvolt cm}^{-1}.$$

From $\nabla \cdot \mathbf{D} = \rho$, Gaussian $\nabla \cdot \mathbf{D} = 4\pi\rho$, one obtains for D:

$$D: \qquad 1 \text{ C m}^{-2} = 4\pi \frac{3 \times 10^9 \text{ statcoulombs}}{10^4 \text{ cm}^2}$$

$$= 12\pi \times 10^5 \text{ statcoulombs cm}^{-2}$$

$$(\text{or statvolt cm}^{-1}).$$

For resistance R:

$$R: \qquad 1 \text{ } \Omega = 1 \text{ V}/1 \text{ A} = \frac{\text{statvolt}}{300 \times 3 \times 10^9 \text{ statamperes}}$$

$$= \frac{1}{9} \times 10^{-11} \text{statvolt sec/statcoulomb}$$

$$(\text{or sec/statfarad or sec cm}^{-1}),$$

and in the case of the conductivity σ:

$$\sigma: \qquad 1 \text{ } \Omega^{-1} \text{ m}^{-1} = 9 \times 10^9 \text{ statamperes cm}^{-1}/\text{statvolt} \quad (\text{or sec}^{-1}).$$

The old e.m.u. (electromagnetic units) are still occasionally mentioned, e.g. also in Jackson [3], and one finds these in older texts. We therefore mention these briefly, since some aspects of their definition are not without interest, even today. These units were introduced via the formula for the field strength H at the centre of a single circular conductor with radius r and current I, i.e. the expression $H \propto 2\pi I/r$, in which the proportionality constant is chosen as 1, i.e. with

$$H = \frac{2\pi I}{r}.$$

For $r = 1\,\text{cm}, I = $ unit of current in e.m.u. (in old literature also called 'abampere') one has $H = 2\pi$ oersted, with the unit oersted given in dyne per magnetic pole. (Thus in the case of the e.m.u. the unit of current is defined first, whereas in the case of the e.s.u. the unit of charge is defined first). Using the expressions $I = dq/dt, dq = Idt$, current I is given in abampere and dt in seconds, and charge dq in abcoulomb, where 1 abcoulomb is defined as that amount of charge which is transported in one second by a current of one abampere.

The relation between e.m.u. and e.s.u. can be determined experimentally. The easiest way is with the help of a parallel plate condenser of known mass. Let q_s be the charge in e.s.u. and q_m its value in e.m.u. What one is looking for is the constant c in the relation $q_m = cq_s$. The capacity of the condenser in vacuum is (let us say)

$$C_s = \frac{q_s}{V_s} \text{ e.s.u.} \quad \text{or} \quad C_m = \frac{q_m}{V_m} \text{ e.m.u.}$$

Since in the case of both kinds of units the energy is measured in ergs, one can equate them, i.e.

$$\frac{1}{2}q_s V_s = \frac{1}{2}q_m V_m \quad \text{or} \quad \frac{q_s}{q_m} = \frac{V_m}{V_s} = \frac{q_m}{q_s}\frac{C_s}{C_m}.$$

If we replace here the ratio of the potentials by that of the capacities from the preceding relation, we obtain

$$\frac{q_s}{q_m} = \sqrt{\frac{C_s}{C_m}}$$

and hence

$$c = \sqrt{\frac{C_m}{C_s}}.$$

The value of C_s can now be given directly in centimeters for the parallel plate condenser (cf. Example 2.9), and C_m can be determined experimentally,

e.g. with the help of a ballistic galvanometer. This interesting experiment was performed in 1907 by Rosa and Dorsey[§] and yielded to a high degree of precision for the constant c the numerical value of the velocity of light. Hence

$$q \text{ e.m.u.} = 3 \times 10^{10} \, q \text{ e.s.u.},$$

or

$$1 \text{ C} = 3 \times 10^9 \text{ e.s.u.} = 10^{-1} \text{ e.m.u.}$$

In this way we obtain the relationships between the systems of units given in Table 4 (here only for a few typical examples). The derivation of such relations can, for instance, also be found in electrodynamics exercises in the literature.[¶]

	Table 4: Relations between e.s.u. and e.m.u.		
Quantity	e.s.u.	e.m.u.	MKSA
charge	$3 \times 10^9 =$	10^{-1}	C
current	$3 \times 10^9 =$	10^{-1}	A
potential	$1/300 =$	10^8	V
electric field strength	$1/300 =$	10^6	V m^{-1}
resistance	$\frac{1}{9} \times 10^{-11} =$	10^{-9}	Ω
capacity	$9 \times 10^{11} =$	10^{-9}	F
inductance	$\frac{1}{9} \times 10^{-11} =$	10^9	H

Finally we collect in Table 5 some useful values of energy-equivalences for conversion from one unit to another.[‖]

[§]See e.g. H. B. Lemon and M. Ference, *Analytical Experimental Physics* (University of Chicago Press, 1943), p. 385.

[¶]See e.g. E. H. Booth and P. M. Nicol, *Physics, Fundamental Laws and Principles with Problems and Worked Solutions*, 12th ed. (Australasian Medical Publ. Co., 1952), pp. 179, 506–507.

[‖]Extracted (with permission gratefully granted by the authors and with notification to the Rights Office, AIP) from P. J. Mohr and B. N. Taylor, "The Fundamental Physical Constants", *Physics Today* **56**, August 2003, BG6. Note that in our brief extract some precise values have been reduced to fewer decimal places.

Table 5: Energy equivalences

	J	kg	m^{-1}	Hz
		Relevant unit		
1 J	$(1\ J) =$	$(1\ J)/c^2 =$	$(1\ J)/hc =$	$(1\ J)/h =$
	1	$1.112\ 650 \times 10^{-17}$	$5.034\ 117\ 6 \times 10^{24}$	$1.509\ 190\ 5 \times 10^{33}$
1 kg	$(1\ kg)c^2 =$	$(1\ kg) =$	$(1\ kg)c/h =$	$(1\ kg)c^2/h =$
	$8.987\ 551\ 787 \times 10^{16}$	1	$4.524\ 439 \times 10^{41}$	$1.356\ 392\ 77 \times 10^{50}$
$1\ m^{-1}$	$(1\ m^{-1})hc =$	$(1\ m^{-1})h/c =$	$(1\ m^{-1}) =$	$(1\ m^{-1})c =$
	$1.986\ 445 \times 10^{-25}$	$2.210\ 218\ 6 \times 10^{-42}$	$1\ m^{-1}$	299 792 458
1 Hz	$(1\ Hz)h =$	$(1\ Hz)h/c^2 =$	$(1\ Hz)/c =$	$(1\ Hz) =$
	$6.626\ 068\ 76 \times 10^{-34}$	$7.372\ 495\ 78 \times 10^{-51}$	$3.335\ 640\ 95 \times 10^{-9}$	1
1 K	$(1\ K)k =$	$(1\ K)k/c^2 =$	$(1\ K)k/hc =$	$(1\ K)k/h =$
	$1.380\ 650\ 3 \times 10^{-23}$	$1.536\ 180\ 7 \times 10^{-40}$	69.503 56	$2.083\ 664\ 4 \times 10^{10}$
1 eV	$(1\ eV) =$	$(1\ eV)/c^2 =$	$(1\ eV)/hc =$	$(1\ eV)/h =$
	$1.602\ 176 \times 10^{-19}$	$1.782\ 661\ 7 \times 10^{-36}$	$8.065\ 544\ 77 \times 10^5$	$2.417\ 989\ 49 \times 10^{14}$

Appendix C

Formulae

C.1 Vector Products

The following relations are frequently used in the text and are therefore summarised here:

$$\mathbf{A} \cdot (\mathbf{B} \times \mathbf{C}) = (\mathbf{A} \times \mathbf{B}) \cdot \mathbf{C},$$

$$(\mathbf{P} \times \mathbf{Q}) \times \mathbf{R} = (\mathbf{R} \cdot \mathbf{P})\mathbf{Q} - (\mathbf{R} \cdot \mathbf{Q})\mathbf{P},$$

$$\mathbf{P} \times (\mathbf{Q} \times \mathbf{R}) = (\mathbf{P} \cdot \mathbf{R})\mathbf{Q} - (\mathbf{P} \cdot \mathbf{Q})\mathbf{R}.$$

For vectorial derivatives one has in particular: *curl curl = grad div − div grad*, i.e.

$$\nabla \times (\nabla \times \mathbf{A}) = \nabla(\nabla \cdot \mathbf{A}) - (\nabla \cdot \nabla)\mathbf{A}.$$

In addition:

$$\nabla \times (\mathbf{u} \times \mathbf{v}) = \mathbf{u}(\nabla \cdot \mathbf{v}) - \mathbf{v}(\nabla \cdot \mathbf{u}) + (\mathbf{v} \cdot \nabla)\mathbf{u} - (\mathbf{u} \cdot \nabla)\mathbf{v},$$

$$\nabla(\mathbf{u} \cdot \mathbf{v}) = (\mathbf{v} \cdot \nabla)\mathbf{u} + (\mathbf{u} \cdot \nabla)\mathbf{v} + \mathbf{v} \times (\nabla \times \mathbf{u}) + \mathbf{u} \times (\nabla \times \mathbf{v}),$$

$$\nabla(u\mathbf{v}) = \nabla u \cdot \mathbf{v} + u\nabla \cdot \mathbf{v},$$

$$\nabla \times (u\mathbf{v}) = \nabla u \times \mathbf{v} + u\nabla \times \mathbf{v},$$

$$\nabla \cdot (\mathbf{u} \times \mathbf{v}) = \mathbf{v} \cdot (\nabla \times \mathbf{u}) - \mathbf{u} \cdot (\nabla \times \mathbf{v}).$$

The following relations hold

$$\text{curl grad } V = 0 \quad \text{and} \quad \text{div curl } \mathbf{F} = 0,$$

i.e.

$$\nabla \times \nabla V = 0 \quad \text{and} \quad \nabla \cdot (\nabla \times \mathbf{F}) = 0.$$

513

C.2   Integral Theorems

Gauss' divergence theorem:

$$\int_V \boldsymbol{\nabla} \cdot \mathbf{E}\, dV = \int_{F(V)} \mathbf{E} \cdot d\mathbf{F},$$

and
Stokes' circulation theorem:

$$\int_F \boldsymbol{\nabla} \times \mathbf{E} \cdot d\mathbf{F} = \oint_{C(F)} \mathbf{E} \cdot d\mathbf{s}.$$

Bibliography

[1] A. Sommerfeld, *Elektrodynamik*, 3rd ed., revised and supplemented by F. Bopp and J. Meixner (Akad. Verlagsgesellschaft, 1961).

[2] Y.K. Lim, *Introduction to Classical Electrodynamics* (World Scientific, 1986).

[3] J.D. Jackson, *Classical Electrodynamics*, Second Edition (Wiley, 1975).

[4] W. Greiner, *Theoretische Physik, Vol. 3: Klassische Elektrodynamik*, 3. Auflage (H. Deutsch, 1982).

[5] L. Page and N. I. Adams, *Electrodynamics* (Dover Publications, 1965); see pp. 154–155.

[6] A. O'Rahilly, *Electromagnetic Theory*, Vols. I, II (Dover Publications, 1965).

[7] E.T. Whittaker and G. N. Watson, *A Course of Modern Analysis*, 4th ed. (Cambridge, 1958).

[8] M. Zahn, *Electromagnetic Field Theory* (Wiley, 1979).

[9] B.H. Chirgwin, C. Plumpton and C. M. Kilmister, *Elementary Electromagnetic Theory, Vol. I* (Pergamon Press, 1971).

[10] A.C. Aitken, *Determinants and Matrices* (Oliver and Boyd, 1954).

[11] LI. G. Chambers, *An Introduction to the Mathematics of Electricity and Magnetism* (Chapman and Hall, 1973).

[12] D.K. Cheng, *Field and Wave Electromagnetics* (Reading, Addison–Wesley, 1983).

[13] P. C. Clemmow, *An Introduction to Electromagnetic Theory* (Cambridge University Press, 1973).

[14] G. P. Harnwell, *Principles of Electricity and Magnetism* (McGraw–Hill, 1949).

[15] E. Stiefel, *Methoden der Mathematischen Physik II* (Verlag der Fachvereine, Zürich, 1980).

[16] F. Schwabl, *Quantenmechanik* (Springer, 1998).

[17] J.A. Stratton, *Electromagnetic Theory* (McGraw-Hill, 1941).

[18] R.B. Dingle, *Electricity and Magnetism*, Mimeographed Lecture Notes, University of Western Australia (1956).

[19] S. Schweber, *An Introduction to Relativistic Quantum Field Theory* (Harper and Row, 1961).

[20] B. Felsager, *Geometry, Particles and Fields* (Odense University Press, 1981).

[21] Staff of Research and Education Association, *The Electromagnetic Problem Solver* (SREA, revised 1988).

Index

517