Chapter 16
Appendix

16.1 Units of Measurement for Electromagnetic Phenomena

The units of measurement relative to electromagnetic phenomena have been introduced through a long and complex historical process. Without going into the details of this process, we try here to summarise the fundamental points in light of a modern vision of the phenomena themselves. We stress that these notes are aimed at a reader who is already familiar with the basic phenomenology of electromagnetism.

Within electrostatics, the fundamental law is Coulomb's law which is written, in vacuum, in the general form

$$\mathbf{F} = k_{\mathrm{C}} \frac{q_1 q_2}{r^2} \operatorname{vers} \mathbf{r},$$

where $\mathbf{F}$ is the force that a point charge q_1 exerts on the point charge q_2 placed at the distance $\mathbf{r} = \mathbf{r}_2 - \mathbf{r}_1$, and where k_{C} is a constant that implicitly defines the unit of measurement of the charge (assuming, of course, that the units of measurement of the mechanical quantities have already been set). We note that k_{C} can be chosen to be dimensional or dimensionless. The electric field vector $\mathbf{E}$ is defined in an arbitrary point using the equation

$$\mathbf{E} = \frac{\mathbf{F}}{q_{\mathrm{p}}},$$

where $\mathbf{F}$ is the electric force exerted on the "test" charge q_{p} placed at the same location. From this definition and Coulomb's law we can deduce the expression of the electric field due to a point charge q, which is

$$\mathbf{E} = k_{\mathrm{C}} \frac{q}{r^2} \operatorname{vers} \mathbf{r},$$

from which the Gauss theorem (in its integral form) follows

$$\Phi(\mathbf{E}) = \int_{\Sigma} \mathbf{E} \cdot \mathbf{n} \, dS = 4\pi k_{\mathrm{C}} Q,$$

where Q is the charge contained within the surface Σ. In its differential form, we have

$$\operatorname{div} \mathbf{E} = 4\pi k_C \rho,$$

where ρ is the density of the electric charge (the charge contained within the unit volume).

Regarding the definition of the electric field vector, we note that it is not the only possible one, since we could have defined the electric field created by the charge q as

$$\mathbf{E} = k_C \delta \frac{q}{r^2} \operatorname{vers} \mathbf{r},$$

with δ an arbitrary constant (possibly dimensional), as long as the electric force that the field exerts on the test charge q_p is written in the form

$$\mathbf{F} = \frac{1}{\delta} q_p \mathbf{E}.$$

Fortunately, the constant δ has (historically) always been set to unity. The same is not true for magnetic phenomena.

Regarding magnetostatics, we have equations that are similar to those of electrostatics. In these equations, for historical reason, the fictitious concept of "magnetic mass" (or "magnetic pole") is introduced. The equations corresponding to those previously written are the following ones (the symbol m denoting the magnetic mass):

Gilbert's law[1] (analogous to Coulomb's law)

$$\mathbf{F} = k_G \frac{m_1 m_2}{r^2} \operatorname{vers} \mathbf{r}.$$

Definition of the vector of the magnetic induction generated from the magnetic mass m

$$\mathbf{B} = k_G \gamma \frac{m}{r^2} \operatorname{vers} \mathbf{r},$$

where γ is an arbitrary constant (possibly dimensional).
Force acting on the test magnetic mass m_p

$$\mathbf{F} = \frac{1}{\gamma} m_p \mathbf{B}. \tag{16.1}$$

Equivalent of the Gauss theorem (integral form)

$$\Phi(\mathbf{B}) = 0,$$

---

[1]This law is not universally attributed to Gilbert. In fact, this law was discovered experimentally by Coulomb himself and could therefore be rightly called the "second Coulomb's law". William Gilbert (1564–1603) was an English physician who lived well before Coulomb. He is remembered for his studies on the terrestrial magnetism and by the fact that he realised that the magnetic force should increase with decreasing distance.

since isolated magnetic masses (magnetic monopoles) do not exist.
Equivalent of the Gauss theorem (differential form)

$$\text{div}\,\mathbf{B} = 0.$$

The first quantitative relations between electric and magnetic phenomena were established with experiments based on electric currents.[2] The intensity of the electric current i flowing in a conductor is defined by the simple equation

$$i = \frac{dq}{dt},$$

from which we can define the current density $\mathbf{j}$ as a vector directed along the direction of motion of the (positive) charges of magnitude

$$j = \frac{i}{\sigma},$$

where σ is the cross-sectional area of the conductor. The experiments performed during the first half of the nineteenth century especially by Ørsted, Ampère, and Faraday, led to the idea that electric currents create magnetic fields in their surroundings and that, at the same time, a magnetic field is able to exert a force on electric currents. During the same period, a new idea clearly emerged: that permanent magnets contain, at the microscopic level, a large number of elementary electric currents. These currents would be responsible, ultimately, for magnetostatic phenomena.

In modern terms, the magnetic properties of the currents can be summarised by a single law that is expressed by saying that, in stationary conditions, the current element of an elementary circuit (microscopic or macroscopic) $i_1\,d\boldsymbol{\ell}_1$ acts on the current element of another elementary circuit, $i_2\,d\boldsymbol{\ell}_2$, with an infinitesimal force $d\mathbf{F}$ given by

$$d\mathbf{F} = k_{\mathrm{A}} i_2\,d\boldsymbol{\ell}_2 \times \left(i_1\,d\boldsymbol{\ell}_1 \times \frac{\text{vers}\,\mathbf{r}}{r^2} \right),$$

where k_{A} is a new constant (which cannot be independent of those already introduced), and where $\mathbf{r}$ is the radius vector that goes from the current element $i_1\,d\boldsymbol{\ell}_1$ to the current element $i_2\,d\boldsymbol{\ell}_2$. This law allows the introduction of the magnetic induction vector. The definition of this vector is somewhat arbitrary and it is assumed, in general, that the current element $i\,d\boldsymbol{\ell}$ creates the elementary induction vector $d\mathbf{B}$ given by (first law of Laplace or Biot and Savart's law)

$$d\mathbf{B} = k_{\mathrm{A}}\beta i\,d\boldsymbol{\ell} \times \frac{\text{vers}\,\mathbf{r}}{r^2}, \tag{16.2}$$

---

[2]These experiments were made possible thanks to the discovery of the electric battery by Alessandro Volta.

and that a current element $i \, \mathrm{d}\boldsymbol{\ell}$ is subject, in the presence of an induction vector $\mathbf{B}$, to a force $\mathrm{d}\mathbf{F}$ given by (second law of Laplace)

$$\mathrm{d}\mathbf{F} = \frac{1}{\beta} i \, \mathrm{d}\boldsymbol{\ell} \times \mathbf{B}.$$

The quantity β introduced in these equations is arbitrary.

Let's see the mathematical consequences of Eq. (16.2) for a closed circuit. The magnetic induction vector is given by

$$\mathbf{B} = k_\mathrm{A}\beta \oint_\mathrm{C} i \, \mathrm{d}\boldsymbol{\ell} \times \frac{\mathrm{vers}\,\mathbf{r}}{r^2},$$

where C is the curve describing the closed circuit. Using standard mathematical methods, one finds the following equations

$$\mathrm{div}\,\mathbf{B} = 0,$$

which confirms the analogous equation for magnetostatics, and

$$\mathrm{rot}\,\mathbf{B} = 4\pi k_\mathrm{A}\beta\mathbf{j},$$

where $\mathbf{j}$ is the current density.

This equation, known as Ampère's law, applies only to stationary phenomena. As shown by Maxwell, it can be transformed into a more general equation that is also valid for phenomena that are variable in time. To do this we observe that, taking the divergence of both sides, we have

$$\mathrm{div}\,\mathbf{j} = 0,$$

while, in general, the continuity equation must hold

$$\mathrm{div}\,\mathbf{j} + \frac{\partial\rho}{\partial t} = 0,$$

ρ being the charge density. In order to rearrange things, we take the derivative (with respect to time) of the differential expression of Coulomb's law

$$\frac{\partial\rho}{\partial t} = \frac{1}{4\pi k_\mathrm{C}} \frac{\partial}{\partial t}(\mathrm{div}\,\mathbf{E}),$$

so that in general the following equation holds

$$\mathrm{div}\left(\mathbf{j} + \frac{1}{4\pi k_\mathrm{C}} \frac{\partial\mathbf{E}}{\partial t}\right) = 0.$$

The second term in parentheses is the so-called displacement current density. With its introduction, the equation for $\mathrm{rot}\,\mathbf{B}$, corrected to include non-stationary phenomena, is

$$\mathrm{rot}\,\mathbf{B} - \frac{k_\mathrm{A}\beta}{k_\mathrm{C}} \frac{\partial\mathbf{E}}{\partial t} = 4\pi k_\mathrm{A}\beta\mathbf{j}.$$

Finally, we need to consider the phenomena of magnetic induction. The law that describes them can be deduced, at least in a particular case, from the second law of Laplace. We have

$$\operatorname{rot} \mathbf{E} = -\frac{1}{\beta} \frac{\partial \mathbf{B}}{\partial t},$$

so that, in summary, the laws governing the electromagnetic phenomena can all be enclosed in the following four Maxwell's equations

$$\operatorname{div} \mathbf{E} = 4\pi k_C \rho, \qquad\qquad \operatorname{div} \mathbf{B} = 0,$$

$$\operatorname{rot} \mathbf{B} - \frac{k_A \beta}{k_C} \frac{\partial \mathbf{E}}{\partial t} = 4\pi k_A \beta \mathbf{j}, \qquad \operatorname{rot} \mathbf{E} + \frac{1}{\beta} \frac{\partial \mathbf{B}}{\partial t} = 0.$$

We now consider Maxwell's equations in vacuum. Taking the curl of the third equation and substituting the fourth, we obtain the wave equation

$$\nabla^2 \mathbf{B} = \frac{k_A}{k_C} \frac{\partial^2 \mathbf{B}}{\partial t^2}.$$

On the other hand we know that electromagnetic waves propagate in vacuum with velocity c, so that we must have

$$\frac{k_A}{k_C} = \frac{1}{c^2},$$

or

$$k_A = \frac{k_C}{c^2},$$

that is, a relation between the quantities k_A and k_C that is independent of the unit system under consideration.

Let's see how we proceed in the two more common systems of units, the cgs system of Gauss (sometimes also called the Gauss-Hertz system) and the International System of Units (SI). In the cgs system, we assume $k_C = 1$, so that the unit of charge is defined as the charge that repels an equal charge, at a distance of one centimeter, with the force of one dyne. Such unit of charge is called Franklin or statcoulomb. Since $k_C = 1$, it follows that $k_A = 1/c^2$. Within this system we also assume that $\beta = c$, so that Maxwell's equations are written as

$$\operatorname{div} \mathbf{E} = 4\pi \rho, \qquad\qquad \operatorname{div} \mathbf{B} = 0,$$

$$\operatorname{rot} \mathbf{B} - \frac{1}{c} \frac{\partial \mathbf{E}}{\partial t} = 4\pi \frac{\mathbf{j}}{c}, \qquad \operatorname{rot} \mathbf{E} + \frac{1}{c} \frac{\partial \mathbf{B}}{\partial t} = 0.$$

Moreover, the first and the second law of Laplace, together with the law that sum-
marises them, can be written in the form

$$dB = \frac{i}{c} d\ell \times \frac{\text{vers } \mathbf{r}}{r^2}, \qquad dF = \frac{i}{c} d\ell \times \mathbf{B},$$

$$dF = \frac{i_2}{c} d\ell_2 \times \left(\frac{i_1}{c} d\ell_1 \times \frac{\text{vers } \mathbf{r}}{r^2} \right).$$

Within the International System, instead, two new constants are introduced. They
are the vacuum permittivity (also called dielectric permittivity of the vacuum) ϵ_0 and
the vacuum permeability (magnetic permeability of the vacuum) μ_0, such that

$$\epsilon_0 \mu_0 = \frac{1}{c^2}.$$

Using these quantities, we put

$$k_C = \frac{1}{4\pi \epsilon_0},$$

so that we have

$$k_A = \frac{k_C}{c^2} = \frac{1}{4\pi \epsilon_0 c^2} = \frac{\mu_0}{4\pi}.$$

Within this system we also put $\beta = 1$, so that Maxwell's equations are written as

$$\text{div } \mathbf{E} = \frac{\rho}{\epsilon_0}, \qquad\qquad \text{div } \mathbf{B} = 0,$$

$$\text{rot } \mathbf{B} - \frac{1}{c^2} \frac{\partial \mathbf{E}}{\partial t} = \mu_0 \mathbf{j}, \qquad \text{rot } \mathbf{E} + \frac{\partial \mathbf{B}}{\partial t} = 0.$$

Moreover, the first and the second law of Laplace and the law that summarises them
are written, respectively, in the form

$$dB = \frac{\mu_0}{4\pi} i \, d\ell \times \frac{\text{vers } \mathbf{r}}{r^2}, \qquad dF = i \, d\ell \times \mathbf{B},$$

$$dF = \frac{\mu_0}{4\pi} i_2 \, d\ell_2 \times \left(i_1 \, d\ell_1 \times \frac{\text{vers } \mathbf{r}}{r^2} \right). \tag{16.3}$$

With respect to the numerical values of ϵ_0 and μ_0, the Ampère (unit of measure-
ment of the current) is defined as the current that, flowing along an infinite straight
wire of negligible thickness in vacuum, attracts an equal wire, located at a distance
of one meter, with a force per unit length equal to $2 \times 10^{-7} \, \text{N m}^{-1}$. Using Eq. (16.3)
we deduce that in such a geometry the force per unit length that acts on one of the
two conductors is attractive and has a magnitude given by the following expression

$$\frac{dF}{dl} = 2 \frac{\mu_0}{4\pi} \frac{i^2}{r},$$

so we must have[3]

$$\mu_0 = 4\pi \times 10^{-7} \, \mathrm{N\,A}^{-2} = 1.256637 \times 10^{-6} \, \mathrm{N\,A}^{-2},$$

and then, recalling that the Coulomb is the charge transported in one second by a current of one Ampère

$$\epsilon_0 = \frac{1}{\mu_0 c^2} = 8.854188 \times 10^{-12} \, \mathrm{C^2\,N}^{-1}\,\mathrm{m}^{-2}.$$

Finally, it remains to analyse the relation between magnetic masses and currents. We can infer from Laplace's laws that a filiform planar circuit of area σ and current i behaves, at distances much larger than its size, as a magnetic dipole directed along the unit vector $\mathbf{n}$ perpendicular to the plane of the circuit. The direction of $\mathbf{n}$ is specified by the rule of the corkscrew (or the right screw). This is the so-called Ampère principle of equivalence, which is expressed by the formula

$$\boldsymbol{\mu} = k_\mathrm{P} i \sigma \mathbf{n},$$

where k_P is a new constant to be related to those previously introduced. To establish this relation, we evaluate, for example, the moment of the forces acting on an elementary dipole located at a point in space where the field $\mathbf{B}$ is present. Using Eq. (16.1), we have

$$\mathbf{M} = \frac{1}{\gamma} \boldsymbol{\mu} \times \mathbf{B} = \frac{1}{\gamma} k_\mathrm{P} i \sigma \mathbf{n} \times \mathbf{B}.$$

Instead, from the second law of Laplace we have

$$\mathbf{M} = \frac{1}{\beta} \oint i \mathbf{r} \times (\mathrm{d}\boldsymbol{\ell} \times \mathbf{B}),$$

which can be rewritten as

$$\mathbf{M} = \frac{1}{\beta} i \sigma \mathbf{n} \times \mathbf{B}.$$

Equating the two expressions for $\mathbf{M}$ we have

$$k_\mathrm{P} = \frac{\gamma}{\beta}.$$

Finally, considering the force exerted between two infinitesimal circuits, treated in the first instance as elementary dipoles and then as coils carrying a current, we obtain the relation

$$k_\mathrm{G} k_\mathrm{P}^2 = k_\mathrm{A},$$

---

[3]With the introduction of capacity and inductance, together with their units, the Farad (F) and the Henry (H), the units in which μ_0 and ϵ_0 are expressed are, respectively, $\mathrm{H\,m}^{-1}$ and $\mathrm{F\,m}^{-1}$.

which allows to write k_G in the form

$$k_G = \frac{k_A}{k_P^2} = \frac{k_C \beta^2}{c^2 \gamma^2}.$$

In the cgs system, since $k_C = 1$ and $\beta = c$, and assuming $\gamma = 1$, we obtain

$$k_P = \frac{1}{c}, \qquad k_G = 1.$$

Ampère's principle of equivalence is therefore

$$\boldsymbol{\mu} = \frac{i}{c} \sigma \mathbf{n},$$

and Gilbert's law

$$\mathbf{F} = \frac{m_1 m_2}{r^2} \text{ vers } \mathbf{r}.$$

In the International System, instead, since $k_C = 1/(4\pi\epsilon_0)$ and $\beta = 1$, assuming[4] $\gamma = \mu_0$ and recalling that $c^2 = 1/(\epsilon_0 \mu_0)$, we obtain

$$k_P = \mu_0, \qquad k_G = \frac{1}{4\pi\mu_0}.$$

In this case Ampère's principle of equivalence is

$$\boldsymbol{\mu} = \mu_0 i \sigma \mathbf{n},$$

and Gilbert's law is

$$\mathbf{F} = \frac{1}{4\pi\mu_0} \frac{m_1 m_2}{r^2} \text{ vers } \mathbf{r}.$$

Finally, we note that, besides the two systems introduced here, there are other ones that have been used for the electromagnetic phenomena. In particular, it is worth mentioning the electrostatic cgs system, the electromagnetic cgs system and the cgs system of Heavyside.

16.2 Tensor Algebra

In this volume, we often need to deal with vectors and tensors, together with their differential expressions such as divergences, curls and gradients. It is therefore useful to give a brief introduction to this topic in order to make the reader familiar

---

[4]This convention is not universally accepted. Some authors prefer to assume $\gamma = 1$ also in the International System. In this case Ampère's principle of equivalence is written as $\boldsymbol{\mu} = i\sigma\mathbf{n}$ while in Gilbert's law the factor μ_0 is to appear in the numerator rather than in the denominator.

with a compact formalism that allows to easily deduce a series of vector and tensor identities, as well as various transformation formulae.

The traditional definition of a tensor that is commonly given in physics is based on the generalisation of the definition of a vector. In a Cartesian orthogonal reference system, the vector $\mathbf{v}$ is defined as an entity with three components (v_x, v_y, v_z) (or v_1, v_2, v_3) which, under an arbitrary rotation of the reference system, are modified according to the law

$$v'_i = \sum_j C_{ij} v_j,$$

where the coefficients C_{ij} are the direction cosines of the new axes with respect to the old ones. In close analogy, we define a tensor $\mathbf{T}$ of rank n as an entity with 3^n components ($T_{i...j}$ with $i, \ldots, j = 1, 3$) which, under a rotation of the reference system, are transformed according to the law

$$T'_{i...j} = \sum_{k,...,l} C_{ik} \cdots C_{jl} T_{k...l}.$$

The tensor most commonly known in physics is the stress tensor that characterises inside an elastic material the force $\mathbf{dF}$ that is exerted on a surface $\mathrm{d}S$ with normal $\mathbf{n}$. In components we have

$$\mathrm{d}F_i = \sum_j T_{ij} n_j \, \mathrm{d}S.$$

In addition to the stress tensor we can also mention, for their importance in various fields of physics, the deformation tensor, the inertia tensor, and the dielectric tensor.

A particular tensor of rank two is the so-called dyad that is obtained from two vectors $\mathbf{u}$ and $\mathbf{v}$ when the direct product of their components is considered. The dyad is indicated simply by the symbol $\mathbf{uv}$, and we have by definition

$$(\mathbf{uv})_{ij} = u_i v_j \quad (i, j = 1, 2, 3).$$

Obviously, in general

$$\mathbf{uv} \neq \mathbf{vu}.$$

A scalar quantity is, by definition, a tensor of rank zero, while a vector is, by definition, a tensor of rank one. Tensors of higher rank may be obtained by considering the direct product of tensors of lower rank. For example, by the direct product of two tensors of rank two a tensor of rank four is obtained.

The tensor algebra covers all operations that can be performed on tensors. We now provide some definitions

1. Given two tensors $\mathbf{T}$ and $\mathbf{V}$, the first of rank n ($n \geq 1$) and the second of rank n' ($n' \geq 1$), we define the scalar product (or internal product) of the two tensors a tensor of rank ($n + n' - 2$) obtained by a sum (or saturation) which operates over the last index of the first tensor and the first index of the second tensor. For

example, if n and n' are both equal to 2, defining $\mathbf{W}$ as the tensor obtained by the scalar product, we have that $\mathbf{W}$ is also a tensor of rank two defined by

$$W_{ij} = \sum_k T_{ik} V_{kj}.$$

2. Given a tensor of rank n (with $n \geq 1$), the divergence of such tensor is a tensor of rank $(n-1)$ obtained by saturating its first component with the formal vector ∇ (called "nabla" operator or "del" operator) defined by

$$\nabla \equiv \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z} \right).$$

For example, for a tensor $\mathbf{T}$ of rank two, div $\mathbf{T}$ is a vector whose components are given by

$$(\text{div } \mathbf{T})_i = \sum_j \frac{\partial}{\partial x_j} T_{ji} = (\nabla \cdot \mathbf{T})_i.$$

3. Given a tensor of rank n (with $n \geq 0$), the gradient of such tensor is a tensor of rank $(n+1)$ obtained by applying to it the formal vector ∇ in such a way that the first index of the resulting tensor is the "derivation one". For example, for a tensor of rank 1, i.e. for a a vector $\mathbf{v}$, we have

$$(\text{grad } \mathbf{v})_{ij} = (\nabla \mathbf{v})_{ij} = \frac{\partial}{\partial x_i} v_j.$$

It should be noted that this convention is not universally adopted. Some authors prefer to indicate with the symbol grad $\mathbf{v}$ the quantity

$$(\text{grad } \mathbf{v})_{ij} = \frac{\partial}{\partial x_j} v_i.$$

The reader should therefore pay attention to the conventions used by each author before using the vector identities that are found in different books. For example, using our conventions, we have

$$\sum_i u_i \frac{\partial v_j}{\partial x_i} = (\mathbf{u} \cdot \text{grad } \mathbf{v})_j, \qquad \sum_i u_i \frac{\partial v_i}{\partial x_j} = \left[(\text{grad } \mathbf{v}) \cdot \mathbf{u} \right]_j.$$

Using the formal vector ∇, the quantities in the right-hand side can also be written, respectively, as

$$\left[(\mathbf{u} \cdot \nabla) \mathbf{v}) \right]_j, \qquad \left[(\nabla \mathbf{v}) \cdot \mathbf{u} \right]_j.$$

4. Given a tensor of rank n ($n \geq 1$), the curl (also known as rotor) of such a tensor is a tensor of the same rank n with the first component being obtained by saturating the first component of the given tensor with the completely antisymmetric tensor

(known as the Ricci, or Ricci-Levi Civita tensor) and with the component of the formal vector ∇. For example, for a vector $\mathbf{v}$ we have

$$(\text{rot}\,\mathbf{v})_i = \sum_{jk} \epsilon_{ijk} \frac{\partial v_k}{\partial x_j} = (\nabla \times \mathbf{v})_i,$$

and for a tensor $\mathbf{T}$ of rank two

$$(\text{rot}\,\mathbf{T})_{ij} = \sum_{kl} \epsilon_{ikl} \frac{\partial T_{lj}}{\partial x_k} = (\nabla \times \mathbf{T})_{ij}.$$

The antisymmetric tensor of rank three ϵ_{ijk}, introduced in these expressions, is defined by the equation $\epsilon_{ijk} = 0$ if at least two of the three indices i, j, k are equal; by the equation $\epsilon_{ijk} = 1$ if the ordered triad (i, j, k) is an even permutation of the fundamental triad $(1, 2, 3)$; and by the equation $\epsilon_{ijk} = -1$ if the ordered triad (i, j, k) is an odd permutation of the fundamental triad $(1, 2, 3)$. Ultimately, only 6 of the 27 components of the tensor are different from zero. Note that the usual vector product between two vectors can be conveniently expressed through the antisymmetric tensor. If $\mathbf{w} = \mathbf{u} \times \mathbf{v}$, we have

$$w_i = \sum_{jk} \epsilon_{ijk} u_j v_k.$$

Note also that the vector product operation and the curl operator (which involve the antisymmetric tensor) imply a choice about the chirality of the Cartesian orthogonal system in which the components of the vectors (and tensors) are defined. The convention that is now almost universally accepted (and that we use) is to choose a right-handed triad, i.e. to suppose that, if the axes x and y are directed respectively along the thumb and index finger of the right hand, the z axis is directed along the middle finger.

The antisymmetric tensor has a number of properties. The first concerns the permutation of its indices. For an even permutation the tensor remains unchanged, while for an odd permutation the tensor changes sign. In formulae

$$\epsilon_{ijk} = \epsilon_{jki} = \epsilon_{kij} = -\epsilon_{jik} = -\epsilon_{ikj} = -\epsilon_{kji}.$$

In addition, the following saturation properties hold

$$\sum_k \epsilon_{ijk} \epsilon_{lmk} = \delta_{il} \delta_{jm} - \delta_{im} \delta_{jl},$$

$$\sum_{jk} \epsilon_{ijk} \epsilon_{ljk} = 2\delta_{il},$$

$$\sum_{ijk} \epsilon_{ijk} \epsilon_{ijk} = 6,$$

where δ_{ij} is the so-called Kronecker delta, i.e. the symbol defined by

$$\delta_{ij} = 1 \quad \text{if } i = j, \qquad \delta_{ij} = 0 \quad \text{if } i \neq j.$$

The above definitions and properties can be used to obtain a number of vector identities that are listed below. In these equations, the quantities f and g are scalars, $\mathbf{a}$ and $\mathbf{b}$ are vectors, and $\mathbf{T}$ is a tensor of rank 2.

$$\bullet \quad \text{div}(f\mathbf{a}) = \mathbf{a} \cdot \text{grad } f + f \text{ div } \mathbf{a}. \qquad (16.4)$$

In fact we have

$$\text{div}(f\mathbf{a}) = \sum_i \frac{\partial}{\partial x_i}(fa_i) = \sum_i a_i \frac{\partial f}{\partial x_i} + f \sum_i \frac{\partial a_i}{\partial x_i}.$$

$$\bullet \quad \text{grad}(fg) = g \text{ grad } f + f \text{ grad } g. \qquad (16.5)$$

In fact we have, for the i-th component

$$\left[\text{grad}(fg)\right]_i = \frac{\partial}{\partial x_i}(fg) = g\frac{\partial f}{\partial x_i} + f\frac{\partial g}{\partial x_i}.$$

$$\bullet \quad \text{rot}(f\mathbf{a}) = \text{grad } f \times \mathbf{a} + f \text{ rot } \mathbf{a}. \qquad (16.6)$$

In fact we have, for the i-th component

$$\left[\text{rot}(f\mathbf{a})\right]_i = \sum_{jk} \epsilon_{ijk}\frac{\partial}{\partial x_j}(fa_k) = \sum_{jk} \epsilon_{ijk}\left[\left(\frac{\partial f}{\partial x_j}\right)a_k + f\frac{\partial a_k}{\partial x_j}\right] =$$

$$= \left[(\text{grad } f) \times \mathbf{a}\right]_i + f[\text{rot}\,\mathbf{a}]_i.$$

$$\bullet \quad \text{div}(\mathbf{a} \times \mathbf{b}) = \mathbf{b} \cdot \text{rot}\,\mathbf{a} - \mathbf{a} \cdot \text{rot}\,\mathbf{b}. \qquad (16.7)$$

In fact we have

$$\text{div}(\mathbf{a} \times \mathbf{b}) = \sum_i \frac{\partial}{\partial x_i}\left(\sum_{jk}\epsilon_{ijk}a_j b_k\right) = \sum_{ijk}\epsilon_{ijk}\left[\left(\frac{\partial a_j}{\partial x_i}\right)b_k + a_j\left(\frac{\partial b_k}{\partial x_i}\right)\right]$$

$$= \sum_{ijk} b_k\epsilon_{kij}\frac{\partial a_j}{\partial x_i} - \sum_{ijk} a_j\epsilon_{jik}\frac{\partial b_k}{\partial x_i} = \sum_k b_k(\text{rot}\,\mathbf{a})_k - \sum_j a_j(\text{rot}\,\mathbf{b})_j.$$

$$\bullet \quad \text{grad}(\mathbf{a} \cdot \mathbf{b}) = (\text{grad}\,\mathbf{a}) \cdot \mathbf{b} + (\text{grad}\,\mathbf{b}) \cdot \mathbf{a}. \qquad (16.8)$$

In fact we have, for the i-th component

$$\left[\text{grad}(\mathbf{a} \cdot \mathbf{b})\right]_i = \frac{\partial}{\partial x_i}\left(\sum_j a_j b_j\right) = \sum_j\left(\frac{\partial a_j}{\partial x_i}\right)b_j + \sum_j a_j\left(\frac{\partial b_j}{\partial x_i}\right)$$

$$= \left[(\text{grad}\,\mathbf{a}) \cdot \mathbf{b}\right]_i + \left[(\text{grad}\,\mathbf{b}) \cdot \mathbf{a}\right]_i.$$

- rot$(\mathbf{a} \times \mathbf{b}) = \mathbf{b} \cdot \mathrm{grad}\,\mathbf{a} - \mathbf{a} \cdot \mathrm{grad}\,\mathbf{b} + \mathbf{a}\,\mathrm{div}\,\mathbf{b} - \mathbf{b}\,\mathrm{div}\,\mathbf{a}.$ (16.9)

In fact we have, for the i-th component

$$\left[\mathrm{rot}(\mathbf{a} \times \mathbf{b})\right]_i = \sum_{jk}\epsilon_{ijk}\frac{\partial}{\partial x_j}(\mathbf{a} \times \mathbf{b})_k = \sum_{jklm}\epsilon_{ijk}\epsilon_{klm}\frac{\partial}{\partial x_j}(a_l b_m)$$

$$= \sum_{jlm}(\delta_{il}\delta_{jm} - \delta_{im}\delta_{jl})\left[\left(\frac{\partial a_l}{\partial x_j}\right)b_m + a_l\frac{\partial b_m}{\partial x_j}\right]$$

$$= \sum_{ij}\left(b_j\frac{\partial a_i}{\partial x_j} - b_i\frac{\partial a_j}{\partial x_j} + a_i\frac{\partial b_j}{\partial x_j} - a_j\frac{\partial b_i}{\partial x_j}\right)$$

$$= [\mathbf{b} \cdot \mathrm{grad}\,\mathbf{a}]_i - b_i\,\mathrm{div}\,\mathbf{a} + a_i\,\mathrm{div}\,\mathbf{b} - [\mathbf{a} \cdot \mathrm{grad}\,\mathbf{b}]_i.$$

- grad$(f\mathbf{a}) = (\mathrm{grad}\,f)\mathbf{a} + f\,\mathrm{grad}\,\mathbf{a}.$ (16.10)

In fact we have, for the ij-th component

$$\left[\mathrm{grad}(f\mathbf{a})\right]_{ij} = \frac{\partial}{\partial x_i}(fa_j) = \left(\frac{\partial f}{\partial x_i}\right)a_j + f\frac{\partial a_j}{\partial x_i} = (\mathrm{grad}\,f)_i a_j + f(\mathrm{grad}\,\mathbf{a})_{ij}.$$

- div$(\mathbf{ab}) = \mathbf{b}\,\mathrm{div}\,\mathbf{a} + \mathbf{a} \cdot \mathrm{grad}\,\mathbf{b}.$ (16.11)

In fact we have, for the i-th component

$$\left[\mathrm{div}(\mathbf{ab})\right]_i = \sum_j\frac{\partial}{\partial x_j}(a_j b_i) = \sum_j\left[\left(\frac{\partial a_j}{\partial x_j}\right)b_i + a_j\frac{\partial b_i}{\partial x_j}\right]$$

$$= b_i\,\mathrm{div}\,\mathbf{a} + [\mathbf{a} \cdot \mathrm{grad}\,\mathbf{b}]_i.$$

- $\mathbf{a} \times \mathrm{rot}\,\mathbf{b} = (\mathrm{grad}\,\mathbf{b}) \cdot \mathbf{a} - \mathbf{a} \cdot \mathrm{grad}\,\mathbf{b}.$ (16.12)

In fact we have, for the i-th component

$$[\mathbf{a} \times \mathrm{rot}\,\mathbf{b}]_i = \sum_{jk}\epsilon_{ijk}a_j(\mathrm{rot}\,\mathbf{b})_k = \sum_{jklm}\epsilon_{ijk}\epsilon_{klm}a_j\frac{\partial b_m}{\partial x_l}$$

$$= \sum_{jlm}(\delta_{il}\delta_{jm} - \delta_{im}\delta_{jl})a_j\frac{\partial b_m}{\partial x_l}$$

$$= \sum_j\left(a_j\frac{\partial b_j}{\partial x_i} - a_j\frac{\partial b_i}{\partial x_j}\right) = \left[(\mathrm{grad}\,\mathbf{b}) \cdot \mathbf{a}\right]_i - [\mathbf{a} \cdot \mathrm{grad}\,\mathbf{b}]_i.$$

- div$(f\mathbf{T}) = (\mathrm{grad}\,f) \cdot \mathbf{T} + f\,\mathrm{div}\,\mathbf{T}.$ (16.13)

In fact we have, for the i-th component

$$\left[\mathrm{div}(f\mathbf{T})\right]_i = \sum_j \frac{\partial}{\partial x_j}(f T_{ji}) = \sum_j \left[\left(\frac{\partial f}{\partial x_j}\right)T_{ji} + f\frac{\partial T_{ji}}{\partial x_j}\right]$$

$$= \left[(\mathrm{grad}\, f)\cdot\mathbf{T}\right]_i + f(\mathrm{div}\,\mathbf{T})_i.$$

- $\mathrm{rot}(\mathrm{rot}\,\mathbf{a}) = \mathrm{grad}\,\mathrm{div}\,\mathbf{a} - \nabla^2\mathbf{a}.$ (16.14)

In fact we have, for the i-th component

$$\left[\mathrm{rot}(\mathrm{rot}\,\mathbf{a})\right]_i = \sum_{jk}\epsilon_{ijk}\frac{\partial}{\partial x_j}(\mathrm{rot}\,\mathbf{a})_k = \sum_{jklm}\epsilon_{ijk}\epsilon_{klm}\frac{\partial}{\partial x_j}\frac{\partial}{\partial x_l}a_m$$

$$= \sum_{jlm}(\delta_{il}\delta_{jm} - \delta_{im}\delta_{jl})\frac{\partial^2 a_m}{\partial x_j\partial x_l}$$

$$= \sum_j\left(\frac{\partial^2 a_j}{\partial x_j\partial x_i} - \frac{\partial^2 a_i}{\partial x_j\partial x_j}\right) = [\mathrm{grad}\,\mathrm{div}\,\mathbf{a}]_i - \left[\nabla^2\mathbf{a}\right]_i.$$

There are also other vector identities that apply only in integral form. They result from the theorems of Gauss and Stokes-Ampère, which we recall now.

Gauss theorem: If Σ is a closed surface enclosing the volume V and if $\mathbf{n}$ is the normal external to the surface, Gauss theorem is expressed by the equation

$$\bullet \quad \int_\Sigma \mathbf{a}\cdot\mathbf{n}\,\mathrm{d}S = \int_V \mathrm{div}\,\mathbf{a}\,\mathrm{d}V,$$

where $\mathbf{a}$ is an arbitrary vector that is a function of the position.

Stokes-Ampère theorem: if ℓ is a closed circuit and if Σ is a surface that is leaning on this circuit, the Stokes-Ampère theorem is stated by the equation

$$\bullet \quad \oint_\ell \mathbf{a}\cdot\mathrm{d}\boldsymbol{\ell} = \int_\Sigma \mathrm{rot}\,\mathbf{a}\cdot\mathbf{n}\,\mathrm{d}S,$$

where $\mathbf{n}$ is the normal external to the surface. We note that the validity of this equation implies a convention about the direction of integration along the circuit, which in turn depends on the implicit convention in the definition of the curl operator. When the (x, y, z) system used to define the vector components is a right-handed system, then the direction of integration along the circuit follows the corkscrew (or the right screw) rule, for which the direction of $\mathbf{n}$ coincides with the direction of advancement of the corkscrew.

Various identities can be obtained from the Gauss and Stokes-Ampère theorems. Some of them are collected below.

$$\bullet \quad \oint_\ell f\,\mathrm{d}\boldsymbol{\ell} = \int_\Sigma \mathbf{n}\times\mathrm{grad}\,f\,\mathrm{d}S.$$

This identity can be proven by noting that, if **c** is an arbitrary constant vector, we have

$$\mathbf{c} \cdot \oint_{\ell} f \, \mathrm{d}\boldsymbol{\ell} = \oint_{\ell} (f\mathbf{c}) \cdot \mathrm{d}\boldsymbol{\ell},$$

and, applying the Stokes-Ampère theorem

$$\mathbf{c} \cdot \oint_{\ell} f \, \mathrm{d}\boldsymbol{\ell} = \int_{\Sigma} \mathrm{rot}(f\mathbf{c}) \cdot \mathbf{n} \, \mathrm{d}S.$$

Recalling the vector identity of Eq. (16.6), and taking into account that **c** is a constant vector, we have

$$\mathbf{c} \cdot \oint_{\ell} f \, \mathrm{d}\boldsymbol{\ell} = \int_{\Sigma} \left[(\mathrm{grad}\, f) \times \mathbf{c} \right] \cdot \mathbf{n} \, \mathrm{d}S = \mathbf{c} \cdot \int_{\Sigma} \mathbf{n} \times \mathrm{grad}\, f \, \mathrm{d}S.$$

The identity therefore follows, because **c** is an arbitrary vector.

With entirely similar procedures and taking into account the vector identities demonstrated previously, we obtain the additional identities

$$\bullet \quad \oint_{\ell} \mathbf{a} \times \mathrm{d}\boldsymbol{\ell} = \int_{\Sigma} \left[\mathbf{n} \, \mathrm{div}\, \mathbf{a} - (\mathrm{grad}\, \mathbf{a}) \cdot \mathbf{n} \right] \mathrm{d}S.$$

$$\bullet \quad \int_{\Sigma} \mathbf{n} \times \mathbf{a} \, \mathrm{d}S = \int_{V} \mathrm{rot}\, \mathbf{a} \, \mathrm{d}V.$$

$$\bullet \quad \int_{\Sigma} f\mathbf{n} \, \mathrm{d}S = \int_{V} \mathrm{grad}\, f \, \mathrm{d}V.$$

In particular, if we put $f = 1$ in this last identity, we get

$$\bullet \quad \int_{\Sigma} \mathbf{n} \, \mathrm{d}S = 0,$$

which is an important geometrical relation valid for an arbitrary closed surface.

16.3 The Dirac Delta Function

The Dirac delta function, traditionally indicated by the symbol $\delta(x)$, can be thought of as a function which is null for any value of x, except for an infinitesimal interval centered at the origin where the function has a very high peak which tends to infinity, but such that the integral of the function in $\mathrm{d}x$ is equal to 1. Obviously, it is not a function in strict mathematical sense, but can be thought of as the limit of a family of functions depending on a suitable parameter. For example, if we consider the family of functions $f(x, a)$

$$f(x, a) = \begin{cases} \frac{1}{a} & \text{for } |x| \leq \frac{a}{2}, \\ 0 & \text{for } |x| > \frac{a}{2}, \end{cases}$$

we have that

$$\delta(x) = \lim_{a \to 0} f(x, a).$$

Similarly, if we consider the family

$$g(x, a) = \frac{1}{\sqrt{2\pi a}} e^{-(x/a)^2},$$

we also have

$$\delta(x) = \lim_{a \to 0} g(x, a).$$

There are endless possibilities to represent the Dirac delta as the limit of suitable families of functions. The most common representations in mathematical physics are the following ones

$$\delta(x) = \lim_{\Omega \to \infty} \frac{1}{\pi} \frac{\sin(\Omega x)}{x},$$

$$\delta(x) = \lim_{\Omega \to \infty} \frac{1}{\pi} \frac{\sin^2(\Omega x)}{\Omega x^2}.$$

The fundamental property of the Dirac delta is summarised in the following expression, which constitutes its formal definition

$$\int_{-\infty}^{\infty} F(x)\delta(x)\,\mathrm{d}x = F(0),$$

and from which, by means of simple changes of variable, the following two relations are found

$$\int_{-\infty}^{\infty} F(x)\delta(x - x_0)\,\mathrm{d}x = F(x_0),$$

$$\int_{-\infty}^{\infty} F(x)\delta(ax)\,\mathrm{d}x = \frac{1}{|a|} F(0),$$

where a is any real number different from zero. From these equations we can get an important generalisation concerning the Dirac delta whose argument is an arbitrary real function $g(x)$. Denoting this quantity by the symbol $\delta[g(x)]$ and denoting by x_i the zeroes (if any) of the function $g(x)$, we have

$$\int_{-\infty}^{\infty} F(x)\delta\big[g(x)\big]\,\mathrm{d}x = \sum_i \frac{1}{|g'(x_i)|} F(x_i),$$

where $g'(x)$ is the derivative of the function $g(x)$ with respect to its argument. Further generalisations to the case of the three-dimensional Dirac delta are described directly in the text (see Sect. 3.2).

Finally, we can give a meaning to the derivative of the Dirac delta function, $\delta'(x)$, defined by the usual relation

$$\delta'(x) = \lim_{\Delta x \to 0} \frac{\delta(x + \Delta x) - \delta(x)}{\Delta x}.$$

Using this definition we have, for an arbitrary function $F(x)$

$$\int_{-\infty}^{\infty} F(x)\delta'(x)\,dx = \lim_{\Delta x \to 0} \int_{-\infty}^{\infty} F(x)\frac{\delta(x + \Delta x) - \delta(x)}{\Delta x}\,dx,$$

from which we obtain

$$\int_{-\infty}^{\infty} F(x)\delta'(x)\,dx = \lim_{\Delta x \to 0} \frac{F(-\Delta x) - F(0)}{\Delta x} = -F'(0).$$

16.4 Recovering the Elementary Laws of Electromagnetism

In Chap. 3, starting from the Liénerd and Wiechart potentials, we calculated the expressions of the electric and magnetic field at an arbitrary point in space, due to a single moving charge. The results are contained in Eqs. (3.19) and (3.20). We are now going to show how the basic equations of electromagnetism valid for stationary phenomena can be derived from these equations in the non-relativistic limit. The purpose of this appendix is a simple consistency check, since it is obvious that the equations from which we start, being a consequence of Maxwell's equations, must already contain those results that, even historically, are the basis of Maxwell's equations themselves.

Consider a particle with electric charge e, moving within an electric conductor having a constant transverse section. Its velocity is much lower than the velocity of light. To fix ideas, we can think that the velocity is of the order of 10^{-2} cm s^{-1}, which represents the order of magnitude of the drift velocities of electrons inside a conductor in a typical macroscopic electric circuit. The corresponding value of β is of the order of 10^{-12}, so that the approximation $\beta^2 \ll 1$ is certainly verified. Furthermore, the effects of the curvature of the conductor (causing very small accelerations) can certainly be neglected so that we can assume that the electric field is given only by the Coulomb term of Eq. (3.19). Neglecting terms of the order of β^2, such field is written in the form

$$\mathbf{E}(\mathbf{r}, t) = \frac{e}{\kappa^3 R^2}(\mathbf{n} - \boldsymbol{\beta}),$$

where κ, R, $\mathbf{n}$ are the quantities introduced in Chap. 3 and that need to be calculated at the retarded time t'. The magnetic field is then given by Eq. (3.20), i.e.

$$B(\mathbf{r}, t) = \mathbf{n} \times \mathbf{E}(\mathbf{r}, t).$$
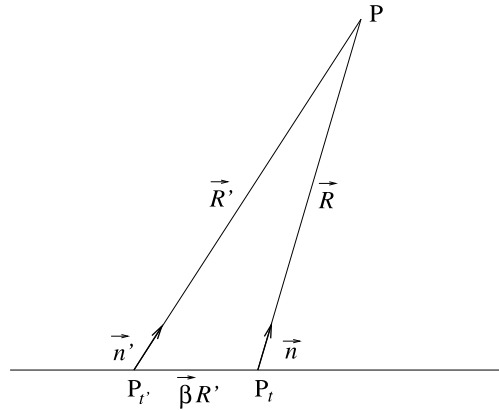
We can immediately notice that, if we put $\beta = 0$ (i.e. we consider an electric
charge at rest), obviously we do not need to consider the difference between real
time and retarded time, so we obtain, being $\kappa = 1$

$$\mathbf{E}(\mathbf{r}) = \frac{e\mathbf{n}}{R^2}, \qquad \mathbf{B}(\mathbf{r}) = 0.$$

These are the ordinary equations of electrostatics which represent, in terms of fields,
Coulomb's law.

We are now going to see what we get at first order in β. With simple consid-
erations it can be shown that the electric field $E(\mathbf{r}, t)$ is exactly equal to what one
would calculate using Coulomb's law and assuming, hypothetically, that the velocity
of light were infinite (i.e. neglecting the difference between real and retarded time).
In fact, referring to Fig. 16.1 and denoting by a single quote the quantities measured
at the retarded time t' and without superscript the same quantities at time t, we have

$$t' = t - \frac{R'}{c}, \qquad \mathbf{R}' = \mathbf{R} + (t - t')\mathbf{v} = \mathbf{R} + R'\boldsymbol{\beta},$$

from which it follows, dividing by R'

$$\mathbf{n}' - \boldsymbol{\beta} = \frac{\mathbf{R}}{R'}. \tag{16.15}$$

Introducing the new notations in the expression for the electric field and recalling
that $\boldsymbol{\beta}$ is constant we obtain

$$\mathbf{E}(\mathbf{r}, t) = \frac{e}{\kappa'^3 R'^2}(\mathbf{n}' - \boldsymbol{\beta}) = \frac{e\mathbf{R}}{\kappa'^3 R'^3}.$$

On the other hand we have by definition that

$$\kappa' = 1 - \boldsymbol{\beta} \cdot \mathbf{n}',$$

and applying Carnot's theorem to the triangle $PP_{t'}P_t$

$$R = R'\sqrt{1 - 2\boldsymbol{\beta} \cdot \mathbf{n'} + \beta^2}. \tag{16.16}$$

Substituting in the expression of the electric field, we obtain the result we anticipated. In fact we obtain, apart from terms of the order of β^2

$$\mathbf{E}(\mathbf{r}, t) = \frac{e\mathbf{R}(1 - 2\boldsymbol{\beta} \cdot \mathbf{n'} + \beta^2)^{3/2}}{R^3(1 - \boldsymbol{\beta} \cdot \mathbf{n'})^3} \simeq \frac{e\mathbf{n}}{R^2}.$$

It remains to evaluate the contribution of the magnetic field. We have

$$\mathbf{B}(\mathbf{r}, t) = \mathbf{n'} \times \mathbf{E}(\mathbf{r}, t) = \frac{e\mathbf{n'} \times \mathbf{n}}{R^2}.$$

On the other hand, also apart from terms of the order of β^2, we have, from Eqs. (16.15) and (16.16)

$$\mathbf{n'} = \boldsymbol{\beta} + \mathbf{n}(1 - \boldsymbol{\beta} \cdot \mathbf{n}),$$

so that

$$\mathbf{B}(\mathbf{r}, t) = \frac{e\boldsymbol{\beta} \times \mathbf{n}}{R^2}.$$

Now we apply this equation to the case of an element of a conductor, of length $d\boldsymbol{\ell}$. Denoting by N the number density of the moving charged particles and with S the transverse section of the conductor, the element contains a number of particles given by $NS\,d\boldsymbol{\ell}$ with velocity $\mathbf{v} = c\boldsymbol{\beta}$ parallel to $d\boldsymbol{\ell}$. There is an equal number of fixed particles of opposite charge, so the resulting electric field is null for the property previously demonstrated. For the magnetic field we have instead

$$B(\mathbf{r}, t) = e\frac{NSv}{c}d\boldsymbol{\ell} \times \frac{\mathbf{n}}{R^2}.$$

On the other hand, if we denote by i the intensity of the current flowing in the conductor

$$i = eNSv,$$

so that the equation for the magnetic field is written

$$B(\mathbf{r}, t) = \frac{i}{c}d\boldsymbol{\ell} \times \frac{\mathbf{n}}{R^2}.$$

This is just the Biot and Savart law expressing the magnetic field generated by a current element. As is clear from our deduction, although electric charges move within the conductor at very low speed, they are nevertheless able to create a relativistic effect which is manifested by the presence of the magnetic field.

16.5 The Relativistic Larmor Equation

Within the radiation zone, Eqs. (3.18) and (3.20) provide the expressions for the
electric and magnetic field due to a moving charge

$$\mathbf{E}(\mathbf{r}, t) = \frac{e}{c^2 \kappa^3 R} \mathbf{n} \times \big[(\mathbf{n} - \boldsymbol{\beta}) \times \mathbf{a} \big], \qquad \mathbf{B}(\mathbf{r}, t) = \mathbf{n} \times \mathbf{E}(\mathbf{r}, t),$$

where e is the value of the electric charge, c is the speed of light, $\mathbf{n}$ is the unit
vector along the direction of $\mathbf{R}$, the vector that goes from the charge to the point of
coordinates $\mathbf{r}$, $\boldsymbol{\beta} = \mathbf{v}/c$ is the velocity of the charge in units of the speed of light, $\mathbf{a}$
is the acceleration, and κ is defined by the equation

$$\kappa = 1 - \mathbf{n} \cdot \boldsymbol{\beta}.$$

We recall that the quantities R, κ, $\mathbf{n}$, $\boldsymbol{\beta}$, and $\mathbf{a}$ that appear in the previous equations
must be evaluated at the retarded time t', related to the time t by the equation

$$t' = t - \frac{R}{c}.$$

Expanding the double vector product, we obtain

$$\mathbf{E}(\mathbf{r}, t) = \frac{e}{c^2 \kappa^3 R} \big[(\mathbf{n} \cdot \mathbf{a})(\mathbf{n} - \boldsymbol{\beta}) - \kappa \mathbf{a} \big].$$

On the other hand, we know that the Poynting vector is given by

$$\mathbf{S}(\mathbf{r}, t) = \frac{c}{4\pi} E^2(\mathbf{r}, t) \mathbf{n},$$

and expanding the square of the electric field we obtain with simple algebra

$$\mathbf{S}(\mathbf{r}, t) = \frac{e^2}{4\pi c^3 R^2} \left[\frac{a^2}{\kappa^4} + 2 \frac{(\mathbf{n} \cdot \mathbf{a})(\boldsymbol{\beta} \cdot \mathbf{a})}{\kappa^5} - \frac{(1 - \beta^2)(\mathbf{n} \cdot \mathbf{a})^2}{\kappa^6} \right] \mathbf{n}.$$

This expression shows that, in the general case, the angular distribution of the emit-
ted radiation (i.e. the radiation diagram) is quite complex. The special cases where
the acceleration is either parallel or perpendicular to the velocity have been dis-
cussed in the text. Here, it is sufficient to emphasize the fact that, for any velocity
and acceleration, there are always two directions where the Poynting vector is zero.
This can be shown simply from the expression of the electric field. The electric field
is obviously zero along the directions characterised by those unit vectors $\mathbf{n}_0$ such
that the vector $\mathbf{n}_0 - \boldsymbol{\beta}$ is parallel to the vector $\mathbf{a}$. The same holds for the Poynting
vector. The directions $\mathbf{n}_0$ are then contained in the plane defined by the vectors $\boldsymbol{\beta}$
and $\mathbf{a}$, and are given by the solutions of the equation

$$(\mathbf{n}_0 - \boldsymbol{\beta}) \times \mathbf{a} = 0.$$

Denoting by α the angle between the velocity and the acceleration vectors, the unit vectors $\mathbf{n}_0$ are defined by the angles $\theta_\pm$ (which start from the acceleration vector and increase in the same direction as α) given by

$$\theta_+ = \arcsin(\beta \sin \alpha), \qquad \theta_- = \pi - \arcsin(\beta \sin \alpha).$$

For example, if $\alpha = 45°$ and $\beta = 0.8$, we have $\theta_+ = 34°.45$ and $\theta_- = 145°.55$.

Let us now move on to the calculation of the power. We note that if the integral

$$\mathcal{I} = \oint \mathbf{S}(\mathbf{r}, t) \cdot \mathbf{n} R^2 \, d\Omega, \tag{16.17}$$

were simply executed over a sphere of radius R centred on the position of the charge at the retarded time $(t - R/c)$, we would obtain the ratio between the energy that flows across the sphere in a time interval dt and the dt itself. This quantity is however of not much interest. It is more interesting to obtain the power emitted by the charged particle. In order to do this, we need to take into account the fact that the energy that flows across the sphere in a time dt was emitted by the particle in the time dt' which depends on the direction and is related to dt by

$$dt = \kappa \, dt'.$$

To find the power W emitted by the charged particle we therefore need to calculate the integral

$$W = \oint \mathbf{S}(\mathbf{r}, t) \cdot \mathbf{n} \frac{dt}{dt'} R^2 \, d\Omega = \oint \mathbf{S}(\mathbf{r}, t) \cdot \mathbf{n} \kappa R^2 \, d\Omega.$$

Substituting the above expression of the Poynting vector, we find

$$W = \frac{e^2}{4\pi c^3} \oint \left[\frac{a^2}{\kappa^3} + 2 \frac{(\mathbf{n} \cdot \mathbf{a})(\boldsymbol{\beta} \cdot \mathbf{a})}{\kappa^4} - \frac{(1 - \beta^2)(\mathbf{n} \cdot \mathbf{a})^2}{\kappa^5} \right] d\Omega.$$

To calculate this integral, we introduce a system of polar coordinates (ψ, χ) with the polar axis directed along the velocity vector and the azimuth χ measured from the plane containing the velocity and the acceleration. With obvious notations, the three vectors $\boldsymbol{\beta}$, $\mathbf{a}$, and $\mathbf{n}$ in this system of coordinates are given by

$$\boldsymbol{\beta} = \beta \mathbf{k}, \qquad \mathbf{a} = a_\perp \mathbf{i} + a_\| \mathbf{k}, \qquad \mathbf{n} = \sin \psi \cos \chi \mathbf{i} + \sin \psi \sin \chi \mathbf{j} + \cos \psi \mathbf{k},$$

so that the integrand can be written in the form

$$\frac{a_\|^2 + a_\perp^2}{(1 - \beta \cos \psi)^3} + 2\beta a_\| \frac{\sin \psi \cos \chi a_\perp + \cos \psi a_\|}{(1 - \beta \cos \psi)^4}$$

$$- \left(1 - \beta^2\right) \frac{\sin^2 \psi \cos^2 \chi a_\perp^2 + 2 \sin \psi \cos \psi \cos \chi a_\perp a_\| + \cos^2 \psi a_\|^2}{(1 - \beta \cos \psi)^5},$$

and $d\Omega$ is given by $\sin\psi\,d\psi\,d\chi$. By integrating in $d\chi$ within the interval $(0, 2\pi)$, the factors that do not contain any function of χ result in 2π, those containing $\cos\chi$ produce zero, while the factor containing $\cos^2\chi$ gives π. In summary we have

$$W = \frac{e^2}{2c^3}\int_0^\pi \left[\frac{a_\parallel^2 + a_\perp^2}{(1-\beta\cos\psi)^3} + \frac{2\beta\cos\psi\,a_\parallel^2}{(1-\beta\cos\psi)^4} \right.$$
$$\left. - (1-\beta^2)\frac{\frac{1}{2}\sin^2\psi\,a_\perp^2 + \cos^2\psi\,a_\parallel^2}{(1-\beta\cos\psi)^5}\right]\sin\psi\,d\psi.$$

The integrals in $d\psi$ appearing in this expression are simple and can be evaluated either by integrating by parts or by changing the integration variable from ψ to $x = 1 - \beta\cos\psi$. We obtain

$$\frac{1}{2}\int_0^\pi \frac{1}{(1-\beta\cos\psi)^3}\sin\psi\,d\psi = \frac{1}{(1-\beta^2)^2},$$

$$\frac{1}{2}\int_0^\pi \frac{\cos\psi}{(1-\beta\cos\psi)^4}\sin\psi\,d\psi = \frac{4}{3}\frac{\beta}{(1-\beta^2)^3},$$

$$\frac{1}{2}\int_0^\pi \frac{\sin^2\psi}{(1-\beta\cos\psi)^5}\sin\psi\,d\psi = \frac{2}{3}\frac{1}{(1-\beta^2)^3},$$

$$\frac{1}{2}\int_0^\pi \frac{\cos^2\psi}{(1-\beta\cos\psi)^5}\sin\psi\,d\psi = \frac{1}{3}\frac{1+5\beta^2}{(1-\beta^2)^4}.$$

Substituting these expressions and grouping separately the terms in $a_\parallel^2$ and in $a_\perp^2$, we get

$$W = \frac{e^2}{2c^3}\left\{a_\parallel^2\left[\frac{1}{(1-\beta^2)^2} + \frac{8}{3}\frac{\beta^2}{(1-\beta^2)^3} - \frac{1}{3}\frac{1+5\beta^2}{(1-\beta^2)^3}\right]\right.$$
$$\left. + a_\perp^2\left[\frac{1}{(1-\beta^2)^2} - \frac{1}{3}\frac{1}{(1-\beta^2)^2}\right]\right\},$$

or, expanding,

$$W = \frac{2e^2}{3c^3}\left[\frac{a_\parallel^2}{(1-\beta^2)^3} + \frac{a_\perp^2}{(1-\beta^2)^2}\right].$$

Recalling the definition of the relativistic factor γ

$$\gamma = \frac{1}{\sqrt{1-\beta^2}},$$

the expression for the power emitted by a relativistic charge in accelerated motion can also be written in the more representative form

$$W = \frac{2e^2}{3c^3}\left(\gamma^6 a_\parallel^2 + \gamma^4 a_\perp^2\right).$$

This formula is a generalisation of the Larmor equation (3.23) to the relativistic case. Obviously, for $\gamma = 1$ we find Larmor equation since $a_\parallel^2 + a_\perp^2 = a^2$.

To conclude, we note that, if we had executed the integral of the Poynting vector on the sphere without taking into account the difference between the dt and the dt' (i.e. the integral $\mathcal{I}$ of Eq. (16.17)), we would have obviously obtained a different expression. Taking into account that

$$\frac{1}{2} \int_0^\pi \frac{1}{(1 - \beta \cos \psi)^4} \sin \psi \, d\psi = \frac{1}{3} \frac{3 + \beta^2}{(1 - \beta^2)^3},$$

$$\frac{1}{2} \int_0^\pi \frac{\cos \psi}{(1 - \beta \cos \psi)^5} \sin \psi \, d\psi = \frac{1}{3} \frac{\beta(5 + \beta^2)}{(1 - \beta^2)^4},$$

$$\frac{1}{2} \int_0^\pi \frac{\sin^2 \psi}{(1 - \beta \cos \psi)^6} \sin \psi \, d\psi = \frac{2}{15} \frac{5 + \beta^2}{(1 - \beta^2)^4},$$

$$\frac{1}{2} \int_0^\pi \frac{\cos^2 \psi}{(1 - \beta \cos \psi)^6} \sin \psi \, d\psi = \frac{1}{15} \frac{5 + 38\beta^2 + 5\beta^4}{(1 - \beta^2)^5},$$

we have, in fact, that

$$\mathcal{I} = \frac{2e^2}{3c^3} \left[\gamma^8 \left(1 + \frac{1}{5}\beta^2 \right) a_\parallel^2 + \gamma^6 \left(1 + \frac{2}{5}\beta^2 \right) a_\perp^2 \right].$$

This difference between the power emitted by the particle (W) and the power received on the sphere ($\mathcal{I}$) is a simple kinematic effect and has nothing to do with relativity. A similar effect occurs in the case of acoustic waves emitted, for example, by an airplane travelling at a speed close to the velocity of sound. While the power emitted by the plane into acoustic waves is fixed, the received power can be very large and, at the limit, almost infinite if the plane travels for a long time at exactly the speed of sound (the so-called sonic bang is precisely due to this phenomenon).

16.6 Gravitational Waves

The equations that we have obtained for the radiation of electromagnetic waves can also be applied, with some slight modifications, to treat gravitational radiation. Obviously, this is not rigorous, since the laws of gravitational radiation should be derived from the general theory of relativity. The approach followed here is however sufficient to describe the fundamental properties of the mechanisms for the generation of gravitational waves and leads to formulae that are substantially correct (as can be verified a posteriori).

We start by performing a formal transformation to the equations for electromagnetic radiation described in Sect. 3.10

$$e_i \rightarrow m_i \quad (i = 1, \ldots, N),$$

i.e. we replace, for each particle, the charge with the mass. Furthermore, in the equations that express the Poynting vector (i.e. in those which express the radiated power), we multiply the right-hand side by the universal gravitational constant G. We note, incidentally, that in these equations the dimensional factor $[e^2]$ is replaced by the dimensional factor $[Gm^2]$ having the same dimensions. The various quantities introduced in Sect. 3.10, i.e. the electric dipole moment **D** (Eq. (3.31)), the magnetic dipole moment **M** (Eq. (3.32)), and the symmetric tensor of order two (related to the electric quadrupole moment) $\mathcal{Q}$ (Eq. (3.33)) are transformed in as many quantities for which we use, respectively, the symbols $\mathbf{D_G}$, $\mathbf{M_G}$, and $\mathcal{Q}_G$, i.e.

$$\mathbf{D} = \sum_{i=1}^{N} e_i \mathbf{s}_i \rightarrow \mathbf{D_G} = \sum_{i=1}^{N} m_i \mathbf{s}_i,$$

$$\mathbf{M} = \frac{1}{2c} \sum_{i=1}^{N} e_i \mathbf{s}_i \times \mathbf{v}_i \rightarrow \mathbf{M_G} = \frac{1}{2c} \sum_{i=1}^{N} m_i \mathbf{s}_i \times \mathbf{v}_i,$$

$$\mathcal{Q} = \sum_{i=1}^{N} e_i \mathbf{s}_i \mathbf{s}_i \rightarrow \mathcal{Q}_G = \sum_{i=1}^{N} m_i \mathbf{s}_i \mathbf{s}_i.$$

We now note that the quantity $\mathbf{D_G}$, the analogous of the electric dipole, is, by definition, the coordinate of the centre of mass of the system of N particles $\mathbf{r_G}$ multiplied by the total mass. We therefore have

$$\mathbf{D_G} = \sum_{i=1}^{N} m_i \mathbf{s}_i = \mathcal{M} \mathbf{r_G},$$

where

$$\mathcal{M} = \sum_{i=1}^{N} m_i.$$

We then have, for an isolated system,

$$\ddot{\mathbf{D}}_G = \mathcal{M} \frac{d^2}{dt^2} \mathbf{r_G} = 0.$$

Furthermore, the analogous of the magnetic dipole, the quantity $\mathbf{M_G}$, is proportional to the total angular momentum of the system **J**, since

$$2c\mathbf{M_G} = \sum_{i=1}^{N} m_i \mathbf{s}_i \times \mathbf{v}_i = \mathbf{J}.$$

We then obtain, for an isolated system,

$$\dot{\mathbf{M}}_G = \frac{d}{dt} \mathbf{J} = 0,$$

and therefore also

$$\ddot{\mathbf{M}}_G = 0.$$

According to our analogy, we can then conclude that for gravitational waves there is neither the analogue of electric dipole radiation, nor the analogue of magnetic dipole radiation. It therefore only remains the analogue of the electric quadrupole radiation (in addition, obviously, to the radiation due to higher multipoles). The tensor $\mathcal{Q}_G$ is traditionally denoted by the symbol $\mathcal{I}$, because it is essentially an inertia tensor. It is however not to be confused with the ordinary inertia tensor $\mathcal{I}$ that is used to describe the dynamics of a rigid body, and that is defined as

$$\mathcal{I} = \sum_{i=1}^{N} m_i \left(s_i^2 \mathbf{U} - \mathbf{s}_i \mathbf{s}_i \right),$$

where $\mathbf{U}$ is the unitary tensor. We have, obviously,

$$\mathcal{I} = -\mathcal{I} + \frac{1}{2}(\mathrm{Tr}\,\mathcal{I})\mathbf{U},$$

since, recalling the definition of the trace of a tensor

$$\mathrm{Tr}\,\mathcal{I} = \sum_i m_i \left(3s_i^2 - x_i^2 - y_i^2 - z_i^2 \right) = 2 \sum_i m_i s_i^2.$$

The two tensors $\mathcal{I}$ and $\mathcal{I}$ differ by a quantity which is proportional to the unitary tensor. This property is strictly analogous to the one that exists between the tensors $\mathcal{Q}$ and $\mathcal{Q}$ in electrodynamics. Therefore, within our analogy, the power emitted in gravitational waves at the lowest order (of the multipolar expansion) can be obtained from Eq. (3.34) and is given by

$$W_G = \frac{G}{20c^5} \sum_{jk} (\dddot{\mathcal{I}}_{jk})^2.$$

This formula is correct in all respects, aside from the numerical factor. The calculations based on general relativity produce a similar result, where the factor $\frac{1}{20}$ is replaced by the factor $\frac{1}{5}$. Intuitively, one can justify this multiplication by a factor of four noting that an electromagnetic wave is described by two vectors $\mathbf{E}$ and $\mathbf{B}$ that are not independent and are perpendicular to the direction of propagation, say z. Only two components of one of the two fields, such as E_x and E_y, are sufficient to describe the wave. A gravitational wave is instead described by two independent tensors also perpendicular to the direction of propagation. If we denote these tensors by the traditional symbols $\mathbf{e}^+$ and $\mathbf{e}^\times$, the wave is described by the eight components $(e_{xx}^+, e_{xy}^+, e_{yx}^+, e_{yy}^+, e_{xx}^\times, e_{xy}^\times, e_{yx}^\times, e_{yy}^\times)$. The factor of four is therefore associated with, say, the degrees of freedom of the polarisation. The correct formula for the power

emitted in gravitational waves is then

$$W_G = \frac{G}{5c^5} \sum_{jk} (\dddot{\mathcal{I}}_{jk})^2.$$

Finally, we note that if we change the origin of coordinates putting

$$\mathbf{s}_i' = \mathbf{b} + \mathbf{s}_i,$$

with $\mathbf{b}$ a constant vector, we obtain that the new inertia tensor $\mathcal{I}'$ is

$$\mathcal{I}' = \mathcal{I} + \mathcal{M}\left[(2\mathbf{b} \cdot \mathbf{r}_G + b^2)\mathbf{U} - \mathbf{b}\mathbf{b} - \mathbf{b}\mathbf{r}_G - \mathbf{r}_G\mathbf{b}\right],$$

so that, for an isolated system,

$$\ddot{\mathcal{I}}' = \ddot{\mathcal{I}},$$

and, all the more so, $\dddot{\mathcal{I}}' = \dddot{\mathcal{I}}$. This equation allows to calculate the inertia tensor in a coordinate system having an arbitrary origin in order to determine the power emitted in gravitational waves.

16.7 Calculation of the Thomas-Fermi Integral

Some applications of atomic physics based on the Thomas-Fermi model require the calculation of the following integral

$$\mathcal{I} = \int_0^\infty \frac{(1 + \chi)\chi^{3/2}}{x^{1/2}} \, dx,$$

where $\chi(x)$ is the solution of the Thomas-Fermi equation

$$x^{1/2}\chi'' = \chi^{3/2},$$

which satisfies the boundary condition

$$\chi(0) = 1, \qquad \lim_{x \to \infty} \chi(x) = 0.$$

The integral is split into the sum of two integrals

$$\mathcal{I} = \mathcal{I}_1 + \mathcal{I}_2, \tag{16.18}$$

where

$$\mathcal{I}_1 = \int_0^\infty \frac{\chi^{3/2}}{x^{1/2}} \, dx, \qquad \mathcal{I}_2 = \int_0^\infty \frac{\chi^{5/2}}{x^{1/2}} \, dx.$$

The first integral is trivial since, taking into account the Thomas-Fermi equation and the boundary conditions of the function χ, we have

$$\mathcal{I}_1 = \int_0^\infty \chi'' \, dx = -\chi'(0). \tag{16.19}$$

The calculation of the second integral is more complex. It can be done in the following way. On one hand, we have

$$\mathcal{I}_2 = \int_0^\infty \frac{\chi^{5/2}}{x^{1/2}} \, dx = \int_0^\infty \chi \chi'' \, dx,$$

and, integrating by parts and taking into account that $\chi(0) = 1$

$$\mathcal{I}_2 = -\chi'(0) - \int_0^\infty \chi'^2 \, dx. \tag{16.20}$$

On the other hand, considering the quantity $x^{-1/2} \, dx$ as a differential factor, by integrating again by parts and recalling the Thomas-Fermi equation, we obtain

$$\mathcal{I}_2 = \int_0^\infty \frac{\chi^{5/2}}{x^{1/2}} \, dx = -5 \int_0^\infty x^{1/2} \chi^{3/2} \chi' \, dx = -5 \int_0^\infty x \chi' \chi'' \, dx.$$

Now we note that the product $\chi' \chi''$ can be expressed in the form

$$\chi' \chi'' = \frac{1}{2} \frac{d\chi'^2}{dx},$$

and integrating again by parts we get

$$\mathcal{I}_2 = \frac{5}{2} \int_0^\infty \chi'^2 \, dx.$$

Comparing this expression with Eq. (16.20), we obtain

$$\int_0^\infty \chi'^2 \, dx = -\frac{2}{7} \chi'(0), \quad \text{or,} \quad \mathcal{I}_2 = -\frac{5}{7} \chi'(0).$$

Finally, recalling Eqs. (16.18) and (16.19) we get

$$\mathcal{I} = \int_0^\infty \frac{(1 + \chi) \chi^{3/2}}{x^{1/2}} \, dx = -\frac{12}{7} \chi'(0). \tag{16.21}$$

16.8 Energy of the Ground Configuration of the Silicon Atom

As an application of the results obtained in Chap. 8, we evaluate the energy of the ground configuration of the silicon atom, i.e. of the $1s^2 2s^2 2p^6 3s^2 3p^2$ configuration.

Before performing these calculations, we need to evaluate some 3-j symbols. By means of the analytical formula given in Eq. (7.16), we have

$$\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}^2 = 1, \qquad \begin{pmatrix} 0 & 1 & 1 \\ 0 & 0 & 0 \end{pmatrix}^2 = \frac{1}{3}, \qquad \begin{pmatrix} 1 & 1 & 2 \\ 0 & 0 & 0 \end{pmatrix}^2 = \frac{2}{15}.$$

The configuration contains four closed subshells and one open subshell. We start by evaluating the degenerate contribution to the energy. The Hamiltonian H_0 (defined in Eq. (7.3)) and the $\mathcal{F}$ part of the Hamiltonian $\mathcal{H}_1$ (defined in Eqs. (8.2) and (8.3)) produce five terms, one for each subshell (open or closed). The corresponding energy (that we denote by $\mathcal{E}_1$) is obtained from Eq. (8.7) and is given by

$$\mathcal{E}_1 = 2W_0(1s) + 2W_0(2s) + 6W_0(2p) + 2W_0(3s) + 2W_0(3p)$$
$$+ 2I(1s) + 2I(2s) + 6I(2p) + 2I(3s) + 2I(3p),$$

where W_0 is defined by Eq. (7.11) and $I(n, l)$ is the integral defined in Eq. (8.6). The energy of the Coulomb interaction (i.e. the part $\mathcal{G}$ of the Hamiltonian $\mathcal{H}_1$) resulting from closed subshells contributes four terms. Denoting by $\mathcal{E}_2$ the corresponding energy, we have, using Eq. (8.17),

$$\mathcal{E}_2 = F^0(1s, 1s) + F^0(2s, 2s) + 15F^0(2p, 2p) - \frac{6}{5}F^2(2p, 2p) + F^0(3s, 3s),$$

where the quantities $F^k(n_a l_a, n_n l_b)$ are defined in Eq. (8.9). Considering the energy of the Coulomb interaction between different closed subshells, we have six contributions, as many as the number of the distinct pairs that can be formed with the four closed subshells. Denoting by $\mathcal{E}_3$ the corresponding energy, we have, using Eqs. (8.14) and (8.16)

$$\mathcal{E}_3 = 4F^0(1s, 2s) - 2G^0(1s, 2s) + 12F^0(1s, 2p) - 2G^1(1s, 2p) + 4F^0(1s, 3s)$$
$$- 2G^0(1s, 3s) + 12F^0(2s, 2p) - 2G^1(2s, 2p) + 4F^0(2s, 3s) - 2G^0(2s, 3s)$$
$$+ 12F^0(2p, 3s) - 2G^1(2p, 3s),$$

where the quantities $G^k(n_a l_a, n_n l_b)$ are defined in Eq. (8.10). Finally, we need to evaluate the contribution of the Coulomb interaction between the open subshell $3p$ and the four closed subshells. Denoting by $\mathcal{E}_4$ the corresponding energy, we have, using Eqs. (8.13) and (8.15),

$$\mathcal{E}_4 = 4F^0(1s, 3p) - \frac{2}{3}G^1(1s, 3p) + 4F^0(2s, 3p) - \frac{2}{3}G^1(2s, 3p) + 12F^0(2p, 3p)$$
$$- 2G^0(2p, 3p) - \frac{4}{5}G^2(2p, 3p) + 4F^0(3s, 3p) - \frac{2}{3}G^1(3s, 3p).$$

The four contributions to the energy that we have calculated are degenerate with respect to all the states of the configuration. For the degenerate part of the energy of

the ground configuration of the silicon atom, $\mathcal{E}$, we then have

$$\mathcal{E} = \mathcal{E}_1 + \mathcal{E}_2 + \mathcal{E}_3 + \mathcal{E}_4.$$

What remains to calculate is given by Eq. (8.11), with the sum extended to the only pair of electrons belonging to the open subshell $3p$. The explicit computation is done in Sect. 8.6. The two $3p$ electrons give rise to three terms that, in order of increasing energy, are ^3P, ^1D, and ^1S. The ratio between the intervals $(^1S - {}^1D)$ and $(^1D - {}^3P)$ is equal to $3/2$.

16.9 Calculation of the Fine-Structure Constant of a Term

The calculation of the constant $\zeta(\alpha, LS)$, which characterises the fine-structure intervals of the terms belonging to a given configuration, can be carried out with a process based on the diagonal sum rule. A similar process was followed in Sect. 8.1 to determine the energy of the terms. The starting point is Eq. (9.8) which, in the case of diagonal matrix elements, is

$$\sum_i \langle \alpha L S M_L M_S | \xi(r_i)\boldsymbol{\ell}_i \cdot \mathbf{s}_i | \alpha L S M_L M_S \rangle = \zeta(\alpha, LS)M_L M_S.$$

On the other hand, for any eigenstate of the configuration of the form $\Psi^A(a_1, a_2, \ldots, a_N)$ of Eq. (7.1), the diagonal matrix element of the same operator is given by

$$\sum_i \langle \Psi^A(a_1, a_2, \ldots, a_N) | \xi(r_i)\boldsymbol{\ell}_i \cdot \mathbf{s}_i | \Psi^A(a_1, a_2, \ldots, a_N) \rangle = \sum_i \zeta_{n_i l_i} m_i m_{si},$$

where $\zeta_{n_i l_i}$ is the quantity defined in Eq. (9.10).

We now consider the particular case of the pf configuration which, as shown in Table 7.3, gives rise to the six terms ^1D, ^1F, ^1G, ^3D, ^3F, and ^3G. We start from a state having the highest values of the quantum numbers M_L and M_S, i.e. $M_L = 4$, $M_S = 1$. This state can only originate from the ^3G term. Considering instead single particle states, this state is of the type $m_1 = 1$, $m_{s1} = \frac{1}{2}$, $m_2 = 3$, $m_{s2} = \frac{1}{2}$, where the indices 1 and 2 refer, respectively, to the p and f electron. Using the same notations as in Sect. 8.1 we can write the equality[5]

$$[4, 1] = \left(1^+, 3^+\right),$$

---

[5]The symbol $[M_L, M_S]$ means the sum of the diagonal matrix elements of the Hamiltonian of the spin-orbit interaction over all the states Ψ^A for which M_L and M_S are the eigenvalues of L_z and S_z, respectively. Similarly, the notation $(m_1^\pm, m_2^\pm)$ is used to denote the diagonal matrix element of the same Hamiltonian on the state where the electron 1 has the magnetic quantum number m_1 and spin quantum number $+1/2$ or $-1/2$ and, similarly, the electron 2 has the magnetic quantum number m_2 and spin quantum number $+1/2$ or $-1/2$.

that, according to the previous equation, is[6]

$$4\zeta\left(^3G\right) = \frac{1}{2}\zeta_{np} + \frac{3}{2}\zeta_{nf}.$$

We therefore obtain the result

$$\zeta\left(^3G\right) = \frac{1}{8}\zeta_{np} + \frac{3}{8}\zeta_{nf}.$$

We then proceed by lowering the value of M_L (maintaining $M_S = 1$). We obtain the equations

$$[3, 1] = \left(0^+, 3^+\right) + \left(1^+, 2^+\right), \qquad [2, 1] = \left(-1^+, 3^+\right) + \left(0^+, 2^+\right) + \left(1^+, 1^+\right),$$

from which we have, noting that the combination $[M_L = 3, M_S = 1]$ can originate from the ^3G and ^3F terms, and that the combination $[M_L = 2, M_S = 1]$ can originate from the ^3G, ^3F, and ^3D terms,

$$3\left[\zeta\left(^3G\right) + \zeta\left(^3F\right)\right] = \frac{3}{2}\zeta_{nf} + \frac{1}{2}\zeta_{np} + \zeta_{nf},$$

$$2\left[\zeta\left(^3G\right) + \zeta\left(^3F\right) + \zeta\left(^3D\right)\right] = -\frac{1}{2}\zeta_{np} + \frac{3}{2}\zeta_{nf} + \zeta_{nf} + \frac{1}{2}\zeta_{np} + \frac{1}{2}\zeta_{nf}.$$

By solving the system, we arrive at the following expressions (which can also be obtained from Eq. (9.11))

$$\zeta\left(^3F\right) = \frac{1}{24}\zeta_{np} + \frac{11}{24}\zeta_{nf}, \qquad \zeta\left(^3D\right) = -\frac{1}{6}\zeta_{np} + \frac{2}{3}\zeta_{nf}.$$

In principle, we could also consider the values of $M_S = 0$. For example,

$$[4, 0] = \left(1^+, 3^-\right) + \left(1^-, 3^+\right).$$

However, in so doing we obtain equations of the form $0 = 0$ and the value of $\zeta\left(^1G\right)$ is undetermined. This is entirely consistent, since the singlet states do not have fine structure and the constant ζ is not defined.

   The cases of the configurations of equivalent electrons are also interesting, because, by repeating the same arguments, we obtain directly the third Hund's rule. For example, consider the configuration p^2 which produces, as shown in Table 7.4, the three terms ^1S, ^1D, and ^3P. For the singlet terms, the fine structure constant remains undetermined, as usual. For the triplet term we have instead

$$[1, 1] = \left(0^+, 1^+\right),$$

---

[6]We note that even if there are electrons in the closed subshells they do not produce any contribution to the equation.

from which we obtain

$$\zeta\left({}^3P\right) = \frac{1}{2}\zeta_{np}.$$

If we consider the complementary configuration p^4, we get the same structure of terms. This time, to find the fine structure constant of the ${}^3P$, term, we need to consider the equation[7]

$$[1, 1] = \left(1^+, 1^-, 0^+, -1^+\right),$$

and we obtain

$$\zeta\left({}^3P\right) = -\frac{1}{2}\zeta_{np},$$

i.e. a value that is exactly the same (but opposite in sign) to the one of the configuration p^2. These arguments can be repeated for any configuration of equivalent electrons and for the corresponding complementary configuration and lead to the third Hund's rule. In the particular case of configurations that fill half of a subshell (such as p^3, d^5, and f^7), the configuration coincides with the complementary one, and the fine structure constant is zero for all the terms.

16.10 The Fundamental Principle of Statistical Thermodynamics

Consider, in all generality, a macroscopic physical system. We suppose that the system is in thermal equilibrium with an ideal heat reservoir having temperature T (canonical ensemble). We also suppose to identify with the index i all possible microscopic states of the system and we denote by E_i the energy of the i-th state. Macroscopically, the system is in a steady state. On the other hand, from the microscopical point of view, we can think that it constantly evolves from one microscopic state to another. We can then introduce a statistical description denoting by p_i the probability that the system is in the i-th microscopic state. The following normalisation property should obviously be valid

$$\sum_i p_i = 1.$$

We now need to relate the probability p_i with the energy E_i. To do so, we give a definition of the entropy by putting, according to an hypothesis originally due to Boltzmann

$$S = -k_{\mathrm{B}} \sum_i p_i \ln p_i,$$

---

[7]The quantity $(1^+, 1^-, 0^+, -1^+)$, relative to the configuration p^4, is obtained from the corresponding quantity $(0^+, 1^+)$, relative to the configuration p^2, taking the "complementary" of the latter, i.e. $(-1^+, -1^-, 0^-, 1^-)$, and then changing sign to all the values of m and m_s.

where k_B is the Boltzmann constant.

This definition can be justified by considering that the entropy of a system measures the amount of "disorder" contained in the system itself and noting that the function defined above has the mathematical property of assuming its maximum value when all the probabilities p_i are equal to one another, and take the minimum value (which is equal to 0) when a single p_i is equal to 1 and all the others are equal to 0. The proof of the second property is trivial. To prove the first property we note that giving an arbitrary variation δp_i to the probabilities, the corresponding change δS of the entropy is

$$\delta S = -k_B \sum_i (\ln p_i + 1)\delta p_i.$$

On the other hand, being

$$\sum_i \delta p_i = 0,$$

it follows that if $\ln p_i$ is constant (i.e. independent of i), δS is null and therefore the entropy presents an extreme. It is then easy to verify that such an extreme is actually a maximum, since

$$\frac{d^2 S}{dp_i^2} = -k_B \frac{1}{p_i} < 0.$$

Having justified the definition of the entropy, we now take into account that the internal energy of the system is given by the expression

$$U = \sum_i p_i E_i.$$

If we consider an infinitesimal thermodynamical transformation of the system, the internal energy will vary, in general, because both the probabilities p_i and the energies E_i change. We then have

$$\delta U = \sum_i (\delta p_i) E_i + \sum_i p_i (\delta E_i).$$

If, however, the external conditions of the system are not varied, the quantities E_i remain fixed to the initial value and the second term of the right hand side is null. On the other hand, to keep constant the external conditions of the system means that the system does not accomplish mechanical work on the ambient medium, so we can write, according to the first principle of thermodynamics

$$\delta U = \sum_i (\delta p_i) E_i = \delta Q = T \delta S,$$

where δQ is the heat exchanged with the reservoir, and taking into account that

$$\delta S = -k_{\mathrm{B}} \sum_i (\delta p_i) \ln p_i,$$

we obtain the equation

$$\sum_i (\delta p_i) E_i = -k_{\mathrm{B}} T \sum_i (\delta p_i) \ln p_i.$$

This equation must be satisfied for an arbitrary thermodynamic transformation (as long as no work is done). The following relation therefore must hold

$$E_i = -k_{\mathrm{B}} T \ln p_i + \text{const.,}$$

which leads to the relation

$$p_i = A e^{-\beta E_i},$$

where A is a constant and where we have put

$$\beta = \frac{1}{k_{\mathrm{B}} T}.$$

The constant A is determined by imposing the normalisation condition. Since we must have

$$\sum_i p_i = \sum_i A e^{-\beta E_i} = 1,$$

it follows that

$$A = \frac{1}{\mathcal{Z}},$$

where the quantity $\mathcal{Z}$, known as the sum over states, is given by

$$\mathcal{Z} = \sum_i e^{-\beta E_i}.$$

The expression of p_i can therefore be written in its final form

$$p_i = \frac{1}{\mathcal{Z}} e^{-\beta E_i} = \frac{e^{-\beta E_i}}{\sum_j e^{-\beta E_j}}. \tag{16.22}$$

This expression, often referred to as Gibbs principle, is of extreme generality and can rightly be considered the basis of all statistical thermodynamics. It can be written in an alternative form by assuming that the microscopic states of the system are not discrete (and therefore countable) but are identified by the representative point in the phase space of the system having dimension $2\mathcal{N}$, where

$\mathcal{N}$ is the number of degrees of freedom of the system itself. In this case, denoting by dP the probability that the representative point of the system is in the cell $d\Gamma = dq_1\, dq_2 \cdots dq_{\mathcal{N}}\, dp_1\, dp_2 \cdots dp_{\mathcal{N}}$ of the phase space centered around the values (q_i, p_i), and denoting by $H(q_i, p_i)$ the Hamiltonian of the system, we have

$$dP = \frac{1}{\mathcal{Z}}e^{-\beta H(q_i, p_i)}\, d\Gamma = \frac{e^{-\beta H(q_i, p_i)}\, d\Gamma}{\int e^{-\beta H(q_i, p_i)}\, d\Gamma}, \qquad (16.23)$$

where the integral is over the entire volume of the phase space available to the system. Equations (16.22) and (16.23) coincide, respectively, with Eqs. (10.2) and (10.1) which we have assumed as the basic principles for the deduction of the various laws of thermodynamical equilibrium in Chap. 10.

16.11 Transition Probability for the Coherences

In Chap. 11, we have introduced the so-called random phase approximation and we have determined the kinetic equations for the diagonal matrix elements ρ_α of the density matrix operator of the physical system. The result that we found is the kinetic equation (9.11), which is interpreted by introducing the transition probability per unit time between different states of the system. This probability is given by Fermi's golden rule, expressed by Eq. (11.10). We now want to generalise these results by determining the kinetic equations for the so-called coherences, i.e. for the non-diagonal matrix elements of the density matrix operator.

   We start again from Eq. (8.11) and introduce the hypothesis, less restrictive than that of the random phases, that in the physical system there might exist coherences, even if only within pairs of states, $|\alpha\rangle$ and $|\alpha'\rangle$, having the same energy eigenvalue (degenerate states) and such that the matrix element of the interaction Hamiltonian between them, $\mathcal{H}^I_{\alpha\alpha'}$, is zero. Taking into account this approximation, when evaluating the product $c_\alpha(t)c^*_{\alpha'}(t)$ we obtain, considering only the terms that are at most quadratic in the matrix elements of $\mathcal{H}^I$,

$$
\begin{aligned}
c_\alpha(t)c^*_{\alpha'}(t) = {}& c_\alpha(0)c^*_{\alpha'}(0) \\
& + \frac{1}{\hbar^2}\sum_{\beta\beta'}\mathcal{H}^I_{\alpha\beta}\mathcal{H}^I_{\beta'\alpha'}c_\beta(0)c^*_{\beta'}(0)\frac{e^{i\omega_{\alpha\beta}t}-1}{\omega_{\alpha\beta}}\frac{e^{-i\omega_{\alpha'\beta'}t}-1}{\omega_{\alpha'\beta'}} \\
& + \frac{1}{\hbar^2}\Bigg[\sum_{\beta\gamma}\mathcal{H}^I_{\alpha\beta}\mathcal{H}^I_{\beta\gamma}c_\gamma(0)c^*_{\alpha'}(0)\left(\frac{e^{i\omega_{\alpha\gamma}t}-1}{\omega_{\alpha\gamma}\omega_{\beta\gamma}}-\frac{e^{i\omega_{\alpha\beta}t}-1}{\omega_{\alpha\beta}\omega_{\beta\gamma}}\right) \\
& + \mathcal{C}.\mathcal{C}.\big(\alpha\leftrightarrow\alpha'\big)\Bigg],
\end{aligned}
$$

where the symbol $[\cdots + \mathcal{C}.\mathcal{C}.(\alpha\leftrightarrow\alpha')]$ means that we need to add to the term in brackets its complex conjugate (with the exchange of the indices α and α').

We now need to recall the approximation we have introduced in that the states between which coherences exist are iso-energetic. As regards the second line of the previous equation, this implies that $\omega_{\alpha\beta} = \omega_{\alpha'\beta'}$ and so the two temporal factors are one the complex conjugate of the other. Regarding the third line, we can consider the limit ($\omega_{\alpha\gamma} \to 0$) and the temporal function between round brackets, which we indicate with $\mathcal{F}(t)$, is equal to

$$\mathcal{F}(t) = \frac{e^{i\omega_{\alpha\gamma}t} - 1}{\omega_{\alpha\gamma}\omega_{\beta\gamma}} - \frac{e^{i\omega_{\alpha\beta}t} - 1}{\omega_{\alpha\beta}\omega_{\beta\gamma}} = -\frac{it}{\omega_{\alpha\beta}} + \frac{e^{i\omega_{\alpha\beta}t} - 1}{\omega_{\alpha\beta}^2}$$

$$= -\frac{2\sin^2(\omega_{\alpha\beta}t/2)}{\omega_{\alpha\beta}^2} + i\frac{\sin(\omega_{\alpha\beta}t) - \omega_{\alpha\beta}t}{\omega_{\alpha\beta}^2}.$$

We now proceed by evaluating the statistical average over the physical system. We introduce the notation of the density matrix by putting[8] $\rho_{\alpha\alpha'} = \langle c_\alpha(t)c_{\alpha'}^*(t)\rangle$. Changing the index of the sum γ in α'', the kinetic equation for the coherences becomes

$$\rho_{\alpha\alpha'}(t) = \rho_{\alpha\alpha'}(0) + \frac{1}{\hbar^2}\sum_{\beta\beta'}\mathcal{H}_{\alpha\beta}^I\mathcal{H}_{\beta'\alpha'}^I\rho_{\beta\beta'}(0)\frac{4\sin^2(\omega_{\alpha\beta}t/2)}{\omega_{\alpha\beta}^2}$$

$$+ \frac{1}{\hbar^2}\left[\sum_{\beta\alpha''}\mathcal{H}_{\alpha\beta}^I\mathcal{H}_{\beta\alpha''}^I\rho_{\alpha''\alpha'}(0)\mathcal{F}(t) + C.C.(\alpha \leftrightarrow \alpha')\right]. \qquad (16.24)$$

We consider the limit of this equation for $t \to \infty$. As we have seen on various occasions within the text (cf. Fig. 11.1)

$$\lim_{t\to\infty}\frac{4\sin^2(\omega_{\alpha\beta}t/2)}{\omega_{\alpha\beta}^2} = 2\pi t\delta(\omega_{\alpha\beta}) = 2\pi\hbar t\delta(E_\alpha - E_\beta),$$

where we have used the definition of the Bohr frequencies in terms of the energies of the states of the physical system. Regarding the function $\mathcal{F}(t)$, while its real part produces again a Dirac delta over the energy, the imaginary part behaves, at the limit of $t \to \infty$, as shown in Fig. 16.2. It can rigorously be shown within the distribution theory that we have

$$\lim_{t\to\infty}\mathcal{F}(t) = -\pi t\delta(\omega_{\alpha\beta}) - it\text{PP}\frac{1}{\omega_{\alpha\beta}} = -\pi\hbar t\left[\delta(E_\alpha - E_\beta) + \frac{i}{\pi}\text{PP}\frac{1}{E_\alpha - E_\beta}\right],$$

where the symbol PP means the Cauchy principal value.

We are now able to write the kinetic equation that generalises Eq. (11.9), valid for the diagonal elements of the density matrix, to the case of coherences. Starting

---

[8]We note that in Chap. 11 we only introduced the diagonal elements of the density matrix, denoted for simplicity by the symbol ρ_α instead of $\rho_{\alpha\alpha}$.

Fig. 16.2 The imaginary part of $\mathcal{F}(t)$ is plotted as a function of $\omega_{\alpha\beta}$ for a fixed time t. As t increases, the behaviour of the function becomes more and more similar to that one of the function $-t/\omega_{\alpha\beta}$ (*dotted line*), except in the origin where it is zero
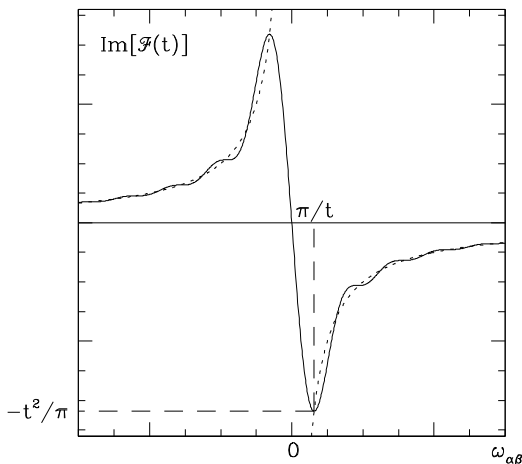
with Eq. (16.24) and noting that all the terms in the right hand side behave linearly with t, we can write, for $t \rightarrow \infty$

$$\frac{\mathrm{d}}{\mathrm{d}t}\rho_{\alpha\alpha'} = \sum_{\beta\beta'} T_{\alpha\alpha'\beta\beta'}\rho_{\beta\beta'} - \sum_{\alpha''}\left(R_{\alpha\alpha''}\rho_{\alpha''\alpha'} + R^*_{\alpha'\alpha''}\rho_{\alpha\alpha''}\right), \qquad (16.25)$$

where $T_{\alpha\alpha'\beta\beta'}$, the rate of transfer from the coherence $\rho_{\beta\beta'}$ to the coherence $\rho_{\alpha\alpha'}$, is given by

$$T_{\alpha\alpha'\beta\beta'} = \frac{2\pi}{\hbar}\mathcal{H}^{\mathrm{I}}_{\alpha\beta}\mathcal{H}^{\mathrm{I}}_{\beta'\alpha'}\delta(E_\alpha - E_\beta),$$

and where $R_{\alpha\alpha''}$, the relaxation rate that relates the coherence $\rho_{\alpha\alpha'}$ to the coherence $\rho_{\alpha''\alpha'}$, is given by

$$R_{\alpha\alpha''} = \frac{\pi}{\hbar}\sum_{\beta}\mathcal{H}^{\mathrm{I}}_{\alpha\beta}\mathcal{H}^{\mathrm{I}}_{\beta\alpha''}\left[\delta(E_\alpha - E_\beta) + \frac{\mathrm{i}}{\pi}\mathrm{PP}\frac{1}{E_\alpha - E_\beta}\right].$$

It is easy to show that Eq. (16.25) coincides with Eq. (11.9) in the case of the random phase approximation, i.e. when we consider only the diagonal elements of the density matrix. We have, in fact,

$$T_{\alpha\alpha\beta\beta} = \frac{2\pi}{\hbar}\left|\mathcal{H}^{\mathrm{I}}_{\alpha\beta}\right|^2\delta(E_\alpha - E_\beta) = P_{\alpha\beta},$$

where $P_{\alpha\beta}$ is the transition probability per unit time between the states $|\alpha\rangle$ and $|\beta\rangle$ (or between the states $|\beta\rangle$ and $|\alpha\rangle$) given by Eq. (11.10) (Fermi's golden rule). Similarly,

$$R_{\alpha\alpha} + R^*_{\alpha\alpha} = \frac{2\pi}{\hbar}\sum_{\beta}\left|\mathcal{H}^{\mathrm{I}}_{\alpha\beta}\right|^2\delta(E_\alpha - E_\beta) = \sum_{\beta}P_{\alpha\beta}.$$

Equation (16.25) can therefore be rightly considered the generalisation of Fermi's golden rule to the case of coherences. It is important to stress the presence of the imaginary factor in the relaxation rates. Such factor is responsible for some phenomena typical of the interaction between material systems and the radiation field such as, in particular, the anomalous dispersion phenomena that occur in the propagation of polarised radiation in an anisotropic medium (Faraday effect, Macaluso-Corbino effect, etc.).

16.12 Sums over the Magnetic Quantum Numbers

Here, we want to prove Eq. (11.17). That is, we want to show that, for any polarisation unit vector $\mathbf{e}$, having defined the averages of the square moduli of the dipole matrix elements $\mathcal{A}$ and $\mathcal{A}'$ over the magnetic quantum numbers by the equations

$$\mathcal{A} = \frac{1}{g_a g_b} \sum_{\alpha, \beta} |\mathbf{r}_{b\beta, a\alpha} \cdot \mathbf{e}|^2,$$

$$\mathcal{A}' = |\mathbf{r}_{ba}|^2 = \frac{1}{g_a g_b} \sum_{\alpha, \beta} |\mathbf{r}_{b\beta, a\alpha}|^2 = \frac{1}{g_a g_b} \sum_{\alpha, \beta} \langle u_{b\beta}|\mathbf{r}|u_{a\alpha}\rangle \cdot \langle u_{a\alpha}|\mathbf{r}|u_{b\beta}\rangle,$$

we have

$$\mathcal{A} = \frac{1}{3} \mathcal{A}'.$$

The indices a and b in the previous equations denote any two energy levels of the atomic system while the indices α and β denote the respective magnetic sublevels, which are degenerate with respect to the energy. To demonstrate the equation, we need to introduce a more detailed notation which takes into account the fact that the atomic levels are normally characterised not only by a set of internal quantum numbers γ (which specify the configuration and the term), but also by the quantum number for the angular momentum J and the magnetic quantum number M. Applying the formal substitutions

$$|u_{a\alpha}\rangle \to |\gamma_a J_a M_a\rangle, \qquad |u_{b\beta}\rangle \to |\gamma_b J_b M_b\rangle,$$
$$g_a \to 2J_a + 1, \qquad g_b \to 2J_b + 1,$$

we obtain

$$\mathcal{A} = \sum_{M_a M_b} \frac{|\langle \gamma_b J_b M_b|\mathbf{r} \cdot \mathbf{e}|\gamma_a J_a M_a\rangle|^2}{(2J_a + 1)(2J_b + 1)},$$

$$\mathcal{A}' = \sum_{M_a M_b} \frac{\langle \gamma_b J_b M_b|\mathbf{r}|\gamma_a J_a M_a\rangle \cdot \langle \gamma_a J_a M_a|\mathbf{r}|\gamma_b J_b M_b\rangle}{(2J_a + 1)(2J_b + 1)}.$$

To calculate $\mathcal{A}$ we apply the Wigner-Eckart theorem, noting that the scalar product $\mathbf{r} \cdot \mathbf{e}$ can be expressed in terms of the spherical components of the two vectors. We have in fact (cf. Eq. (9.5))

$$\mathbf{r} \cdot \mathbf{e} = \sum_q (-1)^q r_q e_{-q}.$$

Using Eq. (9.4) we have

$$\langle \gamma_b J_b M_b | r_q | \gamma_a J_a M_a \rangle$$

$$= (-1)^{J_a + M_b + 1} \sqrt{2 J_b + 1} \begin{pmatrix} J_b & J_a & 1 \\ -M_b & M_a & q \end{pmatrix} \langle \gamma_b J_b \| \mathbf{r} \| \gamma_a J_a \rangle,$$

and we obtain

$$\mathcal{A} = \sum_{qq'} (-1)^{q+q'} e_{-q} (e_{-q'})^*$$

$$\times \sum_{M_a M_b} \begin{pmatrix} J_b & J_a & 1 \\ -M_b & M_a & q \end{pmatrix} \begin{pmatrix} J_b & J_a & 1 \\ -M_b & M_a & q' \end{pmatrix} \frac{|\langle \gamma_b J_b \| \mathbf{r} \| \gamma_a J_a \rangle|^2}{2 J_a + 1}.$$

The sum over M_a and M_b of the product of the two 3-j symbols can be calculated using the property of the 3-j symbols of Eq. (7.18). We have

$$\sum_{M_a M_b} \begin{pmatrix} J_b & J_a & 1 \\ -M_b & M_a & q \end{pmatrix} \begin{pmatrix} J_b & J_a & 1 \\ -M_b & M_a & q' \end{pmatrix} = \frac{1}{3} \delta_{qq'},$$

and we obtain, being $\sum_q e_q (e_q)^* = 1$

$$\mathcal{A} = \frac{1}{3} \frac{|\langle \gamma_b J_b \| \mathbf{r} \| \gamma_a J_a \rangle|^2}{2 J_a + 1}.$$

To calculate the quantity $\mathcal{A}'$ we proceed in a similar way first noting that $\langle \gamma_a J_a M_a | \mathbf{r} | \gamma_b J_b M_b \rangle = \langle \gamma_b J_b M_b | \mathbf{r} | \gamma_a J_a M_a \rangle^*$. We obtain

$$\mathcal{A}' = \sum_q \sum_{M_a M_b} \begin{pmatrix} J_b & J_a & 1 \\ -M_b & M_a & q \end{pmatrix} \begin{pmatrix} J_b & J_a & 1 \\ -M_b & M_a & q' \end{pmatrix} \frac{|\langle \gamma_b J_b \| \mathbf{r} \| \gamma_a J_a \rangle|^2}{2 J_a + 1},$$

and using the same property of the 3-j symbols and summing over q we arrive at the result that we wanted to prove, i.e.

$$\mathcal{A}' = \frac{|\langle \gamma_b J_b \| \mathbf{r} \| \gamma_a J_a \rangle|^2}{2 J_a + 1} = 3 \mathcal{A}.$$

The above results can be used to express the quantity $|\mathbf{r}_{ba}|^2$ that we introduced within the text in terms of the reduced matrix elements of the spherical tensor $\mathbf{r}$.

Since $|\mathbf{r}_{ba}|^2 = \mathcal{A}'$, we have

$$|\mathbf{r}_{ba}|^2 = \frac{|\langle \gamma_b J_b \| \mathbf{r} \| \gamma_a J_a \rangle|^2}{2J_a + 1}.$$

On the other hand, being $|\mathbf{r}_{ba}|^2 = |\mathbf{r}_{ab}|^2$, we obtain by symmetry

$$|\mathbf{r}_{ba}|^2 = |\mathbf{r}_{ab}|^2 = \frac{|\langle \gamma_b J_b \| \mathbf{r} \| \gamma_a J_a \rangle|^2}{2J_a + 1} = \frac{|\langle \gamma_a J_a \| \mathbf{r} \| \gamma_b J_b \rangle|^2}{2J_b + 1},$$

an equation that relates the reduced matrix elements under the exchange of the bra with the ket.

In spectroscopy, the concept of line (or transition) strength is commonly used. Such quantity is invariant with respect to the exchange of the lower and upper level, and is defined by

$$\mathcal{S} = g_b \left| \langle \gamma_b J_b \| \mathbf{d} \| \gamma_a J_a \rangle \right|^2 = g_a \left| \langle \gamma_a J_a \| \mathbf{d} \| \gamma_b J_b \rangle \right|^2,$$

where $\mathbf{d} = -e_0 \mathbf{r}$ is the electric dipole operator. The quantities introduced in the text are therefore related to the line strength via the relation

$$|\mathbf{r}_{ba}|^2 = |\mathbf{r}_{ab}|^2 = \frac{1}{e_0^2} \frac{\mathcal{S}}{g_a g_b}.$$

These relations can then be used to express the Einstein coefficients in terms of the line strength instead of in terms of the dipole matrix elements. For example, recalling Eq. (11.20), the Einstein coefficient A_{ab} can be written in the form

$$g_a A_{ab} = \frac{64\pi^4 v_{ab}^3}{3hc^3} \mathcal{S}.$$

An alternative quantity that is also used to characterise the strength of a line (or a transition) is the so-called oscillator strength. This quantity is introduced in the following way. The absorption coefficient of a plasma of "classical" atoms, described by the Lorentz atomic model and integrated in frequency is given by

$$\left[k_R^{(a)} \right]_{\text{class}} = \mathcal{N} \frac{\pi e_0^2}{mc},$$

where $\mathcal{N}$ is the number density of atoms. Comparing this expression with that one for $k_R^{(a)}$ obtained in Sect. 11.9 (Eq. (11.33)), we see that the two quantities coincide if we identify $\mathcal{N}$ with $\mathcal{N}_b$ and multiply the classical expression for the dimensionless quantity f_{ba}, known as the oscillator strength of the transition, given by

$$f_{ba} = \frac{8\pi^2 m v_{ab}}{3h} g_a |\mathbf{r}_{ba}|^2.$$

The oscillator strength may be considered as a parameter measuring the efficiency of the transition, since it represents a sort of "equivalent number" of classical oscillators. Typically, it is a relatively small number that can reach values of the order of unity only for the strongest spectral lines. The relations between oscillator strength, line strength, and Einstein coefficients are easily obtained using the previous relations. For example, we have

$$g_b f_{ba} = \frac{8\pi^2 m \nu_{ab}}{3 h e_0^2} \mathcal{S}, \qquad g_a A_{ab} = \frac{8\pi^2 e_0^2 \nu_{ab}^2}{m c^3} g_b f_{ba}.$$

16.13   Calculation of a Matrix Element

We wish to calculate the probability per unit time that the following elementary process occurs: a non-relativistic free electron of momentum $\mathbf{q}$ undergoes a transition to a free state of momentum $\mathbf{q}'$ due to absorption of a photon with wave vector $\mathbf{k}$. According to Fermi's golden rule, repeating the arguments presented in Sect. 11.4 but without introducing the dipole approximation,[9] such probability is proportional to the squared modulus of the matrix element $\mathcal{M}$ given by

$$\mathcal{M} = \langle u_{\mathrm{f}} | e^{\mathrm{i} \mathbf{k} \cdot \mathbf{r}} \mathbf{p} \cdot \mathbf{e} | u_{\mathrm{i}} \rangle,$$

where $|u_{\mathrm{i}}\rangle$ and $|u_{\mathrm{f}}\rangle$ are the eigenvectors of the atomic system (in our case of the free electron) in the initial and final state, respectively, $\mathbf{e}$ is the polarisation unit vector of the absorbed photon, and $\mathbf{p}$ is the momentum operator of the electron. Within the representation of the wavefunctions, where the operator $\mathbf{p}$ is given by $-\mathrm{i}\hbar\,\mathrm{grad}$, the matrix element $\mathcal{M}$ is

$$\mathcal{M} = -\mathrm{i}\hbar\mathbf{e} \cdot \int \psi_{\mathrm{f}}^*(\mathbf{r}) e^{\mathrm{i}\mathbf{k}\cdot\mathbf{r}} \,\mathrm{grad}\big[\psi_{\mathrm{i}}(\mathbf{r})\big]\,\mathrm{d}^3\mathbf{r}.$$

On the other hand, the eigenfunctions ψ_{f} and ψ_{i} are of the type of a plane wave, i.e.

$$\psi_{\mathrm{f}}(\mathbf{r}) = \frac{1}{\sqrt{\mathcal{V}}} e^{\mathrm{i}\mathbf{q}'\cdot\mathbf{r}/\hbar}, \qquad \psi_{\mathrm{i}}(\mathbf{r}) = \frac{1}{\sqrt{\mathcal{V}}} e^{\mathrm{i}\mathbf{q}\cdot\mathbf{r}/\hbar},$$

where $\mathcal{V}$ is the normalisation volume. Substituting in the integral we have

$$\mathcal{M} = \frac{\mathbf{e} \cdot \mathbf{q}}{\mathcal{V}} \int e^{-\mathrm{i}(\mathbf{q}' - \hbar\mathbf{k} - \mathbf{q})\cdot\mathbf{r}/\hbar}\,\mathrm{d}^3\mathbf{r}.$$

---

[9]The dipole approximation is appropriate when considering the interaction between radiation and electrons that are bound in an atom. For free electrons, described by eigenfunctions of the type of a plane wave, the approximation cannot be applied.

The integral is null unless the argument of the exponential is zero. This leads to the equality $\mathbf{q}' = (\hbar\mathbf{k} + \mathbf{q})$ which represents the conservation of momentum. In such case, the integral is simply equal to $\mathcal{V}$, so we obtain

$$\mathcal{M} = \mathbf{e} \cdot \mathbf{q} = \mathbf{e} \cdot (\mathbf{q}' - \hbar\mathbf{k}).$$

16.14 Gauge Invariance in Quantum Electrodynamics

Consider the quantity $\mathcal{R}_{\mathrm{fi}}$ defined in Eq. (15.22) of the text that we rewrite here in the form

$$\mathcal{R}_{\mathrm{fi}} = P + Q,$$

where

$$P = W_s^\dagger(\mathbf{p}')(\boldsymbol{\alpha} \cdot \mathbf{e}'^*) \frac{H_{\mathbf{g}} + \epsilon + \hbar\omega}{(\epsilon + \hbar\omega)^2 - \epsilon_{\mathbf{g}}^2} (\boldsymbol{\alpha} \cdot \mathbf{e}) W_r(\mathbf{p}),$$

$$Q = W_s^\dagger(\mathbf{p}')(\boldsymbol{\alpha} \cdot \mathbf{e}) \frac{H_{\mathbf{h}} + \epsilon - \hbar\omega'}{(\epsilon - \hbar\omega')^2 - \epsilon_{\mathbf{h}}^2} (\boldsymbol{\alpha} \cdot \mathbf{e}'^*) W_r(\mathbf{p}).$$

We want to demonstrate that $\mathcal{R}_{\mathrm{fi}}$ is invariant with respect to the transformation

$$\boldsymbol{\alpha} \cdot \mathbf{e} \to \boldsymbol{\alpha} \cdot \mathbf{e} + C(\boldsymbol{\alpha} \cdot \mathbf{u} - 1), \tag{16.26}$$

where C is an arbitrary constant and where $\mathbf{u}$ is the unit vector of the direction of the initial photon ($\mathbf{u} = \mathbf{k}/k$). Performing such transformation, the quantities P and Q are transformed according to the equations

$$P \to P + CP', \qquad Q \to Q + CQ',$$

where

$$P' = W_s^\dagger(\mathbf{p}')(\boldsymbol{\alpha} \cdot \mathbf{e}'^*) \frac{H_{\mathbf{g}} + \epsilon + \hbar\omega}{(\epsilon + \hbar\omega)^2 - \epsilon_{\mathbf{g}}^2} [(\boldsymbol{\alpha} \cdot \mathbf{u}) - 1] W_r(\mathbf{p}),$$

$$Q' = W_s^\dagger(\mathbf{p}')[(\boldsymbol{\alpha} \cdot \mathbf{u}) - 1] \frac{H_{\mathbf{h}} + \epsilon - \hbar\omega'}{(\epsilon - \hbar\omega')^2 - \epsilon_{\mathbf{h}}^2} (\boldsymbol{\alpha} \cdot \mathbf{e}'^*) W_r(\mathbf{p}).$$

We multiply the two quantities P' and Q' by the product $c\hbar k$ and note that

$$c\hbar k[(\boldsymbol{\alpha} \cdot \mathbf{u}) - 1] = c\hbar(\boldsymbol{\alpha} \cdot \mathbf{k}) - \hbar\omega.$$

Recalling the kinematic relations of the Compton effect and noting that the quantities $\mathbf{g}$ and $\mathbf{h}$, contained respectively in P' and Q', are given by (see Eq. (15.15))

$$\mathbf{g} = \mathbf{p} + \hbar\mathbf{k}, \qquad \mathbf{h} = \mathbf{p} - \hbar\mathbf{k}' = \mathbf{p}' - \hbar\mathbf{k},$$

we can perform the following substitution in the expression of P'

$$\hbar \mathbf{k} = \mathbf{g} - \mathbf{p},$$

and in the expression of Q'

$$\hbar \mathbf{k} = \mathbf{p}' - \mathbf{h}.$$

Substituting, and recalling also that $\epsilon - \hbar\omega' = \epsilon' - \hbar\omega$, we obtain

$$c\hbar k\, P' = W_s^\dagger(\mathbf{p}')(\boldsymbol{\alpha} \cdot \mathbf{e}'^*)\frac{H_\mathbf{g} + \epsilon + \hbar\omega}{(\epsilon + \hbar\omega)^2 - \epsilon_\mathbf{g}^2}[c\boldsymbol{\alpha} \cdot \mathbf{g} - c\boldsymbol{\alpha} \cdot \mathbf{p} - \hbar\omega]W_r(\mathbf{p}),$$

$$c\hbar k\, Q' = W_s^\dagger(\mathbf{p}')[c\boldsymbol{\alpha} \cdot \mathbf{p}' - c\boldsymbol{\alpha} \cdot \mathbf{h} - \hbar\omega]\frac{H_\mathbf{h} + \epsilon' - \hbar\omega}{(\epsilon' - \hbar\omega)^2 - \epsilon_\mathbf{h}^2}(\boldsymbol{\alpha} \cdot \mathbf{e}'^*)W_r(\mathbf{p}).$$

Within the square brackets, we add and subtract the factor βmc^2 and recall that an expression of the type $(c\boldsymbol{\alpha} \cdot \mathbf{q} + \beta mc^2)$, with $\mathbf{q}$ arbitrary, is the Dirac Hamiltonian $H_\mathbf{q}$. We obtain

$$c\hbar k\, P' = W_s^\dagger(\mathbf{p}')(\boldsymbol{\alpha} \cdot \mathbf{e}'^*)\frac{H_\mathbf{g} + \epsilon + \hbar\omega}{(\epsilon + \hbar\omega)^2 - \epsilon_\mathbf{g}^2}[H_\mathbf{g} - H_\mathbf{p} - \hbar\omega]W_r(\mathbf{p}),$$

$$c\hbar k\, Q' = W_s^\dagger(\mathbf{p}')[H_{\mathbf{p}'} - H_\mathbf{h} - \hbar\omega]\frac{H_\mathbf{h} + \epsilon' - \hbar\omega}{(\epsilon' - \hbar\omega)^2 - \epsilon_\mathbf{h}^2}(\boldsymbol{\alpha} \cdot \mathbf{e}'^*)W_r(\mathbf{p}).$$

Now, noting that

$$H_\mathbf{p} W_r(\mathbf{p}) = \epsilon\, W_r(\mathbf{p}), \qquad W_s^\dagger(\mathbf{p}')H_{\mathbf{p}'} = \epsilon' W_s^\dagger(\mathbf{p}'),$$

we have

$$c\hbar k\, P' = W_s^\dagger(\mathbf{p}')(\boldsymbol{\alpha} \cdot \mathbf{e}'^*)\frac{H_\mathbf{g} + \epsilon + \hbar\omega}{(\epsilon + \hbar\omega)^2 - \epsilon_\mathbf{g}^2}[H_\mathbf{g} - \epsilon - \hbar\omega]W_r(\mathbf{p}),$$

$$c\hbar k\, Q' = W_s^\dagger(\mathbf{p}')[\epsilon' - H_\mathbf{h} - \hbar\omega]\frac{H_\mathbf{h} + \epsilon' - \hbar\omega}{(\epsilon' - \hbar\omega)^2 - \epsilon_\mathbf{h}^2}(\boldsymbol{\alpha} \cdot \mathbf{e}'^*)W_r(\mathbf{p}).$$

Finally, taking into account that

$$[H_\mathbf{g} + \epsilon + \hbar\omega][H_\mathbf{g} - \epsilon - \hbar\omega] = \epsilon_\mathbf{g}^2 - (\epsilon + \hbar\omega)^2,$$

$$\left[\epsilon' - H_\mathbf{h} - \hbar\omega\right]\left[H_\mathbf{h} + \epsilon' - \hbar\omega\right] = \left(\epsilon' - \hbar\omega\right)^2 - \epsilon_\mathbf{h}^2,$$

we obtain

$$c\hbar k\, P' = -W_s^\dagger(\mathbf{p}')(\boldsymbol{\alpha} \cdot \mathbf{e}'^*)W_r(\mathbf{p}), \qquad c\hbar k\, Q' = W_s^\dagger(\mathbf{p}')(\boldsymbol{\alpha} \cdot \mathbf{e}'^*)W_r(\mathbf{p}),$$

from which it follows that $(P' + Q') = 0$. This shows that the quantity $\mathcal{R}_{\mathrm{fi}}$ is invariant with respect to the transformation (16.26). In all similarity, it can be shown that $\mathcal{R}_{\mathrm{fi}}$ is also invariant with respect to the transformation

$$\boldsymbol{\alpha} \cdot \mathbf{e}'^* \to \boldsymbol{\alpha} \cdot \mathbf{e}'^* + C'\big(\boldsymbol{\alpha} \cdot \mathbf{u}' - 1\big),$$

where C' is an arbitrary constant and where $\mathbf{u}'$ is the unit vector of the direction of the final photon ($\mathbf{u}' = \mathbf{k}'/k'$).

An alternative way to express these invariant properties is to formally consider the quantity $\mathcal{R}_{\mathrm{fi}}$ as a function of the matrix $(\boldsymbol{\alpha} \cdot \mathbf{e})$, or, alternatively, of the matrix $(\boldsymbol{\alpha} \cdot \mathbf{e}'^*)$. From the above proof it follows that

$$\mathcal{R}_{\mathrm{fi}}\{\boldsymbol{\alpha} \cdot \mathbf{e} \to \boldsymbol{\alpha} \cdot \mathbf{u}\} = \mathcal{R}_{\mathrm{fi}}\{\boldsymbol{\alpha} \cdot \mathbf{e} \to 1\},$$
$$\mathcal{R}_{\mathrm{fi}}\big\{\boldsymbol{\alpha} \cdot \mathbf{e}'^* \to \boldsymbol{\alpha} \cdot \mathbf{u}'\big\} = \mathcal{R}_{\mathrm{fi}}\big\{\boldsymbol{\alpha} \cdot \mathbf{e}'^* \to 1\big\}. \tag{16.27}$$

16.15 The Gamma Matrices and the Relativistic Invariants

The relation between energy $\epsilon_{\mathbf{p}}$ and momentum $\mathbf{p}$ of a relativistic particle of mass m is

$$\epsilon_{\mathbf{p}}^2 = c^2 p^2 + m^2 c^4.$$

In particular, for a photon ($m = 0$) we have

$$\epsilon_{\mathbf{p}} = cp, \quad \text{with } p = |\mathbf{p}|,$$

or, in terms of frequency and wavenumber

$$\hbar\omega = c\hbar k, \quad \text{with } k = |\mathbf{k}|.$$

Such relations may be formally simplified if we adopt a unit system in which $\hbar = c = 1$. The introduction of this convention is equivalent to define the unit time interval as the time needed by light to travel the unit of length. With this definition, the energy, the momentum, and the mass (and similarly for a photon, the angular frequency and the wavenumber) all assume the dimensions of the reciprocal of a length (or a time). The relation between momentum and energy is written, in this system of units, in the form

$$\epsilon_{\mathbf{p}}^2 = p^2 + m^2, \quad \text{or} \quad \epsilon_{\mathbf{p}}^2 - p^2 = m^2,$$

and for a photon

$$\epsilon_{\mathbf{p}} = p, \quad \text{or} \quad \omega = k.$$

We now introduce with the symbol $\mathcal{P}_\mu$ ($\mu = 0, 1, 2, 3$) the quadrivector momentum-energy of the particle. It is an entity with four components that are defined in this

way

$$\mathcal{P}_0 = \epsilon_{\mathbf{p}}, \qquad \mathcal{P}_1 = p_1 = p_x, \qquad \mathcal{P}_2 = p_2 = p_y, \qquad \mathcal{P}_3 = p_3 = p_z,$$

or, in a more compact form

$$\mathcal{P} = (\epsilon_{\mathbf{p}}, \mathbf{p}).$$

Defining the metric tensor $g_{\mu\nu}$ as

$$g_{00} = 1, \qquad g_{0i} = g_{i0} = 0 \quad (i = 1, 2, 3), \qquad g_{jk} = -\delta_{ik} \quad (j, k = 1, 2, 3),$$

the scalar product of two quadrivectors $\mathcal{P}$ and $\mathcal{Q}$ is

$$(\mathcal{P}\mathcal{Q}) = \sum_{\mu\nu} g_{\mu\nu} \mathcal{P}_\mu \mathcal{Q}_\nu = \epsilon_{\mathbf{p}} \epsilon_{\mathbf{q}} - \mathbf{p} \cdot \mathbf{q}.$$

In particular we have

$$\mathcal{P}^2 = (\mathcal{P}\mathcal{P}) = \sum_{\mu\nu} g_{\mu\nu} \mathcal{P}_\mu \mathcal{P}_\nu = \epsilon_{\mathbf{p}}^2 - p^2 = m^2.$$

These quantities (the scalar product of two quadrivectors defined by the above metric tensor and, in particular, the square of a quadrivector) are relativistic invariants, i.e. do not change under Lorentz transformations. We are now going to show how the probability amplitudes of Compton scattering can be expressed in terms of these invariants.

Consider the quantity $\mathcal{R}_{\mathrm{fi}}$ defined in Eq. (15.22). This quantity is composed of two terms that we denote by P and Q. For the first one we have, taking into account the system of units we have introduced ($c = \hbar = 1$)

$$P = W_s^\dagger(\mathbf{p}')(\boldsymbol{\alpha} \cdot \mathbf{e}'^*) \frac{H_{\mathbf{g}} + \epsilon + \omega}{(\epsilon + \omega)^2 - \epsilon_{\mathbf{g}}^2} (\boldsymbol{\alpha} \cdot \mathbf{e}) W_r(\mathbf{p}),$$

where

$$H_{\mathbf{g}} = \boldsymbol{\alpha} \cdot \mathbf{g} + \beta m,$$

with

$$\mathbf{g} = \mathbf{p} + \mathbf{k}, \qquad \epsilon_{\mathbf{g}} = \sqrt{g^2 + m^2}.$$

Recalling that the square of the Dirac matrix β is unity, we can write

$$P = W_s^\dagger(\mathbf{p}')(\boldsymbol{\alpha} \cdot \mathbf{e}'^*)\beta^2 \frac{H_{\mathbf{g}} + \epsilon + \omega}{(\epsilon + \omega)^2 - \epsilon_{\mathbf{g}}^2} \beta^2 (\boldsymbol{\alpha} \cdot \mathbf{e}) W_r(\mathbf{p}).$$

If we now also recall that the Dirac matrix β anticommutes with any of the $\boldsymbol{\alpha}$ matrices, we obtain

$$P = W_s^\dagger(\mathbf{p}')\beta(\boldsymbol{\alpha} \cdot \mathbf{e}'^*) \frac{\beta(H_{\mathbf{g}} + \epsilon + \omega)}{(\epsilon + \omega)^2 - \epsilon_{\mathbf{g}}^2} \beta(\boldsymbol{\alpha} \cdot \mathbf{e})\beta W_r(\mathbf{p}).$$

We now define the matrices γ_μ $(\mu = 0, 1, 2, 3)$

$$\gamma_0 = \beta, \qquad \gamma_1 = \beta\alpha_1, \qquad \gamma_2 = \beta\alpha_2, \qquad \gamma_3 = \beta\alpha_3.$$

The fundamental property of these matrices regards their anticommutator which is (as easily derived from the properties of the $\boldsymbol{\alpha}$ and β matrices)

$$\{\gamma_\mu, \gamma_\nu\} = \gamma_\mu\gamma_\nu + \gamma_\nu\gamma_\mu = 2g_{\mu\nu}, \tag{16.28}$$

where $g_{\mu\nu}$ is the metric tensor that we have previously defined. Moreover, given an arbitrary quadrivector $\mathcal{V}$, we define by the symbol $\slashed{V}$ the matrix

$$\slashed{V} = \sum_\mu \gamma_\mu \mathcal{V}_\mu.$$

With these definitions, the quantity P can be written in the form

$$P = W_s^\dagger(\mathbf{p}')\slashed{\mathcal{E}}'^* \frac{\slashed{\mathcal{G}} + m}{G^2 - m^2} \slashed{\mathcal{E}} \gamma_0 W_r(\mathbf{p}), \tag{16.29}$$

where the quadrivectors $\mathcal{G}$, $\mathcal{E}$ and $\mathcal{E}'$ are given by

$$\mathcal{G} = (\omega + \epsilon, \mathbf{g}), \qquad \mathcal{E} = (0, \mathbf{e}), \qquad \mathcal{E}' = (0, \mathbf{e}').$$

We note that the quadrivector $\mathcal{G}$ can also be written in the form

$$\mathcal{G} = \mathcal{P} + \mathcal{K},$$

where

$$\mathcal{K} = (\omega, \mathbf{k}).$$

If we now consider the quantity P^*, complex conjugate of P, we need to proceed carefully because the γ matrices (except γ_0) are not Hermitian. We have in fact

$$\gamma_0^\dagger = \beta^\dagger = \beta = \gamma_0,$$

$$\gamma_i^\dagger = (\beta\alpha_i)^\dagger = \alpha_i^\dagger\beta^\dagger = \alpha_i\beta = -\beta\alpha_i = -\gamma_i \quad (i = 1, 2, 3).$$

These properties can be summarised in only one relation

$$\gamma_\mu^\dagger = \gamma_0\gamma_\mu\gamma_0,$$

which implies, for an arbitrary quadrivector

$$\slashed{V}^\dagger = \gamma_0\slashed{V}^*\gamma_0.$$

We therefore obtain

$$P^* = W_r^\dagger(\mathbf{p})\gamma_0\gamma_0\slashed{\mathcal{E}}^*\gamma_0\frac{\gamma_0\slashed{\mathcal{G}}\gamma_0 + m}{G^2 - m^2}\gamma_0\slashed{\mathcal{E}}'\gamma_0 W_s(\mathbf{p}'),$$

or

$$P^* = W_r^\dagger(\mathbf{p})\slashed{\epsilon}^* \frac{\slashed{G} + m}{G^2 - m^2}\slashed{\epsilon}' \gamma_0 W_s(\mathbf{p}').\tag{16.30}$$

Similar considerations can be repeated for the other term Q of Eq. (15.22) defined by

$$Q = W_s^\dagger(\mathbf{p}')(\boldsymbol{\alpha}\cdot\mathbf{e})\frac{H_\mathbf{h} + \epsilon - \omega'}{(\epsilon - \omega')^2 - \epsilon_\mathbf{h}^2}(\boldsymbol{\alpha}\cdot\mathbf{e}'^*)W_r(\mathbf{p}),$$

where

$$H_\mathbf{h} = \boldsymbol{\alpha}\cdot\mathbf{h} + \beta m,$$

with

$$\mathbf{h} = \mathbf{p} - \mathbf{k}', \qquad \epsilon_\mathbf{h} = \sqrt{h^2 + m^2}.$$

We have

$$Q = W_s^\dagger(\mathbf{p}')\slashed{\epsilon}\frac{\slashed{H} + m}{H^2 - m^2}\slashed{\epsilon}'^* \gamma_0 W_r(\mathbf{p}),\tag{16.31}$$

where the quadrivector $\mathcal{H}$ is defined by

$$\mathcal{H} = (\epsilon - \omega', \mathbf{h}) = \mathcal{P} - \mathcal{K}',$$

being

$$\mathcal{K}' = (\omega', \mathbf{k}').$$

Still in analogy to what discussed before, we also have

$$Q^* = W_r^\dagger(\mathbf{p})\slashed{\epsilon}'\frac{\slashed{H} + m}{H^2 - m^2}\slashed{\epsilon}^* \gamma_0 W_s(\mathbf{p}').\tag{16.32}$$

We can now evaluate the square of the modulus of the quantity $\mathcal{R}_{\text{fi}}$. It is

$$|\mathcal{R}_{\text{fi}}|^2 = (P + Q)(P^* + Q^*) = PP^* + PQ^* + QP^* + QQ^*,$$

where, using Eqs. (16.29)–(16.32), the four terms are given by

$$PP^* = W_s^\dagger(\mathbf{p}')\slashed{\epsilon}'^* \frac{\slashed{G} + m}{G^2 - m^2}\slashed{\epsilon}\gamma_0 W_r(\mathbf{p})W_r^\dagger(\mathbf{p})\slashed{\epsilon}^* \frac{\slashed{G} + m}{G^2 - m^2}\slashed{\epsilon}'\gamma_0 W_s(\mathbf{p}'),$$

$$PQ^* = W_s^\dagger(\mathbf{p}')\slashed{\epsilon}'^* \frac{\slashed{G} + m}{G^2 - m^2}\slashed{\epsilon}\gamma_0 W_r(\mathbf{p})W_r^\dagger(\mathbf{p})\slashed{\epsilon}' \frac{\slashed{H} + m}{H^2 - m^2}\slashed{\epsilon}^* \gamma_0 W_s(\mathbf{p}'),$$

$$QP^* = W_s^\dagger(\mathbf{p}')\slashed{\epsilon}\frac{\slashed{H} + m}{H^2 - m^2}\slashed{\epsilon}'^* \gamma_0 W_r(\mathbf{p})W_r^\dagger(\mathbf{p})\slashed{\epsilon}^* \frac{\slashed{G} + m}{G^2 - m^2}\slashed{\epsilon}'\gamma_0 W_s(\mathbf{p}'),$$

$$QQ^* = W_s^\dagger(\mathbf{p}')\slashed{\epsilon}\frac{\slashed{H} + m}{H^2 - m^2}\slashed{\epsilon}'^* \gamma_0 W_r(\mathbf{p})W_r^\dagger(\mathbf{p})\slashed{\epsilon}' \frac{\slashed{H} + m}{H^2 - m^2}\slashed{\epsilon}^* \gamma_0 W_s(\mathbf{p}').$$

These expressions can be simplified when one considers the average over the initial spin states of the electron and the sum over the final spin states of the electron. Taking into account the results we have obtained in Sect. 15.5 (Eqs. (15.18) and (15.20)) we have

$$\sum_{r=1,2} W_r\left(\mathbf{p}\right) W_r^\dagger(\mathbf{p}) = \frac{\epsilon_\mathbf{p} + H_\mathbf{p}}{2\epsilon_\mathbf{p}}, \qquad \sum_{s=1,2} W_s\left(\mathbf{p}'\right) W_s^\dagger(\mathbf{p}') = \frac{\epsilon_{\mathbf{p}'} + H_{\mathbf{p}'}}{2\epsilon_{\mathbf{p}'}},$$

and defining the quadrivector

$$\mathcal{P} = (\epsilon_\mathbf{p}, \mathbf{p}), \qquad \mathcal{P}' = \left(\epsilon_{\mathbf{p}'}, \mathbf{p}'\right),$$

we obtain

$$\sum_{r=1,2} \gamma_0 W_r(\mathbf{p}) W_r^\dagger(\mathbf{p}) = \frac{\slashed{\mathcal{P}} + m}{2\epsilon_\mathbf{p}}, \qquad \sum_{s=1,2} \gamma_0 W_s\left(\mathbf{p}'\right) W_s^\dagger(\mathbf{p}') = \frac{\slashed{\mathcal{P}}' + m}{2\epsilon_{\mathbf{p}'}}.$$

By denoting with the symbol $\langle \cdots \rangle$ the average over the spin states and using the definition of the trace of a matrix, according to which a scalar product of the form $W_1 \mathcal{X} W_2^\dagger$, with W_1 and W_2 arbitrary spinors and $\mathcal{X}$ an arbitrary matrix, can be written in the form $\text{Tr}(W_2 W_1^\dagger \mathcal{X})$, we have

$$\langle P P^* \rangle = \frac{1}{2} \sum_{r=1,2} \sum_{s=1,2} P P^*$$

$$= \frac{1}{8\epsilon_\mathbf{p}\epsilon_{\mathbf{p}'}} \text{Tr}\left\{ (\slashed{\mathcal{P}}' + m)\slashed{\mathcal{E}}'^* \frac{\slashed{\mathcal{G}} + m}{\mathcal{G}^2 - m^2} \slashed{\mathcal{E}} (\slashed{\mathcal{P}} + m)\slashed{\mathcal{E}}^* \frac{\slashed{\mathcal{G}} + m}{\mathcal{G}^2 - m^2} \slashed{\mathcal{E}}' \right\},$$

with similar expressions for the other three terms $\langle P Q^* \rangle$, $\langle Q P^* \rangle$, and $\langle Q Q^* \rangle$.

This last result can be greatly simplified if we sum over the polarisation states of the final photon and we average over the polarisation states of the initial photon. The average over the polarisation states of the initial photon, for example, is obtained by applying to the previous formula the formal substitution

$$\slashed{\mathcal{E}} (\slashed{\mathcal{P}} + m)\slashed{\mathcal{E}}^* \rightarrow \frac{1}{2} \sum_{i=1,2} \slashed{\mathcal{E}}^{(i)} (\slashed{\mathcal{P}} + m)\slashed{\mathcal{E}}^{(i)},$$

where

$$\mathcal{E}^{(i)} = \left(0, \mathbf{e}^{(i)}\right),$$

$\mathbf{e}^{(i)}$ ($i = 1, 2$) being two unit vectors that we can assume real, perpendicular to each other and perpendicular to the direction of the initial photon. Taking into account the invariance under gauge transformations described in Sect. 16.14 and, in particular, recalling Eq. (16.27), the sum can be modified by extending it to a third "unit quadrivector" (that we denote by $\mathcal{E}^{(3)}$) and subtracting then the contribution

from another unit quadrivector. According to special relativity, this unit quadrivec-
tor, which we denote by $\mathcal{E}^{(0)}$, is of the purely temporal type. Defining

$$\mathcal{E}^{(3)} = \left(0, \mathbf{e}^{(3)}\right), \qquad \mathcal{E}^{(0)} = (1, \mathbf{0}),$$

where $\mathbf{e}^{(3)}$ is an unit vector directed along the direction of the incoming photon
($\mathbf{e}^{(3)} = \mathbf{k}/k$), and recalling the definition of the metric tensor, we apply the following
transformation

$$\not{\mathcal{E}}(\not{P} + m)\not{\mathcal{E}}^* \to \frac{1}{2}\mathcal{S}, \quad \text{where } \mathcal{S} = -\sum_{i,j=0}^{3} g_{ij}\not{\mathcal{E}}^{(i)}(\not{P} + m)\not{\mathcal{E}}^{(j)}.$$

We note that the sum $\mathcal{S}$ can also be written in the form

$$\mathcal{S} = -\gamma_0(\not{P} + m)\gamma_0 + \left(\boldsymbol{\gamma} \cdot \mathbf{e}^{(1)}\right)(\not{P} + m)\left(\boldsymbol{\gamma} \cdot \mathbf{e}^{(1)}\right)$$
$$+ \left(\boldsymbol{\gamma} \cdot \mathbf{e}^{(2)}\right)(\not{P} + m)\left(\boldsymbol{\gamma} \cdot \mathbf{e}^{(2)}\right) + \left(\boldsymbol{\gamma} \cdot \mathbf{e}^{(3)}\right)(\not{P} + m)\left(\boldsymbol{\gamma} \cdot \mathbf{e}^{(3)}\right),$$

where $\boldsymbol{\gamma}$ is the formal vector defined by $\boldsymbol{\gamma} = (\gamma_1, \gamma_2, \gamma_3)$. The right-hand side can
then be transformed to get

$$\mathcal{S} = -\gamma_0(\not{P} + m)\gamma_0 + \sum_{i,j=1}^{3} \gamma_i(\not{P} + m)\gamma_j\left[e_i^{(1)}e_j^{(1)} + e_i^{(2)}e_j^{(2)} + e_i^{(3)}e_j^{(3)}\right].$$

On the other hand, taking into account Eq. (15.7), the quantity in square brackets is
equal to the Kronecker delta δ_{ij}, so we obtain

$$\mathcal{S} = -\gamma_0(\not{P} + m)\gamma_0 + \sum_{i=1}^{3} \gamma_i(\not{P} + m)\gamma_i = -\sum_{\mu,\nu} g_{\mu\nu}\gamma_\mu(\not{P} + m)\gamma_\nu.$$

Finally, we take into account the properties of the γ matrices. From Eq. (16.28) we
have

$$\gamma_\mu\gamma_\nu = -\gamma_\nu\gamma_\mu + 2g_{\mu\nu}.$$

Moreover, it is easy to verify that the following relation holds

$$\sum_{\mu\nu} g_{\mu\nu}\gamma_\mu\gamma_\nu = 4,$$

and that, given the properties of the metric tensor,

$$\sum_{\mu} g_{\mu\nu}g_{\mu\rho} = \delta_{\nu\rho},$$

so that

$$\sum_{\mu\nu} g_{\mu\nu}g_{\mu\rho}\gamma_\nu = \gamma_\rho.$$

Taking advantage of these properties, we get after some algebra

$$\mathcal{S} = 2(\not{P} - 2m).$$

Summarising the foregoing, the average on the states of polarisation of the initial photon is obtained by performing the formal transformation

$$\not{\mathcal{E}}(\not{P} + m)\not{\mathcal{E}}^* \rightarrow \not{P} - 2m.$$

Similarly, the sum over the polarisation states the final photon is obtained by performing the formal transformation

$$\not{\mathcal{E}}'(\not{P}' + m)\not{\mathcal{E}}'^* \rightarrow 2(\not{P}' - 2m).$$

Now we denote by the symbol $\langle\langle PP^* \rangle\rangle$ the quantity obtained by taking the average of $\langle PP^* \rangle$ over the states of initial polarisation and the sum of the same quantity over the states of final polarisation.[10] We have

$$\langle\langle PP^* \rangle\rangle = \frac{1}{4\epsilon_{\mathbf{p}}\epsilon_{\mathbf{p}'}} \, \mathrm{Tr}\left\{ (\not{P}' - 2m)\frac{\not{\mathcal{G}} + m}{\mathcal{G}^2 - m^2}(\not{P} - 2m)\frac{\not{\mathcal{G}} + m}{\mathcal{G}^2 - m^2} \right\}.$$

At this point it is necessary to briefly discuss the traces of the products of the γ matrices. It is easy to verify that the trace of the product of an odd number of γ matrices is null. When instead the number of γ matrices is zero or even, the result is different from zero. Denoting by a an arbitrary constant, with $\mathcal{A}$, $\mathcal{B}$, $\mathcal{C}$, and $\mathcal{D}$ four arbitrary quadrivectors, and recalling the definition of the scalar product of quadrivectors, we have

$$\mathrm{Tr}\{a\} = 4a, \qquad \mathrm{Tr}\{\not{\mathcal{A}}\not{\mathcal{B}}\} = 4(\mathcal{A}\mathcal{B}),$$

$$\mathrm{Tr}\{\not{\mathcal{A}}\not{\mathcal{B}}\not{\mathcal{C}}\not{\mathcal{D}}\} = 4\big[(\mathcal{A}\mathcal{B})(\mathcal{C}\mathcal{D}) - (\mathcal{A}\mathcal{C})(\mathcal{B}\mathcal{D}) + (\mathcal{A}\mathcal{D})(\mathcal{B}\mathcal{C})\big].$$

The first relation is obvious. For the second one we have

$$\mathrm{Tr}\{\not{\mathcal{A}}\not{\mathcal{B}}\} = \sum_{\mu\nu} \mathrm{Tr}\{\gamma_\mu \gamma_\nu\} A_\mu B_\nu,$$

and using the anticommutation property of the γ matrices

$$\mathrm{Tr}\{\gamma_\mu \gamma_\nu\} = 8g_{\mu\nu} - \mathrm{Tr}\{\gamma_\nu \gamma_\mu\}.$$

From the cyclic property of the trace it then follows that

$$\mathrm{Tr}\{\gamma_\mu \gamma_\nu\} = 4g_{\mu\nu},$$

---

[10]Recall that the first average, $\langle PP^* \rangle$, has a similar meaning with respect to the spin states of the electron.

which proves by simple substitution the second relation. For the third relation we have

$$\text{Tr}\{\mathcal{A}\mathcal{B}\mathcal{C}\mathcal{D}\} = \sum_{\mu\nu\rho\sigma} \text{Tr}\{\gamma_\mu\gamma_\nu\gamma_\rho\gamma_\sigma\}\mathcal{A}_\mu\mathcal{B}_\nu\mathcal{C}_\rho\mathcal{D}_\sigma,$$

and, for the anticommutation property of the γ matrices,

$$\begin{aligned}
\text{Tr}\{\gamma_\mu\gamma_\nu\gamma_\rho\gamma_\sigma\} &= 2g_{\mu\nu}\,\text{Tr}\{\gamma_\rho\gamma_\sigma\} - \text{Tr}\{\gamma_\nu\gamma_\mu\gamma_\rho\gamma_\sigma\} \\
&= 2g_{\mu\nu}\,\text{Tr}\{\gamma_\rho\gamma_\sigma\} - 2g_{\mu\rho}\,\text{Tr}\{\gamma_\nu\gamma_\sigma\} + \text{Tr}\{\gamma_\nu\gamma_\rho\gamma_\mu\gamma_\sigma\} \\
&= 2g_{\mu\nu}\,\text{Tr}\{\gamma_\rho\gamma_\sigma\} - 2g_{\mu\rho}\,\text{Tr}\{\gamma_\nu\gamma_\sigma\} + 2g_{\mu\sigma}\,\text{Tr}\{\gamma_\nu\gamma_\rho\} \\
&\quad - \text{Tr}\{\gamma_\nu\gamma_\rho\gamma_\sigma\gamma_\mu\}.
\end{aligned}$$

Using the cyclic property of the trace we then have

$$\text{Tr}\{\gamma_\mu\gamma_\nu\gamma_\rho\gamma_\sigma\} = g_{\mu\nu}\,\text{Tr}\{\gamma_\rho\gamma_\sigma\} - g_{\mu\rho}\,\text{Tr}\{\gamma_\nu\gamma_\sigma\} + g_{\mu\sigma}\,\text{Tr}\{\gamma_\nu\gamma_\rho\},$$

and using of the result previously obtained

$$\text{Tr}\{\gamma_\mu\gamma_\nu\gamma_\rho\gamma_\sigma\} = 4g_{\mu\nu}g_{\rho\sigma} - 4g_{\mu\rho}g_{\nu\sigma} + 4g_{\mu\sigma}g_{\nu\rho}.$$

The third relation is then finally obtained by simple substitution of this identity.

The result obtained for $\langle\!\langle PP^*\rangle\!\rangle$ shows that the trace contained in this quantity can be expressed exclusively in terms of scalar products of quadrivectors, i.e. in terms of relativistic invariants. Similar considerations can then be repeated for the other quantities $\langle\!\langle PQ^*\rangle\!\rangle$, $\langle\!\langle QP^*\rangle\!\rangle$, and $\langle\!\langle QQ^*\rangle\!\rangle$, which, once calculated, allow one to obtain the transition probability per unit time and the cross section. Obviously, in the particular case in which the electron is initially at rest, one finds again for the cross section the Klein-Nishina equation in the form of Eq. (15.37), which refers to the average over the polarisation states of the initial photon and the sum over the polarisation states of the final photon.

The formalism of the γ matrices presented in this chapter is very powerful and elegant. It allows one to deal with relative ease even with the most complex problems in quantum electrodynamics. In any case, we emphasize that the formalism that we have used in the text to deduce the Klein-Nishina equation, which does not make use of the γ matrices, was the first to be used in the applications.

16.16 Physical Constants

The constants are expressed in the cgs system of units with at most six significant digits.

Constant of gravitation: $G = 6.67428 \times 10^{-8}\ \text{cm}^3\,\text{g}^{-1}\,\text{s}^{-2}$
Velocity of light in vacuum: $c = 2.99792 \times 10^{10}\ \text{cm}\,\text{s}^{-1}$
Planck constant: $h = 6.62607 \times 10^{-27}\ \text{erg}\,\text{s}$

Reduced Planck constant: $\hbar = h/(2\pi) = 1.05457 \times 10^{-27}$ erg s

Boltzmann constant: $k_B = 1.38065 \times 10^{-16}$ erg K^{-1}

Charge of the electron (absolute value): $e_0 = 4.80320 \times 10^{-10}$ esu

Electron mass: $m = 9.10938 \times 10^{-28}$ g

Reduced electron mass: $m_r = mM_p/(m + M_p) = 9.10442 \times 10^{-28}$ g

Atomic mass unit (amu): $m_H = 1.66054 \times 10^{-24}$ g

Proton mass: $M_p = 1.67262 \times 10^{-24}$ g

Proton/electron mass ratio: $M_p/m = 1.83615 \times 10^3$

Avogadro constant: $N_A = 6.02214 \times 10^{23}$ mol^{-1}

Fine-structure constant: $\alpha = e_0^2/(\hbar c) = 7.29735 \times 10^{-3}$

Reciprocal of the fine-structure constant: $1/\alpha = \hbar c/e_0^2 = 137.036$

Classical radius of the electron: $r_c = e_0^2/(mc^2) = 2.81794 \times 10^{-13}$ cm

Compton wavelength of the electron: $\lambda_C = h/(mc) = 2.42631 \times 10^{-10}$ cm

Radius of the first Bohr orbit: $a_0 = \hbar^2/(me_0^2) = 5.29177 \times 10^{-9}$ cm

Rydberg constant: $R = me_0^4/(4\pi c\hbar^3) = 1.09737 \times 10^5$ cm^{-1}

Rydberg constant (hydrogen atom): $R_H = m_r e_0^4/(4\pi c\hbar^3) = 1.09677 \times 10^5$ cm^{-1}

Bohr magneton: $\mu_0 = e_0\hbar/(2mc) = 9.27401 \times 10^{-21}$ erg G^{-1}

Thomson cross section: $\sigma_T = 8\pi r_c^2/3 = 6.65246 \times 10^{-25}$ cm^2

Stefan-Boltzmann constant:[11] $\sigma = 5.67040 \times 10^{-5}$ erg cm^{-2} s^{-1} K^{-4}

Radiation density constant:[11] $a = 7.56577 \times 10^{-15}$ erg cm^{-3} K^{-4}

First radiation constant: $c_1 = 2\pi hc^2 = 3.74177 \times 10^{-5}$ erg cm^2 s^{-1}

Second radiation constant: $c_2 = hc/k_B = 1.43877$ cm K

---

[11] $\sigma = \frac{2\pi^5 k_B^4}{15h^3c^2}$, $a = \frac{4\sigma}{c} = \frac{8\pi^5 k_B^4}{15h^3c^3}$.