Relativistic quantum mechanics and statistics

Peter Schlagheck
Université de Liège

Chapter 1

Special relativity

1.1 Space-time

In the framework of the theory of special relativity, space-time is represented by the vector space $\mathbb{R}^4$. An element $x \in \mathbb{R}^4$ has four components and can be written as

$$x = (x^0, x^1, x^2, x^3) \equiv (x^0, \vec{x}) \tag{1.1}$$

where the spatial vector $\vec{x} = (x^1, x^2, x^3) \in \mathbb{R}^3$ contains the spatial coordinates of x and $x^0 = ct$ is its temporal coordinate. x will be called *four-vector* in the following. At this point, we introduce the speed of light c in order to assure that all four components of x possess the same dimensionality. It is after the introduction of the metric tensor, to be defined below, that this constant c acquires a true physical significance.

Within $\mathbb{R}^4$ we can choose a basis of four linearly independent vectors. The most natural choice for this basis is given by the vectors (e_0, e_1, e_2, e_3) where e_0 represents the unit vector along the time axis and e_j with $j = 1, 2, 3$ represent the unit vectors along the spatial x, y, and z axes, respectively. Any vector $x \in \mathbb{R}^4$ can now be decomposed within this standard basis. We obtain

$$x = x^0 e_0 + x^1 e_1 + x^2 e_2 + x^3 e_3 = \sum_{\nu=0}^{3} x^\nu e_\nu \,. \tag{1.2}$$

It is useful at this point to introduce the summation convention of Einstein. This convention simply consists in leaving out the symbol of summation $\sum_{\nu=0}^{3}$ in the above expression. In the following, we therefore write

$$x = x^\nu e_\nu \tag{1.3}$$

instead of $x = \sum_{\nu=0}^{3} x^\nu e_\nu$. This convention will be generalized to more complicated product expressions: if such an expression contains terms in which the same greek letter appears as index in the superscript at some place and as index in the subscript at some other place, we consider this expression to be summed over that very greek letter from 0 to 3.

The choice of the basis is not unique. We can choose another set of four vectors (e'_0, e'_1, e'_2, e'_3) as basis of $\mathbb{R}^4$ provided those new vectors are again linearly independent. Technically, this means that the coefficients $D_\nu{}^\mu \in \mathbb{R}$ describing the representation of the new basis vectors in the old basis according to

$$e'_\nu = \sum_{\mu=0}^{3} D_\nu{}^\mu e_\mu \equiv D_\nu{}^\mu e_\mu \,, \tag{1.4}$$

must form an invertible matrix $D \equiv (D_\nu{}^\mu)_{4\times 4} \in \mathbb{R}^{4\times 4}$. Each four-vector $x \in \mathbb{R}^4$ can now be represented within the old basis (e_0, e_1, e_2, e_3) or within the new basis (e'_0, e'_1, e'_2, e'_3) according to

$$x = x^\nu e_\nu = x'^\nu e'_\nu \tag{1.5}$$

giving rise to different sets of coordinates (x^0, x^1, x^2, x^3) and (x'^0, x'^1, x'^2, x'^3). Inserting Eq. (1.4) in Eq. (1.5), we then obtain the relation

$$x^\mu = D_\nu{}^\mu x'^\nu \tag{1.6}$$

which, after the inversion of the matrix D, yields the corresponding transformation rule for the coordinates:

$$x'^\nu = (D^{-1})_\mu{}^\nu x^\mu \tag{1.7}$$

As a general convention for such transformation matrices $D \equiv (D_\nu{}^\mu)_{4\times 4}$, the element $D_\nu{}^\mu$ refers to the νth row and the μth column (with ν and μ varying from 0 to 3). Eq. (1.7) therefore effectively describes a matrix-vector multiplication of the transpose of the inverse of D with the components of the four-vector x.

In order to properly introduce the notion of distances within the space of four-vectors, we need to define a metric within $\mathbb{R}^4$. This metric is introduced by the definition of a scalar product, *i.e.* by a bilinear transformation $\mathbb{R}^4 \times \mathbb{R}^4 \to \mathbb{R}$, $(x, y) \mapsto xy \in \mathbb{R}$ with

$$xy = g_{\mu\nu} x^\mu y^\nu \equiv \sum_{\mu=0}^{3} \sum_{\nu=0}^{3} g_{\mu\nu} x^\mu y^\nu \tag{1.8}$$

of the two four-vectors $x, y \in \mathbb{R}^4$. The coefficients $g_{\mu\nu} \in \mathbb{R}$ represent a tensorial object with two indices, which can also be presented in the form of a 4×4 matrix $g \equiv (g_{\mu\nu})_{4\times 4} \in \mathbb{R}^{4\times 4}$. We shall name it *metric tensor* in the following. It is generally required that the scalar product be symmetric, *i.e.* $xy = yx$, which implies $g_{\mu\nu} = g_{\nu\mu}$ for all $\mu, \nu = 0, \ldots, 3$, *i.e.* the metric tensor is symmetric as well. Furthermore, we require that $(g_{\mu\nu}) \in \mathbb{R}^{4\times 4}$ be invertible. Generally, a more stringent criterion for the definition of a scalar product is positivity, *i.e.* $xx > 0$

for all $x \neq (0,0,0,0)$. This latter criterion shall not be respected for the definition of the metric tensor within the space-time.

By means of the metric tensor, we can define the *covector* (or *dual vector*) that is associated with a given four-vector $x = x^\nu e_\nu$. This covector is defined by the coordinates

$$x_\nu = g_{\nu\mu} x^\mu \equiv \sum_{\mu=0}^{3} g_{\nu\mu} x^\mu \tag{1.9}$$

which are also named *covariant* coordinates of x, in opposition to the *contravariant* coordinates x^ν. The scalar product between the vectors x and y can then be written as

$$xy = g_{\mu\nu} x^\mu y^\nu = x_\nu y^\nu = y_\mu x^\mu \,. \tag{1.10}$$

It is convenient to specify the inverse of the matrix $(g_{\mu\nu})$ by the matrix elements $g^{\mu\nu} \in \mathbb{R}$, *i.e.* we have

$$g_{\mu\alpha} g^{\alpha\nu} = \delta_\mu^\nu \equiv \begin{cases} 1 : \mu = \nu \\ 0 : \mu \neq \nu \end{cases} \,. \tag{1.11}$$

We can then introduce the *dual basis* (e^0, e^1, e^2, e^3) according to

$$e^\nu = g^{\nu\mu} e_\mu \,, \tag{1.12}$$

which with Eq. (1.11) yields

$$e_\mu = g_{\mu\nu} e^\nu \,. \tag{1.13}$$

It is then straightforward to verify that the covariant coordinates of the four-vector x describe the representation of x within this dual basis, *i.e.*

$$x = x^\nu e_\nu = x_\mu e^\mu \,. \tag{1.14}$$

In the presence of a basis transformation $e_\nu \mapsto e'_\nu = D_\nu{}^\mu e_\mu$, the dual basis transforms according to

$$e^\nu \mapsto e'^\nu = D^\nu{}_\mu e^\mu \tag{1.15}$$

where we define the elements of the dual transformation matrix as

$$D^\nu{}_\mu = g^{\alpha\nu} g_{\beta\mu} D_\alpha{}^\beta \,. \tag{1.16}$$

In analogy with Eq. (1.6), we can then relate the new covariant coordinates x'_ν with the old ones according to

$$x_\mu = D^\nu{}_\mu x'_\nu \tag{1.17}$$

which yields in extension to Eq. (1.5)

$$x = x_\nu e^\nu = x'_\nu e'^\nu \,. \tag{1.18}$$

Correspondingly, the new coefficients $g'_{\mu\nu}$ of the metric tensor under this basis transformation are related to the old ones according to

$$g_{\mu\nu} = D^{\alpha}{}_{\mu} D^{\beta}{}_{\nu} g'_{\alpha\beta} \,. \tag{1.19}$$

In the following, we specifically consider basis transformations that leave the form of the metric tensor invariant, *i.e.* we require $g'_{\mu\nu} = g_{\mu\nu}$ for all $\mu, \nu = 0, \ldots, 3$. Inserting this identification in Eq. (1.19), multiplying this equation on both sides with $g^{\nu\sigma}$, summing the resulting equation over ν (which is automatically implied by Einstein's summation convention), and using Eqs. (1.11) and (1.16), we finally obtain

$$\delta^{\sigma}_{\mu} = D^{\alpha}{}_{\mu} D_{\alpha}{}^{\sigma} \tag{1.20}$$

as necessary and sufficient condition for having $g'_{\mu\nu} = g_{\mu\nu}$ for all μ, ν. We then have

$$(D^{-1})_{\mu}{}^{\nu} = D^{\nu}{}_{\mu} \tag{1.21}$$

for the matrix elements of the inverse of D. This allows us to re-express Eq. (1.7) as

$$x'^{\nu} = D^{\nu}{}_{\mu} x^{\mu} \,. \tag{1.22}$$

The condition (1.20) for keeping $g' = g$ under the basis transformation D can also be reformulated in terms of products of matrices. To this end, we first note that Eq. (1.19) would be equivalent to the relation

$$g = D^T g D \tag{1.23}$$

under the identification $g' = g$, which involves the product of the matrices $D \equiv (D_{\nu}{}^{\mu})_{4\times 4}$, $g \equiv (g_{\mu\nu})_{4\times 4}$, and of the transpose D^T of D. Eq. (1.23) is obviously equivalent to

$$D^{-1} = g^{-1} D^T g \tag{1.24}$$

which directly yields Eq. (1.21) when being expressed in its individual matrix elements.

For the case of a Euclidean space, we would naturally choose the metric tensor as the identity matrix $\mathbb{I}_{4\times 4}$ within $\mathbb{R}^{4\times 4}$. In that case, Eq. (1.24) would simplify to the condition $D^T D = \mathbb{I}_{4\times 4}$. The set of basis transformations that leave the metric invariant is then given by the group of orthogonal transformations $O(4)$ describing rotations of the coordinate system eventually to be combined (in the case $\det D = -1$) with a mirror transformation.

1.2 Lorentz transformations

In the case of space-time, the metric tensor reads

$$g = \mathrm{diag}(1,-1,-1,-1) = \left(\begin{array}{c|ccc} 1 & 0 & 0 & 0 \\ \hline 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{array}\right) \equiv \left(\begin{array}{c|ccc} 1 & & & \\ \hline & -1 & & \\ & & -1 & \\ & & & -1 \end{array}\right) = g^{-1}\,. \tag{1.25}$$

Clearly, the scalar product defined by this metric tensor through Eq. (1.8) is not positive definite. "Distances" between two different four-vectors in space-time can therefore be negative or zero, the latter even if the two concerned four-vectors do not coincide. We also note that the temporal component plays a different role in the definition of the metric than the spatial components, which is the reason why we separate in Eq. (1.25) the first (temporal) row and column from the other three (spatial) rows and columns by means of horizontal and vertical lines, respectively. Equipped with the metric (1.25), the vector space $\mathbb{M} = \mathbb{R}^4$ is named *Minkowski space*.

As in the case of a Euclidean space, the set of basis transformations D that leave the metric invariant still forms a group. Indeed, it is straightforward to verify that $D = D_1 D_2$ satisfies Eq. (1.23) provided the matrices D_1 and D_2 satisfy Eq. (1.23) individually. Moreover, also the inverse matrices of D_1 and D_2 satisfy Eq. (1.23) as well as the identity matrix $\mathbb{I}_{4\times 4}$. We therefore define

$$G = \left\{ D \in \mathbb{R}^{4\times 4} \ : \ D^T g D = g \right\} \tag{1.26}$$

with g given by Eq. (1.25) as the *Lorentz group*. Elements $D \in G$ of the Lorentz group are named *Lorentz transformations*.

Using the fact that $\det g = -1$, we obtain from Eq. (1.23) $(\det D)^2 = 1$, *i.e.*, $\det D = \pm 1$ for any Lorentz transformation $D \in G$. As in the case of a Euclidean space, these two different possibilities correspond to transformations D that keep the orientation of the basis invariant ($\det D = +1$) or involve mirror operations ($\det D = -1$). In contrast to a Euclidean space, the Minkowski space also provides another kind of mirror operation which affects the temporal component only. Considering the uppermost leftmost element of the matrix equation (1.23), we have

$$1 = \left(D^T g D\right)_{00} = ({D_0}^0)^2 - ({D_1}^1)^2 - ({D_2}^2)^2 - ({D_3}^3)^2 \tag{1.27}$$

from which we infer

$$({D_0}^0)^2 = 1 + ({D_1}^1)^2 + ({D_2}^2)^2 + ({D_3}^3)^2 \geq 1\,, \tag{1.28}$$

i.e. we have either ${D_0}^0 \geq 1$ or ${D_0}^0 \leq -1$.

In analogy with the special orthogonal group $SO(4)$ in the Euclidean case, we can define a special sub-group of the Lorentz group, the *restricted Lorentz group*

$$\mathcal{L} = \left\{ D \in G \ : \ \det D = 1 \text{ and } {D_0}^0 \geq 1 \right\} \subset G\,, \tag{1.29}$$

which contains only *proper* Lorentz transformations, *i.e.* Lorentz transformations that keep the orientation of the spatial axes invariant and do not exchange the past with the future. We can then associate with each $D \in G$ a unique proper Lorentz transformation $\tilde{D} \in \mathcal{L}$ such that

- $D = \tilde{D}$ if $\det D = 1$ and ${D_0}^0 \geq 1$,
- $D = P\tilde{D}$ if $\det D = -1$ and ${D_0}^0 \geq 1$,
- $D = T\tilde{D}$ if $\det D = -1$ and ${D_0}^0 \leq -1$ and
- $D = PT\tilde{D}$ if $\det D = 1$ and ${D_0}^0 \leq -1$,

where we define by $P = g = \mathrm{diag}(1, -1, -1, -1)$ and $T = -g = \mathrm{diag}(-1, 1, 1, 1)$ the spatial and temporal mirror operations, respectively.

It therefore remains to characterize all possible proper Lorentz transformations $\tilde{D} \in \mathcal{L}$. As for the case of the special orthogonal group $SO(3)$ describing basis transformations in a three-dimensional Euclidean space, the restricted Lorentz group $\mathcal{L}$ contains all possible geometric rotations of the spatial coordinate system. In general, those rotations can be described by three angle parameters $\omega_1, \omega_2, \omega_3 \in \mathbb{R}$ corresponding, *e.g.*, to the Euler angles associated with the rotation. We adopt a slightly different representation of a rotational Lorentz transformation $D \in \mathcal{L}$ in terms of the angle parameters, namely through

$$D = \exp\left[\sum_{j=1}^{3} \omega_j I_j\right] \tag{1.30}$$

with the matrices

$$I_1 = \left(\begin{array}{c|ccc} 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \end{array}\right), \quad I_2 = \left(\begin{array}{c|ccc} 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{array}\right), \quad I_3 = \left(\begin{array}{c|ccc} 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right) \tag{1.31}$$

which generate rotations around the x, y, and z axis, respectively.

There is, however, another kind of Lorentz transformations which involve *boosts*. The latter refer to transformations that change from a stationary spatial coordinate system to a moving system propagating with constant speed. Consider, *e.g.*, such a boost in the direction of e_1, *i.e.* the new basis vectors e'_1, e'_2, e'_3 keep their spatial orientation, but the origin upon which their are centred moves with constant speed v along e_1, such that it coincides with the stationary origin of the old coordinate system at time $t = 0$. It is natural to state that the longitudinal coordinate x^1 of a spatial vector $\vec{x}$ turns into $x'^1 = x^1 - vt$ in that case, while the other two spatial coordinates do not change. This is, however, no longer true for speeds v that come close to the speed of light. In that case, it is known that the more general expression

$$x'^1 = \frac{x^1 - vt}{\sqrt{1 - \frac{v^2}{c^2}}} \tag{1.32}$$

has to be used for the longitudinal coordinate along e'_1, while we have indeed

$$x'^2 = x^2 \quad \text{and} \quad x'^3 = x^3 \tag{1.33}$$

for the other two spatial coordinates. Moreover, also the notion of time changes in the moving frame, namely according to

$$t' = \frac{t - \frac{v}{c^2} x^1}{\sqrt{1 - \frac{v^2}{c^2}}} \,. \tag{1.34}$$

Defining $\beta = v/c$ and $\gamma = 1/\sqrt{1-\beta^2}$, we can write the transformation of the associated four-vector $x = (x^0, \vec{x})$ with $x^0 \equiv ct$ as $x'^\nu = D^\nu{}_\mu x^\mu$ or $x^\nu = D_\mu{}^\nu x'^\mu$ with the transformation matrix

$$D \equiv \left(D_\mu{}^\nu\right) = \left(\begin{array}{c|ccc} \gamma & \gamma\beta & 0 & 0 \\ \hline \gamma\beta & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{array}\right) . \tag{1.35}$$

Clearly, Eq. (1.35) represents a proper Lorentz transformation as it satisfies Eq. (1.23) together with $\det D = 1$ and $D_0{}^0 \geq 1$.

It is useful to consider once more the nonrelativistic regime $|v| \ll c$. In that case, we can linearize Eq. (1.35) in v/c by approximating $\gamma \simeq 1 + \mathcal{O}(v^2/c^2)$. This yields

$$D \simeq 1 - \beta J_1 \tag{1.36}$$

with

$$J_1 = \left(\begin{array}{c|ccc} 0 & -1 & 0 & 0 \\ \hline -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right) . \tag{1.37}$$

It corresponds to the nonrelativistic Galilei tranformation $x^1 = x'^1 + \beta x'^0$ proposed above as far as the spatial coordinates are concerned; the temporal coordinate, however, changes as well, namely according to $x^0 = x'^0 + \beta x'^1$.

Formally, a boost to a frame moving with a finite speed v coming close to the speed of light can always be represented as a succession of a large number of infinitesimal transformations of the type (1.36) each of them incrementing the speed of the frame by a small amount $\beta \ll c$. Choosing $\beta = \tau/N$ for some finite $\tau \in \mathbb{R}$ where N is the number of such infinitesimal transformations to be performed, we obtain for the resulting Lorentz transformation in the limit $N \to \infty$

$$D = \lim_{N\to\infty} \left(1 - \frac{\tau}{N} J_1\right)^N = \exp\left[-\tau J_1\right] = \left(\begin{array}{c|ccc} \cosh\tau & \sinh\tau & 0 & 0 \\ \hline \sinh\tau & \cosh\tau & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{array}\right) , \quad (1.38)$$

which corresponds to a frame that moves with the speed $v = c \tanh\tau$ in the old coordinate system. The matrix J_1 plays then be role of a generator of boosts along the direction of e_1.

In analogy with rotations, we can define two other generators

$$J_2 = \left(\begin{array}{c|ccc} 0 & 0 & -1 & 0 \\ \hline 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right) \quad \text{and} \quad J_3 = \left(\begin{array}{c|ccc} 0 & 0 & 0 & -1 \\ \hline 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{array}\right) \quad (1.39)$$

corresponding to boosts in the directions of e_2 and e_3, respectively. This altogether yields the insight that 6 parameters $\Omega_1, \Omega_2, \Omega_3, \omega_1, \omega_2, \omega_3 \in \mathbb{R}$ are needed in order to fully characterize a given proper Lorentz transformation $D \in \mathcal{L}$ in a unique manner, namely through

$$D = \exp\left[\sum_{j=1}^{3} (\Omega_j J_j + \omega_j I_j)\right] , \quad (1.40)$$

which generalizes Eq. (1.30).

1.3 Minkowski geometry

After this group theoretical characterization of the Lorentz transformations, we now discuss the physical implications of the form (1.25) of the metric tensor in the Minkowski space $\mathbb{M} = \mathbb{R}^4$. From the physical point of view, we shall name a four-vector $x = (x^0, x^1, x^2, x^3) \in \mathbb{M}$ an *event*, referring to something that happens at some place $\vec{x} = (x^1, x^2, x^3)$ at a given time $t = x^0/c$. Coordinate frames in which the metric tensor has the form (1.25), *i.e.* $g = \mathrm{diag}(1, -1, -1, -1)$, are named *inertial frames*, as they are either stationary or move with constant speed v with respect to a given reference frame, in perfect analogy with the inertial motion of a particle that is not subject to any force. As we have seen in Section 1.2, the notions of "space" and "time" change when we perform a Lorentz transformation to a moving frame. What used to be the temporal coordinate x^0 of the event x in the original reference frame turns into a mixture of temporal and spatial coordinates in the moving frame, and similarly for the spatial coordinates.

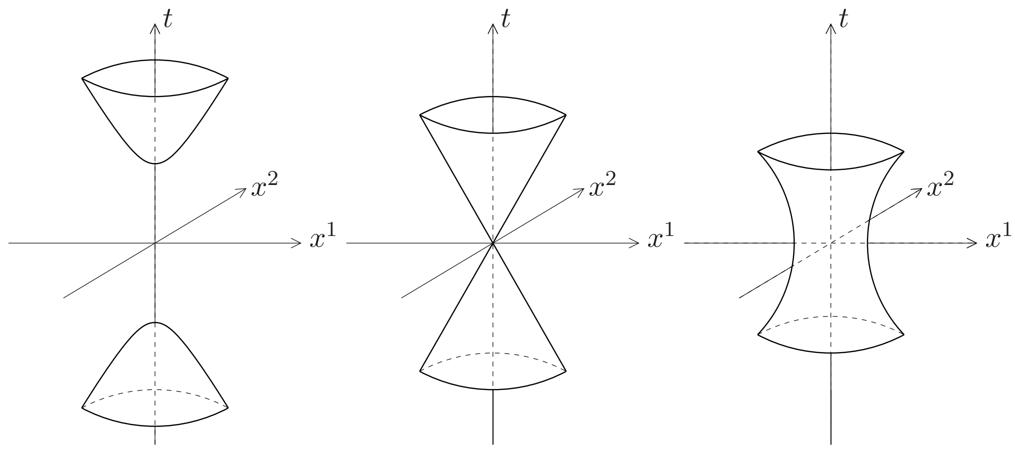

Figure 1.1: Sketches of the invariant manifolds $\mathcal{M}_r^{(\pm)}$ (left panel), $\mathcal{K}^{(\pm)}$ (middle panel), and $\mathcal{R}_\ell$ (right panel) of the Minkowski space, plotted as a function of the time t and the two spatial coordinates x^1 and x^2.

The essence of the theory of relativity is the claim that *there is no "reference" frame that is distinguished with respect to other inertial frames.* "Time" and "space" are therefore no absolute notions, but notions defined *relative* to a given inertial frame. This principle must be respected by physical laws as well. It must therefore be possible to formulate those laws such that *they have the same structural form in each inertial frame.*

It straightforwardly follows from this principle of relativity that information cannot travel faster than the speed of light. To show this particular consequence, we discuss the properties of manifolds in space-time that remain invariant under metric-conserving transformations. For the case of Euclidean spaces, those manifolds are evidently given by spheres characterized by a given radius r. Indeed, any point on such a sphere can be transformed into any other point on the sphere by means of a rotation of the coordinate system, and the distance of the associated vector from the origin remains invariant under such a transformation as it corresponds to the square root of the scalar product of the vector with itself.

The invariance of the scalar product $xx = (x^0)^2 - \vec{x}^2$ of a four-vector $x \equiv (x^0, \vec{x})$ under Lorentz transformations can also be used to classify the invariant manifolds within the Minkowski space. However, this scalar product is no longer positive definite, but can take negative values or be zero for nonzero $x \neq 0$. These three possible cases give rise to qualitatively different invariant manifolds.

Let us, for didactical purpose, begin with the special case $xx = 0$. We then have $(x^0)^2 = \vec{x}^2$, *i.e.* $x^0 = |\vec{x}|$ or $x^0 = -|\vec{x}|$. As shown in Fig. 1.1, these two different possibilities give rise to two different cones in space-time

$$\mathcal{K}^{(+)} = \{x \in \mathbb{M} : xx = (x^0)^2 - \vec{x}^2 = 0 \text{ and } x^0 > 0\} \tag{1.41}$$
$$\mathcal{K}^{(-)} = \{x \in \mathbb{M} : xx = (x^0)^2 - \vec{x}^2 = 0 \text{ and } x^0 < 0\} \tag{1.42}$$

which touch each other at the origin. They are named *light cones* as they contain the trajectories of light particles in space and time. Indeed, to detect a photon at time $t = 0$ at the origin of the spatial coordinate system, this photon has to be launched on the *past light cone* defined by $ct = -|\vec{x}|$, since it travels with the speed of light (provided there is no obstacle in the way). It can later on be observed on the *future light cone* defined by $ct = |\vec{x}|$.

The inertial trajectories of particles with finite mass, which propagate with velocities v that are lower than the speed of light, are contained in the interior of the past and future light cones. To observe a moving massive particle at time $t = 0$ at the spatial origin, it has to be launched in the past within the manifold

$$\mathcal{M}_r^{(-)} = \{x \in \mathbb{M} : xx = (x^0)^2 - \vec{x}^2 = r^2 \text{ and } x^0 < 0\} \tag{1.43}$$

for some $r > 0$. At some later time $t > 0$ in the future, this particle can then be detected within the manifold

$$\mathcal{M}_r^{(+)} = \{x \in \mathbb{M} : xx = (x^0)^2 - \vec{x}^2 = r^2 \text{ and } x^0 > 0\}\,, \tag{1.44}$$

since we have the requirement that $v = |\vec{x}|/t < c$ for such a massive particle. As illustrated in Fig. 1.1, $\mathcal{M}_r^{(+)}$ and $\mathcal{M}_r^{(-)}$ represent two distinct hyperbolic (non-compact) manifolds in space-time.

By means of a suitable Lorentz boost it is possible to transform into the *eigenframe* of the particle in which the latter is permanently at rest. In this eigenframe, the final time of detection is given by $t' = r/c$, which is shorter than the time of detection $t = \sqrt{r^2 + \vec{x}^2}/c$ in the original frame. As no inertial frame is distinguished with respect to another one, we can state that the spatio-temporal difference between the final detection event and the initial launching event of the particle is of *time-like* nature, insofar as one can find a frame in which that difference vector has only a temporal component.

Finally, hypothetical *superluminal* particles that travel with velocities larger than the speed of light would arrive within the hyperbolic manifold

$$\mathcal{R}_\ell = \{x \in \mathbb{M} : xx = (x^0)^2 - \vec{x}^2 = -\ell^2\}\,, \tag{1.45}$$

for some $\ell > 0$ after being launched at time $t = 0$ at the spatial origin. As illustrated in Fig. 1.1, $\mathcal{R}_\ell$ represents, in contrast to $\mathcal{M}_r^{(+)}$ and $\mathcal{M}_r^{(-)}$, a single connected manifold in which the case distinction $x^0 > 0$ or $x^0 < 0$ is no longer meaningful. Indeed, for any four-vector $x \in \mathcal{R}_\ell$ with $x^0 > 0$ it is possible to perform a Lorentz boost into a moving frame in which the temporal component of x can be negative or zero. Physically, this means that the difference vector between the final detection event and the initial launching event of the considered superluminal particle is not time-like but *space-like*, insofar as one can find an inertial frame in which the temporal component of this difference is zero. This, however, poses a fundamental problem from the point of view of causality: While in the original frame the particle is detected well after being launched at the spatial origin, one can perform proper Lorentz transformations into other inertial frames in which the detection event takes place at the same time or even before the launching event. As no inertial frame is distinguished with respect to the other ones, it is therefore impossible to state from a conceptual point of view whether the detection of the particle happens after the launching or the other way round. Those two events are therefore considered to take place *simultaneously*.

In summary, we can conclude that the constant c, which we initially introduced in order to represent the spatial and temporal components of an event with the same units, now acquires a true physical meaning insofar as it represents the ultimate speed limit for any kind of wave packet or particle propagation within the universe. This fundamental property is an immediate consequence of the Lorentz metric (1.25).

Chapter 2

Maxwell's equations

2.1 Vectors and tensors

From a physical point of view, the notion of a four-vector is not restricted to the spatiotemporal coordinates *e.g.* of point particles or of the difference vectors between two individual events. It may also apply to other physical objects and quantities. A prominent example is the *four-momentum* $p = p^\nu e_\nu \equiv (p^\nu)$ of a particle, which contains the momentum components p^j of the particle with $j = 1, 2, 3$ as "spatial" coordinates and its energy E as "temporal" coordinate $p^0 = E/c$. While rotations of the spatial coordinate system do not affect the energy E and the modulus of the momentum vector $\vec{p} = (p^1, p^2, p^3)$, Lorentz boosts generally give rise to a modification of $|\vec{p}|$ which in turn induces a modification of E according to the relativistic energy-momentum relation

$$E = \sqrt{(mc^2)^2 + c^2\vec{p}^2} \tag{2.1}$$

where m is the mass of the particle. This relation can also be expressed as $E^2 - c^2\vec{p}^2 = (mc^2)^2$ or as $p_\nu p^\nu = (mc)^2$, which clearly indicates that the four-momentum $(p^\nu) = (E/c, \vec{p})$ is a time-like vector.

More generally, a physical object $a \equiv (a^\nu)$ with four components a^0, a^1, a^2, a^3 will be named *four-vector* or *Lorentz vector*, or simply *vector* in the following, if its components a^ν are transformed under Lorentz transformations according to the same transformation laws (1.6) and (1.7) as the spatiotemporal coordinates x^ν of any event $x \equiv (x^\nu)$. That is, if we subject our spatiotemporal basis (e_ν) to a Lorentz transformation $e_\nu \mapsto e'_\nu = D_\nu{}^\mu e_\mu$ with $D \equiv (D_\nu{}^\mu) \in G$, we impose the transformation law

$$a^\nu \mapsto a'^\nu = D^\nu{}_\mu a^\mu \tag{2.2}$$

for the components of the four-vector a. This is perfectly analogous to spatial vectors $\vec{a} \equiv (a^1, a^2, a^3)$ in the three-dimensional Euclidean space which are supposed to transform under rotations of the spatial coordinate system in exactly the same manner as any spatial position vector $\vec{x} \equiv (x^2, x^2, x^3)$ *e.g.* of a particle.

It is known in this context that *pseudovectors* can be distinguished from "ordinary" vectors by their behaviour under mirror transformations of the spatial coordinate system: while an ordinary vector transforms as the position vector $\vec{x}$ under such a mirror transformation, the transformation of pseudovectors involves a change of sign in addition. This concept of pseudovectors can indeed be imported to the Minkowski space: $\tilde{a} \equiv (\tilde{a}^\nu)$ is considered to be a *pseudovector* if its components are transformed according to the law

$$\tilde{a}^\nu \mapsto \tilde{a}'^\nu = (\det D) D^\nu{}_\mu \tilde{a}^\mu \tag{2.3}$$

under a general Lorentz transformation $D \in G$. This transformation law involves a change of sign in the case of the case of Lorentz transformations D that include spatial or temporal mirrors for which $\det D = -1$, while it is equivalent to the law (2.2) for vectors in the case of other Lorentz transformations with $\det D = 1$.

Physical objects $t \equiv (t^{\nu\mu})$ with two indices ν, μ are named *tensors of second order* (or, in that case, simply *tensors*) if they transform according to

$$t^{\nu\mu} \mapsto t'^{\nu\mu} = D^\nu{}_\alpha D^\mu{}_\beta t^{\alpha\beta} \tag{2.4}$$

under Lorentz transformations $D \in G$. The metric tensor $(g_{\mu\nu})$ is a prime example in this context, which features the specific property that $g'_{\mu\nu} = g_{\mu\nu}$ for Lorentz transformations. *Pseudotensors* $\tilde{t} \equiv (\tilde{t}^{\nu\mu})$, on the other hand, would satisfy the transformation law

$$\tilde{t}^{\nu\mu} \mapsto \tilde{t}'^{\nu\mu} = (\det D) D^\nu{}_\alpha D^\mu{}_\beta \tilde{t}^{\alpha\beta} \tag{2.5}$$

which again involves a change of sign in the presence of spatial or temporal mirrors such that $\det D = -1$. This concept can be generalized to tensors of Nth order $t \equiv (t^{\nu_1,\ldots,\nu_N})$ and pseudotensors of Nth order $\tilde{t} \equiv (\tilde{t}^{\nu_1,\ldots,\nu_N})$ which respectively satisfy the transformation laws

$$t^{\nu_1,\ldots,\nu_N} \quad\mapsto\quad t'^{\nu_1,\ldots,\nu_N} = D^{\nu_1}{}_{\mu_1} \cdots D^{\nu_N}{}_{\mu_N} t^{\mu_1,\ldots,\mu_N}\,, \tag{2.6}$$

$$\tilde{t}^{\nu_1,\ldots,\nu_N} \quad\mapsto\quad \tilde{t}'^{\nu_1,\ldots,\nu_N} = (\det D) D^{\nu_1}{}_{\mu_1} \cdots D^{\nu_N}{}_{\mu_N} \tilde{t}^{\mu_1,\ldots,\mu_N} \tag{2.7}$$

under Lorentz transformations $D \in G$.

Finally, *Lorentz scalars* are single-component physical objects that do not change at all under Lorentz transformations. The mass m of a particle, for instance, is a scalar object as it appears in Eq. (2.1). *Pseudoscalars* $\tilde{m}$ can also be introduced. They satisfy the transformation law

$$\tilde{m} \mapsto \tilde{m}' = (\det D)\tilde{m} \tag{2.8}$$

under a Lorentz transformation $D \in G$. A prime example of a pseudoscalar in the three-dimensional Euclidean space is the *mixed product* $\vec{a} \cdot (\vec{b} \times \vec{c})$ the modulus of which corresponds to the volume contained within the parallelepiped spanned by the three vectors $\vec{a}, \vec{b}, \vec{c}$ and the sign of which is related to the (right-handed or left-handed) orientation of these three vectors. Using the Levi-Civita symbol

$$\epsilon_{ijk} = \begin{cases} 1 : (i,j,k) \text{ is an even permutation of } (1,2,3) \\ -1 : (i,j,k) \text{ is an odd permutation of } (1,2,3) \\ 0 : \text{otherwise} \end{cases} , \tag{2.9}$$

we can express the mixed product through $\vec{a} \cdot (\vec{b} \times \vec{c}) = \epsilon_{ijk} a^i b^k c^l$ and thereby directly infer its pseudoscalar property.

The notion of the Levi-Civita symbol can be generalized to the Minkowski space giving rise to the *Levi-Civita tensor*

$$\epsilon^{\nu_1\nu_2\nu_3\nu_4} = \begin{cases} 1 : (\nu_1,\nu_2,\nu_3,\nu_4) \text{ is an even permutation of } (0,1,2,3) \\ -1 : (\nu_1,\nu_2,\nu_3,\nu_4) \text{ is an odd permutation of } (0,1,2,3) \\ 0 : \text{otherwise} \end{cases} \tag{2.10}$$

which obviously is a pseudotensor of fourth order. This Levi-Civita tensor will be used in order to define for each antisymmetric tensor (or pseudotensor) $t \equiv (t^{\nu\mu})$ with $t^{\nu\mu} = -t^{\mu\nu}$ for all $\nu, \mu \in \{0,1,2,3\}$ the associated *dual pseudotensor* (or *dual tensor*) $\tilde{t} \equiv (\tilde{t}_{\nu\mu})$, namely through

$$\tilde{t}_{\nu\mu} = \frac{1}{2}\epsilon_{\nu\mu\alpha\beta} t^{\alpha\beta} , \tag{2.11}$$

which also satisfies $\tilde{t}_{\nu\mu} = -\tilde{t}_{\mu\nu}$ for all $\nu, \mu \in \{0,1,2,3\}$. We can straightforwardly verify the duality relation $t^{\nu\mu} = \frac{1}{2}\epsilon^{\nu\mu\alpha\beta}\tilde{t}_{\alpha\beta}$. This notion of a dual tensor will play a role in the subsequent section where we discuss the homogeneous Maxwell equations.

We are now in a position to properly introduce scalar, vector, and tensor fields (as well as pseudoscalar, pseudovector, and pseudotensor fields): They are given by (single- or multicomponent) functions that are defined on the Minkowski space $\mathbb{M}$ and that transform as scalars, vectors, or tensors (or as pseudoscalars, pseudovectors, or pseudotensors) under Lorentz transformations. Assuming that these functions are sufficiently often continously differentiable, we can calculate their partial derivatives through the application of the operators

$$\partial_\nu \equiv \frac{\partial}{\partial x^\nu} \tag{2.12}$$

for $\nu = 0,1,2,3$. These operators form again a vectorial object

$$(\partial_\nu) \equiv (\partial_0, \partial_1, \partial_2, \partial_3) = \left(\frac{1}{c}\frac{\partial}{\partial t}, \vec{\nabla}\right) \quad \text{with } \vec{\nabla} \equiv \frac{\partial}{\partial \vec{r}} \tag{2.13}$$

which transforms under Lorentz transformations $D \in G$ in the same manner as any other (covariant) Lorentz vector: $\partial_\nu \mapsto \partial'_\nu = D_\nu{}^\mu \partial'_\mu$. Hence, the *gradient* of a (complex or real) scalar field $\varphi : \mathbb{M} \to \mathbb{C}, x \mapsto \varphi(x)$, which is defined as $\partial_\nu \varphi : \mathbb{M} \to \mathbb{C}, x \mapsto \partial_\nu \varphi(x)$, represents a *vector field* as its transformation behaviour under Lorentz transformations is determined by the transformation of the partial derivative operator (∂_ν). On the other hand, the *divergence* of a (complex or real) vector field $a^\nu : \mathbb{M} \to \mathbb{C}, x \mapsto a^\nu(x)$ is evaluated as $\partial_\nu a^\nu(x)$ and therefore corresponds to a *scalar field*, in a similar manner as the scalar product of the vector field a^ν with any other Lorentz vector. The divergence of the gradient of a scalar (or vector or tensor) field gives rise to the *d'Alembert operator*

$$\partial_\nu \partial^\nu = \frac{1}{c^2}\frac{\partial^2}{\partial t^2} - \Delta \quad \text{with } \Delta \equiv \vec{\nabla} \cdot \vec{\nabla} = \frac{\partial^2}{\partial \vec{r}^2} \tag{2.14}$$

which clearly corresponds to a scalar object again. Here we make use of the countervariant components

$$(\partial^\nu) = \left(\frac{\partial}{\partial x_\nu}\right) = \left(\frac{1}{c}\frac{\partial}{\partial t}, -\vec{\nabla}\right) \tag{2.15}$$

of the partial derivative operator.

We are now in a position to reformulate in a more precise manner what we meant by stating in Section 1.3 that physical laws must "have the same structural form in each inertial frame" in order to comply with the principle of relativity: If we assume that basic physical laws are generally expressed in terms of partial differential equations (such as Maxwell's equations or Schrödinger's equation), then those partial differential equations must exhibit a well-defined scalar, vectorial, or tensorial (or pseudoscalar, pseudovectorial, or pseudotensorial) transformation behaviour under Lorentz transformations. In other words, we must be able to express those equations in a *covariant* manner using the relativistic index notations and the summation convention of Einstein. *E.g.* a wave equation of the form

$$\left(\frac{1}{c^2}\frac{\partial^2}{\partial t^2} - \Delta\right)\varphi(\vec{r}, t) = 0\,, \tag{2.16}$$

which describes (scalar) waves that propagate with the speed of light c, perfectly satisfies this requirement since we can reformulate it in a covariant manner as $\partial_\nu \partial^\nu \varphi(x) = 0$ using relativistic notations. On the other hand, a diffusion equation of the form

$$\left(\frac{\partial}{\partial t} - D\Delta\right)\varphi(\vec{r}, t) = 0 \tag{2.17}$$

is not in agreement with the principle of relativity (and there is no way to reformulate it in a coherent manner) since spatial and temporal derivatives appear with different orders (first order for the temporal coordinate and second order for the spatial coordinates). Indeed, this equation will have the form (2.17) *only in very specific inertial frames*: Applying a Lorentz boost will mix spatial and temporal coordinates according to Eqs. (1.32) and (1.34), which necessarily implies that

Eq. (2.17) will appear differently in the new frame. The same argument holds for Schrödinger's equation — which means that the latter is not in agreement with the principle of relativity!

2.2 The homogeneous Maxwell equations

In the framework of the theory of special relativity, the electromagnetic field is described by an antisymmetric tensor field of second order $F \equiv (F_{\mu\nu})$ which is defined through

$$(F_{\mu\nu})(\vec{r},t) = \left(\begin{array}{c|ccc} 0 & E_1(\vec{r},t) & E_2(\vec{r},t) & E_3(\vec{r},t) \\ \hline -E_1(\vec{r},t) & 0 & -B_3(\vec{r},t) & B_2(\vec{r},t) \\ -E_2(\vec{r},t) & B_3(\vec{r},t) & 0 & -B_1(\vec{r},t) \\ -E_3(\vec{r},t) & -B_2(\vec{r},t) & B_1(\vec{r},t) & 0 \end{array}\right) \tag{2.18}$$

and which shall be abbreviated as

$$(F_{\mu\nu})(\vec{r},t) \equiv \left(\vec{E}(\vec{r},t), \vec{B}(\vec{r},t)\right) \tag{2.19}$$

in the following, where $\vec{E} \equiv (E_1, E_2, E_3)$ is the electric and $\vec{B} \equiv (B_1, B_2, B_3)$ the magnetic field using cgs units (which are used throughout these lecture notes). The fact that $(F_{\mu\nu})$ is a tensor and not a pseudotensor automatically implies that the magnetic field $\vec{B}$ corresponds to a spatial pseudotensor whereas the electric field $\vec{E}$ is an "ordinary" spatial vector. While F is defined in terms of its covariant coordinates $F_{\mu\nu}$ in Eq. (2.18), we can straightforwardly infer its countervariant representation through

$$(F^{\mu\nu}) = \left(g^{\mu\alpha} g^{\nu\beta} F_{\alpha\beta}\right) = \left(\begin{array}{c|ccc} 0 & -E_1 & -E_2 & -E_3 \\ \hline E_1 & 0 & -B_3 & B_2 \\ E_2 & B_3 & 0 & -B_1 \\ E_3 & -B_2 & B_1 & 0 \end{array}\right) \equiv \left(-\vec{E}, \vec{B}\right) \tag{2.20}$$

as we would do it for any other tensor of second order.

Since F is an antisymmetric tensor with $F_{\mu\nu}(x) = -F_{\nu\mu}(x)$ for all $\mu, \nu \in \{0,1,2,3\}$ and all $x \in \mathbb{M}$, we can, as we pointed out in the previous section 2.1, define its associated dual pseudotensor according to

$$(\tilde{F}^{\mu\nu}) = \left(\frac{1}{2}\epsilon^{\mu\nu\alpha\beta} F_{\alpha\beta}\right) = \left(\begin{array}{c|ccc} 0 & -B_1 & -B_2 & -B_3 \\ \hline B_1 & 0 & E_3 & -E_2 \\ B_2 & -E_3 & 0 & E_1 \\ B_3 & E_2 & -E_1 & 0 \end{array}\right) \equiv \left(-\vec{B}, -\vec{E}\right) . \tag{2.21}$$

The homogeneous Maxwell equations are then written as

$$\partial_\nu \tilde{F}^{\mu\nu}(x) = 0 \tag{2.22}$$

in obvious agreement with the principle of relativity as Eq. (2.22) represents a pseudovectorial object that exhibits a well-defined transformation behaviour under Lorentz transformations. Evaluating $\partial_\nu \tilde{F}^{0\nu}(x) = -\partial_l B_l(x)$ as well as $\partial_\nu \tilde{F}^{l\nu}(x) = \partial_0 B_l(x) + \epsilon_{jkl}\partial_j E_k(x)$ for $l = 1, 2, 3$ where ϵ_{jkl} represents the Levi-Civita symbol (2.9), and expressing the resulting equations in nonrelativistic notation, we recover the standard formulation of the homogeneous Maxwell equations in cgs units:

$$\vec{\nabla} \cdot \vec{B}(\vec{r}, t) = 0\,, \tag{2.23}$$

$$\vec{\nabla} \times \vec{E}(\vec{r}, t) = -\frac{1}{c}\frac{\partial}{\partial t}\vec{B}(\vec{r}, t)\,. \tag{2.24}$$

As is well known from classical electrodynamics, Eqs. (2.23) and (2.24) allow for the introduction of scalar and vector potentials Φ and $\vec{A}$ owing to the decomposition theorem for vector fields (which formally requires that $\vec{E}$ and $\vec{B}$ fall off sufficiently rapidly for $|\vec{r}| \to \infty$). In a similar manner, we can state that Eq. (2.22) is satisfied if and only if we can write

$$F_{\mu\nu}(x) = \partial_\mu A_\nu(x) - \partial_\nu A_\mu(x) \tag{2.25}$$

for a suitable Lorentz vector field (A^ν), which we shall write as

$$(A^\nu)(x) \equiv \left(\Phi(\vec{r}, t), \vec{A}(\vec{r}, t)\right) \tag{2.26}$$

and name as *four-potential* in the following. Indeed, if Eq. (2.25) holds, we can straightforwardly evaluate

$$\partial_\nu \tilde{F}^{\mu\nu} = \frac{1}{2}\epsilon^{\mu\nu\alpha\beta}\partial_\nu(\partial_\alpha A_\beta - \partial_\beta A_\alpha) = \frac{1}{2}\left(\epsilon^{\mu\nu\alpha\beta} - \epsilon^{\mu\nu\beta\alpha}\right)\partial_\nu\partial_\alpha A_\beta = 0 \tag{2.27}$$

using the antisymmetry of the Levi-Civita tensor (2.10). In nonrelativistic terms, Eq. (2.25) is equivalent to the well-known relations

$$\vec{E}(\vec{r}, t) = -\vec{\nabla}\Phi(\vec{r}, t) - \frac{1}{c}\frac{\partial}{\partial t}\vec{A}(\vec{r}, t)\,, \tag{2.28}$$

$$\vec{B}(\vec{r}, t) = \vec{\nabla} \times \vec{A}(\vec{r}, t)\,. \tag{2.29}$$

The choice of the four-potential (A^ν) is not unique for a given electromagnetic field $(F_{\mu\nu}) = (\vec{E}, \vec{B})$. Indeed, applying the *gauge transformation*

$$A_\nu \mapsto A'_\nu = A_\nu + \partial_\nu \chi \tag{2.30}$$

where $\chi : \mathbb{M} \to \mathbb{R}, x \mapsto \chi(x)$ is a Lorentz scalar field yields the same electromagnetic field tensor

$$F_{\mu\nu}(x) = \partial_\mu A_\nu(x) - \partial_\nu A_\mu(x) = \partial_\mu A'_\nu(x) - \partial_\nu A'_\mu(x) \tag{2.31}$$

provided χ is twice continuously differentiable. We shall, in the following, impose the *Lorenz gauge*, *i.e.*, choose a four-potential that satisfies the relation

$$\partial_\nu A^\nu(x) = \frac{1}{c}\frac{\partial}{\partial t}\Phi(\vec{r},t) + \vec{\nabla}\cdot\vec{A}(\vec{r},t) = 0 \tag{2.32}$$

for all $x \equiv (t, \vec{r})$.

2.3 The inhomogeneous Maxwell equations

The inhomogeneous Maxwell equations involve the charge and current densities of the electrically charged matter. The latter can be represented by another Lorentz vector field

$$(j^\nu)(x) \equiv \Big(c\rho(\vec{r},t), \vec{j}(\vec{r},t)\Big) \tag{2.33}$$

which will be named *four-current* in the following, and which has the electrical charge density $\rho(\vec{r},t)$ at position $\vec{r}$ and time t as temporal component and the electrical current density $\vec{j}(\vec{r},t)$ at position $\vec{r}$ and time t as spatial components. The conservation of the total charge can be expressed in terms of the *continuity equation*

$$\partial_\nu j^\nu(x) = \frac{\partial}{\partial t}\rho(\vec{r},t) + \vec{\nabla}\cdot\vec{j}(\vec{r},t) = 0 \tag{2.34}$$

which is valid for all positions $\vec{r}$ and times t.

Using this four-current, we can now express the inhomogeneous Maxwell equations as

$$\partial_\mu F^{\mu\nu}(x) = \frac{4\pi}{c} j^\nu(x)\,. \tag{2.35}$$

Again, this equation is in agreement with the principle of relativity as it exhibits a vectorial transformation behaviour under Lorentz transformations. In nonrelativistic terms, Eq. (2.35) can be rewritten as

$$\vec{\nabla}\cdot\vec{E}(\vec{r},t) = 4\pi\rho(\vec{r},t)\,, \tag{2.36}$$

$$\vec{\nabla}\times\vec{B}(\vec{r},t) = \frac{4\pi}{c}\vec{j}(\vec{r},t) + \frac{1}{c}\frac{\partial}{\partial t}\vec{E}(\vec{r},t)\,, \tag{2.37}$$

which together with Eqs. (2.23) and (2.24) constitutes the well-known set of Maxwell's equations.

Expressing the electromagnetic field tensor $(F^{\mu\nu})$ through the four-potential (A^ν) according to the relation (2.25) yields from Eq. (2.35) the equation

$$\partial_\mu\partial^\mu A^\nu(x) - \partial_\mu\partial^\nu A^\mu(x) = \frac{4\pi}{c} j^\nu(x) \tag{2.38}$$

for the four-potential (A^ν). This equation further simplifies if we impose the Lorenz gauge (2.32): we then obtain that A^ν satisfies an inhomogeneous wave equation of the form

$$\partial_\mu\partial^\mu A^\nu(x) = \frac{4\pi}{c} j^\nu(x)\,. \tag{2.39}$$

2.4 Energy and momentum of the electromagnetic field

As is well known from classical electrodynamics, the electromagnetic field contains energy and momentum, which are conserved in the absence of charged matter. Hence, in close analogy with the charge and current densities of matter, we can again introduce four-currents formed by the densities and fluxes of the total energy as well as of the three spatial components of the total momentum, each of which satisfying a continuity equation of the form (2.34) in the absence of charged matter. As the energy and the three components of the momentum vector constitute a Lorentz vector, too, we thereby obtain a tensorial object of second order, which is generally referred to as *energy-momentum tensor* $T \equiv (T^{\mu\nu})$. For the electromagnetic field its components $T^{\mu\nu} : \mathbb{M} \to \mathbb{R}, x \mapsto T^{\mu\nu}(x)$ are given by

$$T^{\mu\nu}(x) = \frac{1}{4\pi}\left(\frac{1}{4}g^{\mu\nu}F_{\alpha\beta}(x)F^{\alpha\beta}(x) - g_{\alpha\beta}F^{\mu\alpha}(x)F^{\nu\beta}(x)\right) = T^{\nu\mu}(x)\,. \tag{2.40}$$

Evaluating $F_{\alpha\beta}(x)F^{\alpha\beta}(x) = 2(\vec{B}^2(\vec{r},t) - \vec{E}^2(\vec{r},t))$, we specifically obtain

$$T^{00}(x) = \frac{1}{8\pi}\left(\vec{E}^2(\vec{r},t) + \vec{B}^2(\vec{r},t)\right) \tag{2.41}$$

for the energy density and

$$\left(cT^{0l}(x)\right) = \frac{c}{4\pi}\left(\vec{E}(\vec{r},t) \times \vec{B}(\vec{r},t)\right) = \vec{S}(\vec{r},t) \tag{2.42}$$

for the flux or current density of the energy of the electromagnetic field, which is also known as *Poynting vector* $\vec{S}$. The momentum density is given by

$$\left(\frac{1}{c}T^{l0}(x)\right) = \frac{1}{4\pi c}\left(\vec{E}(\vec{r},t) \times \vec{B}(\vec{r},t)\right) \tag{2.43}$$

and its associated flux is described by *Maxwell's stress tensor*

$$T^{ij}(x) = \frac{1}{4\pi}\left(\frac{1}{2}\left(\vec{E}^2(\vec{r},t) + \vec{B}^2(\vec{r},t)\right)\delta_{ij} - E_i(\vec{r},t)E_j(\vec{r},t) - B_i(\vec{r},t)B_j(\vec{r},t)\right)\,. \tag{2.44}$$

By means of the homogeneous and inhomogeneous Maxwell equations (2.22) and (2.35) we can derive

$$\partial_\nu T^{\mu\nu}(x) = -\frac{1}{c}F^{\mu\nu}(x)j_\nu(x) \tag{2.45}$$

where j_ν ($\nu = 0, 1, 2, 3$) are the (covariant) components of the four-current (2.33) describing the charged matter. Conservation of energy and momentum of the electromagnetic field is therefore granted in the absence of matter, while the presence of charged particles will give rise to an exchange of energy and momentum between the electromagnetic field and the particles, which is described by Eq. (2.45). Specifically we have

$$\frac{1}{c}F^{0\nu}(x)j_\nu(x) = \frac{1}{c}\vec{E}(\vec{r},t)\cdot\vec{j}(\vec{r},t) \tag{2.46}$$

as gain of energy density (*i.e.* as power density) of the matter in the presence of the electromagnetic field and, due to conservation of energy, also as loss of energy density of the electromagnetic field due to the interaction with matter. Eq. (2.46) expresses the fact that positively charged particles will be accelerated and thereby gain kinetic energy when moving in the direction of the local electric field $\vec{E}(\vec{r},t)$. The gain of momentum density of charged matter in the presence of the electromagnetic field is given by

$$\left(\frac{1}{c}F^{l\nu}(x)j_\nu(x)\right) = \rho(\vec{r},t)\vec{E}(\vec{r},t) + \frac{1}{c}\vec{j}(\vec{r},t)\times\vec{B}(\vec{r},t)\,, \tag{2.47}$$

which expresses the *Lorentz force* acting on the charged matter.

Problem

2.1 Show the validity of Eq. (2.45).

Chapter 3

Nonrelativistic quantum mechanics

3.1 Classical mechanics

In the framework of the Hamiltonian formalism, a classical point particle is described by its position $\vec{r} \equiv \vec{r}(t) \in \mathbb{R}^3$ and by its momentum $\vec{p} \equiv \vec{p}(t) \in \mathbb{R}^3$ which evolve with time according to Hamilton's equations of motion

$$\frac{d\vec{r}}{dt}(t) = \frac{\partial H}{\partial \vec{p}}\left[\vec{r}(t), \vec{p}(t), t\right] = \frac{1}{m}\vec{p}(t)\,, \tag{3.1}$$

$$\frac{d\vec{p}}{dt}(t) = -\frac{\partial H}{\partial \vec{r}}\left[\vec{r}(t), \vec{p}(t), t\right] = -\frac{\partial}{\partial \vec{r}}V\left[\vec{r}(t), t\right]\,. \tag{3.2}$$

Here,

$$H(\vec{r}, \vec{p}) = \frac{\vec{p}^2}{2m} + V(\vec{r}, t) \tag{3.3}$$

is the classical Hamiltonian, m is the mass of the particle, and $V(\vec{r}, t)$ denotes its potential energy at the position $\vec{r}$ at time t. Differentiating Eq. (3.1) with respect to time and combining it with Eq. (3.2) yields Newton's second law

$$m\frac{d^2\vec{r}}{dt^2}(t) = -\frac{\partial}{\partial \vec{r}}V\left[\vec{r}(t), t\right] \tag{3.4}$$

for the time evolution of the position of the particle.

Let us consider, as an example, a three-dimensional anisotropic harmonic oscillator characterized by the oscillation frequencies $\omega_1, \omega_2, \omega_3$ along the spatial axes $\vec{e}_1, \vec{e}_2, \vec{e}_3$, respectively. The potential energy of the particle is written as

$$V(\vec{r}, t) = \frac{1}{2}m\sum_{l=1}^{3}\omega_l^2 r_l^2 \equiv V(\vec{r}) \tag{3.5}$$

and we obtain the Newtonian equations

$$\frac{d^2 r_l}{dt^2}(t) = -\omega_l^2 r_l(t) \tag{3.6}$$

for the time evolution of the spatial coordinates of the particle ($l = 1, 2, 3$). These equations are generally solved as

$$r_l(t) = \alpha_l e^{-i\omega_l t} + \alpha_l^* e^{i\omega_l t} \tag{3.7}$$

for some complex amplitudes $\alpha_l \in \mathbb{C}$. Evaluating the momentum coordinates of the particle according to Eq. (3.1) as

$$p_l(t) = -im\omega_l \left(\alpha_l e^{-i\omega_l t} - \alpha_l^* e^{i\omega_l t}\right) , \tag{3.8}$$

we then obtain its total energy as

$$H\left[\vec{r}(t), \vec{p}(t)\right] = \sum_{l=1}^{3} 2m\omega_l^2 |\alpha_l|^2 \tag{3.9}$$

which is a constant of motion.

3.2 The Schrödinger equation

In contrast to classical mechanics, it is impossible to simultaneously define with certitude the precise position and momentum of a quantum point particle. This fact is reflected by the presence of fundamental experimental limits in the imaging of a quantum particle using, *e.g.*, a focused laser beam the spatial resolution of which would be given by the chosen wavelength. Indeed, the more we lower this wavelength in order to increase the spatial resolution, the higher will be the recoil that the particle receives when being hit by a laser photon, which enhances the uncertainy concerning the precise momentum of the particle. We therefore give up the attempt to maintain the notions of precise positions and momenta in the context of a quantum particles and accept the fact that such a particle is rather to be described by finite probability distributions $\rho(\vec{r}, t)$ and $\tilde{\rho}(\vec{p}, t)$ that at the time t the particle is at the position $\vec{r}$ and moves with the momentum $\vec{p}$, respectively.

The essential assumption that underlies the theory of quantum mechanics is that these two probability distributions are not independent of each other. They are linked by a complex *wavefunction* $\psi : \mathbb{R}^4 \to \mathbb{R}, (\vec{r}, t) \mapsto \psi(\vec{r}, t)$ such that $\rho(\vec{r}, t) = |\psi(\vec{r}, t)|^2$ is the probability density to find the particle at the position $\vec{r}$ at time t and $\tilde{\rho}(\vec{p}, t) = |\tilde{\psi}(\vec{p}, t)|^2$ with

$$\tilde{\psi}(\vec{p}, t) = \frac{1}{\sqrt{2\pi\hbar}^3} \int d^3r\, \psi(\vec{r}, t) e^{-\frac{i}{\hbar}\vec{p}\cdot\vec{r}} \tag{3.10}$$

is the probability density to detect that the particle moves with the momentum $\vec{p}$ at time t. Equation (3.10) essentially corresponds to the Fourier transform of ψ which is inverted as

$$\psi(\vec{r},t) = \frac{1}{\sqrt{2\pi\hbar}^3} \int d^3p\, \tilde{\psi}(\vec{p},t) e^{\frac{i}{\hbar}\vec{p}\cdot\vec{r}}, \tag{3.11}$$

where we introduce the reduced Planck constant $\hbar = h/(2\pi) \simeq 1.055 \cdot 10^{-34}$ Js as basic unit for action. By definition we have

$$1 = \int d^3r |\psi(\vec{r},t)|^2 = \int d^3p |\tilde{\psi}(\vec{p},t)|^2 \tag{3.12}$$

for the total probability to detect the particle, where the second equality in Eq. (3.12) is granted to hold by construction of the Fourier transform (3.10) and its inversion (3.11) owing to Parseval's theorem.

The impossibility to precisely pinpoint the position and the momentum of a particle emerges now as a straightforward mathematical consequence of the relation between the corresponding probability distributions through Eqs. (3.10) and (3.11). It essentially results from the fact that the Fourier transform of a rather narrowly localized wavefunction ψ in position space yields a rather broad function $\tilde{\psi}$ in momentum space, and *vice versa*, where the characteristic scale that relates the notions of "narrow" and "broad" with each other is provided by the reduced Planck constant $\hbar$. More precisely, one can evaluate that the product of the standard deviations associated with the mean position and momentum coordinates of a particle along the axis $\vec{e}_l$ $(l = 1, 2, 3)$ verifies *Heisenberg's uncertainty principle*

$$\Delta r_l \Delta p_l \geq \frac{\hbar}{2} \tag{3.13}$$

where the lower limit, yielding an equality sign in Eq. (3.13), is attained for the (optimal) case of a Gaussian wavefunction in position and momentum space.

While it would indeed be straightforward to calculate the mean values of arbitrary functions of the position or the momentum of a particle according to

$$\langle f \rangle_t = \int d^3r\, f(\vec{r}) |\psi(\vec{r},t)|^2, \tag{3.14}$$

$$\langle g \rangle_t = \int d^3p\, g(\vec{p}) |\tilde{\psi}(\vec{p},t)|^2, \tag{3.15}$$

respectively, with $f : \mathbb{R}^3 \to \mathbb{R}, \vec{r} \mapsto f(\vec{r})$ and $g : \mathbb{R}^3 \to \mathbb{R}, \vec{p} \mapsto g(\vec{p})$, it turns out to more convenient to privilege only one of the two representations of the wavefunction, namely the one in position space, and to evaluate mean values of functions g defined in momentum space also by using this position representation of the wavefunction. This yields for the mean momentum of the particle

$$\begin{aligned}
\langle\vec{p}\rangle_t &= \int d^3p|\tilde{\psi}(\vec{p},t)|^2\vec{p} && (3.16)\\
&= \frac{1}{(2\pi\hbar)^3}\int d^3p\int d^3r\int d^3r'\psi^*(\vec{r},t)e^{\frac{i}{\hbar}\vec{p}\cdot\vec{r}}\psi(\vec{r}',t)e^{-\frac{i}{\hbar}\vec{p}\cdot\vec{r}'}\vec{p} && (3.17)\\
&= \int d^3r\,\psi^*(\vec{r},t)\frac{\hbar}{i}\vec{\nabla}\psi(\vec{r},t) && (3.18)\\
&= -\int d^3r\,\psi(\vec{r},t)\frac{\hbar}{i}\vec{\nabla}\psi^*(\vec{r},t) && (3.19)
\end{aligned}$$

with $\vec{\nabla}\equiv\frac{\partial}{\partial\vec{r}}$ where we use $\exp(\mp\frac{i}{\hbar}\vec{p}\cdot\vec{r})\vec{p}=\pm i\hbar\vec{\nabla}\exp(\mp\frac{i}{\hbar}\vec{p}\cdot\vec{r})$ and integrate by parts for deriving Eq. (3.18) and Eq. (3.19) from Eq. (3.17), respectively. We also make use of the fact that the integral

$$\frac{1}{(2\pi\hbar)^3}\int d^3p\,e^{\frac{i}{\hbar}\vec{p}\cdot(\vec{r}-\vec{r}')}=\delta(\vec{r}-\vec{r}') \tag{3.20}$$

is a valid representation of Dirac's delta distribution in the space of square integrable test functions $\psi:\mathbb{R}^3\mapsto\mathbb{C},\vec{r}\mapsto\psi(\vec{r})$ that satisfy the normalization condition (3.12).

While Eq. (3.18) is most commonly used to represent the mean value of the momentum of a particle, it is instructive to form the arithmetic average of the equivalent expressions (3.18) and (3.19) yielding

$$\langle\vec{p}\rangle_t=\int d^3r\frac{\hbar}{2i}\left[\psi^*(\vec{r},t)\vec{\nabla}\psi(\vec{r},t)-\psi(\vec{r},t)\vec{\nabla}\psi^*(\vec{r},t)\right]. \tag{3.21}$$

In contrast to Eqs. (3.18) and (3.19), the integrand of this latter expression for the mean momentum is real and can therefore be interpreted as some sort of "momentum density" of the particle, providing the local contribution from the position $\vec{r}$ for the evaluation of $\langle\vec{p}\rangle_t$. Using according to Eq. (3.1) the (nonrelativistic) relation $\vec{p}=m\vec{v}$ between the momentum $\vec{p}$ and the velocity $\vec{v}$ of the particle allows us then to define the density of the probability current or the *flux* of the probability at the position $\vec{r}$ according to

$$\begin{aligned}
\vec{j}(\vec{r},t) &= \frac{\hbar}{2im}\left[\psi^*(\vec{r},t)\vec{\nabla}\psi(\vec{r},t)-\psi(\vec{r},t)\vec{\nabla}\psi^*(\vec{r},t)\right].\\
&= \frac{1}{m}\mathrm{Re}\left[\psi^*(\vec{r},t)\frac{\hbar}{i}\vec{\nabla}\psi(\vec{r},t)\right] && (3.22)
\end{aligned}$$

As already pointed out in Section 2.3 in the context of charged matter, the spatial probability density ρ and its associated flux $\vec{j}$ have to satisfy a continuity equation of the form (2.34) in order to grant the conservation of the total probability according to Eq. (3.12). Using $\rho(\vec{r},t)=|\psi(\vec{r},t)|^2$ and the expression (3.22) for the flux yields the continuity equation

$$\begin{aligned} 0 &= \frac{\partial}{\partial t}|\psi(\vec{r},t)|^2 + \vec{\nabla}\cdot\vec{j}(\vec{r},t) && (3.23)\\ &= 2\mathrm{Re}\left\{\psi^*(\vec{r},t)\left[\frac{\partial}{\partial t}\psi(\vec{r},t) + \frac{\hbar}{2mi}\Delta\psi(\vec{r},t)\right]\right\} && (3.24)\end{aligned}$$

with $\Delta \equiv \vec{\nabla}\cdot\vec{\nabla}$. It is straightforward to verify that this latter equation is satisfied if and only if the expression within the brackets $[\ldots]$ yields $\psi(\vec{r},t)$ multiplied with a purely imaginary prefactor that may depend on position and time. Denoting this prefactor as $-iV(\vec{r},t)/\hbar$ for some real function $V:\mathbb{R}^4\to\mathbb{R}, (\vec{r},t)\mapsto V(\vec{r},t)$ finally yields the celebrated *Schrödinger equation*

$$i\hbar\frac{\partial}{\partial t}\psi(\vec{r},t) = -\frac{\hbar^2}{2m}\Delta\psi(\vec{r},t) + V(\vec{r},t)\psi(\vec{r},t)\,. \tag{3.25}$$

It is instructive to multiply Eq. (3.25) with $\psi^*(\vec{r},t)$ and integrate the resulting equation over the entire space. Using Eq. (3.11) this yields

$$\int d^3r\,\psi^*(\vec{r},t)\left(-\frac{\hbar^2}{2m}\right)\Delta\psi(\vec{r},t) = \int d^3p\,\frac{\vec{p}^2}{2m}|\tilde{\psi}(\vec{p},t)|^2 \tag{3.26}$$

for the first term on the right-hand side of Eq. (3.25), which we immediately recognize as the mean value of the (nonrelativistic) kinetic energy of the particle. It is therefore straightforward to interpret, in close analogy with the classical Hamiltonian (3.3), the expression $\int d^3r\,V(\vec{r},t)|\psi(\vec{r},t)|^2$ that results from the second term on the right-hand side of Eq. (3.25) as mean value of the *potential energy* of the particle and the expression

$$E(t) = \int d^3r\,\psi^*(\vec{r},t)i\hbar\frac{\partial}{\partial t}\psi(\vec{r},t) \tag{3.27}$$

as the mean value of its total energy. We thereby identify $V(\vec{r},t)$ as the external potential energy that the particle is subject to at position $\vec{r}$ and time t, in an identical manner as for the classical Hamiltonian (3.3).

While its validity has been verified in an overwhelming number of experiments in the nonrelativistic context, the Schrödinger equation (3.25) is not in agreement with the principle of relativity. This is straightforwardly seen by reformulating Eq. (3.25) as

$$i\hbar c\partial_0\psi(\vec{r},t) = \frac{\hbar^2}{2m}\partial_l\partial^l\psi(\vec{r},t) + V(\vec{r},t)\psi(\vec{r},t)\,, \tag{3.28}$$

using according to Eq. (2.12) the relativistic notation $\partial_0 \equiv \frac{1}{c}\frac{\partial}{\partial t}$ and $\partial_l \equiv \frac{\partial}{\partial x^l}$ with $l=1,2,3$. Clearly, this equation is not invariant under Lorentz transformations as it does not correspond to any scalar, vectorial, or tensorial object with a well-defined transformation behaviour in the Minkowski space. It would therefore acquire the form (3.28) only in one specific reference frame, what would then distinguish this frame with respect to other possible reference frames.

Obviously, the nonrelativistic association $\vec{p} = m\vec{v}$ between the momentum $\vec{p}$ and the velocity $\vec{v}$ of the particle, which we used in order to infer the expression (3.22) for the probability flux, is at the origin of the disagreement of the Schrödinger equation with the theory of relativity. This association is, according to Eq. (3.1), derived from the nonrelativistic expression $E = p^2/(2m)$ of the kinetic energy of the particle, which is approximately obtained from the relativistic energy-momentum relation (2.1) according to

$$E = \sqrt{(mc^2)^2 + c^2p^2} \simeq mc^2 + \frac{p^2}{2m} - \frac{p^4}{8m^3c^2} + \mathcal{O}\left(\frac{1}{c^4}\right) \tag{3.29}$$

in the formal limit of an infinitely large speed of light c (or, more precisely, for $p \ll mc$). For very large momenta $p \gtrsim mc$ we would rather derive the expression

$$v = \frac{1}{m}\frac{\partial E}{\partial p} = \frac{p/m}{\sqrt{1 + [p/(mc)]^2}} = \frac{c}{\sqrt{1 + (mc/p)^2}} < c\,, \tag{3.30}$$

which, however, cannot be easily exploited to generalize the above line of arguments to the relativistic domain.

3.3 Interaction with an electromagnetic field

While the Schrödinger equation is not in agreement with the principle of relativity, the quantum theory that we developed in the previous section does nevertheless exhibit covariant features that can be reformulated in relativistic terms. This is particularly the case for the expectation value of the four-momentum $(p^\nu) = (E/c, \vec{p})$ which is determined with respect to the wavefunction ψ as

$$\langle p^\nu \rangle_t = \int d^3r\, \psi^*(x) i\hbar\partial^\nu \psi(x) \tag{3.31}$$

according to Eqs. (3.18) and (3.27), where we identify $x \equiv (ct, \vec{r})$ and $(\partial^\nu) \equiv (\frac{1}{c}\frac{\partial}{\partial t}, -\vec{\nabla})$.

Similarly, we can formulate in an analogous covariant manner the interaction of a charged quantum particle with an electromagnetic field. The latter is described by the four-potential $(A^\nu) = (\Phi, \vec{A})$ where $\Phi(\vec{r}, t)$ and $\vec{A}(\vec{r}, t)$ respectively represent the scalar and vector potential of the field at the position $\vec{r}$ at time t. While Φ simply provides a contribution to the external potential energy of the particle — or, more precisely, essentially *constitutes* the potential energy $V(\vec{r}, t)$ after being multiplied with the charge of the particle — the vector potential alters the nonrelativistic relation between its velocity $\vec{v}$ and its momentum $\vec{p}$ such that we have to distinguish between the *canonical momentum* $\vec{p}$ and the *kinetic momentum* $\vec{\pi} = m\vec{v}$ of the particle.

More precisely, the classical Hamiltonian of a point particle with the charge q reads

$$H(\vec{r}, \vec{p}, t) = \frac{1}{2m}\left[\vec{p} - \frac{q}{c}\vec{A}(\vec{r}, t)\right]^2 + q\Phi(\vec{r}, t)\,. \tag{3.32}$$

in the presence of an electromagnetic field. It gives rise to the classical equations of motion

$$\frac{d\vec{r}}{dt}(t) = \frac{\partial H}{\partial \vec{p}}\left[\vec{r}(t), \vec{p}(t), t\right] = \frac{1}{m}\vec{\pi}(t)\,, \tag{3.33}$$

$$\frac{d\vec{p}}{dt}(t) = -\frac{\partial H}{\partial \vec{r}}\left[\vec{r}(t), \vec{p}(t), t\right] = -q\frac{\partial \Phi}{\partial \vec{r}}\left[\vec{r}(t), t\right] + \frac{q}{c}\frac{dr_l}{dt}(t)\frac{\partial A_l}{\partial \vec{r}}\left[\vec{r}(t), t\right] \tag{3.34}$$

where we have introduced the kinetic momentum as

$$\vec{\pi}(t) = \vec{p}(t) - \frac{q}{c}\vec{A}\left[\vec{r}(t), t\right]\,. \tag{3.35}$$

Calculating its time derivative as

$$\frac{d\vec{\pi}}{dt}(t) = \frac{d\vec{p}}{dt}(t) - \frac{q}{c}\left\{\frac{\partial \vec{A}}{\partial r_l}\left[\vec{r}(t), t\right]\frac{dr_l}{dt}(t) + \frac{\partial \vec{A}}{\partial t}\left[\vec{r}(t), t\right]\right\}\,, \tag{3.36}$$

where we implicitly sum over $l = 1, 2, 3$ in this equation [and also in Eq. (3.34)], and using the relation

$$\frac{dr_l}{dt}(t)\left[\frac{\partial A_l}{\partial \vec{r}}(\vec{r}, t) - \frac{\partial \vec{A}}{\partial r_l}(\vec{r}, t)\right] = \frac{d\vec{r}}{dt}(t) \times \left[\vec{\nabla} \times \vec{A}(\vec{r}, t)\right]\,, \tag{3.37}$$

we finally obtain the Newtonian equation

$$m\frac{d^2\vec{r}}{dt^2}(t) = q\vec{E}\left[\vec{r}(t), t\right] + \frac{q}{c}\frac{d\vec{r}}{dt}(t) \times \vec{B}\left[\vec{r}(t), t\right] \tag{3.38}$$

describing the Lorentz force on a charged particle in the presence of the electric field $\vec{E}$ and the magnetic field $\vec{B}$ which are defined through Eqs. (2.28) and (2.29), respectively.

While we still maintain the basic postulate that the probability distributions of a quantum particle in position and momentum space are determined by a complex wavefunction ψ as described in Section 3.2, we now have to modify the definition (3.22) of the probability flux of the particle, which is, in view of Eq. (3.33), now given by the "density of kinetic momentum" divided by the mass m of the particle. This yields the modified expression

$$\vec{J}(\vec{r}, t) = \frac{1}{m}\mathrm{Re}\left[\psi^*(\vec{r}, t)\left(\frac{\hbar}{i}\vec{\nabla} - \frac{q}{c}\vec{A}(\vec{r}, t)\right)\psi(\vec{r}, t)\right] \tag{3.39}$$

for the probability flux in the presence of the electromagnetic field. Following the subsequent steps of argumentation that we developed in Section 3.2, where we set this time $V(\vec{r},t) = q\Phi(\vec{r},t) + q^2[\vec{A}(\vec{r},t)]^2/(2mc^2)$ for the real function $V : \mathbb{R}^4 \to \mathbb{R}$ to be freely choosen, we then obtain the modified Schrödinger equation

$$i\hbar\frac{\partial}{\partial t}\psi(\vec{r},t) = \frac{1}{2m}\left[\frac{\hbar}{i}\vec{\nabla} - \frac{q}{c}\vec{A}(\vec{r},t)\right]^2 \psi(\vec{r},t) + q\Phi(\vec{r},t)\psi(\vec{r},t)\,. \tag{3.40}$$

We note that Eq. (3.40) can formally be obtained from the ordinary Schrödinger equation (3.25) by setting $V \equiv 0$ and by replacing

$$\frac{\partial}{\partial t} \mapsto \frac{\partial}{\partial t} + \frac{iq}{\hbar}\Phi(\vec{r},t)\,, \tag{3.41}$$

$$\vec{\nabla} \mapsto \vec{\nabla} - \frac{iq}{\hbar c}\vec{A}(\vec{r},t)\,, \tag{3.42}$$

in Eq. (3.25). This replacement, which is also referred to as *minimal coupling* of the particle to the electromagnetic field, can be rewritten in a covariant manner according to

$$\partial_\nu \mapsto \partial_\nu + \frac{iq}{\hbar c}A_\nu(x) \tag{3.43}$$

and would therefore be in agreement with the principle of relativity.

It may appear strange that the electromagnetic field is represented in the Schrödinger equation (3.40) by the associated scalar and vector potentials the choice of which is not unique as was pointed out in Section 2.2. Indeed, it seems that a gauge transformation according to Eq. (2.30), which does not change the electric and magnetic fields (and hence does not modify the Newtonian equations of motion (3.37) either), does alter the time evolution of the wavefunction when being incorporated in the Schrödinger equation (3.40). The solution to this problem is that the wavefunction ψ is transformed as well under such a gauge transformation. More precisely, the gauge transformation $A_\nu \mapsto A'_\nu$ of the four-potential with

$$A'_\nu(x) = A_\nu(x) + \partial_\nu\chi(x) \tag{3.44}$$

for a given (twice continuously differentiable) scalar field χ has to be accompanied by an associated gauge transformation $\psi \mapsto \psi'$ of the wavefunction with

$$\psi'(\vec{r},t) = \psi(\vec{r},t)\exp\left[-\frac{iq}{\hbar c}\chi(\vec{r},t)\right]\,. \tag{3.45}$$

It is then straightforward to show that the new wavefunction ψ' verifies the modified Schrödinger equation

$$i\hbar\frac{\partial}{\partial t}\psi'(\vec{r},t) = \frac{1}{2m}\left[\frac{\hbar}{i}\vec{\nabla} - \frac{q}{c}\vec{A}'(\vec{r},t)\right]^2 \psi'(\vec{r},t) + q\Phi'(\vec{r},t)\psi'(\vec{r},t) \tag{3.46}$$

containing the transformed four-potential A'_ν, provided the old wavefunction ψ verified the original Schrödinger equation (3.40).

From a more fundamental point of view, the invariance of the Schrödinger equation with respect to the combined gauge transformation of ψ and A_ν according to Eqs. (3.44) and (3.45) reflects the request that the phase of the wavefunction should not carry any physical information. More specifically, assuming, for the sake of simplicity, that the wavefunction $\psi(\vec{r}, t) \in \mathbb{R}$ is real at time t a given position $\vec{r}$ in some physical laboratory should not have any impact onto the definition of the phase of the wavefunction in other nearby or distant laboratories. We conversely note that this freedom in the choice of the local phase can be imposed if and only if the wavefunction is coupled to an electromagnetic field according to Eq. (3.40). The presence of the latter can then be seen as a necessity to grant such a *local gauge invariance* of the wavefunction, which is the essence of *gauge theory*.

3.4 The Heisenberg picture

Heisenberg undertook a different approach in order to explain the uncertainty principle (3.13). He suggested that position and momentum ought to be non-commuting objects, such as matrices, instead of ordinary numbers. This point of view can indeed be reconciled with the Schrödinger picture of quantum mechanics that we developed in Section 3.2. We note for this purpose that by virtue of Eq. (3.12) the wavefunction ψ describing the state of the particle represents a vector of the Hilbert space

$$\mathcal{H} = \left\{ \psi : \mathbb{R}^3 \to \mathbb{C}, \int |\psi(\vec{r})|^2 d^3 r < \infty \right\} \tag{3.47}$$

of square-integrable (L^2) functions defined on the three-dimensional space. The position and momentum of the particle as well as any other physical observable $A \equiv A(\vec{p}, \vec{q})$ in relation with the particle can then be represented by a linear *operator* $\hat{A} : \mathcal{H} \to \mathcal{H}, \psi \mapsto \hat{A}\psi$ that acts on this Hilbert space. Its mean or *expecation value* with respect to the state ψ of the particle is then given by the expression $\langle \hat{A} \rangle = \langle \psi | \hat{A} | \psi \rangle$, where we define in a more general manner the *matrix element* of the operator $\hat{A}$ with respect to the states $\psi \in \mathcal{H}$ and $\phi \in \mathcal{H}$ by

$$\langle \phi | \hat{A} | \psi \rangle = \int d^3 r \phi^*(\vec{r}) \hat{A} \psi(\vec{r}) \,. \tag{3.48}$$

Physically relevant operators ought to be *hermitian*, *i.e.* identical to their adjoint $\hat{A}^\dagger$ with respect to the standard scalar product $\langle \phi | \psi \rangle = \int d^3 r \phi^*(\vec{r}) \psi(\vec{r})$, such that they satisfy $\langle \phi | \hat{A} | \psi \rangle = \langle \psi | \hat{A} | \phi \rangle^*$ for all $\psi, \phi \in \mathcal{H}$ and exhibit purely real expectation values $\langle \hat{A} \rangle \in \mathbb{R}$.

Most naturally, the position operator $\hat{\vec{r}}$ of the particle acts as

$$\left(\hat{\vec{r}}\psi\right)(\vec{r}) = \vec{r}\psi(\vec{r}) \tag{3.49}$$

on the wavefunction ψ and is straightforwardly shown to be hermitian, while we infer from Eq. (3.18) the definition

$$\left(\hat{\vec{p}}\psi\right)(\vec{r}) = \frac{\hbar}{i}\frac{\partial}{\partial\vec{r}}\psi(\vec{r}) \tag{3.50}$$

of the momentum operator $\hat{\vec{p}}$ the hermiticity of which is demonstrated through integration by parts as done in Eq. (3.19). In perfect accordance with the point of view of Heisenberg, the position and momentum operators do not commute with each other as we evaluate

$$\left(\hat{p}_{l'}\hat{r}_l\psi\right)(\vec{r}) = \frac{\hbar}{i}\frac{\partial}{\partial r_{l'}}\left[r_l\psi(\vec{r})\right] = \frac{\hbar}{i}\left[r_l\frac{\partial}{\partial r_{l'}}\psi(\vec{r}) + \delta_{ll'}\psi(\vec{r})\right] = \left(\hat{r}_l\hat{p}_{l'} - i\hbar\delta_{ll'}\right)\psi(\vec{r}) \tag{3.51}$$

for all $\psi \in \mathcal{H}$ and all $l, l' = 1, 2, 3$ and hence obtain

$$[\hat{r}_l, \hat{p}_{l'}] \equiv \hat{r}_l\hat{p}_{l'} - \hat{p}_{l'}\hat{r}_l = i\hbar\delta_{ll'} \tag{3.52}$$

for the *commutator* of the operators $\hat{r}_l$ and $\hat{p}_{l'}$. Arbitrary functions $f \equiv f(\vec{r})$ and $g \equiv g(\vec{p})$ of the position or momentum of the particle are then represented by the hermitian operators $\hat{f} = f(\hat{\vec{r}})$ and $\hat{g} = g(\hat{\vec{p}})$, respectively, whose expectation values with respect to the state ψ are then indeed given by the expressions (3.14) and (3.15). In particular, the total energy (3.3) of the particle is represented by the *Hamiltonian* operator

$$\hat{H} = \frac{\hat{\vec{p}}^{\,2}}{2m} + V(\hat{\vec{r}}, t) \equiv \hat{H}(t)\,, \tag{3.53}$$

which allows us to rewrite the Schrödinger equation (3.25) as

$$i\hbar\frac{\partial}{\partial t}\psi(\vec{r}, t) = \left(\hat{H}(t)\psi\right)(\vec{r}, t)\,. \tag{3.54}$$

We assume for the following that the potential energy $V(\vec{r}, t) \equiv V(\vec{r})$ is not explicitly dependent on time, which implies that the definition of the Hamiltonian $\hat{H}$ is time-independent. We can then formally integrate Eq. (3.54) yielding

$$\psi(\vec{r}, t) = \left[\hat{U}(t)\psi_0\right](\vec{r})\,, \tag{3.55}$$

where we define by $\psi_0(\vec{r}) = \psi(\vec{r}, 0)$ the initial state of the particle at the time $t = 0$ and by

$$\hat{U}(t) = \exp\left(-\frac{it}{\hbar}\hat{H}\right) = \sum_{k=0}^{\infty}\frac{1}{k!}\left(\frac{-it}{\hbar}\right)^k\hat{H}^k = \left[\hat{U}^\dagger(t)\right]^{-1} \tag{3.56}$$

the unitary *time evolution operator* or *propagator* of the system. It is then straightforward to show that the time-dependent expectation value of the operator $\hat{A}$ with respect to the state $\psi = \hat{U}(t)\psi_0$ can be expressed as

$$\langle \hat{A} \rangle_t = \langle \hat{U}(t)\psi_0 | \hat{A} | \hat{U}(t)\psi_0 \rangle = \langle \psi_0 | \hat{U}^\dagger(t)\hat{A}\hat{U}(t) | \psi_0 \rangle = \langle \psi_0 | \hat{A}(t) | \psi_0 \rangle \tag{3.57}$$

where we use $\langle \phi | \hat{U}(t) | \psi \rangle = \langle \psi | \hat{U}^\dagger(t) | \phi \rangle^*$ for all $\psi, \phi \in \mathcal{H}$ and define the *time-dependent* operator

$$\hat{A}(t) \equiv \hat{U}^\dagger(t)\hat{A}\hat{U}(t) = \exp\left(\frac{it}{\hbar}\hat{H}\right)\hat{A}\exp\left(-\frac{it}{\hbar}\hat{H}\right). \tag{3.58}$$

The essence of the Heisenberg picture is to claim that the state of the system is given by a *time-independent* wavefunction $\psi = \psi_0 \in \mathcal{H}$ while any operator $\hat{A}$ representing a physical observable *evolves with time* according to Eq. (3.58). Deriving this latter equation (3.58) with respect to time yields the *Heisenberg equation of motion*

$$\begin{aligned} \frac{d}{dt}\hat{A}(t) &= \frac{i}{\hbar}\hat{H}\exp\left(\frac{it}{\hbar}\hat{H}\right)\hat{A}\exp\left(-\frac{it}{\hbar}\hat{H}\right) - \exp\left(\frac{it}{\hbar}\hat{H}\right)\hat{A}\exp\left(-\frac{it}{\hbar}\hat{H}\right)\frac{i}{\hbar}\hat{H} \\ &= \frac{i}{\hbar}[\hat{H}, \hat{A}(t)] \end{aligned} \tag{3.59}$$

describing the time evolution of the operator $\hat{A}$. It is straightforward to show that $\hat{A}(t)\hat{B}(t) = (\hat{A}\hat{B})(t)$ for any pair of operators $\hat{A}$ and $\hat{B}$, and that any operator $\hat{A}$ that commutes with the Hamiltonian, $[\hat{H}, \hat{A}] = 0$, is independent of time, $\hat{A}(t) \equiv \hat{A}$, in the Heisenberg representation. This applies, in particular, to the Hamiltonian $\hat{H}$ itself, which we can now express as

$$\hat{H} = \frac{[\hat{\vec{p}}(t)]^2}{2m} + V[\hat{\vec{r}}(t)] \tag{3.60}$$

in the Heisenberg representation. It also applies to the commutator (3.52) of the position and momentum operators which is therefore evaluated as

$$[\hat{r}_l(t), \hat{p}_{l'}(t)] = [\hat{r}_l(0), \hat{p}_{l'}(0)] = i\hbar\delta_{ll'} \tag{3.61}$$

for any time t. Using Eqs. (3.60) and (3.61) we thereby obtain the Heisenberg equations (3.59) for the position and momentum operators as

$$\frac{d}{dt}\hat{\vec{r}}(t) = \frac{i}{\hbar}[\hat{H}, \hat{\vec{r}}(t)] = \frac{1}{m}\hat{\vec{p}}(t), \tag{3.62}$$

$$\frac{d}{dt}\hat{\vec{p}}(t) = \frac{i}{\hbar}[\hat{H}, \hat{\vec{p}}(t)] = -\vec{\nabla}V[\hat{\vec{r}}(t)]. \tag{3.63}$$

We note that these equations are perfectly analogous to the classical equations of motion (3.1) and (3.2) describing the time evolution of the position and momentum coordinates of a classical particle, and can formally be generated from the latter by putting operator hats on top of $\vec{r}$ and $\vec{p}$.

3.5 The harmonic oscillator

Let us now again consider the specific case of the three-dimensional anisotropic harmonic oscillator that is characterized by the potential energy (3.5). The Heisenberg equations (3.62) and (3.63) are then written as

$$\frac{d}{dt}\hat{r}_l(t) = \frac{1}{m}\hat{p}_l(t)\,, \tag{3.64}$$

$$\frac{d}{dt}\hat{p}_l(t) = -m\omega_l^2\hat{r}_l(t) \tag{3.65}$$

for this particular system (with $l = 1, 2, 3$), which yields

$$\frac{d^2}{dt^2}\hat{r}_l(t) = -\omega_l^2\hat{r}_l(t) \tag{3.66}$$

as quantum analog of the classical Newtonian equations (3.6). In perfect analogy with the classical harmonic oscillator [see Eq. (3.7)], these equations would generally be solved as $\hat{r}_l(t) = \hat{\alpha}_l e^{-i\omega_l t} + \hat{\alpha}_l^\dagger e^{i\omega t}$ for some operators $\hat{\alpha}_l$. It is, however, more convenient to factor out $\sqrt{\hbar/(2m\omega_l)}$ from this expression, such that we define the time-dependent operators $\hat{a}_l(t) = \sqrt{2m\omega_l/\hbar}\hat{\alpha}_l e^{-i\omega_l t}$ and consequently write the general solution of the Heisenberg equations (3.64) and (3.65) as

$$\hat{r}_l(t) = \sqrt{\frac{\hbar}{2m\omega_l}}\left[\hat{a}_l(t) + \hat{a}_l(t)^\dagger\right]\,, \tag{3.67}$$

$$\hat{p}_l(t) = i\sqrt{\frac{\hbar m\omega_l}{2}}\left[\hat{a}_l(t)^\dagger - \hat{a}_l(t)\right]\,. \tag{3.68}$$

These equations can be solved for $\hat{a}_l$ yielding

$$\hat{a}_l(t) = \sqrt{\frac{m\omega_l}{2\hbar}}\left[\hat{r}_l(t) + \frac{i}{m\omega_l}\hat{p}_l(t)\right]\,. \tag{3.69}$$

Using Eq. (3.61), we straightforwardly evaluate from this expression (3.69) and its adjoint the commutation rules

$$[\hat{a}_l(t), \hat{a}_{l'}^\dagger(t)] = \delta_{ll'} \tag{3.70}$$

as well as $[\hat{a}_l(t), \hat{a}_{l'}(t)] = 0 = [\hat{a}_l^\dagger(t), \hat{a}_{l'}^\dagger(t)]$ for all $l, l' = 1, 2, 3$. The Hamiltonian of this system can then be expressed as

$$\hat{H} = \frac{1}{2}\sum_{l=1}^{3}\left(\frac{\hat{p}_l^2}{m} + m\omega_l^2\hat{r}_l^2\right) = \sum_{l=1}^{3}\frac{1}{2}\hbar\omega_l\left(\hat{a}_l\hat{a}_l^\dagger + \hat{a}_l^\dagger\hat{a}_l\right) \tag{3.71}$$

$$= \sum_{l=1}^{3}\hbar\omega_l\left(\hat{a}_l^\dagger\hat{a}_l + \frac{1}{2}\right)\,, \tag{3.72}$$

where we used the commutator (3.70) to obtain the expression (3.72) from Eq. (3.71). Deriving from this expression the Heisenberg equation for the operator $\hat{a}_l$ according to Eq. (3.59) yields

$$\frac{d}{dt}\hat{a}_l(t) = \frac{i}{\hbar}[\hat{H}, \hat{a}_l(t)] = -i\omega_l \hat{a}_l(t) \tag{3.73}$$

using again the commutator (3.70). This equation is straightforwardly solved as $\hat{a}_l(t) = e^{-i\omega_l t}\hat{a}_l(0)$, in perfect agreement with the general solution of the Heisenberg equations (3.64) and (3.65) for the position and momentum operators.

$\hat{a}_l$ and $\hat{a}_l^\dagger$ are named *ladder operators*, owing to the specific spectral properties that exhibits the harmonic oscillator Hamiltonian (3.72). Indeed, assuming that $|n\rangle$ is an eigenstate of the (hermitian) operator $\hat{a}_l^\dagger \hat{a}_l$, *i.e.* we have

$$\hat{a}_l^\dagger \hat{a}_l |n\rangle = n|n\rangle \tag{3.74}$$

for some (real) eigenvalue $n \in \mathbb{R}$, we can show through

$$\hat{a}_l^\dagger \hat{a}_l \hat{a}_l^\dagger |n\rangle = \hat{a}_l^\dagger \hat{a}_l^\dagger \hat{a}_l |n\rangle + \hat{a}_l^\dagger |n\rangle = (n+1)\hat{a}_l^\dagger |n\rangle\,, \tag{3.75}$$

$$\hat{a}_l^\dagger \hat{a}_l \hat{a}_l |n\rangle = \hat{a}_l \hat{a}_l^\dagger \hat{a}_l |n\rangle - \hat{a}_l |n\rangle = (n-1)\hat{a}_l |n\rangle \tag{3.76}$$

[where we use again the commutator (3.70)] that $\hat{a}_l^\dagger|n\rangle$ and $\hat{a}_l|n\rangle$ are eigenstates of $\hat{a}_l^\dagger \hat{a}_l$, too, for the eigenvalues $(n+1)$ and $(n-1)$, respectively. Realizing furthermore that $n = \langle n|\hat{a}_l^\dagger \hat{a}_l|n\rangle = \langle \hat{a}_l n|\hat{a}_l n\rangle \geq 0$, we can infer that $n \in \mathbb{N}_0 = \{0, 1, 2, \ldots\}$ has to be a natural number and that we have $\hat{a}_l|0\rangle = 0$.

The spectrum of the Hamiltonian (3.72) is therefore discrete and bounded from below. Its normalized eigenstates can formally be written as

$$|n_1, n_2, n_3\rangle = \left[\prod_{l=1}^{3} \frac{1}{\sqrt{n_l!}} \left(\hat{a}_l^\dagger\right)^{n_l}\right] |0, 0, 0\rangle \tag{3.77}$$

for $n_1, n_2, n_3 \in \mathbb{N}_0$, where we denote by $|0, 0, 0\rangle$ the normalized ground state of the Hamiltonian. The associated eigenvalues read

$$E_{n_1,n_2,n_3} = \sum_{l=1}^{3} \hbar\omega_l \left(n_l + \frac{1}{2}\right)\,. \tag{3.78}$$

We note that these properties can be inferred for any Hamiltonian that can be written in the form (3.72) where $\hat{a}_l$ and $\hat{a}_l^\dagger$ satisfy the commutation relation (3.70), independently of the physical context under consideration.

Problem

3.1 Show that the gauge-transformed wavefunction (3.45) verifies the modified Schrödinger equation (3.46) if the original wavefunction evolves according to Eq. (3.40).

Chapter 4

The quantization of fields

4.1 Classical waves

We consider a field $\phi : \mathbb{R}^4 \to \mathbb{R}, (\vec{r}, t) \mapsto \phi(\vec{r}, t)$ that satisfies the wave equation

$$\left(\frac{1}{\tilde{c}^2} \frac{\partial^2}{\partial t^2} - \Delta \right) \phi(\vec{r}, t) = 0 \tag{4.1}$$

where $\tilde{c}$ is some speed constant. This equation would describe a wave that propagates in three-dimensional space with the speed $\tilde{c}$. The field $\phi(\vec{r}, t)$ could, for instance, represent the local variation of the density of molecules with respect to its average value at the position $\vec{r}$ and the time t in order to model sound waves in the air (in which case we would have $\tilde{c} \simeq 330$ m/s).

Provided that ϕ is integrable (which implies $|\phi(\vec{r}, t)| \to 0$ for $|\vec{r}| \to \infty$), we can solve Eq. (4.1) by performing a Fourier transform. This amounts to introducing another field $\tilde{\phi} : \mathbb{R}^4 \to \mathbb{C}, (\vec{k}, t) \mapsto \tilde{\phi}(\vec{k}, t)$ which is complex-valued and defined by the relation

$$\tilde{\phi}(\vec{k}, t) = \frac{1}{\sqrt{2\pi}^3} \int \phi(\vec{r}, t) e^{-i\vec{k}\cdot\vec{r}} d^3 r \, . \tag{4.2}$$

Subjecting Eq. (4.1) to this Fourier transform yields the now ordinary differential equation

$$\left(\frac{1}{\tilde{c}^2} \frac{\partial^2}{\partial t^2} + \vec{k}^2 \right) \tilde{\phi}(\vec{k}, t) = 0 \, . \tag{4.3}$$

The general solution of this latter equation is straightforwardly written as

$$\tilde{\phi}(\vec{k}, t) = \alpha_{\vec{k}} e^{-i\omega_k t} + \beta_{\vec{k}} e^{i\omega_k t} \tag{4.4}$$

for some complex coefficients $\alpha_{\vec{k}}, \beta_{\vec{k}} \in \mathbb{C}$ where we define $\omega_k = \tilde{c}|\vec{k}|$. Reverting the Fourier transform (4.2) yields then the general solution of the wave equation

(4.1) according to

$$\phi(\vec{r},t) = \frac{1}{\sqrt{2\pi}^3}\int d^3k\tilde{\phi}(\vec{k},t)e^{i\vec{k}\cdot\vec{r}} \tag{4.5}$$

$$= \frac{1}{\sqrt{2\pi}^3}\int d^3k\left(\alpha_{\vec{k}}e^{i(\vec{k}\cdot\vec{r}-\omega_k t)} + \beta_{-\vec{k}}e^{-i(\vec{k}\cdot\vec{r}-\omega_k t)}\right), \tag{4.6}$$

where we have to require $\beta_{-\vec{k}} = \alpha^*_{\vec{k}}$ for all $\vec{k} \in \mathbb{R}^3$ in order to obtain a real-valued wave field.

For technical reasons, we shall, in the following, also consider the solution of wave equations of the above type (4.1) not within the infinitely extended free three-dimensional space $\mathbb{R}^3$ but within a *normalization volume* of finite extent. The latter is most conveniently defined by a cube of length L and volume $V = L^3$ that exhibits periodic boundary conditions at its surfaces. Evidently, it is assumed that L exceeds all relevant length scales of the system, and the limit $L \to \infty$ will be performed at the end of the calculation in order to approach the situation of an infinitely extended space. While the realization of such periodic boundary conditions is hard to conceive in the experimental practice, one could imagine that they might indeed occur within the universe whose spatial volume is believed to be finite.

The integral in the Fourier transform (4.2) is then to be restricted to this particular normalization volume, while the inverse Fourier transform (4.5) has to be restricted to the plane waves $\exp(i\vec{k}\cdot\vec{r})$ that comply with the above periodic boundary conditions, *i.e.*, whose wave vectors satisfy $\vec{k} = (2\pi/L)\vec{l}$ for some $\vec{l} \equiv (l_1, l_2, l_3) \in \mathbb{Z}^3$. The general solution of the wave equation (4.1) is then obtained as the Fourier series

$$\phi(\vec{r},t) = \frac{1}{\sqrt{V}}\sum_{\vec{k}}\left(\alpha_{\vec{k}}e^{i(\vec{k}\cdot\vec{r}-\omega_k t)} + \beta_{-\vec{k}}e^{-i(\vec{k}\cdot\vec{r}-\omega_k t)}\right), \tag{4.7}$$

where we use the notation $\sum_{\vec{k}} \equiv \sum_{l_1=-\infty}^{\infty}\sum_{l_2=-\infty}^{\infty}\sum_{l_3=-\infty}^{\infty}$ and require again $\beta_{-\vec{k}} = \alpha^*_{\vec{k}}$ for all $\vec{k} \in \mathbb{R}^3$ in order to obtain a real-valued wave field.

While Eq. (4.1) represents the Newtonian description of the time evolution of the wave in analogy with Eq. (3.6), a Hamiltonian formulation that generates the propagation of the wave in analogy with Eqs. (3.1–3.3) can be obtained as well. To this end, we interpret the local amplitude $\phi(\vec{r},t)$ of the wave at the place $\vec{r}$ as an effective "position variable" which evolves due to its coupling with other "position variables" defined at nearby places $\vec{r}' \neq \vec{r}$. This interpretation would be rather obvious for sound waves in solids where $\phi(\vec{r},t)$ can be identified with the elongation of an atom located at $\vec{r}$ with respect to its equilibrium position.

Pursuing this analogy, we can formulate the Hamiltonian

$$H = \int d^3r\left\{\frac{1}{2\tilde{m}}\left[\Pi(\vec{r},t)\right]^2 + \frac{1}{2}\tilde{m}\tilde{c}^2\left[\vec{\nabla}\phi(\vec{r},t)\right]^2\right\} \equiv H[\phi,\Pi] \tag{4.8}$$

for some arbitrarily chosen effective mass parameter $\tilde{m}$, which is a functional of the wave field ϕ and of another real-valued field $\Pi : \mathbb{R}^4 \to \mathbb{R}$ representing the effective "momentum variables". These fields evolve according to Hamilton's equations of motion

$$\frac{\partial}{\partial t}\phi(\vec{r},t) = \frac{\delta H}{\delta \Pi(\vec{r},t)} = \frac{1}{\tilde{m}}\Pi(\vec{r},t)\,, \tag{4.9}$$

$$\frac{\partial}{\partial t}\Pi(\vec{r},t) = -\frac{\delta H}{\delta \phi(\vec{r},t)} = \tilde{m}\tilde{c}^2\Delta\phi(\vec{r},t)\,, \tag{4.10}$$

which are defined in perfect analogy with Eqs. (3.1) and (3.2) where we replace the partial derivatives with respect to $\vec{r}$ and $\vec{p}$ by functional derivatives with respect to $\phi(\vec{r},t)$ and $\Pi(\vec{r},t)$. These functional derivatives are evaluated using

$$\frac{\delta}{\delta \Pi(\vec{r},t)}\int \left[\Pi(\vec{r}\,',t)\right]^2 d^3r' = 2\Pi(\vec{r},t)\,, \tag{4.11}$$

$$\frac{\delta}{\delta \phi(\vec{r},t)}\int \left[\vec{\nabla}\phi(\vec{r}\,',t)\right]^2 d^3r' = -2\Delta\phi(\vec{r},t) \tag{4.12}$$

where for the latter functional derivation we can employ the integration by parts

$$\int \vec{\nabla}\chi(\vec{r}\,')\cdot\vec{\nabla}\varphi(\vec{r}\,')d^3r' = -\int \chi(\vec{r}\,')\Delta\varphi(\vec{r}\,')d^3r' = -\int \left[\Delta\chi(\vec{r}\,')\right]\varphi(\vec{r}\,')d^3r' \tag{4.13}$$

which is valid for two integrable fields $\varphi, \chi : \mathbb{R}^3 \to \mathbb{R}$. Combining the Hamiltonian equations (4.9) and (4.10) through

$$\frac{\partial^2}{\partial t^2}\phi(\vec{r},t) = \frac{1}{\tilde{m}}\frac{\partial}{\partial t}\Pi(\vec{r},t) = \tilde{c}^2\Delta\phi(\vec{r},t) \tag{4.14}$$

finally allows us to recover the wave equation (4.1).

The Hamiltonian formalism can also be applied for the case of a wave that is described by a complex-valued field. In that case, we would employ the Hamiltonian functional

$$H = \int \left[\frac{1}{\tilde{m}}\Pi^*(\vec{r},t)\Pi(\vec{r},t) + \tilde{m}\tilde{c}^2\vec{\nabla}\phi^*(\vec{r},t)\cdot\vec{\nabla}\phi(\vec{r},t)\right] d^3r \equiv H[\phi,\phi^*,\Pi,\Pi^*]\,, \tag{4.15}$$

which is formally defined in terms of four independent fields $\phi, \phi^*, \Pi, \Pi^* : \mathbb{R}^4 \to \mathbb{C}$, and obtain Hamilton's equations of motion according to the prescriptions

$$\frac{\partial}{\partial t}\phi(\vec{r},t) = \frac{\delta H}{\delta \Pi^*(\vec{r},t)} = \frac{1}{\tilde{m}}\Pi(\vec{r},t)\,, \tag{4.16}$$

$$\frac{\partial}{\partial t}\Pi(\vec{r},t) = -\frac{\delta H}{\delta \phi^*(\vec{r},t)} = \tilde{m}\tilde{c}^2\Delta\phi(\vec{r},t)\,. \tag{4.17}$$

This yields again the wave equation (4.1).

4.2 Quantization of (real) waves

A quantum theory of sound can be obtained through a quantum description of the underlying material medium that carries the sound waves, which in the case of solids would effectively be constituted by a dense network of coupled oscillators. In contrast, electromagnetic waves in vacuum do not appear to exhibit a material carrier medium, even though the existence of such an *aether* was seriously considered for a while until some one hundred years ago. It is therefore necessary to introduce a scheme how to directly quantize classical waves, without resorting to the existence of a (possibly purely virtual) carrier medium.

We shall employ the Heisenberg picture for this purpose. In close analogy with Section 3.4, the "position" and "momentum variables" $\phi(\vec{r},t)$ and $\Pi(\vec{r},t)$ that describe the classical wave field at the (spatial) position $\vec{r}$ and the time t within the Hamiltonian formalism are replaced by hermitian operators

$$\phi(\vec{r},t) \mapsto \hat{\phi}(\vec{r},t) = \hat{\phi}^\dagger(\vec{r},t)\,, \tag{4.18}$$
$$\Pi(\vec{r},t) \mapsto \hat{\Pi}(\vec{r},t) = \hat{\Pi}^\dagger(\vec{r},t) \tag{4.19}$$

that do not commute with each other. More precisely, these *field operators* are proposed to satisfy the commutation relations

$$\left[\hat{\phi}(\vec{r},t), \hat{\Pi}(\vec{r}',t)\right] = i\hbar\delta(\vec{r}-\vec{r}') \tag{4.20}$$

for all $\vec{r},\vec{r}' \in \mathbb{R}^3$ and all $t \in \mathbb{R}$, in analogy with the commutation relations (3.52) of the ordinary position and momentum operators, where we replace the Kronecker symbol $\delta_{ll'}$ appearing in Eq. (3.52) by Dirac's delta distribution $\delta(\vec{r}-\vec{r}')$. We furthermore impose

$$\left[\hat{\phi}(\vec{r},t), \hat{\phi}(\vec{r}',t)\right] = 0 = \left[\hat{\Pi}(\vec{r},t), \hat{\Pi}(\vec{r}',t)\right] \tag{4.21}$$

for all $\vec{r},\vec{r}' \in \mathbb{R}^3$ and all $t \in \mathbb{R}$, which reflects the fact that different components of the ordinary position (or momentum) operator do also commute with each other.

The quantum analog of the Hamiltonian (4.8) describing the spatiotemporal evolution of the wave field is then straightforwardly obtained by applying the replacements (4.18) and (4.19) within Eq. (4.8). This yields the quantum Hamiltonian operator

$$\hat{H} = \int \left\{ \frac{1}{2\tilde{m}} \left[\hat{\Pi}(\vec{r},t)\right]^2 + \frac{1}{2}\tilde{m}\tilde{c}^2 \left[\vec{\nabla}\hat{\phi}(\vec{r},t)\right]^2 \right\} d^3r\,. \tag{4.22}$$

In perfect analogy with Eqs. (3.62) and (3.63), the time evolution of the quantum wave field operators is described by the Heisenberg equations

$$\frac{\partial}{\partial t}\hat{\phi}(\vec{r},t) = \frac{i}{\hbar}\left[\hat{H}, \hat{\phi}(\vec{r},t)\right] = \frac{1}{\tilde{m}}\hat{\Pi}(\vec{r},t)\,, \tag{4.23}$$
$$\frac{\partial}{\partial t}\hat{\Pi}(\vec{r},t) = \frac{i}{\hbar}\left[\hat{H}, \hat{\Pi}(\vec{r},t)\right] = \tilde{m}\tilde{c}^2\Delta\hat{\phi}(\vec{r},t)\,, \tag{4.24}$$

which are obtained using

$$\int d^3r' \left[\hat{\Pi}(\vec{r}', t)\hat{\Pi}(\vec{r}', t), \hat{\phi}(\vec{r}, t)\right] = -2i\hbar\hat{\Pi}(\vec{r}, t)\,, \tag{4.25}$$

$$\int d^3r' \left[\vec{\nabla}\hat{\phi}(\vec{r}', t)\cdot\vec{\nabla}\hat{\phi}(\vec{r}', t), \hat{\Pi}(\vec{r}, t)\right] = -2i\hbar\Delta\hat{\phi}(\vec{r}, t) \tag{4.26}$$

as a consequence of the commutation relations (4.20) (we again apply an integration by parts in Eq. (4.26)). A second derivation of Eq. (4.23) with respect to time yields the wave equation

$$\frac{1}{\tilde{c}^2}\frac{\partial^2}{\partial t^2}\hat{\phi}(\vec{r}, t) - \Delta\hat{\phi}(\vec{r}, t) = 0 \tag{4.27}$$

for the field operator $\hat{\phi}$, which is perfectly equivalent to its classical counterpart (4.1).

As a consequence, the solution of this quantum wave equation (4.27) can be calculated in exactly the same manner as for classical waves. Let us first consider the conceptually simpler case of a quantum wave field that evolves within a cubic normalization volume of length L. In perfect analogy with Eq. (4.7), we obtain

$$\hat{\phi}(\vec{r}, t) = \frac{1}{\sqrt{V}}\sum_{\vec{k}}\left(\hat{\alpha}_{\vec{k}}e^{i(\vec{k}\cdot\vec{r}-\omega_k t)} + \hat{\alpha}^\dagger_{\vec{k}}e^{-i(\vec{k}\cdot\vec{r}-\omega_k t)}\right) \tag{4.28}$$

as general solution of Eq. (4.27), where the sum $\sum_{\vec{k}} \equiv \sum_{l_1=-\infty}^{\infty}\sum_{l_2=-\infty}^{\infty}\sum_{l_3=-\infty}^{\infty}$ is restricted to wave vectors that satisfy $\vec{k} = (2\pi/L)(l_1, l_2, l_3)$ with $l_j \in \mathbb{Z}$, and where we define again $\omega_k = \tilde{c}|\vec{k}|$. In analogy with the harmonic oscillator, it is convenient to rewrite Eq. (4.28) as well as the corresponding expression for the "momentum" field operator $\hat{\Pi}(\vec{r}, t) = \tilde{m}(\partial/\partial t)\hat{\phi}(\vec{r}, t)$ as

$$\hat{\phi}(\vec{r}, t) = \frac{1}{\sqrt{V}}\sum_{\vec{k}}\sqrt{\frac{\hbar}{2\tilde{m}\omega_k}}\left[\hat{a}_{\vec{k}}(t)e^{i\vec{k}\cdot\vec{r}} + \hat{a}^\dagger_{\vec{k}}(t)e^{-i\vec{k}\cdot\vec{r}}\right]\,, \tag{4.29}$$

$$\hat{\Pi}(\vec{r}, t) = \frac{1}{i\sqrt{V}}\sum_{\vec{k}}\sqrt{\frac{\hbar\tilde{m}\omega_k}{2}}\left[\hat{a}_{\vec{k}}(t)e^{i\vec{k}\cdot\vec{r}} - \hat{a}^\dagger_{\vec{k}}(t)e^{-i\vec{k}\cdot\vec{r}}\right]\,, \tag{4.30}$$

where we introduce the time-dependent operators

$$\hat{a}_{\vec{k}}(t) = \sqrt{\frac{2\tilde{m}\omega_k}{\hbar}}\hat{\alpha}_{\vec{k}}e^{-i\omega_k t}\,. \tag{4.31}$$

The latter can be determined through an inversion of the Fourier series expansions (4.29) and (4.30) yielding

$$\hat{a}_{\vec{k}}(t) = \frac{1}{\sqrt{V}}\int_V\sqrt{\frac{\tilde{m}\omega_k}{2\hbar}}\left[\hat{\phi}(\vec{r}, t) + \frac{i}{\tilde{m}\omega_k}\hat{\Pi}(\vec{r}, t)\right]e^{-i\vec{k}\cdot\vec{r}}d^3r\,, \tag{4.32}$$

$$\hat{a}^\dagger_{\vec{k}}(t) = \frac{1}{\sqrt{V}}\int_V\sqrt{\frac{\tilde{m}\omega_k}{2\hbar}}\left[\hat{\phi}(\vec{r}, t) - \frac{i}{\tilde{m}\omega_k}\hat{\Pi}(\vec{r}, t)\right]e^{i\vec{k}\cdot\vec{r}}d^3r\,. \tag{4.33}$$

Evaluating Eqs. (4.32) and (4.33) at $t = 0$ allows one then to express the solution of the Heisenberg equations (4.23) and (4.24) in terms of the initial definitions of the field operators $\hat{\phi}$ and $\hat{\Pi}$, which in turn can be practically used in order to calculate expectation values of any kind of (one-, two-, or many-body) operators with respect to the initial quantum state of the system.

It is a tedious but straightforward calculation to show that the quantum Hamiltonian (4.22) can be expressed in terms of these operators $\hat{a}_{\vec{k}}$ and $\hat{a}^\dagger_{\vec{k}}$ as

$$\hat{H} = \sum_{\vec{k}} \frac{\hbar\omega_k}{2} \left[\hat{a}_{\vec{k}}(t)\hat{a}^\dagger_{\vec{k}}(t) + \hat{a}^\dagger_{\vec{k}}(t)\hat{a}_{\vec{k}}(t) \right] . \tag{4.34}$$

This expression can be even further simplified by inferring from Eqs. (4.32) and (4.33) the commutators

$$\left[\hat{a}_{\vec{k}}(t), \hat{a}^\dagger_{\vec{k}'}(t) \right] = \delta_{\vec{k}\vec{k}'} \tag{4.35}$$

as well as

$$\left[\hat{a}_{\vec{k}}(t), \hat{a}_{\vec{k}'}(t) \right] = 0 = \left[\hat{a}^\dagger_{\vec{k}}(t), \hat{a}^\dagger_{\vec{k}'}(t) \right] \tag{4.36}$$

for all $\vec{k} = (2\pi/L)\vec{l}$ and all $\vec{k}' = (2\pi/L)\vec{l}'$ with $\vec{l}, \vec{l}' \in \mathbb{Z}^3$, using the proposed commutation rules (4.20) and (4.21) for the field operators $\hat{\phi}$ and $\hat{\Pi}$. We then obtain

$$\hat{H} = \sum_{\vec{k}} \hbar\omega_k \left[\hat{a}_{\vec{k}}(t)\hat{a}^\dagger_{\vec{k}}(t) + \frac{1}{2} \right] . \tag{4.37}$$

This latter Hamiltonian effectively describes a combination of infinitely many (uncoupled) quantum harmonic oscillators that are associated with the eigenmodes $\exp(i\vec{k}\cdot\vec{r})$ and oscillate with the frequencies ω_k. As is seen from the commutation relations (4.35), $\hat{a}_{\vec{k}}$ and $\hat{a}^\dagger_{\vec{k}}$ represent the corresponding ladder operators. In close analogy with the "ordinary" quantum harmonic oscillator discussed in Section 3.5, the application of $\hat{a}^\dagger_{\vec{k}}$ increments by one the quantum number associated with the mode $\exp(i\vec{k}\cdot\vec{r})$ or, in other words, "creates an excitation" within that mode, whereas the application of $\hat{a}_{\vec{k}}$ decrements the mode's quantum number or "removes an excitation" within that mode. These operators are therefore named *creation* and *annihilation* operators, and the "excitations" that they create or annihilate are identified with indistinguishable quantum *particles* that are of *bosonic* nature (such as photons, phonons, or magnons, to mention a few examples), which is in accordance with the fact that two or more such particles can occupy the same mode.

In analogy with Eq. (3.77), the normalized eigenstates of the Hamiltonian (4.37) can formally be written as

$$\left| \{n_{\vec{k}}\}_{\vec{k}=(2\pi/L)\vec{l} \text{ with } \vec{l}\in\mathbb{Z}^3} \right\rangle = \left[\prod_{\vec{k}} \frac{1}{\sqrt{n_{\vec{k}}!}} \left(\hat{a}^\dagger_{\vec{k}} \right)^{n_{\vec{k}}} \right] |-\rangle \tag{4.38}$$

for all non-negative integers $n_{\vec{k}} \in \mathbb{N}_0$, where $|-\rangle \equiv |\ldots, 0, 0, 0, \ldots\rangle$ represents the ground state of the Hamiltonian. As this particular state $|-\rangle$ is characterized by the absence of any excitation or particle within the wave modes, we also refer to it as *vacuum state*. It is interesting to note that the associated ground state energy $E_0 = \langle -|\hat{H}|-\rangle = \frac{1}{2}\sum_{\vec{k}} \hbar\omega_k = \infty$ is infinitely large, due to the fact that the wave equation (4.1) exhibits infinitely many eigenmodes. While such an infinity may eventually give rise to some conceptual trouble in the framework of theories for quantum gravity that rely on absolute energies, it can essentially be discarded in our context of special relativity as E_0 is merely a constant that does not influence the time evolution of the system.

Let us finally discuss the case of a quantum wave that evolves within the infinitely extended three-dimensional space $\mathbb{R}^3$. Inspired from Eq. (4.5), we would now express the solution of the quantum wave equation (4.27) as

$$\hat{\phi}(\vec{r}, t) = \frac{1}{\sqrt{2\pi}^3} \int d^3k \sqrt{\frac{\hbar}{2\tilde{m}\omega_k}} \left[\hat{a}_{\vec{k}}(t) e^{i\vec{k}\cdot\vec{r}} + \hat{a}^\dagger_{\vec{k}}(t) e^{-i\vec{k}\cdot\vec{r}}\right], \tag{4.39}$$

which essentially amounts to replacing within Eq. (4.29) the summation over wave vectors by an integration. This expression implicitly requires an integrability condition to hold concerning the spatial distribution of excitations or particles within the intial quantum state of the system, in order for such a Fourier ansatz to be justified. In analogy with Eq. (4.32), we consequently obtain

$$\hat{a}_{\vec{k}}(t) = \frac{1}{\sqrt{2\pi}^3} \int d^3r \sqrt{\frac{\tilde{m}\omega_k}{2\hbar}} \left[\hat{\phi}(\vec{r}, t) + \frac{i}{\tilde{m}\omega_k}\hat{\Pi}(\vec{r}, t)\right] e^{-i\vec{k}\cdot\vec{r}} \tag{4.40}$$

for the annihilation operator associated with the wave mode $\exp(i\vec{k}\cdot\vec{r})$. We can then derive the commutation relations

$$\left[\hat{a}_{\vec{k}}(t), \hat{a}^\dagger_{\vec{k}'}(t)\right] = \delta(\vec{k} - \vec{k}'), \tag{4.41}$$

which can be obtained from Eq. (4.35) by replacing the Kronecker delta $\delta_{\vec{k}\vec{k}'}$ with Dirac's delta distribution $\delta(\vec{k} - \vec{k}')$, as well as

$$\left[\hat{a}_{\vec{k}}(t), \hat{a}_{\vec{k}'}(t)\right] = 0 = \left[\hat{a}^\dagger_{\vec{k}}(t), \hat{a}^\dagger_{\vec{k}'}(t)\right] \tag{4.42}$$

for all $\vec{k}, \vec{k}' \in \mathbb{R}^3$.

The quantum Hamiltonian (4.22) is now reformulated as

$$\hat{H} = \int d^3k \frac{\hbar\omega_k}{2} \left[\hat{a}_{\vec{k}}(t)\hat{a}^\dagger_{\vec{k}}(t) + \hat{a}^\dagger_{\vec{k}}(t)\hat{a}_{\vec{k}}(t)\right], \tag{4.43}$$

where for obvious reasons we do not attempt to further simplify the expression by applying the commutator (4.41). This latter expression gives rise to the Heisenberg equations

$$\frac{d}{dt}\hat{a}_{\vec{k}}(t) = \frac{i}{\hbar}\left[\hat{H}, \hat{a}_{\vec{k}}(t)\right] = -i\omega_k \hat{a}_{\vec{k}}(t)\,, \tag{4.44}$$

$$\frac{d}{dt}\hat{a}^\dagger_{\vec{k}}(t) = \frac{i}{\hbar}\left[\hat{H}, \hat{a}^\dagger_{\vec{k}}(t)\right] = i\omega_k \hat{a}^\dagger_{\vec{k}}(t) \tag{4.45}$$

owing to the commutation relations (4.41) and (4.42). These equations are solved as $\hat{a}_{\vec{k}}(t) = \hat{a}_{\vec{k}}(0)\exp(-i\omega_k t)$ and $\hat{a}^\dagger_{\vec{k}}(t) = \hat{a}^\dagger_{\vec{k}}(0)\exp(i\omega_k t)$, which is in perfect agreement with the time dependence (4.31) that would be required within Eq. (4.39) in order for that expression to solve the quantum wave equation (4.27).

4.3 The photon

Let uns now specifically discuss the quantization of electromagnetic waves in vacuum. Following the reasoning developed in Section 2.2, we can represent the electromagnetic field according to Eq. (2.25) through the associated four-potential $(A^\nu) = (\Phi, \vec{A})$ which is chosen such that it satisfies the Lorenz gauge (2.32) $\partial_\nu A^\nu(x) = 0$ for all $x \in \mathbb{R}^4$. In the absence of electric charges and currents, this four-potential evolves according to the wave equation

$$\partial_\mu \partial^\mu A^\nu(x) = 0 \tag{4.46}$$

for all $x \in \mathbb{R}^4$, as was shown in Section 2.3.

The absence of charged matter allows us to further simplify the problem by assuming without loss of generality that the scalar potential $\Phi \equiv 0$ vanishes. More precisely, starting from a four-potential A'^ν that satisfies the Lorenz gauge and the wave equation (4.46) but exhibits a nonvanishing scalar potential $\Phi' \not\equiv 0$, we can perform the gauge transformation (2.30) $A'_\nu \mapsto A_\nu = A'_\nu + \partial_\nu \chi$ with the Lorentz scalar field

$$\chi(\vec{r}, t) = -c\int_0^t \Phi'(\vec{r}, t')dt' + \frac{1}{4\pi c}\int d^3r' \frac{1}{|\vec{r} - \vec{r}'|}\frac{\partial \Phi'}{\partial t}(\vec{r}', 0)\,. \tag{4.47}$$

This yields $\Phi(\vec{r}, t) = 0$ and

$$\vec{\nabla} \cdot \vec{A}(\vec{r}, t) = 0 \tag{4.48}$$

for all $(\vec{r}, t) \in \mathbb{R}^4$ for the new scalar and vector potentials, respectively. Hence, the wave equation (4.46) has to be considered for the vector potential $\vec{A}$ only, *i.e.*, we have to solve

$$\left(\frac{1}{c^2}\frac{\partial^2}{\partial t^2} - \Delta\right)\vec{A}(\vec{r}, t) = 0\,. \tag{4.49}$$

Following the lines of reasoning developed within Section 4.1, the general solution of Eq. (4.49) is written as

$$\vec{A}(\vec{r},t) = \frac{1}{\sqrt{2\pi}^3}\int d^3k \left(\vec{A}_{\vec{k}} e^{i(\vec{k}\cdot\vec{r}-\omega_k t)} + \vec{A}^*_{\vec{k}} e^{-i(\vec{k}\cdot\vec{r}-\omega_k t)}\right) \tag{4.50}$$

with $\omega_k = c|\vec{k}|$, provided we can assume that $\vec{A}$ is integrable (which implies $|\vec{A}(\vec{r},t)| \to 0$ for $|\vec{r}| \to \infty$). The gauge condition (4.48) implies that the complex amplitudes $\vec{A}_{\vec{k}}$ are perpendicular to the associated wave vectors, *i.e.*, they satisfy $\vec{k}\cdot\vec{A}_{\vec{k}} = 0$ for all $\vec{k} \in \mathbb{R}^3$. This latter property can be incorporated by defining for each wave vector $\vec{k}$ an orthonormal basis of properly oriented *polarization vectors* $\vec{e}_1(\vec{k}), \vec{e}_2(\vec{k}), \vec{e}_3(\vec{k})$ satisfying the relations $\vec{e}_1(\vec{k}) \times \vec{e}_2(\vec{k}) = \vec{e}_3(\vec{k})$, $\vec{e}_2(\vec{k}) \times \vec{e}_3(\vec{k}) = \vec{e}_1(\vec{k})$, and $\vec{e}_3(\vec{k}) \times \vec{e}_1(\vec{k}) = \vec{e}_2(\vec{k})$, where we choose $\vec{e}_3(\vec{k}) = \vec{e}_{\vec{k}} \equiv \vec{k}/|\vec{k}|$ to be the unit vector in the direction of $\vec{k}$. We can then express

$$\vec{A}_{\vec{k}} = \alpha_{\vec{k}1}\vec{e}_1(\vec{k}) + \alpha_{\vec{k}2}\vec{e}_2(\vec{k}) \tag{4.51}$$

for some complex amplitudes $\alpha_{\vec{k}1}, \alpha_{\vec{k}2} \in \mathbb{C}$ and thereby obtain the expression

$$\vec{A}(\vec{r},t) = \frac{1}{\sqrt{2\pi}^3}\int d^3k \sum_{\sigma=1,2}\left(\alpha_{\vec{k}\sigma} e^{i(\vec{k}\cdot\vec{r}-\omega_k t)} + \alpha^*_{\vec{k}\sigma} e^{-i(\vec{k}\cdot\vec{r}-\omega_k t)}\right)\vec{e}_\sigma(\vec{k}) \tag{4.52}$$

for the electromagnetic vector potential in vacuum. Note that the precise expressions for $\vec{e}_1(\vec{k})$ and $\vec{e}_2(\vec{k})$ are not important for the following. If needed, they could, *e.g.*, be defined through the prescriptions $\vec{e}_1(\vec{k}) = \pm\vec{e}_1$ if $\vec{e}_{\vec{k}} = \pm\vec{e}_3$ and $\vec{e}_1(\vec{k}) = \vec{e}_{\vec{k}} \times \vec{e}_3/|\vec{e}_{\vec{k}} \times \vec{e}_3|$ otherwise, where $\vec{e}_1, \vec{e}_2, \vec{e}_3$ denote the original unit vectors of the spatial coordinate system.

Following Eqs. (2.28) and (2.29) for the special case $\Phi \equiv 0$, the expressions for the associated electric and magnetic fields are obtained through

$$\vec{E}(\vec{r},t) = -\frac{1}{c}\frac{\partial}{\partial t}\vec{A}(\vec{r},t)\,, \tag{4.53}$$

$$\vec{B}(\vec{r},t) = \vec{\nabla}\times\vec{A}(\vec{r},t)\,. \tag{4.54}$$

Using the expression (2.41) for the energy density and combining it with Eqs. (4.53) and (4.54), the total energy contained within the electromagnetic field is then calculated as

$$H = \int d^3r \frac{1}{8\pi}\left[\vec{E}^2(\vec{r},t) + \vec{B}^2(\vec{r},t)\right] \tag{4.55}$$

$$= \frac{1}{8\pi}\int d^3r \left\{\left[\frac{1}{c}\frac{\partial}{\partial t}\vec{A}(\vec{r},t)\right]^2 + \left[\vec{\nabla}\times\vec{A}(\vec{r},t)\right]^2\right\} \tag{4.56}$$

$$= \frac{1}{2\pi}\int d^3k \left(|\alpha_{\vec{k}1}|^2 + |\alpha_{\vec{k}2}|^2\right)k^2\,, \tag{4.57}$$

where for the last calculation step we insert the explicit expression (4.52) for the vector potential. Similarly, we obtain

$$\vec{P} = \int d^3r \frac{1}{4\pi c} \vec{E}(\vec{r},t) \times \vec{B}(\vec{r},t) = \frac{1}{2\pi c} \int d^3k \left(|\alpha_{\vec{k}1}|^2 + |\alpha_{\vec{k}2}|^2 \right) k^2 \vec{e}_{\vec{k}} \tag{4.58}$$

for the total momentum contained within the field, using the expression (2.43) for the momentum density.

The expression (4.56) for the total energy can be interpreted as classical Hamiltonian functional governing the evolution of electromagnetic waves. Indeed, by introducing the associated (pseudo-)"momentum potential"

$$\vec{\Pi}(\vec{r},t) = \tilde{m} \frac{\partial}{\partial t} \vec{A}(\vec{r},t) \tag{4.59}$$

and the effective (pseudo-)"mass parameter"

$$\tilde{m} = \frac{1}{4\pi c^2}\,, \tag{4.60}$$

we can rewrite Eq. (4.56) (using an integration by parts) as

$$H = \int d^3r \left\{ \frac{1}{2\tilde{m}} \left[\vec{\Pi}(\vec{r},t) \right]^2 - \frac{1}{2} \tilde{m} c^2 \vec{A}(\vec{r},t) \cdot \Delta \vec{A}(\vec{r},t) \right\} \tag{4.61}$$

and thereby establish a perfect analogy with the general expression (4.8) for a Hamiltonian that describes (scalar) classical waves.

This analogy can now be exploited in order to quantize the electromagnetic field using the Heisenberg picture. Following the reasoning developed in Section 4.2, the real-valued vector fields $\vec{A}$ and $\vec{\Pi}$ become hermitian operators $\hat{\vec{A}}$ and $\hat{\vec{\Pi}}$ within a quantum theory of electromagnetic radiation, which satisfy the commutation relations

$$\left[\hat{A}_l(\vec{r},t), \hat{\Pi}_{l'}(\vec{r}',t) \right] = i\hbar \delta_{ll'} \delta(\vec{r} - \vec{r}') \tag{4.62}$$

as well as

$$\left[\hat{A}_l(\vec{r},t), \hat{A}_{l'}(\vec{r}',t) \right] = 0 = \left[\hat{\Pi}_l(\vec{r},t), \hat{\Pi}_{l'}(\vec{r}',t) \right] \tag{4.63}$$

for all $l, l' \in \{1,2,3\}$, all $\vec{r}, \vec{r}' \in \mathbb{R}^3$, and all $t \in \mathbb{R}$. This correspondingly implies that the complex amplitudes $\alpha_{\vec{k}\sigma}$ and $\alpha^*_{\vec{k}\sigma}$ arising within the expression (4.52) for the vector potential respectively have to be substituted by operators $\hat{\alpha}_{\vec{k}\sigma}$ and $\hat{\alpha}^\dagger_{\vec{k}\sigma}$ that are hermitian conjugates of each other. To establish again the connection with the quantum harmonic oscillator, we replace them by ladder operators $\hat{a}_{\vec{k}\sigma}$ and $\hat{a}^\dagger_{\vec{k}\sigma}$ that are defined according to

$$\hat{a}_{\vec{k}\sigma}(t) = \sqrt{\frac{2\tilde{m}\omega_k}{\hbar}} \hat{\alpha}_{\vec{k}\sigma} e^{-i\omega_k t} \tag{4.64}$$

in analogy with Eq. (4.31). It is then straightforward to show that these ladder operators satisfy the commutation relations

$$\left[\hat{a}_{\vec{k}\sigma}(t), \hat{a}^\dagger_{\vec{k}'\sigma'}(t)\right] = \delta(\vec{k} - \vec{k}')\delta_{\sigma\sigma'} \tag{4.65}$$

as well as

$$\left[\hat{a}_{\vec{k}\sigma}(t), \hat{a}_{\vec{k}'\sigma'}(t)\right] = 0 = \left[\hat{a}^\dagger_{\vec{k}\sigma}(t), \hat{a}^\dagger_{\vec{k}'\sigma'}(t)\right] \tag{4.66}$$

for all $\vec{k}, \vec{k}' \in \mathbb{R}^3$, all $\sigma, \sigma' \in \{1, 2\}$, and all $t \in \mathbb{R}$.

Expressed in terms of these ladder operators, the quantum Hamiltonian of the electromagnetic field in vacuum is then evaluated as

$$\begin{aligned} \hat{H} &= \frac{1}{2}\tilde{m}c^2 \int d^3r \left[\hat{\vec{E}}^2(\vec{r}, t) + \hat{\vec{B}}^2(\vec{r}, t)\right] && (4.67) \\ &= \int d^3k \sum_{\sigma=1,2} \frac{\hbar\omega_k}{2} \left[\hat{a}_{\vec{k}\sigma}(t)\hat{a}^\dagger_{\vec{k}\sigma}(t) + \hat{a}^\dagger_{\vec{k}\sigma}(t)\hat{a}_{\vec{k}\sigma}(t)\right], && (4.68) \end{aligned}$$

which corresponds again to a combination of infinitely many (uncoupled) quantum harmonic oscillators. The associated excitations are identified with quantum particles of bosonic nature which are termed *photons* in the electromagnetic context. $\hat{a}^\dagger_{\vec{k}\sigma}$ and $\hat{a}_{\vec{k}\sigma}$ are then respectively identifed with the *creation* and *annihilation* operators of a photon within the mode $\exp(i\vec{k} \cdot \vec{r})\vec{e}_\sigma(\vec{k})$. As the application of those creation and annihilation operators to a given quantum state of the electromagnetic field will respectively increase or decrease the total energy contained within that mode by $\hbar\omega_k$, we state that the corresponding photon possesses the energy $\hbar\omega_k$.

Let us finally point out that photons do not only carry energy but also momentum. Indeed, the quantum analog for the expression (4.58) of the total momentum contained within the electromagnetic field is evaluated as

$$\begin{aligned} \hat{\vec{P}} &= \tilde{m}c \int d^3r \hat{\vec{E}}(\vec{r}, t) \times \hat{\vec{B}}(\vec{r}, t) && (4.69) \\ &= \int d^3k \sum_{\sigma=1,2} \frac{\hbar}{2}\vec{k} \left[\hat{a}_{\vec{k}\sigma}(t)\hat{a}^\dagger_{\vec{k}\sigma}(t) + \hat{a}^\dagger_{\vec{k}\sigma}(t)\hat{a}_{\vec{k}\sigma}(t)\right], && (4.70) \end{aligned}$$

which implies that the creation of an excitation within the mode $\exp(i\vec{k} \cdot \vec{r})\vec{e}_\sigma(\vec{k})$ increases the expectation value of the field momentum by $\hbar\vec{k}$. This suggests that the associated photon possesses the momentum $\hbar\vec{k}$.

Problem

4.1 Show the validity of Eq. (4.34) using the definition (4.22) of the quantum Hamiltonian in combination with the expressions (4.29) and (4.30) for the field operators $\hat{\phi}(\vec{r}, t)$ and $\hat{\Pi}(\vec{r}, t)$.

4.2 Show that Eq. (4.48) is satisfied after the gauge transformation defined by Eq. (4.47).

4.3 Show how Eq. (4.68) can be calculated from Eq. (4.67).

Chapter 5

The Klein-Gordon theory

5.1 The Klein-Gordon equation

As we pointed out in Section 3.2, the Schrödinger equation

$$i\hbar\frac{\partial}{\partial t}\psi(\vec{r},t) = -\frac{\hbar^2}{2m}\Delta\psi(\vec{r},t)\,, \tag{5.1}$$

which describes a free particle with mass m in the absence of any external potential, is not in agreement with the principle of relativity as it corresponds to the nonrelativistic approximation $E = p^2/(2m)$ of the particle's energy-momentum relation. The correct expression for the latter is given by Eq. (3.29), namely $E = [(mc^2)^2 + c^2p^2]^{1/2}$, from which follows the relation

$$E^2 = (mc^2)^2 + c^2p^2\,. \tag{5.2}$$

In accordance with the association $p^\nu \longleftrightarrow \partial^\nu$ between the four-momentum $(p^\nu) \equiv (E/c, \vec{p})$ and $(\partial^\nu) \equiv (\frac{1}{c}\frac{\partial}{\partial t}, \vec{\nabla})$ that was discussed at the beginning of Section 3.3, we can obtain a wave equation from Eq. (5.2) through the replacement $E \longrightarrow i\hbar\partial/\partial t$ and $\vec{p} \longrightarrow -i\hbar\partial/\partial\vec{r}$. This wave equation will be considered as relativistic generalization for the Schrödinger equation in this chapter.

We therefore propose to describe a relativistic quantum particle with mass m by a real- or complex-valued scalar field φ that satisfies the equation

$$\left(\frac{1}{c^2}\frac{\partial^2}{\partial t^2} - \Delta + \frac{1}{\lambda^2}\right)\varphi(\vec{r},t) = 0\,, \tag{5.3}$$

where we introduce by

$$\lambda = \frac{\hbar}{mc} \tag{5.4}$$

the *Compton wavelength* of the particle. Using covariant notation, Eq. (5.3) can be rewritten as

$$\left(\partial_\nu\partial^\nu + \frac{1}{\lambda^2}\right)\varphi(x) = 0\,. \tag{5.5}$$

It is clearly in agreement with the principle of relativity as Eq. (5.5) obviously represents a Lorentz scalar.

Equation (5.3) is named after the physicists Oskar Klein and Walter Gordon who proposed it in 1926. It is said that Erwin Schrödinger, too, had seriously considered this equation in 1925 for describing the electron within the hydrogen atom. He finally discarded it, as it didn't seem to correctly reproduce the fine structure of the hydrogen spectrum, and went on instead, in the beginning of 1926, with the publication of what is now known as Schrödinger's equation.

5.2 Conservation of charge

It would be most natural to proceed as in the case of the Schrödinger equation and interpret the field φ as wavefunction of the particle under consideration, meaning that the modulus square of $\varphi(\vec{r}, t)$ should describe the probability that at time t the particle is found at the position $\vec{r}$. It is, however, impossible to derive a continuity equation of the form (3.23) from the Klein-Gordon equation (5.3), simply because the latter represents not a first- but a second-order differential equation in time, in contrast to the Schrödinger equation. We therefore discard this possibility and construct instead a four-current

$$j^\nu(x) = \frac{i\hbar}{2m}\left[\varphi^*(x)\partial^\nu\varphi(x) - \varphi(x)\partial^\nu\varphi^*(x)\right] . \tag{5.6}$$

through the relativistic generalization of the expression (3.22) for the probability flux of a nonrelativistic quantum particle. As is straightforwardly verified, this four-current satisfies the continuity equation $\partial_\nu j^\nu(x) = 0$ provided φ is a solution of the Klein-Gordon equation (5.5).

In nonrelativistic terms, we can express $(j^\nu) = (c\rho, \vec{j})$ with

$$\rho(\vec{r}, t) = \frac{i\hbar}{2mc^2}\left[\varphi^*(\vec{r}, t)\frac{\partial}{\partial t}\varphi(\vec{r}, t) - \varphi(\vec{r}, t)\frac{\partial}{\partial t}\varphi^*(\vec{r}, t)\right] , \tag{5.7}$$

$$\vec{j}(\vec{r}, t) = \frac{\hbar}{2im}\left[\varphi^*(\vec{r}, t)\vec{\nabla}\varphi(\vec{r}, t) - \varphi(\vec{r}, t)\vec{\nabla}\varphi^*(\vec{r}, t)\right] . \tag{5.8}$$

While Eq. (5.8) exactly corresponds to the analogous expression (3.22) for the probability flux in the nonrelativistic quantum theory (which actually motivates the choice of the global prefactor within Eq. (5.6)), the expression (5.7) poses a conceptual problem insofar as it would not necessarily yield a positive definite density. Indeed, it can be easily verified within Eq. (5.6) that j^ν changes sign under complex conjugation of the field φ. Hence, if we happen to find a solution φ of the Klein-Gordon equation (5.3) whose associated density according to Eq. (5.7) satisfies $\rho(\vec{r}, t) > 0$ for all $(\vec{r}, t) \in \mathbb{R}^4$, complex conjugation of φ would yield another solution of Eq. (5.3) whose density would be negative definite within $\mathbb{R}^4$.

We therefore give up the attempt to somehow distill from the Klein-Gordon field φ a conservation law for a positive definite probability to find a particle somewhere in space. Instead, we interpret $(c\rho, \vec{j})$ as *electric* four-current of the particle under consideration, in the sense that $\rho(\vec{r}, t)$ describes the density of the electric charge at position $\vec{r}$ and time t, and $\vec{j}(\vec{r}, t)$ represents the associated current density. As we shall work out in more detail later on within this chapter, a generic complex-valued field φ can then be decomposed into two components $\varphi^{(+)}$ and $\varphi^{(-)}$ that respectively have a purely positive and negative charge density. They can be associated with the *particle* and *antiparticle* components of a given particle species, *e.g.* π^+ and π^- in the case of the π meson. In the case of a purely real-valued field $\varphi \in \mathbb{R}$, on the other hand, we obtain a vanishing four-current $j^\nu \equiv 0$. Consequently, such a field could describe an electrically neutral particle, such as π^0 in the case of the π meson.

5.3 Energy-momentum tensor

As for the theory of electrodynamics (see Section 2.4), the conservation of energy and momentum within the framework of the Klein-Gordon theory has to be described by an energy-momentum tensor. Indeed, as we already pointed out in Section 2.1, energy and momentum constitute a four-vector that has identical transformation properties under Lorentz transformations as the space-time vector of an event. The conservation of any individual component of this four-vector, on the other hand, is governed by another Lorentz vector, namely the four-current consisting of the spatial density and the flux that are associated with the conserved quantity.

The precise form of the energy-momentum tensor can be derived from the Lagrangian formulation of the Klein-Gordon theory, namely by using the Emmy-Noether theorem which relates continuous symmetries that arise within a field theory with conserved quantities. The conservation of energy and momentum is generally obtained in systems that exhibit translational invariance in time and space, respectively. This holds obviously for the Klein-Gordon equation (5.3) which is invariant under displacements of the origin in the Minkowski space.

Specifically, the energy-momentum tensor of the Klein-Gordon theory can be written as

$$\begin{aligned} T^{\mu\nu}(x) &= \frac{\hbar^2}{2m}\Big\{[\partial^\nu\varphi^*(x)]\,[\partial^\mu\varphi(x)] + [\partial^\mu\varphi^*(x)]\,[\partial^\nu\varphi(x)] \\ &\qquad -g^{\mu\nu}\,[\partial^\alpha\varphi^*(x)]\,[\partial_\alpha\varphi(x)] + \frac{1}{\lambda^2}g^{\mu\nu}\varphi^*(x)\varphi(x)\Big\}\,. \end{aligned} \tag{5.9}$$

It is obviously symmetric, *i.e.* $T^{\mu\nu}(x) = T^{\nu\mu}(x)$ for all $x \in \mathbb{R}^4$, and satisfies the continuity equation $\partial_\nu T^{\mu\nu}(x) = 0$ provided φ solves the Klein-Gordon equation

(5.5). The energy density of the Klein-Gordon field is then obtained as

$$u(\vec{r},t) = T^{00}(\vec{r},t) = \frac{\hbar^2}{2m}\left(\left|\frac{1}{c}\frac{\partial\varphi}{\partial t}(\vec{r},t)\right|^2 + \left|\vec{\nabla}\varphi(\vec{r},t)\right|^2 + \left|\frac{1}{\lambda}\varphi(\vec{r},t)\right|^2\right). \tag{5.10}$$

This then yields the total energy

$$H = \frac{\hbar^2}{2m}\int d^3r\left(\left|\frac{1}{c}\frac{\partial\varphi}{\partial t}(\vec{r},t)\right|^2 + \left|\vec{\nabla}\varphi(\vec{r},t)\right|^2 + \left|\frac{1}{\lambda}\varphi(\vec{r},t)\right|^2\right). \tag{5.11}$$

Equation (5.11) is identical with the Hamiltonian functional (4.15) describing a complex-valued wave, except for the last term $\propto |\varphi(\vec{r},t)|^2$. The latter could be interpreted as some sort of eigenfrequency contribution of the effective carrier medium that transports the wave, if such an aether interpretation would make sense in the framework of the Klein-Gordon theory.

The momentum density of the Klein-Gordon field is evaluated as

$$\begin{aligned}\vec{p}(\vec{r},t) &= \left(\frac{1}{c}T^{0i}(\vec{r},t)\right) & (5.12)\\ &= -\frac{\hbar^2}{2mc^2}\left(\frac{\partial\varphi^*}{\partial t}(\vec{r},t)\vec{\nabla}\varphi(\vec{r},t) + \frac{\partial\varphi}{\partial t}(\vec{r},t)\vec{\nabla}\varphi^*(\vec{r},t)\right). & (5.13)\end{aligned}$$

As in electrodynamics, it is identical to the flux of energy $\vec{S}(\vec{r},t) = (cT^{i0}(\vec{r},t)) = c^2\vec{p}(\vec{r},t)$ up to a constant prefactor c^2.

5.4 Coupling to an electromagnetic field

As in the case of nonrelativistic quantum mechanics, the interaction of a Klein-Gordon particle with an electromagnetic field is incorporated through the replacement

$$\partial_\nu \mapsto \partial_\nu + \frac{iq}{\hbar c}A_\nu(x)\,, \tag{5.14}$$

which is referred to as *minimal coupling*. Here, q is the electric charge of the particle and $(A^\nu) = (\Phi, \vec{A})$ represents the four-potential that describes the electromagnetic field. Carrying out this replacement (5.14) within the Klein-Gordon equation (5.5) yields the modified equation

$$\left(\partial_\nu + \frac{iq}{\hbar c}A_\nu(x)\right)\left(\partial^\nu + \frac{iq}{\hbar c}A^\nu(x)\right)\varphi(x) + \frac{1}{\lambda^2}\varphi(x) = 0 \tag{5.15}$$

which is rewritten in nonrelativistic terms as

$$\left[\frac{1}{c^2}\left(\frac{\partial}{\partial t} + \frac{iq}{\hbar}\Phi(\vec{r},t)\right)^2 - \left(\vec{\nabla} - \frac{iq}{\hbar c}\vec{A}(\vec{r},t)\right)^2 + \frac{1}{\lambda^2}\right]\varphi(\vec{r},t) = 0\,. \tag{5.16}$$

Complex conjugation of Eq. (5.15) yields

$$\left(\partial_\nu - \frac{iq}{\hbar c}A_\nu(x)\right)\left(\partial^\nu - \frac{iq}{\hbar c}A^\nu(x)\right)\varphi^*(x) + \frac{1}{\lambda^2}\varphi^*(x) = 0\,, \tag{5.17}$$

which essentially corresponds to a Klein-Gordon equation for a particle that has the opposite charge $-q$. This supports our interpretation that the four-current j^ν defined through Eq. (5.6), which changes sign under complex conjugation of φ, describes the conservation of charge.

As in the case of the Schrödinger equation, the invariance of the Klein-Gordon equation (5.15) under gauge transformations $A_\nu \mapsto A'_\nu$ with

$$A'_\nu(x) = A_\nu(x) + \partial_\nu\chi(x) \tag{5.18}$$

for some (twice continuously differentiable) scalar field χ is assured by performing the associated gauge transformation $\varphi \mapsto \varphi'$ of the Klein-Gordon field with

$$\varphi'(\vec{r},t) = \varphi(\vec{r},t)\exp\left[-\frac{iq}{\hbar c}\chi(\vec{r},t)\right]\,. \tag{5.19}$$

It can then be shown that φ' satisfies the Klein-Gordon equation

$$\left(\partial_\nu + \frac{iq}{\hbar c}A'_\nu(x)\right)\left(\partial^\nu + \frac{iq}{\hbar c}A'^\nu(x)\right)\varphi(x) + \frac{1}{\lambda^2}\varphi'(x) = 0 \tag{5.20}$$

defined with the transformed four-potential A'^ν.

As the minimal coupling (5.14) arises quite frequently in relativistic quantum mechanics, it makes sense to define a modified derivative operator according to

$$D_\nu \equiv \partial_\nu + \frac{iq}{\hbar c}A_\nu(x)\,, \tag{5.21}$$

which is named *covariant derivative*. The Klein-Gordon equation (5.15) in the presence of an electromagnetic field can then be rewritten in a compact manner as

$$\left(D_\nu D^\nu + \frac{1}{\lambda^2}\right)\varphi(x) = 0\,. \tag{5.22}$$

We finally note that the definitions for conserved quantities, as given by the expressions (5.6) and (5.9) for the four-current and the energy-momentum tensor, respectively, are also modified in the presence of a coupling to an electromagnetic field. This is very similar to nonrelativistic quantum mechanics where we already pointed out the necessity to distinguish between a canonical and a kinetic momentum. In practice, this modification can be accomplished through the application of the minimal coupling procedure (5.14), which amounts to replacing derivative operators ∂_ν with the covariant derivative (5.21). This yields the modified four-current

$$J^\nu(x) = \frac{i\hbar}{2m}\left[\varphi^*(x)D^\nu\varphi(x) - \varphi(x)(D^\nu\varphi)^*(x)\right]\,. \tag{5.23}$$

satisfying $\partial_\nu J^\nu(x) = 0$, as well as the modified energy-momentum tensor

$$\begin{aligned}\mathcal{T}^{\mu\nu}(x) &= \frac{\hbar^2}{2m}\Big\{[(D^\nu\varphi)^*(x)]\,[D^\mu\varphi(x)] + [(D^\mu\varphi)^*(x)]\,[D^\nu\varphi(x)] \\ &\quad -g^{\mu\nu}\,[(D^\alpha\varphi)^*(x)]\,[D_\alpha\varphi(x)] + \frac{1}{\lambda^2}g^{\mu\nu}\varphi^*(x)\varphi(x)\Big\}\,. \end{aligned} \tag{5.24}$$

Quite logically, this modified energy-momentum tensor does no longer feature the conservation of energy and momentum of the Klein-Gordon field as those quantities are now exchanged with the electromagnetic field. Indeed, provided φ solves the modified Klein-Gordon equation (5.15), one can straightforwardly derive from Eq. (5.24)

$$\partial_\nu \mathcal{T}^{\mu\nu}(x) = \frac{q}{c}F^{\mu\nu}(x)J_\nu(x)\,, \tag{5.25}$$

where we use the expression (5.23) for the modified four-current and the relation (2.25) for the electromagnetic field tensor. Note that the right-hand side of this equation, which consequently represents the density of sources or sinks of energy and momentum for the Klein-Gordon field, is exactly counterbalanced by the source term appearing on the right-hand side of the analogous energy-momentum balance equation (2.45) for the electromagnetic field. This expresses the fact that the total energy and momentum contained within both the Klein-Gordon field and the electromagnetic field are well conserved.

5.5 Plane waves

In the absence of electromagnetic fields, the Klein-Gordon equation (5.3) can be solved by applying the Fourier transformation

$$\varphi(\vec{r},t) = \frac{1}{\sqrt{2\pi}^3}\int \tilde{\varphi}(\vec{k},t)e^{i\vec{k}\cdot\vec{r}}d^3k\,, \tag{5.26}$$

which can be performed provided the Klein-Gordon field is integrable. This yields the ordinary differential equation

$$\left(\frac{1}{\tilde{c}^2}\frac{\partial^2}{\partial t^2} + \vec{k}^2 + \frac{1}{\lambda^2}\right)\tilde{\varphi}(\vec{k},t) = 0 \tag{5.27}$$

whose general solution can be written as

$$\tilde{\varphi}(\vec{k},t) = \alpha^{(+)}_{\vec{k}}e^{-i\omega_k t} + \alpha^{(-)}_{-\vec{k}}e^{i\omega_k t} \tag{5.28}$$

for some complex coefficients $\alpha^{(\pm)}_{\pm\vec{k}}$, where we define

$$\omega_k = c\sqrt{\vec{k}^2 + 1/\lambda^2}. \tag{5.29}$$

The general solution of the Klein-Gordon equation is then obtained through Eq. (5.26) yielding

$$\varphi(\vec{r},t) = \frac{1}{\sqrt{2\pi}^3}\int d^3k\left(\alpha_{\vec{k}}^{(+)}e^{i(\vec{k}\cdot\vec{r}-\omega_k t)} + \alpha_{\vec{k}}^{(-)}e^{-i(\vec{k}\cdot\vec{r}-\omega_k t)}\right) \tag{5.30}$$

$$= \frac{1}{\sqrt{2\pi}^3}\int d^3k\left(\varphi_{\vec{k}}^{(+)}(\vec{r},t) + \varphi_{\vec{k}}^{(-)}(\vec{r},t)\right)\,, \tag{5.31}$$

where we define by

$$\varphi_{\vec{k}}^{(\pm)}(\vec{r},t) = \alpha_{\vec{k}}^{(\pm)}e^{\pm i(\vec{k}\cdot\vec{r}-\omega_k t)} = \alpha_{\vec{k}}^{(\pm)}e^{\mp ikx} \tag{5.32}$$

the plane monochromatic waves that are associated with the *wave four-vector* $k \equiv (k^\nu) = (\omega_k/c, \vec{k})$.

It is instructive to evaluate the charge density associated with the wave $\varphi_{\vec{k}}^{(\pm)}$. We obtain according to Eq. (5.7)

$$\rho_{\vec{k}}^{(\pm)} = \frac{\hbar}{mc^2}\mathrm{Im}\left[\varphi_{\vec{k}}^{(\pm)}(\vec{r},t)\left(\frac{\partial}{\partial t}\varphi_{\vec{k}}^{(\pm)}(\vec{r},t)\right)^*\right] = \pm\frac{\hbar\omega_k}{mc^2}\left|\alpha_{\vec{k}}^{(\pm)}\right|^2\,. \tag{5.33}$$

Hence, a plane wave that exhibits the temporal oscillation behaviour $e^{-i\omega_k t}$ can be associated with a particle that has positive charge q, while a plane wave that oscillates according to $e^{i\omega_k t}$ would describe the corresponding antiparticle with negative charge $-q$. This interpretation is confirmed by the expression for the total charge that is contained within the general solution (5.30) of the Klein-Gordon equation, which is evaluated as

$$Q = \frac{\hbar}{mc^2}\int d^3r\mathrm{Im}\left[\varphi(\vec{r},t)\frac{\partial}{\partial t}\varphi^*(\vec{r},t)\right] = \int d^3k\frac{\hbar\omega_k}{mc}\left(\left|\alpha_{\vec{k}}^{(+)}\right|^2 - \left|\alpha_{\vec{k}}^{(-)}\right|^2\right)\,. \tag{5.34}$$

In the case of a purely real Klein-Gordon field with $\varphi(\vec{r},t) = \varphi^*(\vec{r},t)$ for all $\vec{r}$ and t, we have to impose the condition $\alpha_{\vec{k}}^{(+)} = (\alpha_{\vec{k}}^{(-)})^*$ within the expression (5.30) for the general solution. This naturally yields $Q = 0$ for the total charge, in accordance with the interpretation that a real-valued Klein-Gordon field describes an electrically neutral particle.

5.6 Quantization

Our discussion of the Klein-Gordon equation and its implications has revealed so far that it essentially corresponds to a classical field theory. It is therefore subject to quantization, as are other field theories such as electrodynamics. Quantizing the Klein-Gordon field according to the scheme that was presented in Chapter 4 will yield particles (and antiparticles) as quantum excitations of the effective harmonic oscillator modes that correspond to the plane waves (5.32), in close analogy with photons in Maxwell's theory.

Let us first discuss the conceptually simpler case of an electrically neutral particle species, which is described by a purely real-valued Klein-Gordon field. In perfect analogy with the quantization of waves discussed in Section 4.2, we employ the Heisenberg picture and replace the Klein-Gordon field by a hermitian field operator:

$$\varphi(\vec{r},t) \mapsto \hat{\varphi}(\vec{r},t) = \hat{\varphi}^{\dagger}(\vec{r},t)\,. \tag{5.35}$$

The associated "momentum" field operator is obtained through the replacement

$$\tilde{m}\frac{\partial}{\partial t}\varphi(\vec{r},t) \mapsto \hat{\Pi}(\vec{r},t) = \hat{\Pi}^{\dagger}(\vec{r},t)\,. \tag{5.36}$$

Here we define

$$\tilde{m} = \frac{\hbar^2}{mc^2} \tag{5.37}$$

such that the expression (5.11) for the total energy corresponds to the Hamiltonian functional that generates the Klein-Gordon equation through Hamilton's equations of motion, in close analogy with Eqs. (4.16) and (4.17). We then impose the commutation rules

$$\left[\hat{\varphi}(\vec{r},t), \hat{\Pi}(\vec{r}',t)\right] = i\hbar\delta(\vec{r}-\vec{r}') \tag{5.38}$$

as well as

$$[\hat{\varphi}(\vec{r},t), \hat{\varphi}(\vec{r}',t)] = 0 = \left[\hat{\Pi}(\vec{r},t), \hat{\Pi}(\vec{r}',t)\right] \tag{5.39}$$

for all $\vec{r}, \vec{r}' \in \mathbb{R}^3$ and all $t \in \mathbb{R}$.

The general solution of the Klein-Gordon equation that governs the time evolution of the field operator can be written in the form (5.30) where we replace the amplitudes by operators and take into account that $\hat{\varphi}(\vec{r},t)$ is a hermitian operator. As in Section 4.2, we perform the substitution

$$\alpha^{(+)}_{\vec{k}} \quad\mapsto\quad \sqrt{\frac{\hbar}{2\tilde{m}\omega_k}}\hat{a}_{\vec{k}}\,, \tag{5.40}$$

$$\alpha^{(-)}_{\vec{k}} \quad\mapsto\quad \sqrt{\frac{\hbar}{2\tilde{m}\omega_k}}\hat{a}^{\dagger}_{\vec{k}}\,, \tag{5.41}$$

where ω_k is defined through Eq. (5.29). This yields ladder operators that satisfy the commutation rules

$$\left[\hat{a}_{\vec{k}}(t), \hat{a}^{\dagger}_{\vec{k}'}(t)\right] = \delta(\vec{k}-\vec{k}') \tag{5.42}$$

as well as

$$\left[\hat{a}_{\vec{k}}(t), \hat{a}_{\vec{k}'}(t)\right] = 0 = \left[\hat{a}^{\dagger}_{\vec{k}}(t), \hat{a}^{\dagger}_{\vec{k}'}(t)\right] \tag{5.43}$$

for all $\vec{k}, \vec{k}' \in \mathbb{R}^3$.

The quantum operators that are associated with the total energy and the total momentum of the Klein-Gordon field are given by

$$\begin{aligned}
\hat{H} &= \frac{\hbar^2}{2m}\int d^3r\left[\frac{1}{c^2}\left(\frac{\partial}{\partial t}\hat{\varphi}^\dagger(\vec{r},t)\right)\left(\frac{\partial}{\partial t}\hat{\varphi}(\vec{r},t)\right)+\left(\vec{\nabla}\hat{\varphi}^\dagger(\vec{r},t)\right)\cdot\left(\vec{\nabla}\hat{\varphi}(\vec{r},t)\right)\right. \\
&\quad \left.+\frac{1}{\lambda^2}\hat{\varphi}^\dagger(\vec{r},t)\hat{\varphi}(\vec{r},t)\right], \qquad (5.44) \\
\hat{\vec{P}} &= -\frac{\hbar^2}{2mc^2}\int d^3r\left[\left(\frac{\partial}{\partial t}\hat{\varphi}^\dagger(\vec{r},t)\right)\left(\vec{\nabla}\hat{\varphi}(\vec{r},t)\right)\right. \\
&\quad \left.+\left(\vec{\nabla}\hat{\varphi}^\dagger(\vec{r},t)\right)\left(\frac{\partial}{\partial t}\hat{\varphi}(\vec{r},t)\right)\right], \qquad (5.45)
\end{aligned}$$

as can be obtained from a direct quantization of the expressions (5.10) and (5.13) for the classical energy and momentum density, respectively. Expressing them in terms of the above ladder operators yields

$$\begin{aligned}
\hat{H} &= \int\frac{\hbar\omega_k}{2}\left(\hat{a}_{\vec{k}}\hat{a}^\dagger_{\vec{k}}+\hat{a}^\dagger_{\vec{k}}\hat{a}_{\vec{k}}\right)d^3k\,, \qquad (5.46) \\
\hat{\vec{P}} &= \int\frac{\hbar}{2}\vec{k}\left(\hat{a}_{\vec{k}}\hat{a}^\dagger_{\vec{k}}+\hat{a}^\dagger_{\vec{k}}\hat{a}_{\vec{k}}\right)d^3k \qquad (5.47)
\end{aligned}$$

for the quantum Hamiltonian and the momentum operator of an electrically neutral particle species. We can therefore infer that $\hat{a}^\dagger_{\vec{k}}$ represents the creation operator of a particle that travels with momentum $\hbar\vec{k}$ and has the energy $\hbar\omega_k = [(mc^2)^2+\hbar^2c^2\vec{k}^2]^{1/2}$, while $\hat{a}_{\vec{k}}$ is the corresponding annihilation operator.

The case of an electrically charged particle species is more involved insofar as it deals with a complex-valued Klein-Gordon field on a classical level. Its quantization consequently yields a nonhermitian field operator

$$\varphi(\vec{r},t)\mapsto\hat{\varphi}(\vec{r},t)\not\equiv\hat{\varphi}^\dagger(\vec{r},t)\,. \qquad (5.48)$$

The commutation rules with the associated "momentum" field operator, which is nonhermitian as well and can be obtained through

$$\tilde{m}\frac{\partial}{\partial t}\varphi(\vec{r},t)\mapsto\hat{\Pi}(\vec{r},t)\not\equiv\hat{\Pi}^\dagger(\vec{r},t)\,, \qquad (5.49)$$

are determined in close analogy with the Hamiltonian formalism for a classical complex-valued field as was discussed at the end of Section 4.1: we formally treat $\hat{\varphi}(\vec{r},t)$ and $\hat{\varphi}^\dagger(\vec{r},t)$ as independent field operators and consider $\hat{\Pi}^\dagger(\vec{r},t)$ to be the conjugate counterpart of $\hat{\varphi}(\vec{r},t)$ while $\hat{\Pi}(\vec{r},t)$ would be the counterpart of $\hat{\varphi}^\dagger(\vec{r},t)$. This yields the commutation relations

$$\left[\hat{\varphi}(\vec{r},t),\hat{\Pi}^\dagger(\vec{r}',t)\right]=i\hbar\delta(\vec{r}-\vec{r}')=\left[\hat{\varphi}^\dagger(\vec{r},t),\hat{\Pi}(\vec{r}',t)\right] \qquad (5.50)$$

as well as

$$\begin{aligned}
[\hat{\varphi}(\vec{r},t),\hat{\varphi}(\vec{r}',t)] &= 0 = \left[\hat{\varphi}^\dagger(\vec{r},t),\hat{\varphi}^\dagger(\vec{r}',t)\right], &(5.51)\\
\left[\hat{\Pi}(\vec{r},t),\hat{\Pi}(\vec{r}',t)\right] &= 0 = \left[\hat{\Pi}^\dagger(\vec{r},t),\hat{\Pi}^\dagger(\vec{r}',t)\right], &(5.52)\\
\left[\hat{\varphi}(\vec{r},t),\hat{\varphi}^\dagger(\vec{r}',t)\right] &= 0 = \left[\hat{\Pi}(\vec{r},t),\hat{\Pi}^\dagger(\vec{r}',t)\right], &(5.53)\\
\left[\hat{\varphi}(\vec{r},t),\hat{\Pi}(\vec{r}',t)\right] &= 0 = \left[\hat{\varphi}^\dagger(\vec{r},t),\hat{\Pi}^\dagger(\vec{r}',t)\right] &(5.54)
\end{aligned}$$

for all $\vec{r},\vec{r}' \in \mathbb{R}^3$ and all $t \in \mathbb{R}$. In close analogy with the classical Hamiltonian functional (4.15) of a complex-valued field, we now have to set

$$\tilde{m} = \frac{\hbar^2}{2mc^2} \tag{5.55}$$

in order to ensure that the Heisenberg equations derived from the Hamiltonian (5.44) in combination with the above commutation rules yield the correct time evolution for the field operators $\hat{\varphi}(\vec{r},t)$ and $\hat{\Pi}(\vec{r},t)$.

The general solution of the Klein-Gordon equation for $\hat{\varphi}(\vec{r},t)$ yields again the expression (5.30) where we have to replace the amplitudes $\alpha_{\vec{k}}^{(\pm)}$ by operators. In analogy with Eqs. (5.40) and (5.41), we perform the substitution

$$\begin{aligned}
\alpha_{\vec{k}}^{(+)} &\mapsto \sqrt{\frac{\hbar}{2\tilde{m}\omega_k}}\hat{a}_{\vec{k}}\,, &(5.56)\\
\alpha_{\vec{k}}^{(-)} &\mapsto \sqrt{\frac{\hbar}{2\tilde{m}\omega_k}}\hat{b}^\dagger_{\vec{k}}\,, &(5.57)
\end{aligned}$$

where contrary to the case of an electrically neutral particle species we cannot apply the identification $\hat{a}_{\vec{k}} \equiv \hat{b}_{\vec{k}}$ here. We thereby obtain two sets of ladder operators $\hat{a}_{\vec{k}}, \hat{a}^\dagger_{\vec{k}}$ and $\hat{b}_{\vec{k}}, \hat{b}^\dagger_{\vec{k}}$ that satisfy the commutation relations

$$\left[\hat{a}_{\vec{k}},\hat{a}^\dagger_{\vec{k}'}\right] = \delta(\vec{k}-\vec{k}') = \left[\hat{b}_{\vec{k}},\hat{b}^\dagger_{\vec{k}'}\right] \tag{5.58}$$

as well as

$$\begin{aligned}
\left[\hat{a}_{\vec{k}},\hat{a}_{\vec{k}'}\right] &= 0 = \left[\hat{a}^\dagger_{\vec{k}},\hat{a}^\dagger_{\vec{k}'}\right], &(5.59)\\
\left[\hat{b}_{\vec{k}},\hat{b}_{\vec{k}'}\right] &= 0 = \left[\hat{b}^\dagger_{\vec{k}},\hat{b}^\dagger_{\vec{k}'}\right], &(5.60)\\
\left[\hat{a}_{\vec{k}},\hat{b}_{\vec{k}'}\right] &= 0 = \left[\hat{a}^\dagger_{\vec{k}},\hat{b}^\dagger_{\vec{k}'}\right] &(5.61)\\
\left[\hat{a}_{\vec{k}},\hat{b}^\dagger_{\vec{k}'}\right] &= 0 = \left[\hat{a}^\dagger_{\vec{k}},\hat{b}_{\vec{k}'}\right] &(5.62)
\end{aligned}$$

for all $\vec{k}, \vec{k}' \in \mathbb{R}^3$. Expressed in terms of these ladder operators, the quantum operators for the total energy and the total momentum are evaluated as

$$\hat{H} = \int \hbar\omega_k \left(\hat{a}^\dagger_{\vec{k}}\hat{a}_{\vec{k}} + \hat{b}_{\vec{k}}\hat{b}^\dagger_{\vec{k}}\right) d^3k\,, \tag{5.63}$$

$$\hat{\vec{P}} = \int \hbar\vec{k} \left(\hat{a}^\dagger_{\vec{k}}\hat{a}_{\vec{k}} + \hat{b}_{\vec{k}}\hat{b}^\dagger_{\vec{k}}\right) d^3k\,. \tag{5.64}$$

The total charge contained within the Klein-Gordon field corresponds to an operator as well, which can be obtained from the direct quantization of the expression (5.34). This yields the charge operator

$$\begin{aligned}\hat{Q} &= \int d^3r \frac{i\hbar}{2mc^2}\left[\hat{\varphi}^\dagger(\vec{r},t)\frac{\partial}{\partial t}\hat{\varphi}(\vec{r},t) - \left(\frac{\partial}{\partial t}\hat{\varphi}^\dagger(\vec{r},t)\right)\hat{\varphi}(\vec{r},t)\right] \\ &= \int \left(\hat{a}^\dagger_{\vec{k}}\hat{a}_{\vec{k}} - \hat{b}_{\vec{k}}\hat{b}^\dagger_{\vec{k}}\right) d^3k \end{aligned} \tag{5.65}$$

in terms of the above ladder operators. From the expressions (5.63 – 5.65) we infer that $\hat{a}^\dagger_{\vec{k}}$ and $\hat{a}_{\vec{k}}$ respectively represent the creation and annihilation operators of a particle with positive charge — or, more precisely, with charge q — that travels with momentum $\hbar\vec{k}$ and has the energy $\hbar\omega_k$, while $\hat{b}^\dagger_{\vec{k}}$ and $\hat{b}_{\vec{k}}$ are the analogous creation and annihilation operators for the corresponding antiparticle with charge $-q$. Note that the latter appear in antinormal order within the above expressions for the total energy, momentum, and charge operators, with $\hat{b}^\dagger_{\vec{k}}$ being executed before $\hat{b}_{\vec{k}}$, while the particle creation and annihilation operators $\hat{a}^\dagger_{\vec{k}}, \hat{a}_{\vec{k}}$ are normally ordered within Eqs. (5.63 – 5.65). Enforcing a normal order also for the antiparticle operators $\hat{b}^\dagger_{\vec{k}}, \hat{b}_{\vec{k}}$ will give rise to infinite constants within the expressions (5.63), (5.65) for the total energy and charge, which one can associate with the presence of an infinitely extended background "sea" of negatively charged antiparticles within the universe.

In practice, these constants can be eliminated by a simple redefinition of the zero levels of energy and charge. They do not have any impact on the time evolution of the system. Indeed, the Heisenberg equations governing the evolution of the creation and annihilation operators associated with Klein-Gordon particles and antiparticles are derived from Eq. (5.63) as

$$\frac{d}{dt}\hat{a}_{\vec{k}}(t) = \frac{i}{\hbar}\left[\hat{H}, \hat{a}_{\vec{k}}(t)\right] = -i\omega_k\hat{a}_{\vec{k}}(t)\,, \tag{5.66}$$

$$\frac{d}{dt}\hat{a}^\dagger_{\vec{k}}(t) = \frac{i}{\hbar}\left[\hat{H}, \hat{a}^\dagger_{\vec{k}}(t)\right] = i\omega_k\hat{a}^\dagger_{\vec{k}}(t)\,, \tag{5.67}$$

$$\frac{d}{dt}\hat{b}_{\vec{k}}(t) = \frac{i}{\hbar}\left[\hat{H}, \hat{b}_{\vec{k}}(t)\right] = -i\omega_k\hat{b}_{\vec{k}}(t)\,, \tag{5.68}$$

$$\frac{d}{dt}\hat{b}^\dagger_{\vec{k}}(t) = \frac{i}{\hbar}\left[\hat{H}, \hat{b}^\dagger_{\vec{k}}(t)\right] = i\omega_k\hat{b}^\dagger_{\vec{k}}(t) \tag{5.69}$$

using the commutation relations $[\hat{b}_{\vec{k}'}\hat{b}^\dagger_{\vec{k}'}, \hat{b}_{\vec{k}}] = [\hat{b}^\dagger_{\vec{k}'}\hat{b}_{\vec{k}'}, \hat{b}_{\vec{k}}] = -\delta(\vec{k}-\vec{k}')\hat{b}_{\vec{k}}$ and $[\hat{b}_{\vec{k}'}\hat{b}^\dagger_{\vec{k}'}, \hat{b}^\dagger_{\vec{k}}] = [\hat{b}^\dagger_{\vec{k}'}\hat{b}_{\vec{k}'}, \hat{b}^\dagger_{\vec{k}}] = \delta(\vec{k}-\vec{k}')\hat{b}^\dagger_{\vec{k}}$. Equations (5.66 – 5.69) are solved as

$$\hat{a}_{\vec{k}}(t) = e^{-i\omega_k t}\hat{a}_{\vec{k}}(0)\,, \tag{5.70}$$

$$\hat{a}^\dagger_{\vec{k}}(t) = e^{i\omega_k t}\hat{a}^\dagger_{\vec{k}}(0)\,, \tag{5.71}$$

$$\hat{b}_{\vec{k}}(t) = e^{-i\omega_k t}\hat{b}_{\vec{k}}(0)\,, \tag{5.72}$$

$$\hat{b}^\dagger_{\vec{k}}(t) = e^{i\omega_k t}\hat{b}^\dagger_{\vec{k}}(0)\,, \tag{5.73}$$

respectively, which is in perfect consistency with the expression (5.30) for the general solution of the Klein-Gordon equation.

5.7 The nonrelativistic limit

As it is supposed to represent a relativistic generalization of standard quantum mechanics, the Klein-Gordon theory has to somehow yield the Schrödinger equation (3.25) in the formal nonrelativistic limit $c \to \infty$. To evidence this claim, we first recall that the latter represents a first-order differential equation in time, whereas the time evolution of the Klein-Gordon equation (5.3) is governed by a second time derivative alike Newton's second law (3.4). It appears therefore appropriate to decompose Eq. (5.3) into a set of two first-order differential equations in time. We could in principle use the Hamiltonian formalism for this purpose, which would yield two first-order equations for the Klein-Gordon field and its conjugate "momentum" counterpart in close analogy with Eqs. (4.9) and (4.10). However, it turns out to be more convenient to operate with complex linear combinations of those two fields, which then can be directly associated with the two wavefunctions that respectively describe the particle and antiparticle components contained within the Klein-Gordon field.

Let us specifically consider the Klein-Gordon equation (5.15) describing a charged particle in the presence of an electromagnetic field, which can be written as

$$\left(D_\nu D^\nu + \frac{1}{\lambda^2}\right)\varphi(x) = \left(D_0^2 - \vec{D}^2 + \frac{1}{\lambda^2}\right)\varphi(x) = 0 \tag{5.74}$$

in a covariant manner. Here, $(D_\nu) = (D_0, \vec{D})$ represents the four-vector of the covariant derivative (5.21) with the temporal and spatial components

$$D_0 = \frac{1}{c}\frac{\partial}{\partial t} + \frac{iq}{\hbar c}\Phi(\vec{r},t)\,, \tag{5.75}$$

$$\vec{D} = \vec{\nabla} - \frac{iq}{\hbar c}\vec{A}(\vec{r},t)\,, \tag{5.76}$$

where $(A^\nu) = (\Phi, \vec{A})$ is the four-potential that describes the electromagnetic field.

We now define two complex-valued fields

$$\varphi_\pm(x) = \frac{1}{2}\left[\varphi(x) \pm i\lambda D_0\varphi(x)\right] \tag{5.77}$$

such that we can decompose

$$\begin{aligned} \varphi(x) &= \varphi_+(x) + \varphi_-(x)\,, &(5.78)\\ i\lambda D_0\varphi(x) &= \varphi_+(x) - \varphi_-(x)\,. &(5.79)\end{aligned}$$

It is straightforward to show that the temporal covariant derivative of $\varphi_\pm$ satisfies

$$\pm i\lambda D_0\varphi_\pm(x) = \varphi_\pm(x) - \frac{\lambda^2}{2}\vec{D}^2\left[\varphi_+(x) + \varphi_-(x)\right] \tag{5.80}$$

provided φ is a solution of Eq. (5.74). This yields a set of two first-order differential equations in time which are perfectly equivalent to the Klein-Gordon equation (5.74). Using $\lambda = \hbar/(mc)$, they can be expressed in nonrelativistic terms as

$$\begin{aligned} i\hbar\frac{\partial}{\partial t}\varphi_+(\vec{r},t) &= \left[mc^2 + q\Phi(\vec{r},t)\right]\varphi_+(\vec{r},t) + \frac{\vec{\pi}^2}{2m}\left[\varphi_+(\vec{r},t) + \varphi_-(\vec{r},t)\right], &(5.81)\\ i\hbar\frac{\partial}{\partial t}\varphi_-^*(\vec{r},t) &= \left[mc^2 - q\Phi(\vec{r},t)\right]\varphi_-^*(\vec{r},t) + \frac{\vec{\pi}^{*2}}{2m}\left[\varphi_+^*(\vec{r},t) + \varphi_-^*(\vec{r},t)\right] &(5.82)\end{aligned}$$

where we define the kinetic momentum operator and its complex conjugate as

$$\begin{aligned} \vec{\pi} &= \frac{\hbar}{i}\vec{\nabla} - \frac{q}{c}\vec{A}(\vec{r},t)\,, &(5.83)\\ \vec{\pi}^* &= -\left(\frac{\hbar}{i}\vec{\nabla} + \frac{q}{c}\vec{A}(\vec{r},t)\right)\,, &(5.84)\end{aligned}$$

in perfect analogy with the corresponding classical expression (3.35).

Formally, Eqs. (5.81) and (5.82) form a set of two Schrödinger-like equations that are coupled to each other. Indeed, we recognize within these equations the same kinetic and potential energy terms that arise also within the nonrelativistic Schrödinger equation (3.40) describing a particle with charge $\pm q$ in the presence of an electromagnetic field. The rest energy $E_0 = mc^2$ of the particle under consideration appears as additional contribution to the total energy within Eqs. (5.81) and (5.82), which appears perfectly logical from a relativistic point of view. As Eq. (5.82) can effectively be obtained from Eq. (5.81) by inverting the sign of the charge q, it makes sense to consider φ_+ to be the wavefunction associated with a particle of charge q, while φ_-^* would be the wavefunction of the corresponding antiparticle with charge $-q$. This interpretation is indeed confirmed by evaluating the charge density according to Eq. (5.23). We obtain

$$\rho(x) = \frac{i\hbar}{2mc}\left[\varphi^*(x)D_0\varphi(x) - \varphi(x)(D_0\varphi)^*(x)\right] = |\varphi_+(x)|^2 - |\varphi_-(x)|^2\,, \tag{5.85}$$

which implies that φ_+ and φ_-^* respectively contribute purely positively and purely negatively to the total charge.

Let us now specifically elaborate the Schrödinger equation that describes the time evolution of the wavefunction associated with a Klein-Gordon particle of charge q in the nonrelativistic limit. To this end, we first eliminate the presence of the rest energy mc^2 within Eq. (5.81) by introducing rescaled wavefunctions according to

$$\psi_\pm(\vec{r},t) = \varphi_\pm(\vec{r},t)e^{imc^2t/\hbar}\,. \tag{5.86}$$

This essentially amounts to redefining the zero level of energy, namely such that a particle with charge q at rest has no energy. The rescaled wavefunctions $\psi_\pm$ then evolve according to

$$\begin{aligned} i\hbar\frac{\partial}{\partial t}\psi_+(\vec{r},t) &= \left(\frac{\vec{\pi}^2}{2m} + q\Phi(\vec{r},t)\right)\psi_+(\vec{r},t) + \frac{\vec{\pi}^2}{2m}\psi_-(\vec{r},t)\,, & (5.87)\\ -i\hbar\frac{\partial}{\partial t}\psi_-(\vec{r},t) &= \left(\frac{\vec{\pi}^2}{2m} - q\Phi(\vec{r},t)\right)\psi_-(\vec{r},t) + \frac{\vec{\pi}^2}{2m}\psi_+(\vec{r},t) \\ &\quad +2mc^2\psi_-(\vec{r},t)\,. & (5.88) \end{aligned}$$

Equation (5.88) can be rewritten as

$$\psi_-(\vec{r},t) = -\frac{\vec{\pi}^2}{4m^2c^2}\psi_+(\vec{r},t) + \frac{1}{2mc^2}\left(-i\hbar\frac{\partial}{\partial t} - \frac{\vec{\pi}^2}{2m} - q\Phi(\vec{r},t)\right)\psi_-(\vec{r},t) \tag{5.89}$$

and is formally solved by recursively replacing $\psi_-(\vec{r},t)$ on the right-hand side of Eq. (5.89) with the entire self-consistent expression (5.89) for $\psi_-(\vec{r},t)$. This yields the series

$$\begin{aligned} \psi_-(\vec{r},t) &= -\frac{\vec{\pi}^2}{4m^2c^2}\psi_+(\vec{r},t) \\ &\quad +\frac{1}{2mc^2}\left(-i\hbar\frac{\partial}{\partial t} - \frac{\vec{\pi}^2}{2m} - q\Phi(\vec{r},t)\right)\left(-\frac{\vec{\pi}^2}{4m^2c^2}\psi_+(\vec{r},t)\right) \\ &\quad +\left[\frac{1}{2mc^2}\left(-i\hbar\frac{\partial}{\partial t} - \frac{\vec{\pi}^2}{2m} - q\Phi(\vec{r},t)\right)\right]^2\left(-\frac{\vec{\pi}^2}{4m^2c^2}\psi_+(\vec{r},t)\right) \\ &\quad +\ldots\,, \end{aligned} \tag{5.90}$$

where the terms in the first, second, and third line of Eq. (5.90) scale as $1/c^2$, $1/c^4$, and $1/c^6$, respectively, provided $\psi_+(\vec{r},t)$ does not exhibit any particular scaling with c (which holds, *e.g.*, if we start the time evolution with a vanishing antiparticle component, *i.e.*, we have $\psi_-(\vec{r},t_0) = 0$ for all $\vec{r}$ at the initial time t_0). Inserting this expression (5.90) into Eq. (5.87) finally yields

$$i\hbar\frac{\partial}{\partial t}\psi_+(\vec{r},t) = \left(\frac{\vec{\pi}^2}{2m} - \frac{\vec{\pi}^4}{8m^3c^2} + q\Phi(\vec{r},t)\right)\psi_+(\vec{r},t) + \mathcal{O}\left(1/c^4\right) \tag{5.91}$$

up to corrections that scale as $1/c^4$. The wavefunction ψ_+ therefore evolves in the nonrelativistic limit $c \to \infty$ according to the Schrödinger equation (3.40) that describes a quantum particle with charge q in the presence of an electromagnetic field. The additional term $\propto \vec{\pi}^4$ arising within Eq. (5.91) corresponds to a relativistic correction to the kinetic energy, which is perfectly consistent with the second-order term in the nonrelativistic expansion (3.29) of the energy-momentum relation. As ψ_- scales at most as $1/c^2$ according to Eq. (5.90), the charge density (5.85) is given by

$$\rho(\vec{r},t) = |\psi_+(\vec{r},t)|^2 + \mathcal{O}\left(1/c^4\right) \tag{5.92}$$

and can therefore be identified with the probability density that results from the wavefunction ψ_+, up to corrections of the order of $1/c^4$.

The Schrödinger equation that describes the time evolution of the wavefunction φ_-^* associated with the corresponding antiparticle can be obtained in a perfectly analogous manner. We start by performing the rescaling procedure

$$\tilde{\psi}_\pm^*(\vec{r},t) = \varphi_\pm^*(\vec{r},t) e^{imc^2 t/\hbar}\,, \tag{5.93}$$

which effectively amounts to redefining the zero level of energy such that an antiparticle with charge $-q$ at rest has no energy. Equations (5.81) and (5.82) are then rewritten as

$$\begin{aligned} i\hbar\frac{\partial}{\partial t}\tilde{\psi}_-^*(\vec{r},t) &= \left(\frac{\vec{\pi}^{*2}}{2m} - q\Phi(\vec{r},t)\right)\tilde{\psi}_-^*(\vec{r},t) + \frac{\vec{\pi}^{*2}}{2m}\tilde{\psi}_+^*(\vec{r},t)\,, &(5.94)\\ -i\hbar\frac{\partial}{\partial t}\tilde{\psi}_+^*(\vec{r},t) &= \left(\frac{\vec{\pi}^{*2}}{2m} + q\Phi(\vec{r},t)\right)\tilde{\psi}_+^*(\vec{r},t) + \frac{\vec{\pi}^{*2}}{2m}\tilde{\psi}_-^*(\vec{r},t)\\ &\quad +2mc^2\tilde{\psi}_+^*(\vec{r},t)\,. &(5.95)\end{aligned}$$

in terms of these rescaled wavefunctions $\tilde{\psi}_\pm^*$. Equation (5.95) is formally solved as

$$\tilde{\psi}_+^*(\vec{r},t) = -\frac{\vec{\pi}^{*2}}{4m^2c^2}\tilde{\psi}_-^*(\vec{r},t) + \mathcal{O}\left(1/c^4\right) \tag{5.96}$$

up to corrections that scale as $1/c^4$, provided $\tilde{\psi}_-^*$ does not exhibit any particular scaling with c. Inserting this latter expression (5.96) into Eq. (5.94) yields the modified Schrödinger equation

$$i\hbar\frac{\partial}{\partial t}\tilde{\psi}_-^*(\vec{r},t) = \left(\frac{\vec{\pi}^{*2}}{2m} - \frac{\vec{\pi}^{*4}}{8m^3c^2} - q\Phi(\vec{r},t)\right)\tilde{\psi}_-^*(\vec{r},t) + \mathcal{O}\left(1/c^4\right)\,, \tag{5.97}$$

which can formally be obtained from Eq. (5.91) by inverting the sign of the charge. Correspondingly, the charge density (5.85) is now given by

$$\rho(\vec{r},t) = -\left|\tilde{\psi}_-^*(\vec{r},t)\right|^2 + \mathcal{O}\left(1/c^4\right)\,, \tag{5.98}$$

which is consistent with the assertion that $|\tilde{\psi}_-^*(\vec{r},t)|^2$ represents the probability density of the antiparticle.

Chapter 6

The Dirac theory

6.1 Construction of the theory

Despite the fact that it contains the Schrödinger equation in the nonrelativistic limit, the Klein-Gordon theory falls short of describing an electron. Indeed, it can be evaluated that the Klein-Gordon equation does not correctly reproduce the relativistic fine-structure corrections to the spectrum of the hydrogen atom, which can be attributed to the fact that the electron spin and its interaction with a magnetic field do not emerge from this theory. From a more conceptual point of view, a major drawback of the Klein-Gordon equation is that it fails to deliver a conservation law for a probability as requested by the concept of a wavefunction, since it represents a second-order differential equation in time which, as we discussed in Section 5.2, does not allow for deriving a continuity equation that involves a positive definite density. Such a conservation law can be approximately established in the nonrelativistic regime [see Eq. (5.92)] but is no intrinsic property of the theory.

To remedy this shortcoming, Paul Dirac had the ingenious idea to seek a *square root* of the Klein-Gordon operator $\partial_\nu \partial^\nu + \lambda^{-2}$, which by construction would yield a first-order differential equation in time. While this would be impossible to achieve in the framework of a commutative algebra, such a square root can indeed be defined using non-commuting matrices instead of ordinary real or complex numbers as prefactors of the involved derivative operators. This implies that the wavefunction ψ of the relativistic particle would have to exhibit several components, *i.e.* $\psi \equiv (\psi_1, \ldots, \psi_N)^T$ with $N > 1$, which would be coupled with each other in the course of the time evolution that is generated by the resulting differential equation.

In practice, this amounts to stating that the wavefunction evolves according to a multi-component linear differential equation which contains at most a first-order derivative in time. Three basic requirements are formulated in order to construct that equation and to determine the dimension N of the vector space in which the wavefunction is defined, namely:

(a) Relativistic covariance must be respected by this equation.

(b) Since ψ is supposed to be a wavefunction in the sense of Schrödinger, a conservation law for the total probability must follow from this equation, where the probability density is given by

$$\rho(\vec{r},t) = \sum_{s=1}^{N} |\psi_s(\vec{r},t)|^2 \,. \tag{6.1}$$

(c) Each component ψ_s of the wavefunction must fulfill the Klein-Gordon equation (5.3), *i.e.*

$$\left(\frac{1}{c^2}\frac{\partial^2}{\partial t^2} - \Delta + \frac{1}{\lambda^2}\right)\psi_s(\vec{r},t) = 0 \tag{6.2}$$

with $\lambda = \hbar/(mc)$.

It straightforwardly follows from the requirement (a) that the equation to be constructed contains at most first-order derivatives with respect to the spatial coordinates. Indeed, if we had a second-order spatial derivative operator in this equation, a Lorentz boost to a moving reference frame would then give rise to second-order temporal derivatives as well, which would then imply that the equation would not have the same form in every inertial frame, in contradiction to the basic principle of relativity. Consequently, we can explicitly write this differntial equation through the most general ansatz

$$\frac{1}{c}\frac{\partial}{\partial t}\psi_s(\vec{r},t) + \sum_{l=1}^{3}\sum_{s'=1}^{N}\alpha^l_{ss'}\frac{\partial}{\partial r_l}\psi_{s'}(\vec{r},t) + \sum_{s'=1}^{N}\Gamma_{ss'}\psi_{s'}(\vec{r},t) = 0 \tag{6.3}$$

with some constants $\alpha^l_{ss'}, \Gamma_{ss'} \in \mathbb{C}$ that are yet to be determined, or equivalently

$$i\hbar\frac{\partial}{\partial t}\psi_s(\vec{r},t) = \frac{\hbar c}{i}\sum_{l=1}^{3}\sum_{s'=1}^{N}\alpha^l_{ss'}\frac{\partial}{\partial r_l}\psi_{s'}(\vec{r},t) + E_0\sum_{s'=1}^{N}\beta_{ss'}\psi_{s'}(\vec{r},t) \tag{6.4}$$

where we identify $\hbar c\Gamma_{ss'} = iE_0\beta_{ss'}$ with E_0 being a real-valued energy. The coefficients $\beta_{ss'}$ are then dimensionless, as are the coefficients $\alpha^l_{ss'}$. We can group them in the $N\times N$ matrices $\alpha^l \equiv (\alpha^l_{ss'})_{N\times N}$ and $\beta = (\beta_{ss'})_{N\times N}$. Defining by $\vec{\alpha} \equiv (\alpha^1,\alpha^2,\alpha^3)$ an object that can be colloquially interpreted as a spatial vector of $N\times N$ matrices, even though it does not undergo any transformation under rotations or mirror operations of the spatial coordinate system, we can rewrite Eq. (6.4) in a more compact manner as

$$i\hbar\frac{\partial}{\partial t}\psi(\vec{r},t) = c\vec{\alpha}\cdot\hat{\vec{p}}\psi(\vec{r},t) + E_0\beta\psi(\vec{r},t) \tag{6.5}$$

with $\hat{p} = -i\hbar\vec{\nabla}$ the momentum operator and $\vec{\alpha}\cdot\hat{\vec{p}} \equiv \sum_{l=1}^{3}\alpha^l \hat{p}_l$.

In view of the requirement (b), we now determine the conditions under which a continuity equation

$$\frac{\partial}{\partial t}\rho(\vec{r},t) + \vec{\nabla}\cdot\vec{j}(\vec{r},t) = 0 \tag{6.6}$$

describing the conservation of the total probability can be derived from Eq. (6.4), with the probability density being given by the expression (6.1). Deriving this latter expression with respect to time yields

$$\begin{aligned}\frac{\partial}{\partial t}\rho(\vec{r},t) &= \sum_{s=1}^{N}\left[\psi_s^*(\vec{r},t)\frac{\partial}{\partial t}\psi_s(\vec{r},t) + \psi_s(\vec{r},t)\frac{\partial}{\partial t}\psi_s^*(\vec{r},t)\right] \\ &= -c\sum_{l=1}^{3}\sum_{s,s'=1}^{N}\left[\psi_s^*(\vec{r},t)\alpha_{ss'}^l\frac{\partial}{\partial r_l}\psi_{s'}(\vec{r},t) + \psi_s(\vec{r},t)(\alpha_{ss'}^l)^*\frac{\partial}{\partial r_l}\psi_{s'}^*(\vec{r},t)\right] \\ &\quad -\frac{i}{\hbar}E_0\sum_{s,s'=1}^{N}\left[\psi_s^*(\vec{r},t)\beta_{ss'}\psi_{s'}(\vec{r},t) - \psi_s(\vec{r},t)\beta_{ss'}^*\psi_{s'}^*(\vec{r},t)\right]\end{aligned} \tag{6.7}$$

in combination with Eq. (6.4). This equation is equivalent to the continuity equation (6.6) if and only if $\beta_{ss'} = \beta_{s's}^*$ and $\alpha_{ss'}^l = (\alpha_{s's}^l)^*$ for all $s, s' = 1, \ldots, N$ and $l = 1, 2, 3$, *i.e.* if the matrices β, α^1, α^2, and α^3 are all hermitian. The probability current density $\vec{j} = (j_1, j_2, j_3)$ appearing in Eq. (6.6) is then given by the components

$$j_l(\vec{r},t) = c\sum_{s,s'=1}^{N}\psi_s(\vec{r},t)\alpha_{ss'}^l\psi_{s'}(\vec{r},t) \tag{6.8}$$

and can be most compactly expressed as

$$\vec{j}(\vec{r},t) = c\psi^\dagger(\vec{r},t)\vec{\alpha}\psi(\vec{r},t) \tag{6.9}$$

where we introduce by $\psi^\dagger = (\psi_1^*, \ldots, \psi_N^*)$ the hermitian conjugate of the wavefunction ψ.

Finally, to verify under which conditions the requirement (c) is valid, we apply an additional time derivative to Eq. (6.4) and use this equation recursively in order to eliminate the resulting time derivatives on its right-hand side. This yields

$$\begin{aligned}\frac{1}{c^2}\frac{\partial^2}{\partial t^2}\psi_s(\vec{r},t) &= \sum_{s',s''=1}^{N}\left[\sum_{l,l'=1}^{3}\alpha_{ss'}^l\alpha_{s's''}^{l'}\frac{\partial^2}{\partial r_l\partial r_l'} - \frac{E_0^2}{\hbar^2c^2}\beta_{ss'}\beta_{s's''}\right. \\ &\quad \left. +\frac{iE_0}{\hbar c}\sum_{l=1}^{3}\left(\alpha_{ss'}^l\beta_{s's''} + \beta_{ss'}\alpha_{s's''}^l\right)\frac{\partial}{\partial r_l}\right]\psi_{s''}(\vec{r},t)\,.\end{aligned} \tag{6.10}$$

Choosing $E_0 = mc^2$ as natural energy scale for a particle with mass m, it is straightforward to figure out that this latter equation is equivalent to the Klein-Gordon equation (6.2) if and only if the matrices β and α^l for $l = 1, 2, 3$ fulfill the conditions

$$\begin{aligned}
\sum_{s'=1}^{N} \left(\alpha^l_{ss'} \beta_{s's''} + \beta_{ss'} \alpha^l_{s's''} \right) &= 0 \,, &(6.11)\\
\sum_{s'=1}^{N} \beta_{ss'} \beta_{s's''} &= \delta_{ss'} \,, &(6.12)\\
\sum_{s'=1}^{N} \left(\alpha^l_{ss'} \alpha^{l'}_{s's''} + \alpha^{l'}_{ss'} \alpha^l_{s's''} \right) &= 2\delta_{ll'} \delta_{ss'} &(6.13)
\end{aligned}$$

for all $l, l' = 1, 2, 3$ and all $s, s'' = 1, \ldots, N$, where for the last condition (6.13) we implicitly assume that the wavefunction ψ be twice continuously differentiable with respect to its spatial coordinates, such that the order in which the second derivatives with respect to r_l and $r_{l'}$ are taken in Eq. (6.10) does not matter. Equations (6.11–6.13) are rewritten in matrix notation as

$$\begin{aligned}
\alpha^l \beta + \beta \alpha^l &= 0 \,, &(6.14)\\
\beta\beta &= \mathbb{I}_{N\times N} \,, &(6.15)\\
\alpha^l \alpha^{l'} + \alpha^{l'} \alpha^l &= 2\delta_{ll'} \mathbb{I}_{N\times N} &(6.16)
\end{aligned}$$

with $\mathbb{I}_{N\times N}$ denoting the $N \times N$ unit matrix — or, more compactly,

$$\alpha^l \alpha^{l'} + \alpha^{l'} \alpha^l = 2\delta_{ll'} \mathbb{I}_{N\times N} \tag{6.17}$$

for all $l, l' = 0, 1, 2, 3$ where we formally define $\alpha^0 \equiv \beta$.

Setting $l' = l$ in Eq. (6.17) yields

$$\alpha^l \alpha^l = \mathbb{I}_{N\times N} \,, \tag{6.18}$$

which implies that the matrices α^l are square roots of the $N \times N$ unit matrix. They can therefore have the eigenvalues $+1$ and -1, which becomes obvious when representing this matrix equation in the eigenbasis of α^l. Different matrices $\alpha^l, \alpha^{l'}$ anticommute with each other, since we obtain from Eq. (6.17)

$$\alpha^l \alpha^{l'} + \alpha^{l'} \alpha^l = 0 \tag{6.19}$$

for $l \neq l'$. Multiplying this latter equation (6.19) onto $\alpha^{l'}$ and taking the trace of the resulting matrix equation, where we use the cyclic relation $\mathrm{Tr}[\alpha^{l'} \alpha^l \alpha^{l'}] = \mathrm{Tr}[\alpha^{l'} \alpha^{l'} \alpha^l]$ as well as the identity (6.18) evaluated for $l = l'$, we obtain the insight that the matrices α^l are traceless, *i.e.* $\mathrm{Tr}[\alpha^l] = 0$ for all $l = 0, 1, 2, , 3$. In combination with the fact that they are diagonalisable (as they are hermitian) and can have the eigenvalues ± 1, we can infer from this property that the dimension N of the vector space in which the wavefunction is defined must be an *even number*, such that α^l has $N/2$ eigenvalues $+1$ and $N/2$ eigenvalues -1.

To summarize, the partial differential equation (6.5) fulfils the requirements (b) and (c) if and only if the involved matrices α^l and β are hermitian, represent square roots of the unit matrix, and anticommute with each other. We note that these properties as well as the expressions (6.1) and (6.9) for the probability density and its associated flux are invariant under unitary transformations

$$\psi \mapsto U\psi\,, \tag{6.20}$$
$$\psi^\dagger \mapsto \psi^\dagger U^\dagger\,, \tag{6.21}$$
$$\alpha^l \mapsto U\alpha^l U^\dagger\,, \tag{6.22}$$
$$\beta \mapsto U\beta U^\dagger \tag{6.23}$$

for $U \in SU(N)$, satisfying $U^\dagger U = \mathbb{I}_{N\times N}$. As we show below, this intrinsic ambiguity in the choice of the matrices can be exploited in order to facilitate the task of determining a specific representation of the evolution equation (6.5).

6.2 The Majorana equation

As we inferred in the previous section, the dimension N of the vector space in which the wavefunction is defined has to be an even number. The simplest possible choice for this dimension is therefore $N = 2$, in which case we are left with the task of determining four 2×2 matrices. In view of the above unitary ambiguity expressed by the equations (6.20–6.23), we can opt for a representation in which one of those matrices, say α^3, is diagonal and reads

$$\alpha^3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{6.24}$$

Being hermitian and traceless, the other matrices are then most generally expressed as

$$\alpha^l = \begin{pmatrix} a_l & b_l \\ b_l^* & -a_l \end{pmatrix} \tag{6.25}$$

for $l < 3$, with $a_l \in \mathbb{R}$ and $b_l \in \mathbb{C}$. The anticommutation of those matrices with α^3 yields the relation

$$\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} = \alpha^l\alpha^3 + \alpha^3\alpha^l = \begin{pmatrix} 2a_l & 0 \\ 0 & 2a_l \end{pmatrix}, \tag{6.26}$$

from which we infer $a_l = 0$. Since the eigenvalues of all α^l matrices are ± 1, we have $-1 = \det \alpha^l = -|b_l|^2$, which implies $|b_l| = 1$.

We now exploit one more freedom that is still left in the choice of the unitary basis in which these matrices are represented, namely concerning the phases of the eigenvectors of α^3. Specifically, we can opt for mulitplying those eigenvectors with suitable phase factors $\exp(i\varphi_j)$ (with $j = 1, 2$) which are such that we have $b_1 = 1$, *i.e.*

$$\alpha^1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} . \tag{6.27}$$

The anticommutation of α^1 and α^2 according to Eq. (6.19) straightforwardly yields $b_2 + b_2^* = 0$, which leaves us with the possibilities $b_2 = i$ or $-i$. Choosing this second option for b_2, we finally obtain the *Pauli matrices* $\vec{\alpha} = \vec{\sigma} = (\sigma_1, \sigma_2, \sigma_3)$ defined by

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} , \qquad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} , \qquad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} , \tag{6.28}$$

which satisfy the cyclic relations $\sigma_1\sigma_2 = i\sigma_3$, $\sigma_2\sigma_3 = i\sigma_1$, and $\sigma_3\sigma_1 = i\sigma_2$.

While α^1, α^2, and α^3 are now well defined, there is no option left for choosing β such that it anticommutes with the other matrices. The choice $N = 2$ for the dimension of the vector space is therefore insufficient for representing the wavefunction. An exception arises if particle under consideration has no mass, which implies $E_0 = mc^2 = 0$. In this latter case, we obtain the *Majorana equation*

$$i\hbar\frac{\partial}{\partial t}\psi(\vec{r}, t) = c\vec{\alpha} \cdot \hat{\vec{p}}\psi(\vec{r}, t) \tag{6.29}$$

with the most general choice $\alpha^l = \pm U\sigma_l U^\dagger$ for $l = 1, 2, 3$ where $U \in SU(2)$ is a unitary matrix. For a long time, neutrinos were believed to be massless and evolve according to such a Majorana equation. This hypothesis was recently ruled out through the detection of neutrino oscillations.

6.3 The Dirac equation

As $N = 2$ does not work out for massive particles, the next possible choice for the dimension of the vector space in view of the requirement that it has to be an even number is $N = 4$, which implies that β, α^1, α^2, and α^3 are 4×4 matrices. As in the previous section, we opt for a representation in which one of the matrices, this time β, is diagonal. Recalling that it has the eigenvalues 1, 1, −1, −1, we write it in block notation as

$$\beta = \left(\begin{array}{c|c} \mathbb{I} & \mathbb{O} \\ \hline \mathbb{O} & -\mathbb{I} \end{array}\right) \tag{6.30}$$

with the 2×2 unit and zero matrices

$$\mathbb{I} \equiv \mathbb{I}_{2\times2} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} , \qquad \mathbb{O} \equiv \mathbb{O}_{2\times2} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} . \tag{6.31}$$

For the (hermitian) matrices α^l we make the general ansatz

$$\alpha^l = \left(\begin{array}{c|c} a_l & b_l \\ \hline b_l^\dagger & c_l \end{array}\right) \tag{6.32}$$

with $a_l, b_l, c_l \in \mathbb{C}^{2\times 2}$ for $l = 1, 2, 3$. The anticommutation property (6.14) yields

$$\left(\begin{array}{c|c} \mathbb{O} & \mathbb{O} \\ \hline \mathbb{O} & \mathbb{O} \end{array}\right) = \alpha^l\beta + \beta\alpha^l = \left(\begin{array}{c|c} 2a_l & \mathbb{O} \\ \hline \mathbb{O} & -2c_l \end{array}\right) , \tag{6.33}$$

from which we infer $a_l = c_l = \mathbb{O}$ for $l = 1, 2, 3$. Mutual anticommutation of the α^l matrices according to Eq. (6.16) yields the relation

$$2\delta_{ll'} \left(\begin{array}{c|c} \mathbb{I} & \mathbb{O} \\ \hline \mathbb{O} & \mathbb{I} \end{array}\right) = \alpha^l\alpha^{l'} + \alpha^{l'}\alpha^l = \left(\begin{array}{c|c} b_l b_{l'}^\dagger + b_{l'} b_l^\dagger & \mathbb{O} \\ \hline \mathbb{O} & b_l^\dagger b_{l'} + b_{l'}^\dagger b_l \end{array}\right) \tag{6.34}$$

for $l, l' = 1, 2, 3$, which implies the properties

$$b_l^\dagger b_{l'} + b_{l'}^\dagger b_l = b_l b_{l'}^\dagger + b_{l'} b_l^\dagger = 2\delta_{ll'}\mathbb{I} \tag{6.35}$$

for the b_l matrices. The latter are unitary, as is inferred from setting $l' = l$ in the above equation.

The fact that we can perform independent unitary basis transformations within the upper and lower 2×2 block without altering the form (6.30) of the matrix β gives us a lot of freedom to choose the matrices b_l. In particular, we can choose them to be the Pauli matrices (6.28), *i.e.* $b_l = \sigma_l$ for all $l = 1, 2, 3$, which obviously satisfy the relations (6.35). This gives rise to the so-called *standard* or *Dirac representation* of the Dirac equation

$$i\hbar\frac{\partial}{\partial t}\psi(\vec{r}, t) = c\vec{\alpha} \cdot \hat{\vec{p}}\psi(\vec{r}, t) + mc^2\beta\psi(\vec{r}, t) \tag{6.36}$$

with $\vec{\alpha} \equiv (\alpha^1, \alpha^2, \alpha^3)$, which is defined in terms of the matrices

$$\beta = \left(\begin{array}{c|c} \mathbb{I} & \mathbb{O} \\ \hline \mathbb{O} & -\mathbb{I} \end{array}\right) = \left(\begin{array}{cc|cc} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ \hline 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{array}\right) \tag{6.37}$$

and

$$\alpha^l = \left(\begin{array}{c|c} \mathbb{O} & \sigma_l \\ \hline \sigma_l & \mathbb{O} \end{array}\right) \tag{6.38}$$

for $l = 1, 2, 3$, *i.e.*

$$\alpha^1 = \left(\begin{array}{cc|cc} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ \hline 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{array}\right) , \ \alpha^2 = \left(\begin{array}{cc|cc} 0 & 0 & 0 & -i \\ 0 & 0 & i & 0 \\ \hline 0 & -i & 0 & 0 \\ i & 0 & 0 & 0 \end{array}\right) , \ \alpha^3 = \left(\begin{array}{cc|cc} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \\ \hline 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{array}\right) .$$

Note that all other possible choices for the matrices β, α^1, α^2, and α^3 correspond to representations that are equivalent to the above standard representation in the sense that they can be obtained from the latter via the application of unitary transformations. If we subject, for instance, the wavefunction to the transformation $\psi \mapsto U\psi$ with the unitary matrix

$$U = \frac{1}{\sqrt{2}} \left(\begin{array}{c|c} \mathbb{I} & \mathbb{I} \\ \hline \mathbb{I} & -\mathbb{I} \end{array} \right) = U^\dagger = U^{-1}\,, \tag{6.39}$$

the subsequent transformations $\beta \mapsto U\beta U^\dagger$ and $\alpha^l \mapsto U\alpha^l U^\dagger$ for $l = 1, 2, 3$ give rise to the so-called *chiral* or *Weyl representation* defined by the choice

$$\alpha^l = \left(\begin{array}{c|c} \sigma_l & \mathbb{O} \\ \hline \mathbb{O} & -\sigma_l \end{array} \right) \tag{6.40}$$

for $l = 1, 2, 3$ and

$$\beta = \left(\begin{array}{c|c} \mathbb{O} & \mathbb{I} \\ \hline \mathbb{I} & \mathbb{O} \end{array} \right) . \tag{6.41}$$

6.4 Relativistic covariance

The Dirac equation (6.36) in the above standard or Weyl representations obviously fulfils the requirements (b) and (c) that we formulated at the beginning of Section 6.1 for the differential equation to be constructed. As far as the requirement (a) is concerned, we inferred from it that this equation must contain at most first-order spatial derivatives, which is a necessary condition for relativistic covariance to hold. However, this latter property was not yet fully demonstrated, *i.e.*, we do not yet know if Eq. (6.36) has the same form in every intertial frame.

Technically, it is not very complicated to reformulate Eq. (6.36) such that it acquires a covariant appearance. We introduce to this end a set of four matrices $(\gamma^\nu) \equiv (\gamma^0, \gamma^1, \gamma^2, \gamma^3)$ which are defined as

$$\gamma^0 = \beta\,, \tag{6.42}$$

$$\gamma^l = \beta\alpha^l \tag{6.43}$$

for $l = 1, 2, 3$, such that they read

$$\gamma^0 = \left(\begin{array}{c|c} \mathbb{I} & \mathbb{O} \\ \hline \mathbb{O} & -\mathbb{I} \end{array} \right) \qquad \text{and} \qquad \gamma^l = \left(\begin{array}{c|c} \mathbb{O} & \sigma_l \\ \hline -\sigma_l & \mathbb{O} \end{array} \right) . \tag{6.44}$$

for $l = 1, 2, 3$ in the standard representation, and which form what one could colloquially call a four-vector of matrices. Multiplying the matrix β onto Eq. (6.36) and dividing it by c yields then a relativistic formulation

$$(i\hbar\gamma^\nu\partial_\nu - mc)\,\psi(x) = 0 \tag{6.45}$$

of the Dirac equation. Note, however, that this latter equation does not represent a Lorentz scalar as it looks like, simply because the "four-vector" (γ^ν) does not undergo any modification under Lorentz transformations, in agreement with the requirement that the Dirac equation ought to have the same form in every inertial frame.

Nevertheless, the matrices γ^ν play a key role in the discussion of the relativistic covariance of the Direac equation, and it is useful for this purpose to evaluate their basic commutation and anticommutation rules. Using their definitions (6.42) and (6.43), we obtain in combination with the anticommutation properties (6.17) of the matrices α^l and β the relations

$$\gamma^0\gamma^l \pm \gamma^l\gamma^0 = \alpha^l \mp \alpha^l\,, \tag{6.46}$$
$$\gamma^k\gamma^l \pm \gamma^l\gamma^k = -(\alpha^k\alpha^l \pm \alpha^l\alpha^k) \tag{6.47}$$

for all $k, l = 1, 2, 3$. This yields the anticommutation and commutation rules

$$\{\gamma^\nu, \gamma^\mu\} \equiv \gamma^\nu\gamma^\mu + \gamma^\mu\gamma^\nu = 2g^{\mu\nu}\mathbb{I}_{4\times 4}\,, \tag{6.48}$$
$$[\gamma^\nu, \gamma^\mu] \equiv \gamma^\nu\gamma^\mu - \gamma^\mu\gamma^\nu = 2i\sigma^{\nu\mu} \tag{6.49}$$

for all $\mu, \nu = 0, 1, 2, 3$, with the 4×4 matrices $\sigma^{\mu\nu} = -\sigma^{\nu\mu}$ being defined as $\sigma^{00} = \mathbb{O}_{4\times 4}$,

$$\sigma^{0l} = i\alpha^l\,, \tag{6.50}$$
$$\sigma^{kl} = -\frac{i}{2}[\alpha^k, \alpha^l] \tag{6.51}$$

for all $k, l = 1, 2, 3$. Using the well-known commutation rules $[\sigma_k, \sigma_l] = 2i\epsilon_{klm}\sigma_m$ for the Pauli matrices (6.28) with ϵ_{klm} the Levi-Civita symbol (2.9), we can explicitly evaluate those matrices in the standard representation of the Dirac equation as

$$\sigma^{0l} = i\left(\begin{array}{c|c} \mathbb{O} & \sigma_l \\ \hline \sigma_l & \mathbb{O} \end{array}\right) \qquad \text{and} \qquad \sigma^{kl} = \epsilon_{klm}\left(\begin{array}{c|c} \sigma_m & \mathbb{O} \\ \hline \mathbb{O} & \sigma_m \end{array}\right) \tag{6.52}$$

for all $k, l = 1, 2, 3$.

The problem of relativistic covariance is solved by admitting that the wave-function ψ may change under Lorentz transformations. To work this out, let us consider a general Lorentz transformation $e_\nu \mapsto e'_\nu = D_\nu{}^\mu e_\mu$ of the basis of the Minkowski space, with the transformation matrix $D = (D_\nu{}^\mu) \in G$ that satisfies $gD^TgD = 1$ or, equivalently,

$$D^\nu{}_\rho D_\mu{}^\rho = \delta^\nu_\mu\,. \tag{6.53}$$

We assume that the wavefunction correspondingly undergoes the linear transformation

$$\psi \mapsto \psi' = S\psi \tag{6.54}$$

with the invertible matrix $S \in \mathbb{C}^{4\times 4}$. Requiring relativistic covariance implies that ψ' evolves according to the transformed Dirac equation

$$(i\hbar\gamma^\nu \partial'_\nu - mc)\,\psi'(x) = 0\,, \tag{6.55}$$

with

$$\partial'_\nu = D_\nu{}^\mu \partial_\mu \tag{6.56}$$

the transformed partial derivative operators, if and only if the original wavefunction ψ satisfies Eq. (6.45). Multiplying S^{-1} onto Eq. (6.55) yields with Eqs. (6.54) and (6.56)

$$\big(i\hbar S^{-1}\gamma^\nu S D_\nu{}^\mu \partial_\mu - mc\big)\,\psi(x) = 0\,, \tag{6.57}$$

which is identical to the original Dirac equation (6.45) if and only if

$$S^{-1}\gamma^\nu S D_\nu{}^\mu = \gamma^\mu\,, \tag{6.58}$$

or equivalently, using Eq. (6.53),

$$S^{-1}\gamma^\nu S = D^\nu{}_\mu \gamma^\mu\,. \tag{6.59}$$

Let us now focus more specifically on *proper* Lorentz transformations $D \in \mathcal{L}$ which can, as was discussed in Section 1.2, be expressed as a succession of infinitesimal transformations. In view of Eq. (1.40) we can therefore write

$$D = \exp(\tau I) \tag{6.60}$$

for some real parameter $\tau \in \mathbb{R}$ and some generating matrix $I \in \mathbb{R}^{4\times 4}$. Inserting this expression into the relation (6.53), taking the derivative of the resulting equation with respect to τ, and evaluating it at $\tau = 0$ yields the relation

$$I^{\nu\mu} + I^{\mu\nu} = 0\,, \tag{6.61}$$

which implies that the generator I corresponds to an antisymmetric tensor. The corresponding ansatz

$$S = \exp(\tau A) \tag{6.62}$$

with $A \in \mathbb{C}^{4\times 4}$ can be made for the transformation matrix of the wavefunction, since we most naturally require that an infinitesimal Lorentz transformation can only result in an infinitesimal change of the wavefunction. Equation (6.59) is then rewritten as

$$e^{-\tau A}\gamma^\nu e^{\tau A} = \big(e^{\tau I}\big)^\nu{}_\mu \gamma^\mu\,. \tag{6.63}$$

Taking the derivative of this latter equation with respect to τ and evaluating it at $\tau = 0$ yields the relation

$$\gamma^\nu A - A\gamma^\nu = I^\nu{}_\mu \gamma^\mu \tag{6.64}$$

between the generators I and A.

In order to explicitly solve this equation for A, we rewrite its right-hand side as

$$\begin{aligned} I^{\nu}{}_{\mu}\gamma^{\mu} = g^{\nu\sigma}I_{\sigma\mu}\gamma^{\mu} &= \frac{1}{2}\left(g^{\nu\sigma}I_{\sigma\mu}\gamma^{\mu} + g^{\nu\mu}I_{\mu\sigma}\gamma^{\sigma}\right) \\ &= \frac{1}{2}I_{\sigma\mu}\left(g^{\nu\sigma}\gamma^{\mu} - \gamma^{\sigma}g^{\nu\mu}\right) \\ &= \frac{1}{4}I_{\sigma\mu}\left(\{\gamma^{\nu},\gamma^{\sigma}\}\gamma^{\mu} - \gamma^{\sigma}\{\gamma^{\nu},\gamma^{\mu}\}\right) \\ &= \frac{1}{4}I_{\sigma\mu}\left(\gamma^{\nu}\gamma^{\sigma}\gamma^{\mu} - \gamma^{\sigma}\gamma^{\mu}\gamma^{\nu}\right) , \end{aligned} \tag{6.65}$$

using the antisymmetry property (6.61) and the anticommutation rules (6.48). We therefore infer that we have to choose

$$A = \frac{1}{4}I_{\nu\mu}\gamma^{\nu}\gamma^{\mu} = \frac{1}{8}I_{\nu\mu}\left(\gamma^{\nu}\gamma^{\mu} - \gamma^{\mu}\gamma^{\nu}\right) = -\frac{i}{4}I_{\nu\mu}\sigma^{\nu\mu} \tag{6.66}$$

in order to satisfy the relation (6.64), where we use again Eq. (6.61) as well as the expression (6.49) for the commutation rules of the matrices γ^{ν} and γ^{μ}. The transformation matrix (6.62) is then written as

$$S = \exp\left(\frac{\tau}{4}I_{\nu\mu}\gamma^{\nu}\gamma^{\mu}\right) = \exp\left(-i\frac{\tau}{4}I_{\nu\mu}\sigma^{\nu\mu}\right) . \tag{6.67}$$

Wavefunction vectors $\psi \in \mathbb{C}^4$ that are transformed according to $\psi \mapsto \psi' = S\psi$ with the above matrix S in the presence of a proper Lorentz transformation (6.60) will be called *Dirac spinors* in the following.

Let us now explicitly evaluate this transformation for the basic generators $I = -I_l$ and $-J_l$ that correspond to rotations and boosts, respectively. According to Eq. (1.31), a rotation about the axis $l = 1, 2, 3$ of the spatial coordinate system is generated by the matrix $I = -I_l = (I^{\nu}{}_{\mu})$ with the elements

$$I^{\nu}{}_{\mu} = \epsilon_{jkl}\delta^{\nu}_{j}\delta^{k}_{\mu} , \tag{6.68}$$

where ϵ_{jkl} represents again the Levi-Civita tensor (2.9) (and where a sum is implicitly performed over the spatial indices $j, k = 1, 2, 3$). The covariant components of the associated tensor of second order are constructed as

$$I_{\nu\mu} = g_{\nu\alpha}I^{\alpha}{}_{\mu} = -\epsilon_{jkl}\delta^{j}_{\nu}\delta^{k}_{\mu} , \tag{6.69}$$

since $g_{j\alpha} = -\delta_{j\alpha}$ for $j = 1, 2, 3$. We then obtain from Eq. (6.66)

$$A = \frac{i}{4}\epsilon_{jkl}\sigma^{jk} = \frac{1}{8}\epsilon_{jkl}[\alpha^{j}, \alpha^{k}] \tag{6.70}$$

with the help of the expression (6.51) for σ^{jk}, which explicitly reads

$$A = \frac{i}{2}\left(\begin{array}{c|c} \sigma_l & \mathbb{O} \\ \hline \mathbb{O} & \sigma_l \end{array}\right) \tag{6.71}$$

in the standard representation, using $\epsilon_{jkl}\epsilon_{jkm} = 2\delta_{lm}$ for all $l, m = 1, 2, 3$ (with implicit summation over $j, k = 1, 2, 3$). The exponential of this matrix is straightforwardly calculated using the property $\sigma_l^2 = \mathbb{I}$ for all $l = 1, 2, 3$ and hence $\sigma_l^{2k} = \mathbb{I}$ and $\sigma_l^{2k+1} = \sigma_l$ for all integer $k \in \mathbb{N}_0$. We therefore obtain

$$\exp\left(i\frac{\tau}{2}\sigma_l\right) = \cos(\tau/2)\mathbb{I} + i\sin(\tau/2)\sigma_l\,, \tag{6.72}$$

from which we infer the transformation matrix

$$S = e^{\tau A} = \cos(\tau/2)\left(\begin{array}{c|c} \mathbb{I} & \mathbb{O} \\ \hline \mathbb{O} & \mathbb{I} \end{array}\right) + i\sin(\tau/2)\left(\begin{array}{c|c} \sigma_l & \mathbb{O} \\ \hline \mathbb{O} & \sigma_l \end{array}\right) \tag{6.73}$$

in the standard representation. A rotation of the coordinate system about a given axis gives then rise to separate unitary mixings of the two upper and the two lower coordinates of the Dirac spinor, generated by the Pauli matrix that corresponds to this axis. This is already a first indication that those two components are to be associated with the two *spin states* of the fermionic elementary particle under consideration.

Accroding to Eqs. (1.37) and (1.39), a boost along the axis $l = 1, 2, 3$ is generated by the matrix $I = -J_l = (I^\nu{}_\mu)$ with the elements

$$I^\nu{}_\mu = -\delta_0^\nu\delta_\mu^l - \delta_l^\nu\delta_\mu^0\,. \tag{6.74}$$

We therefore have

$$I_{\nu\mu} = g_{\nu\alpha}I^\alpha{}_\mu = \delta_\nu^l\delta_\mu^0 - \delta_\nu^0\delta_\mu^l\,, \tag{6.75}$$

since $g_{0\alpha} = \delta_{0\alpha}$ and $g_{j\alpha} = -\delta_{j\alpha}$ for $j = 1, 2, 3$. The generator of the corresponding transformation of the Dirac spinor is then given by

$$A = -\frac{i}{4}\left(\sigma^{l0} - \sigma^{0l}\right) = -\frac{1}{2}\alpha^l \tag{6.76}$$

using the expression (6.50) for σ^{0l}, which reads in the standard representation

$$A = -\frac{1}{2}\left(\begin{array}{c|c} \mathbb{O} & \sigma_l \\ \hline \sigma_l & \mathbb{O} \end{array}\right)\,. \tag{6.77}$$

Using $\alpha^l\alpha^l = \mathbb{I}_{4\times 4}$ we obtain the transformation matrix

$$S = e^{\tau A} = \cosh(\tau/2)\left(\begin{array}{c|c} \mathbb{I} & \mathbb{O} \\ \hline \mathbb{O} & \mathbb{I} \end{array}\right) - \sinh(\tau/2)\left(\begin{array}{c|c} \mathbb{O} & \sigma_l \\ \hline \sigma_l & \mathbb{O} \end{array}\right) \tag{6.78}$$

for the Dirac spinor in the standard representation. A Lorentz transformation to a moving inertial frame gives therefore rise to a non-unitary mixing of the upper and lower two-component blocks of the Dirac spinor.

Evidently, the above transformation law for Dirac spinors, as expressed by the matrix (6.67), can become rather involved for more general Lorentz transformations and lacks the transparency and simplicity of Lorentz vectors or tensors that we introduced in Section 2.1. It is to be noted, however, that the latter notions can indeed be restored in the framework of the Dirac equation, and expressed in terms of the matrices γ^ν as we were tempted to do so in the beginning of this section, namely by introducing the concept of the *adjoint spinor*. We first point out, to this end, that the commutator (6.49) of the matrices γ^ν and γ^μ satisfies the property

$$(\sigma^{\nu\mu})^\dagger \gamma^0 - \gamma^0 \sigma^{\nu\mu} = 0 \tag{6.79}$$

for all $\nu, \mu = 0, 1, 2, 3$, which can be shown from the expressions (6.50) and (6.51), the hermiticity of the matrices $\alpha^l = (\alpha^l)^\dagger$, as well as their anticommutation with the matrix $\beta = \gamma^0$. From Eq. (6.79) we infer

$$A^\dagger \gamma^0 = -\gamma^0 A \tag{6.80}$$

for the matrix (6.66) that generates the transformation S of the Dirac spinor according to Eq. (6.62), and hence

$$\exp\left(\tau A^\dagger\right)\gamma^0 = \sum_{k=0}^{\infty} \frac{\tau^k}{k!} \left(A^\dagger\right)^k \gamma^0 = \gamma^0 \sum_{k=0}^{\infty} \frac{\tau^k}{k!} (-A)^k = \gamma^0 \exp\left(-\tau A\right) . \tag{6.81}$$

The *adjoint spinor* is then defined by the complex line vector

$$\bar\psi \equiv \left(\bar\psi_1, \bar\psi_2, \bar\psi_3, \bar\psi_4\right) = \psi^\dagger \gamma^0 . \tag{6.82}$$

Owing to Eq. (6.81) it is transformed according to

$$\bar\psi \mapsto \psi^\dagger S^\dagger \gamma^0 = \psi^\dagger \gamma^0 S^{-1} = \bar\psi S^{-1} \tag{6.83}$$

under Lorentz transformations that induce the mapping $\psi \mapsto S\psi$ of the Dirac spinor. Hence, we obtain by means of Eq. (6.59) the transformation laws

$$\bar\psi \gamma^\nu \psi \mapsto \bar\psi S^{-1} \gamma^\nu S \psi = D^\nu{}_\mu \bar\psi \gamma^\mu \psi \tag{6.84}$$

for all $\nu = 0, 1, 2, 3$ as well as, with a similar reasoning,

$$\bar\psi \gamma^{\nu_1} \cdots \gamma^{\nu_N} \psi \mapsto D^{\nu_1}{}_{\mu_1} \cdots D^{\nu_N}{}_{\mu_N} \bar\psi \gamma^{\mu_1} \cdots \gamma^{\mu_N} \psi \tag{6.85}$$

for all $\nu_1, \ldots, \nu_N$. Those two expressions (6.84) and (6.85) therefore represent a Lorentz vector and a Lorentz tensor of Nth order, respectively. A Lorentz scalar is obtained by $\bar\psi\psi$ since we have

$$\bar\psi\psi \mapsto \bar\psi S^{-1} S \psi = \bar\psi\psi . \tag{6.86}$$

In the standard representation, this particularly implies that not the probability density (6.1) but rather the expression $|\psi_1(\vec r, t)|^2 + |\psi_2(\vec r, t)|^2 - |\psi_3(\vec r, t)|^2 - |\psi_4(\vec r, t)|^2$ is invariant under Lorentz transformations.

6.5 Conservation laws

Thanks to the adjoint spinor (6.82) it is now straightforward to formulate conservation laws in terms of continuity equations that involve relativistically covariant four-currents. The continuity equation (6.6) describing the conservation of the total probability can be rewritten as

$$\partial_\nu j^\nu(x) = 0 \tag{6.87}$$

in a covariant manner, where according to Eqs. (6.1) and (6.9) we have the probability density

$$\rho(\vec{r},t) = \frac{1}{c} j^0(\vec{r},t) = \psi^\dagger(\vec{r},t)\psi(\vec{r},t) = \bar{\psi}(\vec{r},t)\gamma^0\psi(\vec{r},t) \tag{6.88}$$

and the associated current density components

$$j^l(\vec{r},t) = c\psi^\dagger(\vec{r},t)\alpha^l\psi(\vec{r},t) = c\bar{\psi}(\vec{r},t)\gamma^l\psi(\vec{r},t) \tag{6.89}$$

using $\gamma^l = \gamma^0\alpha^l$. This yields the covariant expression

$$j^\nu(x) = c\bar{\psi}(x)\gamma^\nu\psi(x) \tag{6.90}$$

for the four-current describing the conservation of probability.

As usual for relativistic theories, the conservation of energy and momentum in the framework of the Dirac equation is described by an energy-momentum tensor. Its components are written as

$$T^{\nu\mu}(x) = \frac{i\hbar c}{2}\left[\bar{\psi}(x)\gamma^\mu\left(\partial^\nu\psi\right)(x) - \left(\partial^\nu\bar{\psi}\right)(x)\gamma^\mu\psi(x)\right], \tag{6.91}$$

as can be derived from a Lagrangian formulation of this theory. The energy density contained within the Dirac spinor ψ is evaluated as

$$\begin{aligned} T^{00}(\vec{r},t) &= \frac{i\hbar}{2}\left[\psi^\dagger(\vec{r},t)\frac{\partial}{\partial t}\psi(\vec{r},t) - \left(\frac{\partial}{\partial t}\psi^\dagger(\vec{r},t)\right)\psi(\vec{r},t)\right] \\ &= \frac{c}{2}\left[\psi^\dagger(\vec{r},t)\vec{\alpha}\cdot\hat{\vec{p}}\psi(\vec{r},t) + \left(\hat{\vec{p}}\psi(\vec{r},t)\right)^\dagger\cdot\vec{\alpha}\psi(\vec{r},t)\right] \\ &\quad +mc^2\psi^\dagger(\vec{r},t)\beta\psi(\vec{r},t), \end{aligned} \tag{6.92}$$

where we use the fact that ψ satisfies the Dirac equation (6.36). This expression is perfectly consistent with the Dirac Hamiltonian $\hat{H} = c\vec{\alpha}\cdot\hat{\vec{p}} + mc^2\beta$ appearing on the right-hand side of Eq. (6.36) and can be identified with the expectation value of the symmetrized (and therefore hermitian) operator $[\hat{H}\delta(\hat{\vec{r}}-\vec{r}) + \delta(\hat{\vec{r}}-\vec{r})\hat{H}]/2$ with respect to the wavefunction ψ. Similarly, the components of the momentum density are evaluated as

$$T^{l0}(\vec{r},t) = \frac{1}{2}\left[\psi^\dagger(\vec{r},t)\hat{p}_l\psi(\vec{r},t) + \left(\hat{p}_l\psi^\dagger(\vec{r},t)\right)\psi(\vec{r},t)\right], \tag{6.93}$$

corresponding to the expectation value of $[\hat{p}_l\delta(\hat{\vec{r}}-\vec{r})+\delta(\hat{\vec{r}}-\vec{r})\hat{p}_l]/2$. The continuity equations

$$\partial_\mu T^{\nu\mu}(x) = 0 \tag{6.94}$$

for $\nu = 0, 1, 2, 3$ can be straightforwardly obtained from the covariant formulation (6.45) of the Dirac equation and its counterpart for the adjoint spinor. The latter reads

$$i\hbar\partial_\nu\bar{\psi}\gamma^\nu + mc\bar{\psi}(x) = 0\,, \tag{6.95}$$

as is shown from the hermitian conjugation of the Dirac equation (6.45) in combination with the property $(\gamma^\nu)^\dagger\gamma^0 = \gamma^0\gamma^\nu$ for all $\nu = 0, 1, 2, 3$.

As it contains spatiotemporal derivatives, the expression (6.91) for the energy-momentum tensor of the Dirac theory will change if the (charged) particle described by this theory is exposed to an electromagnetic field. The latter is, as usual, represented in terms of the four-potential $(A^\nu) \equiv (\Phi, \vec{A})$ which is incorporated into the Dirac equation through the minimal coupling procedure $\partial_\nu \mapsto D_\nu$ giving rise to the covariant derivative

$$D_\nu = \partial_\nu + \frac{iq}{\hbar c}A_\nu(x)\,. \tag{6.96}$$

This yields the modified Dirac equation

$$(i\hbar\gamma^\nu D_\nu - mc)\,\psi(x) = 0\,, \tag{6.97}$$

which reads in nonrelativistic terms

$$i\hbar\frac{\partial}{\partial t}\psi(\vec{r},t) = c\vec{\alpha}\cdot\hat{\vec{\pi}}\psi(\vec{r},t) + mc^2\beta\psi(\vec{r},t) + q\Phi(\vec{r},t)\psi(\vec{r},t)\,, \tag{6.98}$$

where we use again the kinetic momentum operator

$$\hat{\vec{\pi}} = \frac{\hbar}{i}\vec{\nabla} - \frac{q}{c}\vec{A}(\vec{r},t)\,. \tag{6.99}$$

Consequently, we obtain the accordingly modified energy-momentum tensor

$$T^{\nu\mu}(x) = \frac{i\hbar c}{2}\left[\bar{\psi}(x)\gamma^\mu\left(D^\nu\psi\right)(x) - \left(\overline{D^\nu\psi}\right)(x)\gamma^\mu\psi(x)\right] \tag{6.100}$$

with $\overline{D^\nu\psi} \equiv (D^\nu)^*\bar{\psi} = [\partial^\nu - (iq/\hbar c)A^\nu]\bar{\psi}$. Note that the latter does no longer satisfy the continuity equations (6.94) since the energy and the momentum of the particle are not necessarily conserved in the presence of an electromagnetic field but can be exchanged with the latter. More precisely, we obtain from Eq. (6.100)

$$\partial_\mu T^{\nu\mu}(x) = q\left[\partial^\nu A_\mu(x) - \partial_\mu A^\nu(x)\right]\bar{\psi}(x)\gamma^\mu\psi(x) = \frac{q}{c}F^{\nu\mu}(x)j_\mu(x) \tag{6.101}$$

using the expressions (6.90) for the four-current and (2.25) for the electromagnetic field tensor. This equation for the energy-momentum balance is precisely compensated by the analogous equation (2.45) for the electromagnetic field, which means that the total energy and momentum contained within the particle and the electromagnetic field are well conserved.

6.6 Plane waves

More insight into the physical nature of the Dirac spinor and its components is obtained from the general solution of the Dirac equation (6.36) in the absence of electromagnetic fields. To facilitate the interpretation of the results, we calculate this solution in the presence of a normalization volume corresponding to a cube of length L with periodic boundary conditions. The Fourier series expansion of the Dirac spinor is then written as

$$\psi(\vec{r},t) = \frac{1}{\sqrt{V}} \sum_{\vec{k}} \tilde{\psi}_{\vec{k}}(t) e^{i\vec{k}\cdot\vec{r}} \tag{6.102}$$

with $V = L^3$, where the summation involves all possible wave vectors $\vec{k}$ for which the corresponding plane wave $\exp(i\vec{k}\cdot\vec{r})$ fulfils the periodic bounday conditions, *i.e.*, $\vec{k} = (2\pi/L)\vec{l}$ for all $\vec{l} \in \mathbb{Z}$. The application of the momentum operator $\hat{\vec{p}} = -i\hbar\vec{\nabla}$ onto Eq. (6.102) straightforwardly yields

$$\hat{\vec{p}}\psi(\vec{r},t) = \frac{1}{\sqrt{V}} \sum_{\vec{k}} \tilde{\psi}_{\vec{k}}(t) e^{i\vec{k}\cdot\vec{r}} \hbar\vec{k}\,. \tag{6.103}$$

We thereby obtain the multicomponent ordinary differential equation

$$i\hbar\frac{\partial}{\partial t}\tilde{\psi}_{\vec{k}}(t) = H_{\vec{k}}\tilde{\psi}_{\vec{k}}(t) \tag{6.104}$$

for the Fourier components of the Dirac spinor, with the Hamiltonian matrix

$$H_{\vec{k}} = \hbar c\vec{k}\cdot\vec{\alpha} + mc^2\beta \tag{6.105}$$

which is written in the standard representation as

$$H_{\vec{k}} = \left(\begin{array}{c|c} mc^2\mathbb{I} & \hbar c\vec{k}\cdot\vec{\sigma} \\ \hline \hbar c\vec{k}\cdot\vec{\sigma} & -mc^2\mathbb{I} \end{array}\right) \tag{6.106}$$

where $\sigma = (\sigma_x, \sigma_y, \sigma_z)$ is the vector of the Pauil matrices.

The general solution of the Dirac equation in Fourier space (6.104) is then written as $\tilde{\psi}_{\vec{k}}(t) = \exp(-itH_{\vec{k}}/\hbar)\tilde{\psi}_{\vec{k}}(0)$ and can therefore be explicitly obtained from a diagonalization of the 4×4 matrix $H_{\vec{k}}$. This diagonalization is most easily achieved through forming the square of this matrix. Using the anticommutation properties (6.14–6.16) of the matrices α^l and β, we straighforwardly evaluate

$$\left(\vec{k}\cdot\vec{\alpha}\right)^2 = \vec{k}^2 \tag{6.107}$$

and obtain

$$H_{\vec{k}}H_{\vec{k}} = (\hbar\omega_k)^2\,\mathbb{I}_{4\times 4} \tag{6.108}$$

where we define again, as for the Klein-Gordon theory,

$$\omega_k = c\sqrt{\vec{k}^2 + 1/\lambda^2}\,. \tag{6.109}$$

As the matrices α^l and β are traceless, $H_{\vec{k}}$ is traceless, too. We therefore infer that it has the eigenvalues $\hbar\omega_k$ and $-\hbar\omega_k$ of which each one has a two-fold degeneracy.

The associated eigenvectors are determined through the straightforward calculation

$$H_{\vec{k}}\left(\hbar\omega_k\mathbb{I}_{4\times4} \pm H_{\vec{k}}\right) = \hbar\omega_k H_{\vec{k}} \pm \left(\hbar\omega_k\right)^2\mathbb{I}_{4\times4} = \pm\hbar\omega_k\left(\hbar\omega_k\mathbb{I}_{4\times4} \pm H_{\vec{k}}\right) \tag{6.110}$$

where we use Eq. (6.108). Hence, two linearly independent column vectors of the rank-two matrix $\hbar\omega_k\mathbb{I}_{4\times4} \pm H_{\vec{k}}$ can be selected to span the eigenspace associated with the eigenvalues $\pm\hbar\omega_k$. Within the standard representation, we choose the first two columns of the matrix $\hbar\omega_k\mathbb{I}_{4\times4} + H_{\vec{k}}$ for the eigenvalue $\hbar\omega_k$ and the second pair of columns of the matrix $\hbar\omega_k\mathbb{I}_{4\times4} - H_{\vec{k}}$ for the eigenvalue $-\hbar\omega_k$. This yields the matrix of the eigenvectors of $H_{\vec{k}}$ as

$$U_{\vec{k}} = \left(\begin{array}{c|c} (mc^2+\hbar\omega_k)\mathbb{I} & -\hbar c\vec{k}\cdot\vec{\sigma} \\ \hline \hbar c\vec{k}\cdot\vec{\sigma} & (mc^2+\hbar\omega_k)\mathbb{I} \end{array}\right). \tag{6.111}$$

Owing to the hermiticity of the Pauli matrices and their anticommutation property $\sigma_k\sigma_l + \sigma_l\sigma_k = 2\delta_{kl}\mathbb{I}$, we have

$$(\vec{k}\cdot\vec{\sigma})^\dagger(\vec{k}\cdot\vec{\sigma}) = \vec{k}^2\mathbb{I}\,, \tag{6.112}$$

from which we straightforwardly infer that the column vectors forming the matrix $U_{\vec{k}}$ are mutually orthogonal. Their scalar products with themselves are evaluated as

$$\left(mc^2+\hbar\omega_k\right)^2 + \left(\hbar c\vec{k}\right)^2 = 2\hbar\omega_k\left(\hbar\omega_k + mc^2\right) \tag{6.113}$$

for all four column vectors, where we use Eq. (6.112) as well as the expression (6.109) for ω_k. To normalize those column vectors, we perform the straightforward calculations

$$\frac{\hbar ck}{\sqrt{2\hbar\omega_k(\hbar\omega_k+mc^2)}} = \frac{1}{\sqrt{2}}\sqrt{1-\frac{mc^2}{\hbar\omega_k}} \;\equiv\; \delta_k\,, \tag{6.114}$$

$$\frac{mc^2+\hbar\omega_k}{\sqrt{2\hbar\omega_k(\hbar\omega_k+mc^2)}} = \frac{1}{\sqrt{2}}\sqrt{1+\frac{mc^2}{\hbar\omega_k}} \;=\; \sqrt{1-\delta_k^2}\,. \tag{6.115}$$

This yields the orthonormal basis $(U^{(1)}_{\vec{k}}, U^{(2)}_{\vec{k}}, U^{(-1)}_{-\vec{k}}, U^{(-2)}_{-\vec{k}})$ of the normalized eigenvectors

$$U_{\vec{k}}^{(1)} = \begin{pmatrix} \sqrt{1-\delta_k^2} \\ 0 \\ \hline \delta_k \hat{k}_3 \\ \delta_k(\hat{k}_1 + i\hat{k}_2) \end{pmatrix}, \qquad U_{\vec{k}}^{(2)} = \begin{pmatrix} 0 \\ \sqrt{1-\delta_k^2} \\ \hline \delta_k(\hat{k}_1 - i\hat{k}_2) \\ -\delta_k \hat{k}_3 \end{pmatrix},$$
$$U_{-\vec{k}}^{(-1)} = \begin{pmatrix} -\delta_k \hat{k}_3 \\ -\delta_k(\hat{k}_1 + i\hat{k}_2) \\ \hline \sqrt{1-\delta_k^2} \\ 0 \end{pmatrix}, \qquad U_{-\vec{k}}^{(-2)} = \begin{pmatrix} -\delta_k(\hat{k}_1 - i\hat{k}_2) \\ \delta_k \hat{k}_3 \\ \hline 0 \\ \sqrt{1-\delta_k^2} \end{pmatrix} \tag{6.116}$$

that are associated with the eigenvalues $(\hbar\omega_k, \hbar\omega_k, -\hbar\omega_k, -\hbar\omega_k)$, where we define $\hat{k}_l \equiv k_l/k$ for all $l = 1, 2, 3$. δ_k is defined according to Eq. (6.114) and is therefore approximately evaluated as

$$\delta_k \simeq \frac{\hbar k}{2mc} + \mathcal{O}\left(\left(\frac{\hbar k}{mc}\right)^3\right) \tag{6.117}$$

in the nonrelativistic regime $\hbar k/(mc) \to 0$. We therefore obtain in this regime

$$\left(U_{\vec{k}}^{(1)}, U_{\vec{k}}^{(2)}, U_{-\vec{k}}^{(-1)}, U_{-\vec{k}}^{(-2)}\right) \simeq \mathbb{I}_{4\times 4} + \frac{\hbar}{2mc} \left(\begin{array}{c|c} \mathbb{O} & -\vec{k}\cdot\vec{\sigma} \\ \hline \vec{k}\cdot\vec{\sigma} & \mathbb{O} \end{array}\right) \tag{6.118}$$

up to corrections that scale quadratically with $\hbar k/(mc)$.

The general solution of the differential equation (6.104) can then be written as

$$\tilde{\psi}_{\vec{k}}(t) = \sum_{\sigma=1,2} \left(\alpha_{\vec{k}\sigma}^{(+)} U_{\vec{k}}^{(\sigma)} e^{-i\omega_k t} + \alpha_{-\vec{k}\sigma}^{(-)} U_{-\vec{k}}^{(-\sigma)} e^{i\omega_k t}\right) \tag{6.119}$$

for complex coefficients $\alpha_{\vec{k}\sigma}^{(\pm)} \in \mathbb{C}$ that are defined through the initial values of the vectors $\tilde{\psi}_{\vec{k}}$. The general solution of the Dirac equation (6.36) is then expressed as

$$\psi(\vec{r}, t) = \frac{1}{\sqrt{V}} \sum_{\vec{k}} \sum_{\sigma=1,2} \left(\psi_{\vec{k}\sigma}^{(+)}(\vec{r}, t) + \psi_{\vec{k}\sigma}^{(-)}(\vec{r}, t)\right) \tag{6.120}$$

in terms of the plane-wave spinors

$$\psi_{\vec{k}\sigma}^{(\pm)}(\vec{r}, t) = \alpha_{\vec{k}\sigma}^{(\pm)} U_{\vec{k}}^{(\pm\sigma)} e^{\pm i(\vec{k}\cdot\vec{r} - \omega_k t)}\,. \tag{6.121}$$

Evaluating the total probability yields

$$\begin{aligned} 1 &= \int_V d^3r \psi^\dagger(\vec{r}, t)\psi(\vec{r}, t) = \sum_{\vec{k}} \tilde{\psi}_{\vec{k}}^\dagger(t)\tilde{\psi}_{\vec{k}}(t) \\ &= \sum_{\vec{k}} \sum_{\sigma=1,2} \left(\left|\alpha_{\vec{k}\sigma}^{(+)}\right|^2 + \left|\alpha_{\vec{k}\sigma}^{(-)}\right|^2\right) \end{aligned} \tag{6.122}$$

as normalization condition that the coefficients $\alpha^{(\pm)}_{\vec{k}\sigma}$ must fulfil.

It is very instructive to also evaluate the total energy that is contained within the spinor (6.120) solving the Dirac equation. Using the expression (6.92) for the energy density, we straightforwardly calculate by means of integration by parts

$$\begin{aligned} E &= \int_V d^3r\psi^\dagger(\vec{r},t)\left(c\vec{\alpha}\cdot\hat{\vec{p}}+mc^2\beta\right)\psi(\vec{r},t) = \sum_{\vec{k}}\tilde{\psi}^\dagger_{\vec{k}}(t)H_{\vec{k}}\tilde{\psi}_{\vec{k}}(t) \\ &= \sum_{\vec{k}}\sum_{\sigma=1,2}\hbar\omega_k\left(\left|\alpha^{(+)}_{\vec{k}\sigma}\right|^2-\left|\alpha^{(-)}_{\vec{k}\sigma}\right|^2\right). \end{aligned} \tag{6.123}$$

This expression for the total energy poses a serious conceptual problem insofar as it is unbounded from below and can attain arbitrarily negative values. This would consequently imply that the Dirac Hamiltonian does not feature a well-defined ground state, contrary to the nonrelativistic Schrödinger equation. However, the existence of such a ground state is a necessary ingredient of quantum statistical physics, as it defines the absolute zero of temperature, and is well confirmed by experimental evidence.

There is one possible resort to this problem. It requires to assume that the particles described by the Dirac theory are of *fermionic* nature and satisfy Pauli's exclusion principle stating that two such particles cannot occupy the same quantum state. In that case, we can assume that there are half as many particles of this type in the universe as there are plane-wave eigenstates of the Dirac Hamiltonian. The ground state of this many-particle system is then realized by occupying each eigenstate of negative energy with exactly one particle. While its associated energy would still have an infinitely low negative value, excitations with respect to this ground state are then well defined. For instance, the first excited may-particle eigenstate is realized by removing a particle that occupies a negative-energy eigenstate with the wave vector $\vec{k}=\vec{0}$ (*i.e.*, a state that is associated with the one of the two spinors $U^{(-1)}_{\vec{0}}$ and $U^{(-2)}_{\vec{0}}$) and by placing it instead into one of the two positive-energy eigenstates with $\vec{k}=\vec{0}$ (associated with the spinors $U^{(1)}_{\vec{0}}$ and $U^{(2)}_{\vec{0}}$). As is illustrated in Fig. 6.1, this excitation creates a particle in the manifold of positive energy and leaves behind a hole in the manifold of negative energy, in perfect analogy with the excitation of an electron from the valence band to the conduction band in a semiconductor. The excitation energy that is required to create such a particle-hole pair equals the energetic gap $2mc^2$ between the positive- and negative-energy manifolds. It is therefore suggestive to interpret this hole as *antiparticle*, having the same mass and reacting to the application of an electromagnetic field in exactly the opposite manner as the particle, and to consider the transition from the ground state to this first excited state as *particle-antiparticle creation* out of the *vacuum*, the latter being defined by this many-body ground state where all Dirac particles occupy negative-energy states.

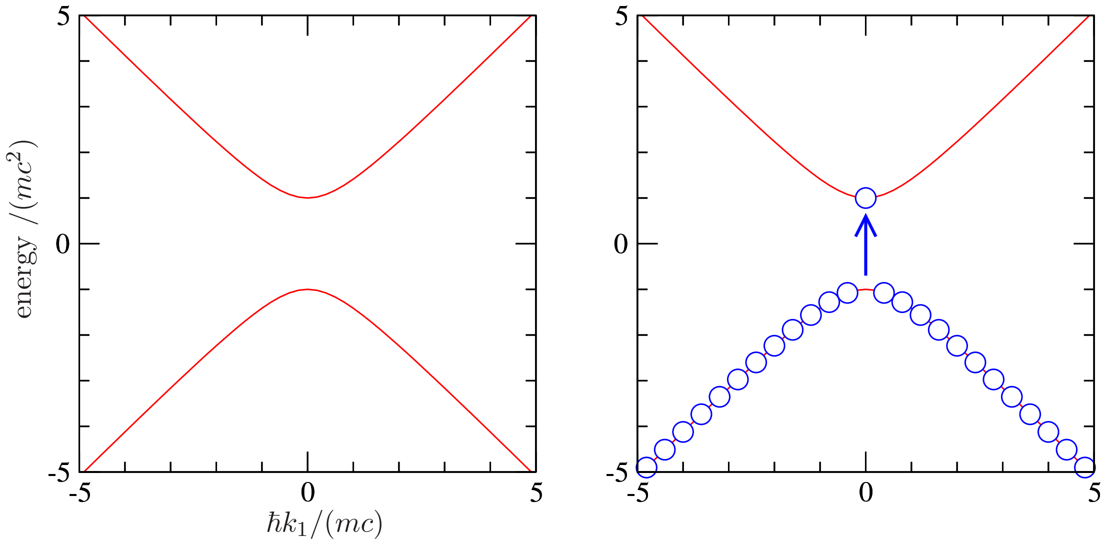

Figure 6.1: Left panel: Eigenenergies $\pm\hbar\omega_k$ of the Dirac Hamiltonian, plotted as a function of k_1 for $k_2 = k_3 = 0$, according to Eq. (6.109). The eigenspectrum is clearly unbounded from below and features a gap of size $2mc^2$ between the positive- and the negative-energy manifold. Right panel: A well-defined ground state is achieved by assuming that particles described by the Dirac theory are of fermionic nature and satisfy Pauli's exclusion principle. Assuming that there are half as many particles as there are eigenstates of the Dirac Hamiltonian, the ground state is realized by occupying each eigenstate on the negative-energy manifold with exactly one particle. The excitation of one particle to a state on the positive-energy manifold leaves then behind a hole in the negative-energy manifold which can be associated with an antiparticle.

6.7 Second Quantization

The above interpretation in terms of particles and antiparticles that are created out of the vacuum can be concretized and corroborated from an analytical point of view by means of a quantization procedure of the Dirac spinor, in analogy with the quantization of the Klein-Gordon field discussed in Section 5.6. This sounds rather inappropriate at first glance, given the fact that the spinor ψ in the Dirac theory was specifically conceived to represent the wavefunction of a quantum particle, which obviously is a concept that already belongs to the quantum realm and admits no physical interpretation in terms of a classical field. From a mathematical point of view, however, we are nevertheless perfectly free to treat the Dirac spinor $\psi : \mathbb{R}^4 \to \mathbb{C}^4$ as a complex-valued vector field defined on space-time and subject it to a quantization procedure, as we did with the Klein-Gordon field in Section 5.6. Quite consequently, that procedure is dubbed *second quantization* since is applied to a concept (the wavefunction) that is the result of a "first" quantization procedure of the particle's classical dynamics.

In contrast to the Klein-Gordon theory, we cannot make use of the results obtained in the framework of Chapter 4 in order to determine how this second quantization is to be done in practice. As a matter of fact, even though it admits plane waves as solutions, the Dirac equation describing the time evolution of the quantum particle's wavefunction is technically not a wave equation given in terms of the d'Alembert operator $\partial_\nu \partial^\nu$ and thus cannot be subjected to the same quantization procedure as classical waves. Moreover, as was developed in Sections 4.2 and 5.6, that procedure would inevitably yield bosonic quantum particles, whereas a particle described by the Dirac theory would, as we pointed out in the previous section, need to be of fermionic nature to assure the existence of a well-defined ground state.

In view of this situation, we have, at that stage in this course, no alternative to introducing the quantization rules of the Dirac spinor in a heuristic manner, in form of a basic postulate. We shall do this such that we obtain analogous implications for the quantization of plane-wave modes as in the case of the Klein-Gordon theory: similarly to Eqs. (5.56) and (5.57), plane-wave amplitudes that oscillate as $\propto \exp(-i\omega_k t)$ within the expression (6.120) for the general solution of the Dirac equation will be substituted by annihilation operators that are associated with the primary particle species, *i.e.*, the "particle" of the corresponding theory, whereas amplitudes oscillating as $\propto \exp(i\omega_k t)$ within Eq. (6.120) will be replaced by creation operators associated with the secondary particle species, to be identified with the "antiparticle". The prefactors that are involved in these substitutions will be chosen such that we get a similar expression for the quantum Hamiltonian as in Eq. (5.63), where a particle or antiparticle associated with the plane-wave mode that is characterized by the wave vektor $\vec{k}$ will contribute with the energy quantum $\hbar\omega_k$ to the total energy.

Most specifically, this prescription amounts to imposing the quantization procedure

$$\alpha^{(+)}_{\vec{k}\sigma} e^{-i\omega_k t} \quad \mapsto \quad \hat{a}_{\vec{k}\sigma}(t)\,, \tag{6.124}$$

$$\alpha^{(-)}_{\vec{k}\sigma} e^{i\omega_k t} \quad \mapsto \quad \hat{b}^\dagger_{\vec{k}\sigma}(t)\,, \tag{6.125}$$

$$(\alpha^{(+)}_{\vec{k}\sigma})^* e^{i\omega_k t} \quad \mapsto \quad \hat{a}^\dagger_{\vec{k}\sigma}(t)\,, \tag{6.126}$$

$$(\alpha^{(-)}_{\vec{k}\sigma})^* e^{-i\omega_k t} \quad \mapsto \quad \hat{b}_{\vec{k}\sigma}(t) \tag{6.127}$$

for the plane-wave amplitudes $\alpha^{(\pm)}_{\vec{k}\sigma}$ that characterize the general solution (6.120) and their complex conjugates, where $\hat{a}^\dagger_{\vec{k}\sigma}(t)$ and $\hat{b}^\dagger_{\vec{k}\sigma}(t)$ represent the hermitian conjugates of the operators $\hat{a}_{\vec{k}\sigma}(t)$ and $\hat{b}_{\vec{k}\sigma}(t)$, respectively. Considering, as in the previous section, the presence of a normalization volume, these creation and annihilation operators are postulated to satify the fermionic anticommutation rules

$$\left\{\hat{a}_{\vec{k}\sigma}, \hat{a}^\dagger_{\vec{k}'\sigma'}\right\} \equiv \hat{a}_{\vec{k}\sigma}\hat{a}^\dagger_{\vec{k}'\sigma'} + \hat{a}^\dagger_{\vec{k}'\sigma'}\hat{a}_{\vec{k}\sigma} = \delta_{\vec{k}\vec{k}'}\delta_{\sigma\sigma'}\,, \tag{6.128}$$

$$\left\{\hat{b}_{\vec{k}\sigma}, \hat{b}^\dagger_{\vec{k}'\sigma'}\right\} \equiv \hat{b}_{\vec{k}\sigma}\hat{b}^\dagger_{\vec{k}'\sigma'} + \hat{b}^\dagger_{\vec{k}'\sigma'}\hat{b}_{\vec{k}\sigma} = \delta_{\vec{k}\vec{k}'}\delta_{\sigma\sigma'} \tag{6.129}$$

for all $\sigma, \sigma' \in \{-1/2, 1/2\}$ and all wave vectors $\vec{k}, \vec{k}'$ that are consistent with the periodic boundary conditions of the normalization volume. We furthermore impose perfect anticommutation of all other possible combinations of those creation and annihilation operators, *i.e.*,

$$\begin{aligned} 0 \;=\; & \left\{\hat{a}_{\vec{k}\sigma}, \hat{a}_{\vec{k}'\sigma'}\right\} = \left\{\hat{b}_{\vec{k}\sigma}, \hat{b}_{\vec{k}'\sigma'}\right\} = \left\{\hat{a}_{\vec{k}\sigma}, \hat{b}_{\vec{k}'\sigma'}\right\} = \left\{\hat{a}_{\vec{k}\sigma}, \hat{b}^\dagger_{\vec{k}'\sigma'}\right\} = \\ & \left\{\hat{a}^\dagger_{\vec{k}\sigma}, \hat{a}^\dagger_{\vec{k}'\sigma'}\right\} = \left\{\hat{b}^\dagger_{\vec{k}\sigma}, \hat{b}^\dagger_{\vec{k}'\sigma'}\right\} = \left\{\hat{a}^\dagger_{\vec{k}\sigma}, \hat{b}^\dagger_{\vec{k}'\sigma'}\right\} = \left\{\hat{a}^\dagger_{\vec{k}\sigma}, \hat{b}_{\vec{k}'\sigma'}\right\} \end{aligned} \tag{6.130}$$

for all possible $\vec{k}, \vec{k}'$ and σ, σ'.

A straightforward consequence of the application of the above set of identities (6.130) to the special case $\vec{k}' = \vec{k}$ and $\sigma' = \sigma$ is *Pauli's exclusion principle* stating that it is impossible to have more than one fermionic particle within the same eigenmode of the Dirac Hamiltonian. Indeed, we have in that case $0 = \{\hat{a}^\dagger_{\vec{k}\sigma}, \hat{a}^\dagger_{\vec{k}\sigma}\} = \hat{a}^\dagger_{\vec{k}\sigma}\hat{a}^\dagger_{\vec{k}\sigma} + \hat{a}^\dagger_{\vec{k}\sigma}\hat{a}^\dagger_{\vec{k}\sigma} = 2\hat{a}^\dagger_{\vec{k}\sigma}\hat{a}^\dagger_{\vec{k}\sigma}$, as well as $\hat{b}^\dagger_{\vec{k}\sigma}\hat{b}^\dagger_{\vec{k}\sigma} = 0$ and $\hat{a}_{\vec{k}\sigma}\hat{a}_{\vec{k}\sigma} = \hat{b}_{\vec{k}\sigma}\hat{b}_{\vec{k}\sigma} = 0$, which implies that the successive application of two identical creation or annihilation operators to any state within the many-body Hilbert space that is generated by those operators yields the null vector in that space. Pauli's exclusion principle is also valid in position space where the above quantization procedure can be introduced via the quantization of the Dirac spinor

$$\psi(\vec{r}, t) \mapsto \hat{\psi}(\vec{r}, t) \equiv \left(\hat{\psi}_1(\vec{r}, t), \ldots, \hat{\psi}_4(\vec{r}, t)\right)^T \tag{6.131}$$

and the associated adjoint spinor (6.82)

$$\bar{\psi}(\vec{r}, t) \mapsto \hat{\bar{\psi}}(\vec{r}, t) \equiv \left(\hat{\bar{\psi}}_1(\vec{r}, t), \ldots, \hat{\bar{\psi}}_4(\vec{r}, t)\right) \tag{6.132}$$

whose components are postulated to satisfy the anticommutation rules

$$\left\{\hat{\psi}_s(\vec{r}, t), \hat{\bar{\psi}}_{s'}(\vec{r}', t)\right\} = \delta_{ss'}\delta(\vec{r} - \vec{r}') \tag{6.133}$$

as well as

$$\left\{\hat{\psi}_s(\vec{r}, t), \hat{\psi}_{s'}(\vec{r}', t)\right\} = \left\{\hat{\bar{\psi}}_s(\vec{r}, t), \hat{\bar{\psi}}_{s'}(\vec{r}', t)\right\} = 0 \tag{6.134}$$

for all $\vec{r}, \vec{r}' \in \mathbb{R}^3$ and all $s, s' \in \{1, \ldots, 4\}$.

The quantum Hamiltonian, which is obtained from the quantized analog of the expression (6.123) for the total energy of the Dirac spinor, is then straightforwardly calculated as

$$\begin{aligned}\hat{H} &= \int_V d^3r \hat{\psi}^\dagger(\vec{r},t)\left(c\vec{\alpha}\cdot\hat{\vec{p}} + mc^2\beta\right)\hat{\psi}(\vec{r},t) \\ &= \sum_{\vec{k}}\sum_{\sigma=1,2}\hbar\omega_k\left(\hat{a}^\dagger_{\vec{k}\sigma}\hat{a}_{\vec{k}\sigma} - \hat{b}_{\vec{k}\sigma}\hat{b}^\dagger_{\vec{k}\sigma}\right) \\ &= \sum_{\vec{k}}\sum_{\sigma=1,2}\hbar\omega_k\left(\hat{a}^\dagger_{\vec{k}\sigma}\hat{a}_{\vec{k}\sigma} + \hat{b}^\dagger_{\vec{k}\sigma}\hat{b}_{\vec{k}\sigma}\right) - \sum_{\vec{k}} 2\hbar\omega_k\,, \end{aligned} \tag{6.135}$$

where we made use of the anticommutation law (6.129) in order to obtain the expression (6.135). This Hamiltonian has the asset that it features a well-defined ground state $|-\rangle$ corresponding to the absence of particles, *i.e.* $\hat{a}_{\vec{k}\sigma}|-\rangle = \hat{b}_{\vec{k}\sigma}|-\rangle = 0$ for all $\vec{k},\sigma$, even though the associated ground state energy $E_0 = -\sum_{\vec{k}} 2\hbar\omega_k$ is infinitely negative. As in the case of the quantization of electromagnetic waves and the Klein-Gordon equation, we can treat this infinitely negative energy as an unimportant constant that can be formally eliminated by a proper renormalization of the energy scale. The presence of this constant in the expression (6.135) is nevertheless rather instructive as it quantitatively corroborates the interpretation that we put forward at the end of the previous section 6.6 and that is illustrated in the right panel of Fig. 6.1: the ground state of this quantum many-body system corresponds to the situation where all one-particle eigenmodes of negative energy are occupied with fermionic particles while all eigenmodes of positive energy are unoccupied. Clearly, this particular many-body state, which is also termed *Dirac sea*, has the total energy $-\sum_{\vec{k}} 2\hbar\omega_k$.

From this reasoning, it becomes clear that $\hat{a}^\dagger_{\vec{k}\sigma}$ is the creation operator of a particle within the positive-energy plane-wave state that is given by the wavefunction $U^{(\sigma)}_{\vec{k}} e^{i\vec{k}\cdot\vec{r}}$, while $\hat{b}^\dagger_{\vec{k}\sigma}$ has to be understood as the creation operator of a *hole* in the corresponding negative-energy plane-wave state, which can then be associated with an antiparticle in the mode that is characterized by the wavefunction $(U^{(-\sigma)}_{\vec{k}})^* e^{i\vec{k}\cdot\vec{r}}$. This antiparticle has exactly the same basic properties as the particle, except for its charge which is opposited to the one of the particle. As is illustrated in the right panel of Fig. 6.1, the minimal energy that is needed to place a particle from the negative-energy branch to the positive-energy branch of the spectrum is given by the gap $2mc^2$ between those two branches. It is by no means a coincidene that this gap also corresponds to the fundamental energy that one would need to generate a particle-antiparticle pair for a particle species with mass m.

The quantum operator associated with total momentum is obtained from the quantization and integration of the expression (6.93) for the momentum density contained in the Dirac spinor. This yields

$$\begin{aligned}\hat{\vec{P}} &= \int_V d^3r \hat{\psi}^\dagger(\vec{r},t)\frac{\hbar}{i}\vec{\nabla}\hat{\psi}(\vec{r},t) = \sum_{\vec{k}}\sum_{\sigma=1,2}\hbar\vec{k}\left(\hat{a}^\dagger_{\vec{k}\sigma}\hat{a}_{\vec{k}\sigma} - \hat{b}_{\vec{k}\sigma}\hat{b}^\dagger_{\vec{k}\sigma}\right) \\ &= \sum_{\vec{k}}\sum_{\sigma=1,2}\hbar\vec{k}\left(\hat{a}^\dagger_{\vec{k}\sigma}\hat{a}_{\vec{k}\sigma} + \hat{b}^\dagger_{\vec{k}\sigma}\hat{b}_{\vec{k}\sigma}\right) , \end{aligned} \qquad (6.136)$$

where owing to $\sum_{\vec{k}}\vec{k} = 0$ the anticommutator of $\hat{b}_{\vec{k}\sigma}$ and $\hat{b}^\dagger_{\vec{k}\sigma}$ does not give rise to any additional constant. We can also define a quantum operator that is associated with the total probability of the Dirac spinor. This operator is evaluated as

$$\begin{aligned}\hat{Q} &= \int_V d^3r \hat{\psi}^\dagger(\vec{r},t)\hat{\psi}(\vec{r},t) = \sum_{\vec{k}}\sum_{\sigma=1,2}\left(\hat{a}^\dagger_{\vec{k}\sigma}\hat{a}_{\vec{k}\sigma} + \hat{b}_{\vec{k}\sigma}\hat{b}^\dagger_{\vec{k}\sigma}\right) \\ &= \sum_{\vec{k}}\sum_{\sigma=1,2}\left(\hat{a}^\dagger_{\vec{k}\sigma}\hat{a}_{\vec{k}\sigma} - \hat{b}^\dagger_{\vec{k}\sigma}\hat{b}_{\vec{k}\sigma}\right) + \sum_{\vec{k}} 2 . \end{aligned} \qquad (6.137)$$

It features again an infinite constant which can be attributed to the presence of the Dirac sea, where every single-particle plane-wave state that is consistent with the periodic boundary conditions of the nomalization volume is populated with exactly two particles. Quite logically, one has to add to this constant the total population of positive-energy modes and subtract the total number of holes in the Dirac sea, in order to obtain the total population of the system. This is exactly what is expressed by Eq. (6.137). Leaving out the unimportant constant, we effectively obtain again a charge operator, assigning a positive (or negative) charge to "particles", *i.e.*, to populations of positive-energy modes, and a charge with opposite sign to "antiparticles", *i.e.*, to holes in the populations of negative-energy modes. This results in a conservation law for the total charge, as in the case of the Klein-Gordon theory, while the total number of of "particles" and "antiparticles" in the above sense is not conserved and can vary, *e.g.* due to the creation (or annihilation) of a particle-antiparticle pair that occurs together with the absorption (or emission) of a photon with a sufficiently high energy ($> 2mc^2$).

We should note in this context that the association of $\hat{b}^\dagger_{\vec{k}\sigma}$ with a creation operator of a fermionic particle species and of $\hat{b}_{\vec{k}\sigma}$ with the corresponding annihilation operator is a matter of interpretation and can be swapped, owing to the symmetry of the anticommutation rule (6.129). Nothing would formally prevent us to interpret $\hat{b}_{\vec{k}\sigma}$ as creation operator and $\hat{b}^\dagger_{\vec{k}\sigma}$ as annihilation operator, in which case those operators would refer to the same particle species (with the same charge) as the $\hat{a}_{\vec{k}\sigma}, \hat{a}^\dagger_{\vec{k}\sigma}$ operators. The price that one would have to pay for adopting this particular choice is that the physical vacuum of the fermionic system does not correspond to the absence of particles but rather to the situation that all negative-energy modes are populated with exactly one particle. This latter aspect makes it preferable to abandon this choice and work with the notion of antiparticles instead.

6.8 The Pauli equation

Owing to the insights that were obtained in the two preceeding sections 6.6 and 6.7, we can safely claim that in the nonrelativistic limit the upper two components of the Dirac spinor describe the particle component of the wavefunction while its lower two components correspond to the associated antiparticle. This association allows us to elaborate the nonrelativistic limit of the Dirac theory in a rather straightforward manner. To this end, let us start with the Dirac equation (6.97) in the presence of an electromagnetic field described in terms of the four-potential $(A^\nu) = (\Phi, \vec{A})$. This equation can be written as

$$i\hbar D_t \psi(\vec{r}, t) = c\vec{\alpha} \cdot \vec{\pi} \psi(\vec{r}, t) + mc^2 \beta \psi(\vec{r}, t) \tag{6.138}$$

with the covariant time derivative operator

$$D_t = \frac{\partial}{\partial t} + \frac{iq}{\hbar} \Phi(\vec{r}, t) \tag{6.139}$$

and the kinetic momentum operator

$$\vec{\pi} = \frac{\hbar}{i} \vec{\nabla} - \frac{q}{c} \vec{A}(\vec{r}, t)\,. \tag{6.140}$$

Employing the standard representation, with the choices

$$\beta = \left(\begin{array}{c|c} \mathbb{I} & \mathbb{O} \\ \hline \mathbb{O} & -\mathbb{I} \end{array} \right), \quad \alpha^l = \left(\begin{array}{c|c} \mathbb{O} & \sigma_l \\ \hline \sigma_l & \mathbb{O} \end{array} \right) \tag{6.141}$$

for $l = 1, 2, 3$, and decomposing the Dirac spinor according to

$$\psi = \begin{pmatrix} \psi_1 \\ \psi_2 \\ \psi_3 \\ \psi_4 \end{pmatrix} \equiv \left(\begin{array}{c} \psi^{(+)} \\ \hline \psi^{(-)} \end{array} \right) \tag{6.142}$$

with the two two-component spinors

$$\psi^{(\pm)} = \begin{pmatrix} \psi_1^{(\pm)} \\ \psi_2^{(\pm)} \end{pmatrix}, \tag{6.143}$$

we can rewrite the Dirac equation (6.138) in terms of the two coupled equations

$$i\hbar D_t \psi^{(+)}(\vec{r}, t) = c\vec{\sigma} \cdot \vec{\pi} \psi^{(-)}(\vec{r}, t) + mc^2 \psi^{(+)}(\vec{r}, t)\,, \tag{6.144}$$

$$i\hbar D_t \psi^{(-)}(\vec{r}, t) = c\vec{\sigma} \cdot \vec{\pi} \psi^{(+)}(\vec{r}, t) - mc^2 \psi^{(-)}(\vec{r}, t)\,. \tag{6.145}$$

As in the case of the Klein-Gordon theory (see Eqs. (5.81) and (5.82) in Section 5.7), we thereby obtain a system of two coupled Schrödinger-like equations for the particle and the antiparticle component, featuring the rest energies mc^2 and $-mc^2$, respectively.

Let us first focus on the development of a nonrelativistic theory for the particle component. To get rid of the rest energy term mc^2 in the equation (6.144) for $\psi^{(+)}$, we first apply the gauge transformation $\psi^{(\pm)} \mapsto \phi^{(\pm)}$ defined through

$$\psi^{(\pm)}(\vec{r},t) \equiv \phi^{(\pm)}(\vec{r},t) e^{-imc^2 t/\hbar}\,, \tag{6.146}$$

which effectively redefines the origin of energy such that it corresponds to the rest energy mc^2. Equations (6.144) and (6.145) are then rewritten as

$$i\hbar D_t \phi^{(+)}(\vec{r},t) = c\vec{\sigma}\cdot\vec{\pi}\phi^{(-)}(\vec{r},t)\,, \tag{6.147}$$

$$i\hbar D_t \phi^{(-)}(\vec{r},t) = c\vec{\sigma}\cdot\vec{\pi}\phi^{(+)}(\vec{r},t) - 2mc^2\phi^{(-)}(\vec{r},t)\,. \tag{6.148}$$

in terms of these redefined two-component spinors $\phi^{(\pm)}$.

As in the discussion of the nonrelativistic limit of the Klein-Gordon theory (see Section 5.7), we make the assumption that the great majority of norm is contained within $\phi^{(+)}$ (which will be fulfilled *e.g.* if for some initial time t_0 we have $\phi^{(-)}(\vec{r},t_0) = 0$ for all $\vec{r}$). Equation (6.148) can then be formally solved through the iterative procedure

$$\begin{aligned} \phi^{(-)} &= \frac{1}{2mc^2}\left(\vec{\sigma}\cdot\vec{\pi}\phi^{(+)} - i\hbar D_t\phi^{(-)}\right) \qquad (6.149)\\ &= \frac{1}{2mc^2}\left[\vec{\sigma}\cdot\vec{\pi}\phi^{(+)} - i\hbar D_t \frac{1}{2mc^2}\left(\vec{\sigma}\cdot\vec{\pi}\phi^{(+)} - i\hbar D_t\phi^{(-)}\right)\right]\\ &= \ldots \end{aligned}$$

where $\phi^{(-)}$ on the right-hand side of this equation is recursively replaced by the expression (6.149). We can thereby express $\phi^{(-)}$ in terms of $\phi^{(+)}$ via the series

$$\phi^{(-)}(\vec{r},t) = \frac{\vec{\sigma}\cdot\vec{\pi}}{2mc}\phi^{(+)}(\vec{r},t) - \frac{1}{4m^2c^3} i\hbar D_t\,\vec{\sigma}\cdot\vec{\pi}\phi^{(+)}(\vec{r},t) + \ldots \tag{6.150}$$

featuring terms that decrease more and more for asymptotically large c. Inserting this expression into Eq. (6.147) yields the equation

$$\begin{aligned} i\hbar D_t\phi^{(+)}(\vec{r},t) &= \frac{(\vec{\sigma}\cdot\vec{\pi})^2}{2m}\phi^{(+)}(\vec{r},t) - \frac{1}{4m^2c^2}\vec{\sigma}\cdot\vec{\pi}\, i\hbar D_t\,\vec{\sigma}\cdot\vec{\pi}\phi^{(+)}(\vec{r},t) + \ldots\\ &= \frac{(\vec{\sigma}\cdot\vec{\pi})^2}{2m}\phi^{(+)}(\vec{r},t) + \mathcal{O}(1/c^2) \qquad (6.151) \end{aligned}$$

in the nonrelativistic limit $c \to \infty$.

To elaborate the kinetic term of this equation in more detail, we make use of the general relation

$$\sigma_l\sigma_{l'} = \delta_{ll'}\mathbb{I} + i\sum_{k=1}^{3}\epsilon_{ll'k}\sigma_k \tag{6.152}$$

characterizing the Pauli matrices σ_l. This yields

$$(\vec{\sigma}\cdot\vec{\pi})^2 = \sum_{l,l'=1}^{3} \sigma_l \sigma_{l'} \pi_l \pi_{l'} = \vec{\pi}^2 + i\,(\vec{\pi}\times\vec{\pi})\cdot\vec{\sigma}\,. \tag{6.153}$$

Contrary to what one would naively expect, the vector product of the kinetic momentum operator with itself is not zero; using its defintion (6.140) we evaluate it as

$$\vec{\pi}\times\vec{\pi} = \left(\frac{\hbar}{i}\vec{\nabla} - \frac{q}{c}\vec{A}\right)\times\left(\frac{\hbar}{i}\vec{\nabla} - \frac{q}{c}\vec{A}\right) = -\frac{\hbar q}{ic}\nabla\times\vec{A} = -\frac{\hbar q}{ic}\vec{B} \tag{6.154}$$

with $\vec{B} = \nabla\times\vec{A}$ the magnetic field. Inserting this expression into Eq. (6.151) and using the definitions (6.139) and (6.140) of the covariant time derivative and the kinetic momentum, respectively, yields

$$\begin{aligned} i\hbar\frac{\partial}{\partial t}\phi^{(+)}(\vec{r},t) &= \frac{1}{2m}\left(\frac{\hbar}{i}\vec{\nabla} - \frac{q}{c}\vec{A}(\vec{r},t)\right)^2 \phi^{(+)}(\vec{r},t) + q\Phi(\vec{r},t)\phi^{(+)}(\vec{r},t) \\ &\quad -\frac{q\hbar}{2mc}\vec{B}(\vec{r},t)\cdot\vec{\sigma}\phi^{(+)}(\vec{r},t)\,, \end{aligned} \tag{6.155}$$

which is the *Pauli equation* describing a particle with the charge q in the presence of an electromagnetic field. This Pauli equation can be seen as a generalization of the Schrödinger equation (3.40) for a spin 1/2 particle, whose wavefunction is represented in terms of a two-component *Pauli spinor*

$$\psi^{(+)}(\vec{r},t) = \begin{pmatrix} \psi_1^{(+)}(\vec{r},t) \\ \psi_2^{(+)}(\vec{r},t) \end{pmatrix} \equiv \begin{pmatrix} \psi_\uparrow^{(+)}(\vec{r},t) \\ \psi_\downarrow^{(+)}(\vec{r},t) \end{pmatrix}. \tag{6.156}$$

Remarkably, the existence of the two "spin up" and "spin down" components of the wavefunction does not need to be separately postulated but is a direct consequence of the Dirac equation. The latter also correctly yields the associated magnetic moment and its interaction with an external magnetic field, corresponding to the last term in Eq. (6.155), which in the case of an electron with the charge $q = -e$ reads $\mu_B\vec{B}(\vec{r},t)\cdot\vec{\sigma}\phi^{(+)}(\vec{r},t)$ with $\mu_B = e\hbar/(2mc)$ the Bohr magneton.

Similarly, a nonrelativistic theory for the antiparticle can be obtained by applying the gauge transformation $\psi^{(\pm)} \mapsto \tilde{\phi}^{(\pm)}$ with

$$\psi^{(\pm)}(\vec{r},t) \equiv \tilde{\phi}^{(\pm)}(\vec{r},t)e^{imc^2t/\hbar}\,, \tag{6.157}$$

by which the origin of energy is effectively set to $-mc^2$. This yields from Eqs. (6.144) and (6.145) the system of equations

$$i\hbar D_t\tilde{\phi}^{(-)}(\vec{r},t) = c\vec{\sigma}\cdot\vec{\pi}\tilde{\phi}^{(+)}(\vec{r},t)\,, \tag{6.158}$$

$$i\hbar D_t\tilde{\phi}^{(+)}(\vec{r},t) = c\vec{\sigma}\cdot\vec{\pi}\tilde{\phi}^{(-)}(\vec{r},t) + 2mc^2\tilde{\phi}^{(+)}(\vec{r},t)\,. \tag{6.159}$$

In analogy with Eq. (6.150), the approximate solution of Eq. (6.159) in the nonrelativistic limit $c \to \infty$ is given by

$$\tilde{\phi}^{(+)}(\vec{r},t) \simeq -\frac{\vec{\sigma}\cdot\vec{\pi}}{2mc}\tilde{\phi}^{(-)}(\vec{r},t)\,, \tag{6.160}$$

from wich we obtain, after insertion into Eq. (6.158),

$$i\hbar D_t\tilde{\phi}^{(-)}(\vec{r},t) \simeq -\frac{(\vec{\sigma}\cdot\vec{\pi})^2}{2m}\tilde{\phi}^{(-)}(\vec{r},t)\,, \tag{6.161}$$

or, rewritten in a more explicit manner,

$$\begin{aligned} i\hbar\frac{\partial}{\partial t}\tilde{\phi}^{(-)}(\vec{r},t) &= -\frac{1}{2m}\left(\frac{\hbar}{i}\vec{\nabla}-\frac{q}{c}\vec{A}(\vec{r},t)\right)^2\tilde{\phi}^{(-)}(\vec{r},t)+q\Phi(\vec{r},t)\tilde{\phi}^{(-)}(\vec{r},t) \\ &\quad +\frac{q\hbar}{2mc}\vec{B}(\vec{r},t)\cdot\vec{\sigma}\tilde{\phi}^{(-)}(\vec{r},t)\,. \end{aligned} \tag{6.162}$$

The opposite-charge counterpart of Eq. (6.155),

$$\begin{aligned} i\hbar\frac{\partial}{\partial t}\phi_c(\vec{r},t) &= \frac{1}{2m}\left(\frac{\hbar}{i}\vec{\nabla}+\frac{q}{c}\vec{A}(\vec{r},t)\right)^2\phi_c(\vec{r},t)-q\Phi(\vec{r},t)\phi_c(\vec{r},t) \\ &\quad +\frac{q\hbar}{2mc}\vec{B}(\vec{r},t)\cdot\vec{\sigma}\phi_c(\vec{r},t)\,, \end{aligned} \tag{6.163}$$

is then obtained for the *conjugated Pauli spinor*, defined by

$$\phi_c(\vec{r},t) = \sigma_2\left(\tilde{\phi}^{(-)})(\vec{r},t)\right)^*\,, \tag{6.164}$$

as can be shown using the properties $\sigma_2\sigma_l^* = -\sigma_l\sigma_2$ of the Pauli matrices σ_l for $l = 1,2,3$, which result from the anticommutation relations $\{\sigma_l,\sigma_{l'}\} = 2\delta_{ll'}$ as well as from the identities $\sigma_1^* = \sigma_1$, $\sigma_2^* = -\sigma_2$, $\sigma_3^* = \sigma_3$. Charge inversion in the framework of Pauli spinors is therefore associated with complex conjugation combined with an exchange of the spin-up and the spin-down components as induced by the application of the σ_2 matrix.

6.9 Relativistic corrections

At first glance, it seems rather straightforward how leading-order relativistic corrections are to be obtained with respect to the Pauli equation (6.155), namely by the evaluation of the second term, scaling as $1/c^2$, on the right-hand side of Eq. (6.151). However, this approach would not account for a subtle but nevertheless important aspect that could be neglected so far, namely the fact that the association of the upper two components of the Dirac spinor with the particle components of the species under consideration represents a nonrelativistic approximation, which is strictly valid only in the asymptotic limit $c \to \infty$. For large but finite c, the particle components feature perturbative admixtures of the lower two components of the Dirac spinor, and *vice versa* for the antiparticle components, as was already worked out in Section 6.6. An appropriate unitary basis transformation of the Dirac spinor, also known as *Foldy-Wouthuysen transformation*, would therefore have to be applied before starting to evaluate relativistic correction terms.

This transformation can be straightforwardly evaluated in the absence of the electromagnetic field, where the Dirac equation can be analytically solved. Using the expressions (6.116) for the eigenspinors of the Dirac Hamiltonian matrix (6.106), we can express the Fourier transform of the Pauli spinor associated with the particle component as

$$\tilde{\phi}_{\vec{k}} = \sqrt{1-\delta_k^2}\tilde{\phi}^{(+)}_{\vec{k}} + \frac{\delta_k}{k}\vec{k}\cdot\vec{\sigma}\tilde{\phi}^{(-)}_{\vec{k}} \tag{6.165}$$

in terms of the Fourier transforms of the upper and lower block $\phi^{(+)}$, $\phi^{(-)}$ of the Dirac spinor, with

$$\delta_k = \frac{1}{\sqrt{2}}\sqrt{1-\frac{mc^2}{\hbar\omega_k}} \simeq \frac{\hbar k}{2mc} + \mathcal{O}\left(1/c^3\right) . \tag{6.166}$$

Using

$$\sqrt{1-\delta_k^2} \simeq 1 - \frac{\delta_k^2}{2} + \mathcal{O}\left(1/c^4\right) , \tag{6.167}$$

we obtain the Foldy-Wouthuysen transformation in configuration space as

$$\phi(\vec{r},t) \simeq \left(1-\frac{\vec{p}^2}{8m^2c^2}\right)\phi^{(+)}(\vec{r},t) + \frac{\vec{\sigma}\cdot\vec{p}}{2mc}\phi^{(-)}(\vec{r},t) + \mathcal{O}\left(1/c^4\right) \tag{6.168}$$

up to corrections that scale as $1/c^4$, with $\vec{p} = -i\hbar\vec{\nabla}$ the momentum operator.

In view of Eq. (6.168), we make the ansatz

$$\phi(\vec{r},t) \simeq \left(1-\frac{1}{2}\hat{K}^2\right)\phi^{(+)}(\vec{r},t) + \hat{K}\phi^{(-)}(\vec{r},t) + \mathcal{O}\left(1/c^4\right) \tag{6.169}$$

for the Pauli spinor representing the particle component in the presence of the electromagnetic field, where we define the operator

$$\hat{K} = \frac{\vec{\sigma}\cdot\vec{\pi}}{2mc} . \tag{6.170}$$

Using Eq. (6.150), which is rewritten as

$$\phi^{(-)}(\vec{r},t) = \hat{K}\phi^{(+)}(\vec{r},t) + \mathcal{O}\left(1/c^3\right) \tag{6.171}$$

with this new operator, we obtain from Eq. (6.169)

$$\phi(\vec{r},t) \simeq \left(1+\frac{1}{2}\hat{K}^2\right)\phi^{(+)}(\vec{r},t)+\mathcal{O}\left(1/c^4\right) . \tag{6.172}$$

As both the Pauli matrices and the components of the kinetic momentum operator $\vec{\pi}$ are hermitian, the operator $\hat{K}$ is hermitian, too. The norm of the particle component spinor is then evaluated as

$$\begin{aligned}\int d^3r\phi^\dagger(\vec{r},t)\phi(\vec{r},t) &\simeq \int d^3r\left[\left(1+\frac{1}{2}\hat{K}^2\right)\phi^{(+)}\right]^\dagger(\vec{r},t)\left(1+\frac{1}{2}\hat{K}^2\right)\phi^{(+)}(\vec{r},t)\\ &\simeq \int d^3r\left(\phi^{(+)^\dagger}(\vec{r},t)\phi^{(+)}(\vec{r},t)+(\hat{K}\phi^{(+)})^\dagger(\vec{r},t)\hat{K}\phi^{(+)}(\vec{r},t)\right)\\ &\simeq \int d^3r\int d^3r\left(\phi^{(+)^\dagger}(\vec{r},t)\phi^{(+)}(\vec{r},t)+\phi^{(-)^\dagger}(\vec{r},t)\phi^{(-)}(\vec{r},t)\right)\\ &= 1+\mathcal{O}\left(1/c^4\right) ,\end{aligned} \tag{6.173}$$

where at each step of the calculation we neglecte terms that scale as $1/c^4$. This proves the validity of the ansatz (6.169) giving rise to a Pauli spinor whose norm is preserved in the course of time evolution up to corrections of the order of $1/c^2$.

Relativistic corrections to the Pauli equation can then be properly obtained from Eq. (6.151), rewritten in terms of the operator (6.170) as

$$i\hbar D_t\phi^{(+)}(\vec{r},t)=2mc^2\hat{K}^2\phi^{(+)}(\vec{r},t)-\hat{K}i\hbar D_t\hat{K}\phi^{(+)}(\vec{r},t)+\mathcal{O}\left(1/c^4\right) , \tag{6.174}$$

by substituting $\phi^{(+)}$ with the true particle component spinor ϕ via the inversion of Eq. (6.172), which is approximately evaluated as

$$\phi^{(+)}(\vec{r},t) \simeq \left(1-\frac{1}{2}\hat{K}^2\right)\phi(\vec{r},t)+\mathcal{O}\left(1/c^4\right) . \tag{6.175}$$

This yields, up to corrections of the order $1/c^4$, the equation

$$\left(i\hbar D_t-2mc^2\hat{K}^2\right)\phi(\vec{r},t) \simeq \left(\frac{1}{2}i\hbar D_t\hat{K}^2-mc^2\hat{K}^4-\hat{K}i\hbar D_t\hat{K}\right)\phi(\vec{r},t) \tag{6.176}$$

for the Pauli spinor ϕ, which is written such that the terms constituting the Pauli equation are gathered on its left-hand side while the right-hand side contains the relativistic corrections scaling as $1/c^2$. We directly infer

$$mc^2\hat{K}^2\phi(\vec{r},t) \simeq \frac{1}{2}i\hbar D_t\phi(\vec{r},t)+\mathcal{O}\left(1/c^2\right) \tag{6.177}$$

from Eq. (6.176), which is another way to write the Pauli equation. Hence, we can substitute

$$-mc^2\hat{K}^4\phi(\vec{r},t) \simeq \frac{1}{2}\hat{K}^2i\hbar D_t\phi(\vec{r},t)-2mc^2\hat{K}^4\phi(\vec{r},t)+\mathcal{O}\left(1/c^4\right) \tag{6.178}$$

in Eq. (6.176), which allows us to rewrite this equation as

$$\left(i\hbar D_t - 2mc^2\hat{K}^2\right)\phi(\vec{r},t) \simeq \frac{1}{2}\left[[i\hbar D_t,\hat{K}],\hat{K}\right]\phi(\vec{r},t) - 2mc^2\hat{K}^4\phi(\vec{r},t) \quad (6.179)$$

up to corrections of the order $1/c^4$.

This latter form of the amended Pauli equation is particularly advantageous for further evaluation since the presence of the double commutator allows for drastic simplifications. Using Eq. (6.170) and the definitions (6.139) and (6.140) of the covariant time derivative and the kinetic momentum operator, we calculate

$$\begin{aligned}[i\hbar D_t,\hat{K}] &= \left[i\hbar\frac{\partial}{\partial t} - q\Phi(\vec{r},t), \frac{1}{2mc}\vec{\sigma}\cdot\left(\frac{\hbar}{i}\vec{\nabla} - \frac{q}{c}\vec{A}(\vec{r},t)\right)\right] \\ &= \frac{i\hbar q}{2mc}\vec{\sigma}\cdot\vec{E}(\vec{r},t) \end{aligned} \quad (6.180)$$

where

$$\vec{E}(\vec{r},t) = -\vec{\nabla}\Phi(\vec{r},t) - \frac{1}{c}\frac{\partial}{\partial t}\vec{A}(\vec{r},t) \quad (6.181)$$

is the external electric field. This yields

$$\left[[i\hbar D_t,\hat{K}],\hat{K}\right] = \frac{i\hbar q}{4m^2c^2}[\vec{\sigma}\cdot\vec{E}(\vec{r},t),\vec{\sigma}\cdot\vec{\pi}]\,. \quad (6.182)$$

The relation (6.152) characterizing the Pauli matrices allows us to express the commutator appearing in Eq. (6.182) as

$$[\vec{\sigma}\cdot\vec{E}(\vec{r},t),\vec{\sigma}\cdot\vec{\pi}] = \sum_{l=1}^{3}[E_l(\vec{r},t),\pi_l] + i\vec{\sigma}\cdot[\vec{E}(\vec{r},t)\times\vec{\pi} - \vec{\pi}\times\vec{E}(\vec{r},t)]\,. \quad (6.183)$$

Using Maxwell's equations (2.24) and (2.36), we evaluate

$$\sum_{l=1}^{3}[E_l(\vec{r},t),\pi_l] = i\hbar\vec{\nabla}\cdot\vec{E}(\vec{r},t) = 4\pi i\hbar\rho_{\rm ext}(\vec{r},t)\,, \quad (6.184)$$

with $\rho_{\rm ext}(\vec{r},t)$ the charge density that generates the external electric field (*i.e.*, which is associated with the presence of other charged particles in the system, besides the Dirac particle under study), as well as

$$\begin{aligned}\vec{\sigma}\cdot[\vec{E}(\vec{r},t)\times\vec{\pi} - \vec{\pi}\times\vec{E}(\vec{r},t)] &= 2\vec{\sigma}\cdot[\vec{E}(\vec{r},t)\times\vec{\pi}] + \frac{i\hbar}{c}\vec{\sigma}\cdot[\vec{\nabla}\times\vec{E}(\vec{r},t)] \\ &= 2\vec{\sigma}\cdot[\vec{E}(\vec{r},t)\times\vec{\pi}] - \frac{i\hbar}{c}\vec{\sigma}\cdot\frac{\partial}{\partial t}\vec{B}(\vec{r},t) \end{aligned} \quad (6.185)$$

where $\vec{B}(\vec{r},t)$ is the external magnetic field.

Inserting these expressions into Eq. (6.179) yields the equation

$$i\hbar\frac{\partial}{\partial t}\phi(\vec{r},t) = (H_P + \delta H_{\rm rel} + \delta H_D + \delta H_{LS})\,\phi(\vec{r},t) + \mathcal{O}\left(1/c^4\right) \tag{6.186}$$

for the Pauli spinor characterizing the particle component, where

$$H_P = \frac{1}{2m}\left(\frac{\hbar}{i}\vec{\nabla} - \frac{q}{c}\vec{A}(\vec{r},t)\right)^2 + q\Phi(\vec{r},t) - \frac{q\hbar}{2mc}\vec{B}(\vec{r},t)\cdot\vec{\sigma} \tag{6.187}$$

is the Hamiltonian generating the Pauli equation (6.155),

$$\delta H_{\rm rel} = -\frac{(\vec{\sigma}\cdot\vec{\pi})^4}{8m^3c^2} = -\frac{1}{8m^3c^2}\left[\left(\frac{\hbar}{i}\vec{\nabla} - \frac{q}{c}\vec{A}(\vec{r},t)\right)^2 - \frac{q\hbar}{c}\vec{B}(\vec{r},t)\cdot\vec{\sigma}\right]^2 \tag{6.188}$$

is the Hamiltonian describing relativistic corrections to the kinetic energy (in analogy with Eq. (5.91) in the framework of the Klein-Gordon theory),

$$\delta H_D = -\frac{\pi q\hbar^2}{2m^2c^2}\rho_{\rm ext}(\vec{r},t) \tag{6.189}$$

is the so-called *Darwin term*, and

$$\delta H_{LS} = -\frac{q\hbar}{4m^2c^2}\vec{\sigma}\cdot\left[\vec{E}(\vec{r},t)\times\left(\frac{\hbar}{i}\vec{\nabla} - \frac{q}{c}\vec{A}(\vec{r},t)\right)\right] + \frac{iq\hbar^2}{8m^2c^3}\vec{\sigma}\cdot\frac{\partial}{\partial t}\vec{B}(\vec{r},t) \tag{6.190}$$

is the term generating spin-orbit coupling. Equations (6.188) and (6.190) further simplify to

$$\delta H_{\rm rel} \simeq -\frac{\hbar^4\Delta^2}{8m^3c^2} + \mathcal{O}\left(1/c^4\right)\,, \tag{6.191}$$

(with Δ the Laplacian) as well as

$$\delta H_{LS} \simeq -\frac{q\hbar}{4m^2c^2}\vec{\sigma}\cdot\left(\vec{E}(\vec{r},t)\times\hat{\vec{p}}\right) + \mathcal{O}\left(1/c^4\right) \tag{6.192}$$

in a physical context that is essentially electrostatic, *i.e.*, where we can assume that the presence of a magnetic field is a relativistic effect such that we have $\vec{A},\vec{B} \propto \mathcal{O}(1/c)$. This situation is clearly encountered in a one-electron atom where the external electric field is generated by the presence of the nucleus. Placing the latter at the origin of the spatial coordinate system, we obtain the time-independent electric field

$$\vec{E}(\vec{r}) = -\vec{\nabla}\Phi(r) = -\frac{1}{r}\Phi'(r)\vec{r} \tag{6.193}$$

expressed in terms of the spherically symmetric scalar potential Φ that would be given by

$$\Phi(r) = \frac{e}{r} \tag{6.194}$$

in the case of the hydrogen atom, with e the elementary charge. The spin-orbit coupling term (6.192) then simplifies to

$$\delta H_{LS} \simeq \frac{q\hbar}{4m^2c^2r}\Phi'(r)\vec{\sigma}\cdot\hat{\vec{L}} + \mathcal{O}\left(1/c^4\right) \tag{6.195}$$

with $\hat{\vec{L}} = \vec{r} \times \vec{p}$ the angular momentum operator of the electron.

The fine structure of the atomic spectrum is then evaluated from the relativistic corrections (6.189), (6.191), and (6.195) using $q = -e$ for the charge of the electron. Those corrections represent an integral part of the Dirac equation and do not need to be postulated separately. The spectrum of hydrogen including its fine structure can therefore be also obtained via a direct solution of the Dirac equation in the presence of the scalar potential (6.194). More generally, combining Dirac's theory in its second quantized formulation (see Section 6.7) with the quantized version of electromagnetism (see Section 4.3) gives rise to the framework of *quantum electrodynamics* within which higher order corrections, of the order of $1/c^4$ and even beyond, can be quantitatively calculated with great accuracy, yielding excellent agreement with high-resolution measurements of atomic spectra. This certainly constitutes one of the greatest success stories in the history of theoretical physics.

Problem

6.1 Show that the time dependence implied by the definitions (6.124)–(6.127) of the particle and antiparticle creation and annihilation operators is consistent with the Heisenberg equations for the time evolution of those operators generated by the Hamiltonian (6.135).

Appendix A

Solutions to the problems

Problems of Chapter 2

2.1 We start from the definition (2.40) of the energy-momentum tensor. Using the product rule as well as the relation $\partial_\nu = g_{\nu\mu}\partial^\mu$, we derive

$$\partial_\nu T^{\mu\nu} = \frac{1}{4\pi}\left(\frac{1}{2}F_{\alpha\beta}\partial^\mu F^{\alpha\beta} - F^{\mu\beta}\partial^\nu F_{\nu\beta} - F_{\nu\beta}\partial^\nu F^{\mu\beta}\right) \tag{A.1}$$

$$= -\frac{1}{c}F^{\mu\nu}j_\nu + \frac{1}{4\pi}F_{\alpha\beta}\left(\frac{1}{2}\partial^\mu F^{\alpha\beta} - \partial^\alpha F^{\mu\beta}\right), \tag{A.2}$$

where in the step from Eq. (A.1) to Eq. (A.2) we make use of the inhomogeneous Maxwell equations (2.35). We now insert into Eq. (A.2) the expression (2.25) of the electromagnetic field tensor in terms of the four-potential. Assuming that the latter is twice continuously differentiable, we can use $\partial^\alpha\partial^\nu A^\beta = \partial^\nu\partial^\alpha A^\beta$ for all α, β, ν and hence rewrite Eq. (A.2) as

$$\partial_\nu T^{\mu\nu} + \frac{1}{c}F^{\mu\nu}j_\nu = \frac{1}{4\pi}\left(\partial_\alpha A_\beta - \partial_\beta A_\alpha\right)\left(\partial^\alpha\partial^\beta A^\nu - \frac{1}{2}\partial^\nu\left(\partial^\alpha A^\beta + \partial^\beta A^\alpha\right)\right)$$

$$= \frac{1}{4\pi}\left(\partial_\alpha A_\beta\right)\left(\partial^\alpha\partial^\beta - \partial^\beta\partial^\alpha\right)A^\nu = 0\,. \tag{A.3}$$

This proves Eq. (2.45).

Problems of Chapter 3

3.1 Equation (3.46) is most straightforwardly verified by inserting the inversion of the expression (3.45), namely

$$\psi(\vec{r}, t) = \psi'(\vec{r}, t)\exp\left[\frac{iq}{\hbar c}\chi(\vec{r}, t)\right], \tag{A.4}$$

into the Schrödinger equation (3.40). By applying the product rule we obtain

$$i\hbar\frac{\partial}{\partial t}\psi(\vec{r},t) = \left(i\hbar\frac{\partial}{\partial t}\psi'(\vec{r},t) - \frac{q}{c}\frac{\partial}{\partial t}\chi(\vec{r},t)\psi'(\vec{r},t)\right)\exp\left[\frac{iq}{\hbar c}\chi(\vec{r},t)\right], \quad \text{(A.5)}$$

from which follows

$$\begin{aligned}\left(i\hbar\frac{\partial}{\partial t} - q\Phi(\vec{r},t)\right)\psi(\vec{r},t) &= \exp\left[\frac{iq}{\hbar c}\chi(\vec{r},t)\right] \\ &\quad \times \left(i\hbar\frac{\partial}{\partial t} - q\Phi'(\vec{r},t)\right)\psi'(\vec{r},t) \qquad \text{(A.6)}\end{aligned}$$

using the transformed scalar potential

$$\Phi'(\vec{r},t) = \Phi(\vec{r},t) + \frac{1}{c}\frac{\partial}{\partial t}\chi(\vec{r},t) \qquad \text{(A.7)}$$

according to Eq. (3.44). Similarly, we calculate

$$\frac{\hbar}{i}\vec{\nabla}\psi(\vec{r},t) = \left(\frac{\hbar}{i}\vec{\nabla}\psi'(\vec{r},t) + \frac{q}{c}\vec{\nabla}\chi(\vec{r},t)\psi'(\vec{r},t)\right)\exp\left[\frac{iq}{\hbar c}\chi(\vec{r},t)\right], \qquad \text{(A.8)}$$

from which we infer

$$\begin{aligned}\left[\frac{\hbar}{i}\vec{\nabla} - \frac{q}{c}\vec{A}(\vec{r},t)\right]\psi(\vec{r},t) &= \exp\left[\frac{iq}{\hbar c}\chi(\vec{r},t)\right] \\ &\quad \times \left[\frac{\hbar}{i}\vec{\nabla} - \frac{q}{c}\vec{A}'(\vec{r},t)\right]\psi'(\vec{r},t) \qquad \text{(A.9)}\end{aligned}$$

and hence

$$\begin{aligned}\left[\frac{\hbar}{i}\vec{\nabla} - \frac{q}{c}\vec{A}(\vec{r},t)\right]^2\psi(\vec{r},t) &= \exp\left[\frac{iq}{\hbar c}\chi(\vec{r},t)\right] \\ &\quad \times \left[\frac{\hbar}{i}\vec{\nabla} - \frac{q}{c}\vec{A}'(\vec{r},t)\right]^2\psi'(\vec{r},t)\,, \qquad \text{(A.10)}\end{aligned}$$

using

$$\vec{A}'(\vec{r},t) = \vec{A}(\vec{r},t) - \vec{\nabla}\chi(\vec{r},t) \qquad \text{(A.11)}$$

according to Eq. (3.44). This proves Eq. (3.46).

Problems of Chapter 4

4.1 We first calculate via integration by parts

$$\int\left[\vec{\nabla}\hat{\phi}(\vec{r},t)\right]^2 d^3r = -\int\hat{\phi}(\vec{r},t)\Delta\hat{\phi}(\vec{r},t)d^3r\,, \qquad \text{(A.12)}$$

using the periodicity of the boundary conditions. From

$$\Delta e^{\pm i\vec{k}\cdot\vec{r}} = -\vec{k}^2 e^{\pm i\vec{k}\cdot\vec{r}} = -\frac{\omega_k^2}{\tilde{c}^2} e^{\pm i\vec{k}\cdot\vec{r}} \tag{A.13}$$

we obtain

$$\Delta\hat{\phi}(\vec{r},t) = -\frac{1}{\sqrt{V}} \sum_{\vec{k}} \sqrt{\frac{\hbar}{2\tilde{m}\omega_k}} \frac{\omega_k^2}{\tilde{c}^2} \left[\hat{a}_{\vec{k}}(t) e^{i\vec{k}\cdot\vec{r}} + \hat{a}^\dagger_{\vec{k}}(t) e^{-i\vec{k}\cdot\vec{r}}\right] \tag{A.14}$$

and hence

$$\begin{aligned}\hat{\phi}(\vec{r},t)\Delta\hat{\phi}(\vec{r},t) &= -\frac{\hbar}{2\tilde{m}\tilde{c}^2 V} \sum_{\vec{k},\vec{k}'} \sqrt{\frac{\omega_k^3}{\omega_{k'}}} \left[\hat{a}_{\vec{k}'}(t) e^{i\vec{k}'\cdot\vec{r}} + \hat{a}^\dagger_{\vec{k}'}(t) e^{-i\vec{k}'\cdot\vec{r}}\right] \\ &\quad \times \left[\hat{a}_{\vec{k}}(t) e^{i\vec{k}\cdot\vec{r}} + \hat{a}^\dagger_{\vec{k}}(t) e^{-i\vec{k}\cdot\vec{r}}\right] . \end{aligned} \tag{A.15}$$

Using $\int_V e^{i(\vec{k}-\vec{k}')\cdot\vec{r}} d^3r = V\delta_{\vec{k}\vec{k}'}$ for all wave vectors $\vec{k}, \vec{k}'$ whose associated plane waves satisfy the periodic boundary conditions, the integration of this latter expression over the renormalization volume yields

$$\begin{aligned}\int d^3r \hat{\phi}(\vec{r},t)\Delta\hat{\phi}(\vec{r},t) &= -\sum_{\vec{k}} \frac{\hbar\omega_k}{2\tilde{m}\tilde{c}^2} \left[\hat{a}_{\vec{k}}(t)\hat{a}^\dagger_{\vec{k}}(t) + \hat{a}^\dagger_{\vec{k}}(t)\hat{a}_{\vec{k}}(t)\right. \\ &\quad \left. + \hat{a}_{\vec{k}}(t)\hat{a}_{-\vec{k}}(t) + \hat{a}^\dagger_{\vec{k}}(t)\hat{a}^\dagger_{-\vec{k}}(t)\right] . \end{aligned} \tag{A.16}$$

Similarly, we calculate

$$\begin{aligned}\left[\hat{\Pi}(\vec{r},t)\right]^2 &= -\frac{\hbar\tilde{m}}{2V} \sum_{\vec{k},\vec{k}'} \sqrt{\omega_k\omega_{k'}} \left[\hat{a}_{\vec{k}}(t) e^{i\vec{k}\cdot\vec{r}} - \hat{a}^\dagger_{\vec{k}}(t) e^{-i\vec{k}\cdot\vec{r}}\right] \\ &\quad \times \left[\hat{a}_{\vec{k}'}(t) e^{i\vec{k}'\cdot\vec{r}} - \hat{a}^\dagger_{\vec{k}'}(t) e^{-i\vec{k}'\cdot\vec{r}}\right] , \end{aligned} \tag{A.17}$$

from which follows

$$\begin{aligned}\int d^3r \left[\hat{\Pi}(\vec{r},t)\right]^2 &= \sum_{\vec{k}} \frac{\hbar\tilde{m}\omega_k}{2} \left[\hat{a}_{\vec{k}}(t)\hat{a}^\dagger_{\vec{k}}(t) + \hat{a}^\dagger_{\vec{k}}(t)\hat{a}_{\vec{k}}(t)\right. \\ &\quad \left. - \hat{a}_{\vec{k}}(t)\hat{a}_{-\vec{k}}(t) - \hat{a}^\dagger_{\vec{k}}(t)\hat{a}^\dagger_{-\vec{k}}(t)\right] . \end{aligned} \tag{A.18}$$

This altogether yields

$$\begin{aligned}\hat{H} &= \frac{1}{2\tilde{m}} \int d^3r \left[\hat{\Pi}(\vec{r},t)\right]^2 - \frac{1}{2}\tilde{m}\tilde{c}^2 \int d^3r \hat{\phi}(\vec{r},t)\Delta\hat{\phi}(\vec{r},t) \\ &= \sum_{\vec{k}} \frac{\hbar\omega_k}{2} \left[\hat{a}_{\vec{k}}(t)\hat{a}^\dagger_{\vec{k}}(t) + \hat{a}^\dagger_{\vec{k}}(t)\hat{a}_{\vec{k}}(t)\right] \end{aligned} \tag{A.19}$$

4.2 Since the old four-potential A'^{ν} satisfies the Lorenz gauge

$$\frac{1}{c}\frac{\partial}{\partial t}\Phi'(\vec{r},t) + \vec{\nabla}\cdot\vec{A}'(\vec{r},t) = 0\,, \tag{A.20}$$

the divergence of the new vector potential, defined through

$$\vec{A}(\vec{r},t) = \vec{A}'(\vec{r},t) - \vec{\nabla}\chi(\vec{r},t)\,, \tag{A.21}$$

can be expressed as

$$\vec{\nabla}\cdot\vec{A}(\vec{r},t) = \vec{\nabla}\cdot\vec{A}'(\vec{r},t) - \Delta\chi(\vec{r},t) = -\frac{1}{c}\frac{\partial}{\partial t}\Phi'(\vec{r},t) - \Delta\chi(\vec{r},t)\,. \tag{A.22}$$

Using the fact that the old scalar potential Φ' is a solution of the wave equation (4.46)

$$\frac{1}{c^2}\frac{\partial^2}{\partial t^2}\Phi'(\vec{r},t) = \Delta\Phi'(\vec{r},t)\,, \tag{A.23}$$

we evaluate

$$\int_0^t \Delta\Phi'(\vec{r},t')dt' = \int_0^t \frac{1}{c^2}\frac{\partial^2}{\partial t'^2}\Phi'(\vec{r},t')dt' = \frac{1}{c^2}\frac{\partial\Phi'}{\partial t}(\vec{r},t) - \frac{1}{c^2}\frac{\partial\Phi'}{\partial t}(\vec{r},0) \tag{A.24}$$

and hence obtain from Eq. (4.47)

$$\begin{aligned}\Delta\chi(\vec{r},t) &= -\frac{1}{c}\frac{\partial\Phi'}{\partial t}(\vec{r},t) + \frac{1}{c}\frac{\partial\Phi'}{\partial t}(\vec{r},0) + \frac{1}{4\pi c}\int d^3r'\frac{\partial^2}{\partial\vec{r}^2}\frac{1}{|\vec{r}-\vec{r}'|}\frac{\partial\Phi'}{\partial t}(\vec{r}',0)\\ &= -\frac{1}{c}\frac{\partial\Phi'}{\partial t}(\vec{r},t)\,,\end{aligned} \tag{A.25}$$

where we use the identity $\Delta|\vec{r}|^{-1} = -4\pi\delta(\vec{r})$. This yields $\vec{\nabla}\cdot\vec{A}(\vec{r},t) = 0$.

4.3 From the quantized version of Eq. (4.54) we have $\hat{\vec{B}}(\vec{r},t) = \vec{\nabla}\times\hat{\vec{A}}(\vec{r},t)$ and can therefore express the second term of Eq. (4.67) as

$$\begin{aligned}\int d^3r\,\hat{\vec{B}}^2(\vec{r},t) &= \int d^3r\left(\vec{\nabla}\times\hat{\vec{A}}(\vec{r},t)\right)\cdot\left(\vec{\nabla}\times\hat{\vec{A}}(\vec{r},t)\right)\\ &= \int d^3r\left(\hat{\vec{A}}(\vec{r},t)\times\vec{\nabla}\right)\cdot\left(\vec{\nabla}\times\hat{\vec{A}}(\vec{r},t)\right)\end{aligned} \tag{A.26}$$

using integration by parts (and assuming that boundary terms at infinity do not play any role). We then straightforwardly rewrite

$$\begin{aligned}\left(\hat{\vec{A}}(\vec{r},t)\times\vec{\nabla}\right)\cdot\left(\vec{\nabla}\times\hat{\vec{A}}(\vec{r},t)\right) &= \hat{\vec{A}}(\vec{r},t)\cdot\left[\vec{\nabla}\times\left(\vec{\nabla}\times\hat{\vec{A}}(\vec{r},t)\right)\right]\\ &= \hat{\vec{A}}(\vec{r},t)\cdot\left[\vec{\nabla}\left(\vec{\nabla}\cdot\hat{\vec{A}}(\vec{r},t)\right) - \Delta\hat{\vec{A}}(\vec{r},t)\right]\\ &= -\hat{\vec{A}}(\vec{r},t)\cdot\Delta\hat{\vec{A}}(\vec{r},t)\end{aligned} \tag{A.27}$$

because of the gauge (4.48).

The quantized version of Eq. (4.52) is given by

$$\hat{\vec{A}}(\vec{r},t) = \frac{c}{2\pi}\int d^3k\sqrt{\frac{\hbar}{\omega_k}}\sum_{\sigma=1,2}\left(\hat{a}_{\vec{k}\sigma}(t)e^{i\vec{k}\cdot\vec{r}} + \hat{a}^\dagger_{\vec{k}\sigma}(t)e^{-i\vec{k}\cdot\vec{r}}\right)\vec{e}_\sigma(\vec{k})\,, \tag{A.28}$$

where we inserted into Eq. (4.52) the expression (4.64) for the photonic annihilation operator, its adjoint, as well as the definition (4.60) of the mass parameter $\tilde{m}$. From

$$\Delta e^{\pm i\vec{k}\cdot\vec{r}} = -\vec{k}^2 e^{\pm i\vec{k}\cdot\vec{r}} = -\frac{\omega_k^2}{c^2}e^{\pm i\vec{k}\cdot\vec{r}} \tag{A.29}$$

we calculate

$$\Delta\hat{\vec{A}}(\vec{r},t) = -\int d^3k\frac{\sqrt{\hbar\omega_k^3}}{2\pi c}\sum_{\sigma=1,2}\left(\hat{a}_{\vec{k}\sigma}(t)e^{i\vec{k}\cdot\vec{r}} + \hat{a}^\dagger_{\vec{k}\sigma}(t)e^{-i\vec{k}\cdot\vec{r}}\right)\vec{e}_\sigma(\vec{k})\,, \tag{A.30}$$

which yields

$$\begin{aligned}\hat{\vec{A}}(\vec{r},t)\cdot\Delta\hat{\vec{A}}(\vec{r},t) &= -\frac{\hbar}{(2\pi)^2}\iint d^3k\,d^3k'\sqrt{\frac{\omega_k^3}{\omega_{k'}}}\sum_{\sigma,\sigma'=1,2}\vec{e}_{\sigma'}(\vec{k}')\cdot\vec{e}_\sigma(\vec{k})\\ &\quad\times\left(\hat{a}_{\vec{k}'\sigma'}(t)e^{i\vec{k}'\cdot\vec{r}} + \hat{a}^\dagger_{\vec{k}'\sigma'}(t)e^{-i\vec{k}'\cdot\vec{r}}\right)\\ &\quad\times\left(\hat{a}_{\vec{k}\sigma}(t)e^{i\vec{k}\cdot\vec{r}} + \hat{a}^\dagger_{\vec{k}\sigma}(t)e^{-i\vec{k}\cdot\vec{r}}\right)\,.\end{aligned} \tag{A.31}$$

We then obtain

$$\begin{aligned}\int d^3r\,\hat{\vec{B}}^2(\vec{r},t) &= -\int d^3r\,\hat{\vec{A}}(\vec{r},t)\cdot\Delta\hat{\vec{A}}(\vec{r},t)\\ &= \int d^3k\,2\pi\hbar\omega_k\left[\sum_{\sigma=1,2}\left(\hat{a}_{\vec{k}\sigma}(t)\hat{a}^\dagger_{\vec{k}\sigma}(t) + \hat{a}^\dagger_{\vec{k}\sigma}(t)\hat{a}_{\vec{k}\sigma}(t)\right)\right. \qquad \text{(A.32)}\\ &\quad+\left.\sum_{\sigma,\sigma'=1,2}\vec{e}_{\sigma'}(-\vec{k})\cdot\vec{e}_\sigma(\vec{k})\left(\hat{a}_{-\vec{k}\sigma'}(t)\hat{a}_{\vec{k}\sigma}(t) + \hat{a}^\dagger_{-\vec{k}\sigma'}(t)\hat{a}^\dagger_{\vec{k}\sigma}(t)\right)\right]\end{aligned}$$

using $\int d^3r e^{i(\vec{k}-\vec{k}')\cdot\vec{r}} = (2\pi)^3\delta(\vec{k}-\vec{k}')$ and the orthogonality of the polarization vectors.

Using

$$\frac{d}{dt}\hat{a}_{\vec{k}\sigma}(t) = -i\omega_k\hat{a}_{\vec{k}\sigma}(t) \tag{A.33}$$

according to Eq. (4.64), we calculate from Eq. (A.28)

$$\frac{1}{c}\frac{\partial}{\partial t}\hat{\vec{A}}(\vec{r},t) = \int d^3k \frac{\sqrt{\hbar\omega_k}}{2\pi i} \sum_{\sigma=1,2} \left(\hat{a}_{\vec{k}\sigma}(t)e^{i\vec{k}\cdot\vec{r}} - \hat{a}^\dagger_{\vec{k}\sigma}(t)e^{-i\vec{k}\cdot\vec{r}}\right)\vec{e}_\sigma(\vec{k})\,, \quad \text{(A.34)}$$

from which follows

$$\begin{aligned}\left(\frac{1}{c}\frac{\partial}{\partial t}\hat{\vec{A}}(\vec{r},t)\right)^2 &= -\frac{\hbar}{(2\pi)^2}\iint d^3k\, d^3k' \sqrt{\omega_k\omega_{k'}} \sum_{\sigma,\sigma'=1,2} \vec{e}_{\sigma'}(\vec{k}')\cdot\vec{e}_\sigma(\vec{k}) \\ &\quad \times\left(\hat{a}_{\vec{k}'\sigma'}(t)e^{i\vec{k}'\cdot\vec{r}} - \hat{a}^\dagger_{\vec{k}'\sigma'}(t)e^{-i\vec{k}'\cdot\vec{r}}\right) \\ &\quad \times\left(\hat{a}_{\vec{k}\sigma}(t)e^{i\vec{k}\cdot\vec{r}} - \hat{a}^\dagger_{\vec{k}\sigma}(t)e^{-i\vec{k}\cdot\vec{r}}\right)\,. \qquad \text{(A.35)}\end{aligned}$$

We then calculate

$$\begin{aligned}\int d^3r\, \hat{\vec{E}}^2(\vec{r},t) &= \int d^3r \left(\frac{1}{c}\frac{\partial}{\partial t}\hat{\vec{A}}(\vec{r},t)\right)^2 \\ &= \int d^3k\, 2\pi\hbar\omega_k \left[\sum_{\sigma=1,2}\left(\hat{a}_{\vec{k}\sigma}(t)\hat{a}^\dagger_{\vec{k}\sigma}(t) + \hat{a}^\dagger_{\vec{k}\sigma}(t)\hat{a}_{\vec{k}\sigma}(t)\right)\right. \qquad \text{(A.36)} \\ &\quad \left. - \sum_{\sigma,\sigma'=1,2}\vec{e}_{\sigma'}(-\vec{k})\cdot\vec{e}_\sigma(\vec{k})\left(\hat{a}_{-\vec{k}\sigma'}(t)\hat{a}_{\vec{k}\sigma}(t) + \hat{a}^\dagger_{-\vec{k}\sigma'}(t)\hat{a}^\dagger_{\vec{k}\sigma}(t)\right)\right]\end{aligned}$$

according to the quantized version of Eq. (4.53). In combination with Eq. (A.33) and the expression (4.60) for the mass parameter, this altogether yields

$$\begin{aligned}\hat{H} &= \frac{1}{8\pi}\int d^3r\left(\hat{\vec{E}}^2(\vec{r},t) + \hat{\vec{B}}^2(\vec{r},t)\right) \\ &= \int d^3k \sum_{\sigma=1,2}\frac{\hbar\omega_k}{2}\left(\hat{a}_{\vec{k}\sigma}(t)\hat{a}^\dagger_{\vec{k}\sigma}(t) + \hat{a}^\dagger_{\vec{k}\sigma}(t)\hat{a}_{\vec{k}\sigma}(t)\right)\,. \qquad \text{(A.37)}\end{aligned}$$

Problems of Chapter 6

6.1 We start with the identities

$$\hat{b}^\dagger_{\vec{k}'\sigma'}\hat{b}_{\vec{k}'\sigma'}\hat{a}_{\vec{k}\sigma} = -\hat{b}^\dagger_{\vec{k}'\sigma'}\hat{a}_{\vec{k}\sigma}\hat{b}_{\vec{k}'\sigma'} = \hat{a}_{\vec{k}\sigma}\hat{b}^\dagger_{\vec{k}'\sigma'}\hat{b}_{\vec{k}'\sigma'}\,, \qquad \text{(A.38)}$$

$$\hat{b}^\dagger_{\vec{k}'\sigma'}\hat{b}_{\vec{k}'\sigma'}\hat{a}^\dagger_{\vec{k}\sigma} = -\hat{b}^\dagger_{\vec{k}'\sigma'}\hat{a}^\dagger_{\vec{k}\sigma}\hat{b}_{\vec{k}'\sigma'} = \hat{a}^\dagger_{\vec{k}\sigma}\hat{b}^\dagger_{\vec{k}'\sigma'}\hat{b}_{\vec{k}'\sigma'}\,, \qquad \text{(A.39)}$$

which result from the anticommutation rules (6.130). Similarly, using also Eq. (6.128), we obtain

$$\begin{aligned} \hat{a}^\dagger_{\vec{k}'\sigma'}\hat{a}_{\vec{k}'\sigma'}\hat{a}_{\vec{k}\sigma} &= -\hat{a}^\dagger_{\vec{k}'\sigma'}\hat{a}_{\vec{k}\sigma}\hat{a}_{\vec{k}'\sigma'} \\ &= \left(\hat{a}_{\vec{k}\sigma}\hat{a}^\dagger_{\vec{k}'\sigma'} - \delta_{\vec{k}\vec{k}'}\delta_{\sigma\sigma'}\right)\hat{a}_{\vec{k}'\sigma'}\,, \qquad \text{(A.40)} \\ \hat{a}^\dagger_{\vec{k}'\sigma'}\hat{a}_{\vec{k}'\sigma'}\hat{a}^\dagger_{\vec{k}\sigma} &= -\hat{a}^\dagger_{\vec{k}'\sigma'}\left(\hat{a}^\dagger_{\vec{k}\sigma}\hat{a}_{\vec{k}'\sigma'} - \delta_{\vec{k}\vec{k}'}\delta_{\sigma\sigma'}\right) \\ &= \hat{a}^\dagger_{\vec{k}\sigma}\hat{a}^\dagger_{\vec{k}'\sigma'}\hat{a}_{\vec{k}'\sigma'} + \delta_{\vec{k}\vec{k}'}\delta_{\sigma\sigma'}\hat{a}^\dagger_{\vec{k}'\sigma'}\,. \qquad \text{(A.41)} \end{aligned}$$

This yields the commutators

$$\begin{aligned} \left[\hat{a}^\dagger_{\vec{k}'\sigma'}\hat{a}_{\vec{k}'\sigma'},\hat{a}_{\vec{k}\sigma}\right] &= -\delta_{\vec{k}\vec{k}'}\delta_{\sigma\sigma'}\hat{a}_{\vec{k}\sigma}\,, \qquad \text{(A.42)} \\ \left[\hat{a}^\dagger_{\vec{k}'\sigma'}\hat{a}_{\vec{k}'\sigma'},\hat{a}^\dagger_{\vec{k}\sigma}\right] &= \delta_{\vec{k}\vec{k}'}\delta_{\sigma\sigma'}\hat{a}^\dagger_{\vec{k}\sigma}\,, \qquad \text{(A.43)} \\ \left[\hat{b}^\dagger_{\vec{k}'\sigma'}\hat{b}_{\vec{k}'\sigma'},\hat{a}_{\vec{k}\sigma}\right] = \left[\hat{b}^\dagger_{\vec{k}'\sigma'}\hat{b}_{\vec{k}'\sigma'},\hat{a}^\dagger_{\vec{k}\sigma}\right] &= 0 \qquad \text{(A.44)} \end{aligned}$$

for all $\vec{k},\vec{k}',\sigma,\sigma'$. Using the expression (6.135) for the Hamiltonian, the Heisenberg equations describing the time evolution of the operators $\hat{a}_{\vec{k}\sigma}$ and $\hat{a}^\dagger_{\vec{k}\sigma}$ are then simplified as

$$\begin{aligned} \frac{d}{dt}\hat{a}_{\vec{k}\sigma} = \frac{i}{\hbar}\left[\hat{H},\hat{a}_{\vec{k}\sigma}\right] &= \sum_{\vec{k}'}\sum_{\sigma'=1,2} i\omega_{k'}\left[\hat{a}^\dagger_{\vec{k}'\sigma'}\hat{a}_{\vec{k}'\sigma'} + \hat{b}^\dagger_{\vec{k}'\sigma'}\hat{b}_{\vec{k}'\sigma'},\hat{a}_{\vec{k}\sigma}\right] \\ &= -i\omega_k\hat{a}_{\vec{k}\sigma}\,, \qquad \text{(A.45)} \\ \frac{d}{dt}\hat{a}^\dagger_{\vec{k}\sigma} = \frac{i}{\hbar}\left[\hat{H},\hat{a}^\dagger_{\vec{k}\sigma}\right] &= \sum_{\vec{k}'}\sum_{\sigma'=1,2} i\omega_{k'}\left[\hat{a}^\dagger_{\vec{k}'\sigma'}\hat{a}_{\vec{k}'\sigma'} + \hat{b}^\dagger_{\vec{k}'\sigma'}\hat{b}_{\vec{k}'\sigma'},\hat{a}^\dagger_{\vec{k}\sigma}\right] \\ &= i\omega_k\hat{a}^\dagger_{\vec{k}\sigma}\,. \qquad \text{(A.46)} \end{aligned}$$

Their general solutions are straightforwardly calculated and agree perfectly with the time dependence that was involved in the definitions (6.124) and (6.125) of those creation and annihilation operators. This agreement can also be shown for the antiparticle operators $\hat{b}_{\vec{k}\sigma},\hat{b}^\dagger_{\vec{k}\sigma}$ in a perfectly analogous manner.