2 Boson Fields

Nonrelativistic quantum mechanics, useful as it is in the formulation of all fields of modern physics and in their applications, nevertheless has limitations. In particular, it is not generally applicable to the study of elementary particles because it cannot properly predict the dynamic behavior of systems evolving at high velocities and, in its usual formulation, cannot account for the phenomena of creation and annihilation of particles that regularly occur at high energies. Two concepts – special relativity and field – have crucially contributed to the creation of relativistic quantum field theory, which has unquestionably become the foundation of particle physics. It is then possible, for example, to have a natural explanation for the existence of spins of particles, to make the connection between particles and interactions, and to associate to each particle a charge conjugate particle.

In nonrelativistic quantum mechanics, the wave function $\phi(t, \boldsymbol{x})$ of a particle of mass m in the absence of any interaction obeys the Schrödinger equation, written in natural units with $\hbar = c = 1$,

$$\mathrm{i}\frac{\partial}{\partial t}\phi(t, \boldsymbol{x}) = -\frac{1}{2m}\nabla^2\,\phi(t, \boldsymbol{x})\,. \tag{2.1}$$

Comparison with the energy-momentum relation for a nonrelativistic particle

$$E = \boldsymbol{p}^2/2m \tag{2.2}$$

suggests the correspondence rules

$$E \to\ \mathrm{i}\frac{\partial}{\partial t}\,; \qquad \boldsymbol{p} \to -\mathrm{i}\nabla = -\mathrm{i}\left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}\right)\,. \tag{2.3}$$

An equation that generalizes (1) to the relativistic regime must have at least a homogeneous coordinate dependence and be symmetric in space and time, a requirement that suggests two possibilities; one, involving only second derivatives and of the general form

$$a\frac{\partial^2}{\partial t^2}\phi(t, \boldsymbol{x}) = \left(\sum_{i=1}^{3} b^i \frac{\partial^2}{\partial x^{i2}} + m^2\right)\phi(t, \boldsymbol{x})\,, \tag{2.4}$$

and the other, involving only first derivatives,

$$\mathrm{i}\gamma^0 \frac{\partial}{\partial t}\phi(t,\boldsymbol{x}) = \left(-\mathrm{i}\sum_{i=1}^{3}\gamma^i \frac{\partial}{\partial x^i} + m\right)\phi(t,\boldsymbol{x})\,. \tag{2.5}$$

In these equations, a, b^i, γ^0, γ^i, and m are *a priori* complex constants to be determined in accordance with the relativistic energy-momentum relation for a free particle,

$$E^2 = \boldsymbol{p}^2 + m^2\,. \tag{2.6}$$

So it becomes apparent that relativity will be called on to play a key role in the solution to this problem and justifies a careful examination of Lorentz invariance. It forms the subject of the following section. In the next four sections we discuss the second-quantized scalar and vector field solutions to (4), called the Klein–Gordon equation. We close the chapter with considerations of the action function and of Noether's theorem.

2.1 Lorentz Symmetry

In Newtonian physics, physical laws are stated in terms of equations that preserve their forms when coordinates are changed by a Galilean transformation (one that independently changes the position vector and the time parameter by constant amounts). This invariance, called the *Galilean invariance*, characterizes Newtonian mechanics. We now know, since Einstein, that nature is endowed with a higher symmetry, including Galilean symmetry as a special case. The transformations that define this symmetry are called the Lorentz transformations; they leave invariant (unchanged in magnitude) the speed of light, and covariant (unchanged in form) the Maxwell equations and, in general, all physical equations. Therefore, it is important to have a detailed look at what constitutes the fundamental principle of special relativity and indeed of all contemporary physics.

2.1.1 Lorentz Transformations

As the coordinates of space and time are to be treated on the same footing, it is convenient to gather them into a single four-component object which behaves by its transformation properties as a vector in a four-dimensional space that one may call *space-time:*

$$x^\mu = (x^0,\, x^1,\, x^2,\, x^3) = (t, \boldsymbol{x})\,. \tag{2.7}$$

In this notation, μ and other Greek indices take on the values $0, 1, 2, 3$; the coordinate x^0 stands for the time parameter t, and x^1, x^2, x^3 (or x, y, z) are the Cartesian components of the usual position vector $\boldsymbol{x}$. We will use Latin indices $i, j, \ldots$, restricted to values $1, 2, 3$, to indicate space components.

A Lorentz transformation is a transformation of some coordinate system $\{x^\mu\}$ into another system $\{x'^\mu\}$ such that

$$\begin{aligned} x'^\mu &= a^\mu{}_0\, x^0 + a^\mu{}_1\, x^1 + a^\mu{}_2\, x^2 + a^\mu{}_3\, x^3 \\ &= a^\mu{}_\nu\, x^\nu \,. \end{aligned} \tag{2.8}$$

We follow here the usual convention that a summation over repeated indices (taking all admissible values) is understood. The parameters of the transformation, $a^\mu{}_\nu$, are *real constants* that specify the relative spatial orientations and the relative velocities of the two reference frames. One can recognize in (8) a generalization to four dimensions of the familiar rotation laws relating to cartesian coordinates. Let us consider some examples.

Example 2.1 Space Rotation
A rotation of a coordinate system about the 3 or z axis by a positive angle θ in the *counterclockwise* direction is defined by the real coefficients $a^\mu{}_\nu$ arranged as the elements of a 4×4 matrix, where the first index $\mu = 0, 1, 2, 3$ labels the columns and the second index $\nu = 0, 1, 2, 3$ labels the rows,

$$a^\mu{}_\nu = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\theta & \sin\theta & 0 \\ 0 & -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} , \tag{2.9}$$

where $0 \le \theta < 2\pi$. Note that the matrix is unimodular, $\det a = 1$. As with any other space rotation, this particular rotation mixes space coordinates, leaving untouched the time component:

$$\begin{aligned} x'^0 &= x^0 \,, \\ x'^1 &= \ x^1 \cos\theta + x^2 \sin\theta \,, \\ x'^2 &= -x^1 \sin\theta + x^2 \cos\theta \,, \\ x'^3 &= x^3 \,. \end{aligned} \tag{2.10}$$

Occasionally, it is useful to define the complex combinations

$$x^{(\pm)} = \mp \frac{1}{\sqrt{2}} (x^1 \pm \mathrm{i} x^2) \tag{2.11}$$

(called the circular or spherical components), which transform as

$$x'^{(\pm)} = \mathrm{e}^{\mp \mathrm{i}\theta}\, x^{(\pm)} \,. \tag{2.12}$$

When θ is very small we may use a linear approximation in which $a^\mu{}_\nu = \delta^\mu{}_\nu + \epsilon^\mu{}_\nu$ for

$$\epsilon^\mu{}_\nu = \theta \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} . \tag{2.13}$$

As defined in (10), $a^{\mu}{}_{\nu}$ describes a *passive rotation*, which leaves the physical system unchanged. In an *active rotation*, where the physical system (rather than the reference system) is rotated (clockwise about the z axis), the coordinates of the transformed object and of the original object in the same reference system are still related by (10) but with θ replaced by $-\theta$ (see Fig. 2.1).

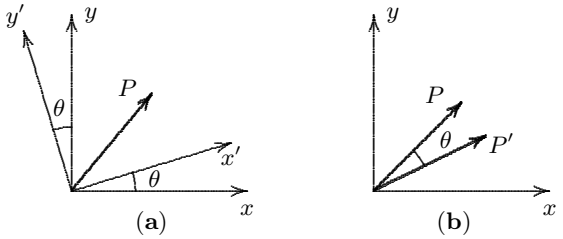

Fig. 2.1. Equivalent rotations about the $\hat{z}$ axis: (**a**) reference axes are rotated; (**b**) physical system P is rotated

Example 2.2 Pure Lorentz Transformation
A pure Lorentz transformation (Lorentz boost) relates two reference frames which differ only by a uniform relative motion of velocity v. When the motion is in the positive x direction, the transformation is given by the matrix

$$a^{\mu}{}_{\nu} = \begin{pmatrix} \cosh\omega & -\sinh\omega & 0 & 0 \\ -\sinh\omega & \cosh\omega & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \tag{2.14}$$

having determinant $\det a = 1$. It is a kind of rotation that mixes space coordinates with the time parameter:

$$\begin{aligned} x'^0 &= \cosh\omega\,(\,x^0 - x^1\tanh\omega) = \gamma\,(\,x^0 - vx^1)\,, \\ x'^1 &= \cosh\omega\,(-x^0\tanh\omega + x^1) = \gamma\,(-vx^0 + x^1)\,, \\ x'^2 &= x^2\,, \\ x'^3 &= x^3\,, \end{aligned} \tag{2.15}$$

where $-\infty < \omega < \infty$, and $\cosh\omega = \gamma = 1/\sqrt{1-v^2}$, or $\tanh\omega = v$.

If an *active* transformation is considered, where a particle of mass m at rest is boosted to velocity v in the positive x direction, the coordinates of the two particle states, all measured in the same, unchanged reference, are related by

$$\begin{aligned} x'^0 &= x^0\cosh\omega + x^1\sinh\omega\,, \\ x'^1 &= x^0\sinh\omega + x^1\cosh\omega\,, \\ x'^2 &= x^2\,, \\ x'^3 &= x^3\,. \end{aligned} \tag{2.16}$$

It will be seen below that the energy and momentum form a four-vector $p^\mu = (E, \boldsymbol{p})$ which transforms as x^μ, so that the two vectors $(m\,, 0, 0, 0)$ and $(E\,, p, 0, 0\,)$ are related by

$$\begin{aligned} E &= m\,\cosh\omega\,, \\ p &= m\,\sinh\omega\,, \end{aligned} \tag{2.17}$$

from which $\tanh\omega = p/E = v$, $\cosh\omega = 1/\sqrt{1-v^2}$, and $\sinh\omega = v/\sqrt{1-v^2}$. A particle of mass m at rest acquires through a Lorentz boost an energy $E = m\,\cosh\omega$ and momentum $p = m\,\sinh\omega$. The parameter ω is called the particle *rapidity*.

Example 2.3 Space Inversion
An inversion in space is defined by the matrix

$$a^\mu{}_\nu = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} . \tag{2.18}$$

Note that in this case $\det a = -1$. The coordinates then transform as $x'^0 = x^0$, $x'^1 = -x^1$, $x'^2 = -x^2$, and $x'^3 = -x^3$. ■

The distinctive property of Lorentz transformations is that they leave invariant the proper time interval $\mathrm{d}\tau$, defined by

$$\mathrm{d}\tau^2 \equiv \mathrm{d}t^2 \,-\, \mathrm{d}\boldsymbol{x}^2 = g_{\mu\nu}\,\mathrm{d}x^\mu\,\mathrm{d}x^\nu\,. \tag{2.19}$$

The symbol $g_{\mu\nu}$ stands for a tensor, called the space-time metric, with components $g_{00} = 1$, $g_{11} = g_{22} = g_{33} = -1$, and $g_{\mu\nu} = 0$ for $\mu \neq \nu$, which can be represented by a matrix,

$$g_{\mu\nu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} . \tag{2.20}$$

It is numerically equal to its inverse, $g^{\mu\nu}$, which carries upper indices.

In another Lorentz frame, the infinitesimal elements of the coordinates are given, according to (8) with constant a^μ_ν, by

$$\mathrm{d}x'^\mu = a^\mu{}_\nu\,\mathrm{d}x^\nu\,, \tag{2.21}$$

and the proper time interval is given by

$$\mathrm{d}\tau'^2 = g_{\mu\nu}\,\mathrm{d}x'^\mu\,\mathrm{d}x'^\nu = g_{\mu\nu}\,a^\mu{}_\rho\,a^\nu{}_\sigma\,\mathrm{d}x^\rho\,\mathrm{d}x^\sigma\,. \tag{2.22}$$

Invariance of proper time,

$$\mathrm{d}\tau'^2 = \mathrm{d}\tau^2\,, \tag{2.23}$$

which expresses the experimental observation that the speed of light in the vacuum is the same in all inertial frames, imposes a condition on the Lorentz transformation matrix similar to the orthogonality relation for the space rotation matrix,

$$g_{\mu\nu}\, a^{\mu}{}_{\lambda}\, a^{\nu}{}_{\kappa} = g_{\lambda\kappa}\,. \tag{2.24}$$

We limit ourselves to *real* Lorentz transformations. This is in fact the case of physical transformations, which map real (coordinate) space into real space; but complex extensions are also possible. From (24), two conditions on a^{μ}_{ν} may be written down:

$$(\det\, a)^2 = 1\,; \tag{2.25}$$

$$a_{00}^2 - \sum_i a_{ii}^2 = 1\,. \tag{2.26}$$

They divide the real Lorentz transformations into four classes, namely,

(1) $L_+^{\uparrow}$	$\det a = +1$	$a_{00} \geq 1$	proper orthochronous	(1),
(2) $L_+^{\downarrow}$	$\det a = +1$	$a_{00} \leq -1$	proper nonorthochronous	(TP),
(3) $L_-^{\uparrow}$	$\det a = -1$	$a_{00} \geq 1$	improper orthochronous	(P),
(4) $L_-^{\downarrow}$	$\det a = -1$	$a_{00} \leq -1$	improper nonorthochronous	(T).

These four classes are disconnected because neither a_{00} nor $\det a$ can be changed continuously from a value less than 1 to a value greater than 1. But a transformation in each class can be continuously deformed into any other transformation of that class and in particular to the basic transformation characteristic of the class, namely, 1, P, T, or TP, where P is space inversion (parity) and T is time inversion.

Remarks. The set of all Lorentz transformations constitute an algebraic structure, called the *Lorentz group*. It has the three key defining properties of a group; namely,

(a) there exists an identity transformation (which effects no changes at all);
(b) to each transformation there corresponds an inverse which is also a member of the set;
(c) two transformations successively applied are equivalent to some element of the set.

In particular, the subset of proper orthochronous Lorentz transformations form a subgroup of the Lorentz group, called the *s*pecial *o*rthogonal group SO(3,1), where the notation reflects the condition $\det a = +1$ and the asymmetry between space and time as manifest in the metric. In general, it is this specific group one refers to when one speaks of 'Lorentz invariance', and it is to this group that we will limit our discussion for the rest of the chapter.

2.1.2 Tensor Algebra

Any vector that transforms as x^μ according to (8) is said to be a *contravariant* (Lorentz) vector,

$$V^\mu \to V'^\mu = a^\mu{}_\nu \, V^\nu \,. \tag{2.27}$$

A *covariant* vector U_μ is one that transforms as

$$U_\mu \to U'_\mu = a_\mu{}^\nu \, U_\nu \,, \tag{2.28}$$

that is, with the matrix inverse of $a^\mu{}_\nu$:

$$a_\mu{}^\nu \equiv (a^{-1})^\nu{}_\mu = g_{\mu\lambda} \, g^{\nu\kappa} \, a^\lambda{}_\kappa \,. \tag{2.29}$$

From these definitions and (24) follow several useful relations. First, to each contravariant vector corresponds a covariant vector, and inversely,

$$V_\mu \equiv g_{\mu\nu} \, V^\nu \,, \qquad U^\mu \equiv g^{\mu\nu} \, U_\nu \,. \tag{2.30}$$

Note in particular that the sign of the space components changes when indices change positions, reflecting the presence of both signs in $g_{\mu\nu}$:

$$V_0 = +V^0, \;\; V_1 = -V^1, \;\; V_2 = -V^2, \;\; V_3 = -V^3 \,. \tag{2.31}$$

Moreover, the scalar product of a covariant vector and a contravariant vector

$$U_\nu V^\nu = U^0 V^0 - \boldsymbol{U} \cdot \boldsymbol{V} \tag{2.32}$$

is a Lorentz-invariant scalar:

$$U'_\mu V'^\mu = a_\mu{}^\lambda \, a^\mu{}_\nu \, U_\lambda V^\nu = U_\nu V^\nu . \tag{2.33}$$

There exist objects, neither vectors nor scalars, that also transform in a well-defined though complicated manner and that may carry several upper or lower indices. They are called (Lorentz) *tensors*. For example, a rank-3 mixed tensor transforms according to

$$T^\lambda{}_{\mu\nu} \to T'^\lambda{}_{\mu\nu} = a^\lambda{}_\alpha \, T^\alpha{}_{\beta\gamma} \, a_\mu{}^\beta \, a_\nu{}^\gamma \,. \tag{2.34}$$

This rule can be readily extended to tensors of any rank.

The four-gradient, $\partial_\mu \equiv \partial/\partial x^\mu$, is a covariant vector:

$$\frac{\partial}{\partial x^\mu} \to \frac{\partial}{\partial x'^\mu} = \frac{\partial x^\nu}{\partial x'^\mu} \, \frac{\partial}{\partial x^\nu} \,, \tag{2.35}$$

where $\partial x^\nu/\partial x'^\mu = a_\mu{}^\nu$, from (28). One can similarly prove that $\partial^\mu \equiv \partial/\partial x_\mu$ is a contravariant vector. It is important to note the sign difference in the following two formulas:

$$\partial_\mu \equiv \frac{\partial}{\partial x^\mu} = \left(\frac{\partial}{\partial t}, \nabla\right), \qquad \partial^\mu \equiv \frac{\partial}{\partial x_\mu} = \left(\frac{\partial}{\partial t}, -\nabla\right). \tag{2.36}$$

It follows, for example, that the divergence of an arbitrary contravariant vector, $\partial V^\mu/\partial x^\mu$, is invariant, as are also the d'Alembertian $\Box = \partial^\mu\partial_\mu$ and any operator of the form $V^\mu\partial_\mu$:

$$V^\mu\partial_\mu = V^0\frac{\partial}{\partial t} + \boldsymbol{V}\cdot\nabla, \tag{2.37}$$

$$\partial^\mu\partial_\mu = \frac{\partial^2}{\partial t^2} - \nabla^2. \tag{2.38}$$

Another example of physical interest is the energy-momentum vector of a particle with mass m, defined by

$$p^\mu \equiv m\frac{\mathrm{d}x^\mu}{\mathrm{d}\tau}. \tag{2.39}$$

It is evidently a vector because $\mathrm{d}x^\mu$ is one while m and $\mathrm{d}\tau$ are both invariants. The particle energy and momentum can then be identified with the time and space components of p^μ:

$$p^0 = E = m\gamma, \tag{2.40}$$

$$\boldsymbol{p} = m\gamma\boldsymbol{v}, \tag{2.41}$$

where

$$\boldsymbol{v} \equiv \frac{\mathrm{d}\boldsymbol{x}}{\mathrm{d}t}, \qquad \gamma \equiv \frac{\mathrm{d}t}{\mathrm{d}\tau} = (1-\boldsymbol{v}^2)^{-1/2}. \tag{2.42}$$

If the total energy-momentum is conserved in a reaction i $\to$ f in some reference frame (i.e. $P_\mathrm{i}^\mu = P_\mathrm{f}^\mu$, where P_i^μ and P_f^μ denote the total energy-momentum in the initial and final states), it is also conserved in any other frame related to the first by a Lorentz transformation (i.e. $P_\mathrm{i}'^\mu = P_\mathrm{f}'^\mu$). This simple yet significant result follows directly from (8) because a four-vector that vanishes in a given frame necessarily vanishes in any other Lorentz frame.

2.1.3 Tensor Fields

Consider two observers $\mathcal{O}$ and $\mathcal{O}'$ moving in two different inertial reference frames related by a Lorentz transformation. If observer $\mathcal{O}$ describes a field by a certain function $\varphi(x) = \varphi(t, \boldsymbol{x})$ using the coordinates of her own frame, then observer $\mathcal{O}'$ will describe the same field by another function $\varphi'(x') = \varphi'(t', \boldsymbol{x}')$ in terms of the transformed coordinates $x'^\mu = a^\mu{}_\nu x^\nu$. The question is, how

are $\varphi(x)$ and $\varphi'(x')$ related? A theory consistent with relativistic principles can only contain fields that have well-defined transformation properties. They include

(a) *scalar fields*, that remain invariant in every Lorentz transformation: $\phi'(x') = \phi(x)$;

(b) *vector fields*, such as the electromagnetic field; a Lorentz transformation acts on both the field and its arguments such that $A'^{\mu}(x') = a^{\mu}{}_{\nu}\, A^{\nu}(x)$;

(c) *tensor fields*, such as the gravitational field or the electromagnetic field tensor, for which the transformation rule is $F'^{\mu\nu}(x') = a^{\mu}{}_{\rho}\, a^{\nu}{}_{\sigma}\, F^{\rho\sigma}(x)$.

One characterizes these rules by saying that ϕ, A^{μ}, and $F^{\mu\nu}$ belong to different *representations* of the Lorentz group. There also exist many other representations that cannot be constructed in such a simple manner. For example, the *spinor representation*, to which belong fields of spin-$1/2$ particles in four-dimensional space-time, also transforms with a 4×4 matrix with elements given by nontrivial functions of $a^{\mu}{}_{\nu}$. We will examine this case in the next chapter. The rest of this chapter is devoted to the study of the two simplest representations of the Lorentz group, the scalar and the vector fields.

2.2 Scalar Fields

The observed mesons π, K, and η, and the postulated Higgs bosons of the standard model are all spin-0 particles, described by scalar fields generically represented by the symbol ϕ. Suppose an observer $\mathcal{O}$ constructs in her frame the coordinates x^{μ} and describes a certain scalar field by the space-time function $\phi(x)$ and the state vector of the observed physical system by Φ_A. Suppose also a second observer $\mathcal{O}'$, with similarly constructed coordinates x'^{μ}, describes the same field by $\phi'(x')$ and the same physical state by the vector $\Phi'_{A'}$. How are the corresponding objects in the two frames related? What observables do such relations imply?

2.2.1 Space-time Translation of a Scalar Field

Let us illustrate the way symmetry arguments are applied to fields by the simple example of space-time translation, which is defined by

$$x^{\mu} \to x'^{\mu} = x^{\mu} - a^{\mu}\,, \tag{2.43}$$

where a^{μ} is the constant displacement parameter. As translation is a continuous transformation and can be constructed by a succession of small translations, it suffices to consider an infinitesimal transformation,

$$x^{\mu} \to x'^{\mu} = x^{\mu} - \delta a^{\mu}\,, \tag{2.44}$$

where δa^μ is a very small constant vector. If the *total variation* of the field is defined as

$$\delta\phi(x) = \phi'(x') - \phi(x)\,, \tag{2.45}$$

then $\delta\phi(x) = 0$, since $\phi(x)$ is by definition a scalar, invariant field. The question then is, if $\mathcal{O}'$ is given $\phi(x)$, how will he obtain from this information the field in his own frame but written in terms of x? What he wants is an expression of the form

$$\phi'(x) = \phi(x) + \delta_0\phi(x)\,, \tag{2.46}$$

where $\delta_0\phi$ defines the variation of the field alone, keeping the argument fixed. Setting $\delta\phi = 0$ in (45) and calling x' simply x, observer $\mathcal{O}'$ will obtain by expanding $\phi(x + \delta a)$ up to terms linear in δa^μ

$$\phi'(x) = \phi(x^\mu + \delta a^\mu) \approx (1 + \delta a^\mu \partial_\mu)\phi(x)\,. \tag{2.47}$$

Therefore,

$$\delta_0\phi = \mathrm{i}\delta a^\mu(-\mathrm{i}\partial_\mu\phi(x))\,, \tag{2.48}$$

where $-\mathrm{i}\partial_\mu$ induces an infinitesimal variation of the field and is called, for this reason, the *generator* of infinitesimal translations. A finite translation is obtained by replication,

$$\phi'(x) = \mathrm{limit}_{\,n\to\infty}\left(1 + \frac{a^\mu}{n}\partial_\mu\right)^n \phi(x) = \exp(a^\mu\partial_\mu)\,\phi(x)\,. \tag{2.49}$$

The operator $U(a) \equiv \exp(a^\mu\partial_\mu)$ is unitary: by Hermitian conjugation one has $\partial_\mu^\dagger = -\partial_\mu$, hence $U^\dagger(a) = \exp(-a^\mu\partial_\mu) = U^{-1}(a) = U(-a)$, which can also be written as $U^\dagger U = 1$. The set of all $U(a)$ form a group, the *translation group*. This group is *Abelian*, meaning that the result of two successive translations does not depend on the order in which they are applied:

$$U(a)U(a') = U(a')U(a)\,. \tag{2.50}$$

This property is equivalent to commutativity of the generators,

$$\partial_\mu\partial_\nu - \partial_\nu\partial_\mu \equiv [\partial_\mu, \partial_\nu] = 0\,, \tag{2.51}$$

and $\{\partial_\mu\}$ is said to form an *Abelian algebra*.

At this point it is useful to introduce a Hermitian operator, $P_\mu = P_\mu^\dagger$, realized in x space by $-\mathrm{i}\partial_\mu$, and interpreted, just as in quantum mechanics, as the total energy-momentum operator of the system. The transformation operator, similarly abstracted, is then given by $U(a) = \exp(\mathrm{i}a^\mu P_\mu)$. Two

vectors in the Hilbert space constructed respectively by the two observers to describe the same physical state are then related by

$$\Phi'_{A'} = U(a)\,\Phi_A\,, \tag{2.52}$$

and the operators representing the same observable are similarly related:

$$X' = U(a)XU^{-1}(a)\,. \tag{2.53}$$

These transformation rules are motivated by the general physical condition that a scalar product in the Hilbert space, interpreted as usual as a probability amplitude, remains unchanged in any symmetry transformation,

$$\langle\,\Phi'_{B'}\,|\,X'\,|\,\Phi'_{A'}\rangle\;=\;\langle\,\Phi_B\,|\,X\,|\,\Phi_A\rangle\;. \tag{2.54}$$

Now, X is said to be invariant to the transformation if

$$X' = X \quad \text{or} \quad UX = XU\,; \tag{2.55}$$

that is, it commutes with all the generators of the transformation group,

$$[X, P_\mu] = 0\,, \qquad \mu = 0, 1, 2, 3\,. \tag{2.56}$$

In particular, the physical system described by a Hamiltonian H is invariant to translations if $[H, P_\mu] = 0$, for $\mu = 0, 1, 2, 3$. Considered as a Heisenberg operator, P_μ is a constant of the motion; that is, the total energy and momentum of the system are conserved. This conservation law, which is a direct consequence of the invariance under constant displacements of space-time coordinates, is valid whenever cosmological effects are negligible. The energy-momentum conservation law turns out to be among the most useful tools in particle physics.

Remarks. (a) Relation (54) can be understood as follows. Let a, b be complex numbers, and $|\varphi_i\rangle$ arbitrary vectors. An operator U is said to be *unitary* if it is *linear*,

$$U(a\;|\varphi_1\rangle + b\;|\varphi_2\rangle) = \;a\,U\;|\varphi_1\rangle + b\,U\;|\varphi_2\rangle\;, \tag{2.57}$$

and *preserves the norm* of every vector

$$\langle U\varphi\,|U\varphi\,\rangle = \langle\varphi\,|\varphi\,\rangle\;. \tag{2.58}$$

It immediately follows that the scalar product is also preserved:

$$\langle U\varphi_1\,|U\varphi_2\,\rangle = \langle\varphi_1\,|\varphi_2\,\rangle\;. \tag{2.59}$$

(b) In quantum mechanics, a physical system is most often described by a wave function which evolves according to the Schrödinger equation which, for a time-independent Hamiltonian, can be formally solved to yield

$$\phi(t,\boldsymbol{x}) = \mathrm{e}^{-\mathrm{i}Ht}\phi(0,\boldsymbol{x})\,. \tag{2.60}$$

The matrix element of an arbitrary operator A (assumed for simplicity to be time independent) may be written as

$$\int \mathrm{d}^3x\,\phi_f^*(t,\boldsymbol{x})A\,\phi_i(t,\boldsymbol{x}) = \int \mathrm{d}^3x\,\phi_f^*(0,\boldsymbol{x})\,A(t)\,\phi_i(0,\boldsymbol{x})\,. \tag{2.61}$$

The time-dependent operator thus defined,

$$A(t) \equiv \mathrm{e}^{\mathrm{i}Ht}A\,\mathrm{e}^{-\mathrm{i}Ht}\,, \tag{2.62}$$

satisfies what is known as the Heisenberg equation:

$$\mathrm{i}\,\frac{\mathrm{d}A(t)}{\mathrm{d}t} = [A(t),\,H]\,. \tag{2.63}$$

Thus, a quantum system may be described either by time-dependent wave functions of the Schrödinger representation or, alternatively, by fixed state vectors and time-varying operators, like $A(t)$, of the Heisenberg representation. When $[A(t),H]=0$, the operator $A(t)$ does not depend on time and, according to (61), its expectation value in an arbitrary state is constant. We shall return to this subject in Chap. 4.

2.2.2 Lorentz Transformation of a Scalar Field

We proceed now to the study of the Lorentz transformation properties of a scalar field. For this purpose, it suffices to consider infinitesimal transformations, defined by

$$x^\mu \to x'^\mu = a^\mu{}_\nu\, x^\nu \approx (\delta^\mu{}_\nu + \epsilon^\mu{}_\nu)\, x^\nu\,, \tag{2.64}$$

where $\epsilon_{\mu\nu} = g_{\mu\lambda}\epsilon^\lambda{}_\nu$ are all very small constants. The basic condition (24) on the transformation matrix tells us that $\epsilon_{\mu\nu}$ is antisymmetric: $\epsilon_{\mu\nu} = -\epsilon_{\nu\mu}$. And so in the examples considered above the infinitesimal rotation $(\hat{z},\delta\theta)$ is defined by the nonvanishing matrix elements $\epsilon_{12} = -\epsilon_{21} = -\delta\theta$, while the Lorentz boost $(\hat{x},\delta\omega)$ is defined by the nonzero elements $\epsilon_{01} = -\epsilon_{10} = -\delta\omega$.

Now, from the defining property of a scalar field, namely

$$\phi'(x') = \phi(x) = \phi(a^{-1}x')\,, \tag{2.65}$$

one obtains to the first-order terms in the variation

$$\begin{aligned}\phi'(x) &= \phi(a^{-1}x) = \phi(x^\mu - \epsilon^\mu{}_\nu\, x^\nu)\\ &\approx (1 - \epsilon^\mu{}_\nu\, x^\nu\partial_\mu)\,\phi(x)\,.\end{aligned} \tag{2.66}$$

The intrinsic variations of the field immediately follow:

$$\begin{aligned}\delta_0\phi &= \phi'(x) - \phi(x) \\ &= \tfrac{1}{2}\epsilon^{\mu\nu}\,(x_\mu\partial_\nu - x_\nu\partial_\mu)\,\phi(x) \equiv -\tfrac{\mathrm{i}}{2}\epsilon^{\mu\nu}\,L_{\mu\nu}\,\phi(x)\,. \end{aligned} \tag{2.67}$$

This result shows that the infinitesimal Lorentz transformations of scalar fields are essentially described by the operators

$$L_{\mu\nu} = \mathrm{i}(x_\mu\partial_\nu - x_\nu\partial_\mu)\,, \tag{2.68}$$

called the generators of infinitesimal transformations. They are Hermitian, ${L^\dagger}_{\mu\nu} = L_{\mu\nu}$, antisymmetric, $L_{\mu\nu} = -L_{\nu\mu}$, and form a complete set in the sense that

$$[L_{\mu\nu}, L_{\rho\sigma}] = -\mathrm{i}g_{\mu\rho}\,L_{\nu\sigma} - \mathrm{i}g_{\nu\sigma}\,L_{\mu\rho} + \mathrm{i}g_{\mu\sigma}\,L_{\nu\rho} + \mathrm{i}g_{\nu\rho}\,L_{\mu\sigma}\,. \tag{2.69}$$

The lack in commutativity of these generators reflects the fact that the order of successive applications of two arbitrary Lorentz transformations is important. The result of the application of a Lorentz boost followed by a rotation, for example, is not the same as that obtained when the order of applications is reversed. Hence the Lorentz group and the corresponding algebra of generators are *non-Abelian*.

An observable X is said to be invariant to the Lorentz group when it commutes with all generators $L_{\mu\nu}$. In particular, a physical system is said to be invariant if the Hamiltonian that describes it satisfies

$$[H, L_{\mu\nu}] = 0 \quad \text{for all } \mu, \nu\,. \tag{2.70}$$

Space rotations are generated by the operators

$$\begin{aligned} L^1 \equiv L_{23} &= -\mathrm{i}\left(y\frac{\partial}{\partial z} - z\frac{\partial}{\partial y}\right)\,, \\ L^2 \equiv L_{31} &= -\mathrm{i}\left(z\frac{\partial}{\partial x} - x\frac{\partial}{\partial z}\right)\,, \\ L^3 \equiv L_{12} &= -\mathrm{i}\left(x\frac{\partial}{\partial y} - y\frac{\partial}{\partial x}\right)\,, \end{aligned} \tag{2.71}$$

which represent the total angular momentum of a scalar field. They obey the familiar commutation relations

$$[L^i,\, L^j] = \mathrm{i}\epsilon^{ijk}\,L^k\,, \tag{2.72}$$

where ϵ^{ijk} is the completely antisymmetric Levi-Cività tensor, such that $\epsilon^{123} = +1$. We learn from these results that the orbital angular momentum is the whole contribution to the total angular momentum. For a translational invariant field,

$$\phi(x) = \mathrm{e}^{\mathrm{i}x^\mu P_\mu}\,\phi(0) = \mathrm{e}^{\mathrm{i}P_0 t - \mathrm{i}\boldsymbol{P}\cdot\boldsymbol{x}}\,\phi(0)\,.$$

In its rest frame, where $\boldsymbol{P} = 0$, the field is independent of space coordinates and the total angular momentum of a particle at rest described by ϕ vanishes, $L^i\phi = 0$ for $i = 1, 2, 3$. The *intrinsic spin* of the scalar field is zero, $s = 0$.

2.3 Vector Fields

A vector field $A^\alpha(x)$ carries a Lorentz index and is defined by its characteristic Lorentz transformation property

$$A'^\alpha(x') = a^\alpha{}_\beta\, A^\beta(x) = a^\alpha{}_\beta\, A^\beta\left(a^{-1}x'\right)\,. \tag{2.73}$$

For an infinitesimal transformation, in which $a^\alpha{}_\beta = \delta^\alpha{}_\beta + \epsilon^\alpha{}_\beta$, we expand the right-hand side of the equation, keeping terms up to the first order in the transformation parameters, to get

$$\begin{aligned} A'^\alpha(x) &= a^\alpha{}_\beta\, A^\beta(a_\rho{}^\sigma\, x_\sigma) = (\delta^\alpha{}_\beta + \epsilon^\alpha{}_\beta)\left[A^\beta(x) - \epsilon^\mu{}_\nu\, x^\nu\, \partial_\mu A^\beta(x)\right] \\ &= \left[\delta^\alpha{}_\beta - \tfrac{\mathrm{i}}{2}\,\epsilon^{\mu\nu}(\Sigma_{\mu\nu})^\alpha{}_\beta\right]\left[A^\beta(x) - \tfrac{\mathrm{i}}{2}\,\epsilon^{\mu\nu}\, L_{\mu\nu}\, A^\beta(x)\right] \\ &= A^\alpha(x) - \tfrac{\mathrm{i}}{2}\,\epsilon^{\mu\nu}\,(L_{\mu\nu} + \Sigma_{\mu\nu})\, A^\alpha(x)\,. \end{aligned} \tag{2.74}$$

Here we have introduced the matrix operator

$$(\Sigma_{\mu\nu})^\alpha{}_\beta \;=\; \mathrm{i}(\delta^\alpha{}_\mu\, g_{\nu\beta} - \delta^\alpha{}_\nu\, g_{\beta\mu}\,) \tag{2.75}$$

and used the simplified notation $\Sigma_{\mu\nu}\, A^\alpha = (\Sigma_{\mu\nu}\,)^\alpha{}_\beta\, A^\beta$. This operator Σ acts on the vector label, i.e. generates the variations stemming from the vectorial character of the field, while $L_{\mu\nu}$, given by the same expression as in (68), generates rather the variations due to the functional field dependence. The antisymmetric tensor operator

$$J_{\mu\nu} = L_{\mu\nu} + \Sigma_{\mu\nu} \tag{2.76}$$

is the full generator of infinitesimal Lorentz transformations of vector fields:

$$\delta_0 A^\alpha(x) = -\tfrac{\mathrm{i}}{2}\,\epsilon^{\mu\nu}\, J_{\mu\nu}\, A^\alpha(x)\,. \tag{2.77}$$

The pure space components, J_{ij} for $i, j = 1, 2, 3$, act nontrivially on the space field components to generate space rotations and represent the total angular momentum of the field. Let us denote the three independent space components by

$$J^i \equiv L^i + S^i \equiv \tfrac{1}{2}\,\epsilon^{ijk}\, J_{jk} \tag{2.78}$$

with

$$(L^i)^a_b = \tfrac{\mathrm{i}}{2}\,\epsilon^{ijk}\,(x_j\partial_k - x_k\partial_j)\,\delta^a_b\,, \tag{2.79}$$

$$(S^i)^a_b = -\mathrm{i}\epsilon^{iab}\,. \tag{2.80}$$

It can be checked that L^i, S^i and J^i satisfy the usual commutation relations of the algebra of angular momenta. For example,

$$[L^1, L^2] = \mathrm{i}L^3, \qquad [S^1, S^2] = \mathrm{i}S^3, \qquad [J^1, J^2] = \mathrm{i}J^3\,; \tag{2.81}$$

together with relations obtained by permuting indices, and $[L^i, S^j] = 0$ for all i, j .

The square of the operator $\boldsymbol{S}$ is

$$(\boldsymbol{S}^2)^a_b = \sum_{k,c}(S^k)^a_c(S^k)^c_b = \sum_{k,c}\epsilon^{kca}\,\epsilon^{kcb} \;=\; 2\,\delta^a_b\,, \tag{2.82}$$

which simply shows that $\boldsymbol{S}^2 = s(s+1) = 1(1+1)$, or that the *intrinsic spin* of a vector field is equal to $s = 1$.

2.4 The Klein–Gordon Equation

Let us return now to the relativistic generalizations of the equation of motion for a free particle which we started to consider in the introduction. We shall focus for now on the first of the two possibilities,

$$a\,\frac{\partial^2}{\partial t^2}\phi(t,\boldsymbol{x}) = \left(\sum_{i=1}^{3} b^i \frac{\partial^2}{\partial x^{i2}} + m^2\right)\phi(t,\boldsymbol{x})\,. \tag{2.83}$$

To be consistent with the relativistic energy-momentum relation (6), the coefficients must have the values $a = b^i = -1$ for $i = 1,2,3$. From this choice follows the *Klein–Gordon equation* for a free particle:

$$\left(\frac{\partial^2}{\partial t^2} - \nabla^2 + m^2\right)\phi(x) = 0\,, \tag{2.84}$$

which in relativistic notations reads

$$(\Box + m^2)\,\phi(x) = 0\,. \tag{2.85}$$

As the operator $\Box = \partial^\mu\partial_\mu$, the complex function ϕ, and the mass parameter m are all Lorentz scalars, the equation (85) is manifestly invariant in the Lorentz group, retaining the same form in any Lorentz frame.

2.4.1 Free-Particle Solutions

For any given momentum $\boldsymbol{p}$, (85) admits a solution

$$\phi(x) = f_{\boldsymbol{p}}(\boldsymbol{x})\mathrm{e}^{-\mathrm{i}Wt} \tag{2.86}$$

provided $W^2 = \boldsymbol{p}^2 + m^2$. The simplest is a plane wave, $f_{\boldsymbol{p}}(\boldsymbol{x}) = C\exp(\mathrm{i}\boldsymbol{p}\cdot\boldsymbol{x})$, with normalization constant C. For a given momentum vector $\boldsymbol{p}$, two energy values are allowed; one, $W = E = \sqrt{\boldsymbol{p}^2 + m^2}$, associated with the *positive-energy solution*

$$\phi^{(+)}(x) = f_{\boldsymbol{p}}\mathrm{e}^{-\mathrm{i}Et}\,; \tag{2.87}$$

and the other, $W = -E = -\sqrt{\boldsymbol{p}^2 + m^2}$, with the *negative-energy solution*

$$\phi^{(-)}(x) = f_{\boldsymbol{p}}\mathrm{e}^{\mathrm{i}Et}\,. \tag{2.88}$$

Note that now, unlike in the nonrelativistic case, the particle energy may take all possible values, positive as well as negative, and the energy spectrum extends from $-\infty$ to $+\infty$ just like the momentum. Thus, the symmetry between energy and momentum is restored.

2.4.2 Particle Probability

To understand the physical meaning of the new, negative-energy solution (88), it is useful to examine its implications on the particle probability. It is recalled that in nonrelativistic quantum mechanics the probability density is defined by $\phi^*\phi$ and is related to a current density via a continuity equation. But now the presence of a second-order time derivative in (85) will force us to reconsider the definition of the probability density. To find out the required changes, let us look at the following equations:

$$\phi^*(\partial^\mu\partial_\mu + m^2)\phi = 0\,,$$
$$\phi(\partial^\mu\partial_\mu + m^2)\phi^* = 0\,.$$

Taking the difference of the two equations, we obtain a local conservation equation

$$\partial_\mu j^\mu = 0\,, \tag{2.89}$$

for the four-current

$$j^\mu \equiv \mathrm{i}[\phi^*\partial^\mu\phi - (\partial^\mu\phi^*)\phi]\,, \tag{2.90}$$

which has the components

$$j^0 = \mathrm{i}\left(\phi^*\frac{\partial\phi}{\partial t} - \frac{\partial\phi^*}{\partial t}\phi\right) \equiv \mathrm{i}\phi^*\,\overset{\leftrightarrow}{\frac{\partial}{\partial t}}\,\phi\,, \tag{2.91}$$

$$\boldsymbol{j} = -\mathrm{i}\,[\phi^*\nabla\phi - (\nabla\phi^*)\phi] \equiv -\mathrm{i}\phi^*\,\overset{\leftrightarrow}{\nabla}\,\phi\,. \tag{2.92}$$

Here, we have used the notation

$$\overset{\leftrightarrow}{\frac{\partial}{\partial t}} \equiv \overrightarrow{\frac{\partial}{\partial t}} - \overleftarrow{\frac{\partial}{\partial t}}\,, \qquad \overset{\leftrightarrow}{\nabla} \equiv \overrightarrow{\nabla} - \overleftarrow{\nabla}\,.$$

Now, in the nonrelativistic limit $E \approx m$, the positive-energy solution yields $j^0 \approx 2m|f_{\boldsymbol{p}}|^2$, which is indeed proportional to the expected nonrelativistic probability density. However, as we presently see, this naive interpretation is not justified. In fact, from (87)–(88) and (91)–(92), one deduces that

$$j^0 = \pm 2E\,|f_{\boldsymbol{p}}|^2\,, \qquad \boldsymbol{j} = 2\boldsymbol{p}\,|f_{\boldsymbol{p}}|^2\,, \tag{2.93}$$

where $\pm$ refer respectively to $\phi^{(+)}$ and $\phi^{(-)}$. It follows that the particle probability, given by the integral $\int \mathrm{d}^3x\, j^0$ up to a constant, is *positive* for the positive-energy solution, but *negative* for the negative-energy solution – a rather unsatisfactory result. To avoid this awkward situation, could one simply omit the embarrassing negative-energy solution? No, because the

wave functions $\phi = f_{\boldsymbol{p}} \exp(-\mathrm{i}Wt)$, with $W > 0$, do not by themselves form a complete set, and a description of a physical system in an arbitrary state in terms of these functions alone is not possible. Both positive- and negative-energy solutions are indispensable. Thus, either the Klein–Gordon equation should be abandoned, as it had effectively been for a time following 1926, or the interpretation of j^0 as a probability density should be given up.

In 1934, six years after P. A. M. Dirac's success in discovering another relativistic equation which now carries his name and in giving the correct physical interpretation of its solutions, the interest for the Klein–Gordon equation was revived by W. Pauli and V. F. Weisskopf's original idea that j^ν should not be interpreted as a *current of probability density*, as nonrelativistic quantum mechanics would suggest, but rather as a *current of charge density*. The reinterpretation is crucial. The sign change of j^0 that occurs when $\phi^{(+)}$ is replaced by $\phi^{(-)}$ would come simply from the fact that the two solutions describe states of electric charges of opposite signs. In the light of this suggestion, let us again examine the two solutions (87)–(88). The function

$$\phi^{(+)}_{\boldsymbol{p}}(x) = C_{\boldsymbol{p}} \mathrm{e}^{\mathrm{i}(\boldsymbol{p}\cdot\boldsymbol{x} - Et)} = C_{\boldsymbol{p}} \mathrm{e}^{-\mathrm{i}p\cdot x} \tag{2.94}$$

describes a state of a particle moving in the normal time direction, whereas

$$\phi^{(-)}_{\boldsymbol{p}}(x) = C_{\boldsymbol{p}} \mathrm{e}^{\mathrm{i}[\boldsymbol{p}\cdot\boldsymbol{x} - E(-t)]} \tag{2.95}$$

can be viewed as a state of a particle moving in the reverse time direction. Now, if the charge density j^0 differs in sign for the two states because of their opposite charges, the current density $\boldsymbol{j}$ must, for the very same reason, carry opposite signs for the two cases; in other words, one ought to have $\boldsymbol{j} = \pm 2\boldsymbol{p}\,|f_{\pm\boldsymbol{p}}|^2$, which means that the momentum direction in the negative-energy solution should be reversed:

$$\phi^{(-)}_{-\boldsymbol{p}}(x) = C_{\boldsymbol{p}} \mathrm{e}^{\mathrm{i}(-\boldsymbol{p}\cdot\boldsymbol{x} + Et)} = C_{\boldsymbol{p}} \mathrm{e}^{\mathrm{i}p\cdot x}\,. \tag{2.96}$$

The interpretation of $\phi^{(-)}$ as a state of negative energy propagating in the reverse time direction is certainly counterintuitive and rather difficult to visualize. It would be conceptually clearer to replace it with the picture intuitively more familiar to our senses of a state of *antiparticle of positive energy* propagating in the *normal time direction*. By antiparticle, we mean a *charge conjugate* state to the positive-energy state, with the same mass, but with opposite charge and reversed momentum direction. In this interpretation, the antiparticle is treated as any other particle, with the appropriate quantum numbers. No more negative energies, no more reversed time arrow. This interpretation will become even more transparent when the wave function itself is quantized, that is, made into a quantum operator that depends on space-time parameters in a procedure called the *second quantization*.

The eigensolutions of the Klein–Gordon equation $\{\phi^{(+)}_{\boldsymbol{p}}\}$ and $\{\phi^{(-)}_{-\boldsymbol{p}}\}$ form a complete set. With normalization $C_{\boldsymbol{p}} = 1/\sqrt{(2\pi)^3 2E_p}$, they satisfy the orthonormality relations

$$\begin{aligned}
&\mathrm{i}\int \mathrm{d}^3x\, \phi^{(+)*}_{\boldsymbol{p}'} \overleftrightarrow{\frac{\partial}{\partial t}} \phi^{(-)}_{\boldsymbol{p}} = 0\,,\\
&\mathrm{i}\int \mathrm{d}^3x\, \phi^{(+)*}_{\boldsymbol{p}'} \overleftrightarrow{\frac{\partial}{\partial t}} \phi^{(+)}_{\boldsymbol{p}} = \delta(\boldsymbol{p}'-\boldsymbol{p})\,,\\
&\mathrm{i}\int \mathrm{d}^3x\, \phi^{(-)*}_{\boldsymbol{p}'} \overleftrightarrow{\frac{\partial}{\partial t}} \phi^{(-)}_{\boldsymbol{p}} = -\delta(\boldsymbol{p}'-\boldsymbol{p})\,.
\end{aligned} \tag{2.97}$$

It is important to note the sign on the right-hand side of the last equation. An arbitrary wave function of the particle can then be written in this basis:

$$\phi(x) = \sum_p \left[a_{\boldsymbol{p}}\phi^{(+)}_{\boldsymbol{p}}(x) + b^*_{\boldsymbol{p}}\phi^{(-)}_{-\boldsymbol{p}}(x)\right], \tag{2.98}$$

where $a_{\boldsymbol{p}}$ and $b^*_{\boldsymbol{p}}$ are the complex amplitudes of the eigenmodes of ϕ. We have used here, as we often will in the following, the shorthand notation $\sum_p = \int \mathrm{d}^3p$.

2.4.3 Second Quantization

In the formalism of quantum field theories, the function $\phi(x)$ becomes a quantum operator on the Hilbert space of state vectors. The algebra of this field operator and its conjugate momentum density operator is defined by a set of quantization rules. The operator $\phi(x)$ still obeys the same dynamic equation as the classical field, but the coefficients $a_{\boldsymbol{p}}$ and $b_{\boldsymbol{p}}$ in the series (98) now become quantum operators. Using the same notations as for the classical field (98), we have the following expressions for the scalar field operator and its Hermitian adjoint:

$$\begin{aligned}
\phi(x) &= \sum_p \left[a_{\boldsymbol{p}}\phi^{(+)}_{\boldsymbol{p}}(x) + b^\dagger_{\boldsymbol{p}}\phi^{(-)}_{-\boldsymbol{p}}(x)\right],\\
\phi^\dagger(x) &= \sum_p \left[a^\dagger_{\boldsymbol{p}}\phi^{(+)*}_{\boldsymbol{p}}(x) + b_{\boldsymbol{p}}\phi^{(-)*}_{-\boldsymbol{p}}(x)\right].
\end{aligned} \tag{2.99}$$

We will show in what follows that

$a_{\boldsymbol{p}}$ destroys a particle of momentum $\boldsymbol{p}$ and positive (by convention) unit charge;
$a^\dagger_{\boldsymbol{p}}$ creates a particle of momentum $\boldsymbol{p}$ and positive unit charge;
$b_{\boldsymbol{p}}$ destroys an antiparticle of momentum $\boldsymbol{p}$ and negative unit charge;
$b^\dagger_{\boldsymbol{p}}$ creates a particle of momentum $\boldsymbol{p}$ and negative unit charge.

Thus, the field operator ϕ reduces the total charge of the system by one unit, either by destroying a (say, positive) charge with the operator $a_{\boldsymbol{p}}$ or

by creating an antiparticle of opposite (negative) charge with $b^\dagger_{\boldsymbol{p}}$. Similarly, the operator $\phi^\dagger$ increases the total charge by one unit, either by creating a particle with $a^\dagger_{\boldsymbol{p}}$ or by destroying an antiparticle with $b_{\boldsymbol{p}}$. In general, the field operator ϕ can operate on a state vector containing an arbitrary number of particles and antiparticles, and so the Klein–Gordon equation which determines the evolution of the field ϕ is also an equation that governs the dynamics of the whole system of particles and antiparticles. A quantum field theory is, for this reason, necessarily a many-particle theory. Evidently, other more complex processes are possible. For example, the quadratic operator $\phi^\dagger(x)\phi(x)$ represents any one of several charge-conserving processes: the creation and the annihilation of a particle, or of an antiparticle, at space-time point x, or the creation or destruction of a particle–antiparticle pair. In any of these events, there are no changes in the net charge and the total charge of the system remains unchanged.

Finally, if the field is Hermitian, $\phi = \phi^\dagger$, then $a_{\boldsymbol{p}} = b_{\boldsymbol{p}}$, which means that the particle is identical to its antiparticle.

2.4.4 Operator Algebra

In this section, we give a physical motivation for the quantization rules of the scalar field operators. For this purpose, let us introduce (see Sect. 2.6 below) the Hamiltonian operator for a complex scalar field ϕ,

$$H = \int \mathrm{d}^3x \left[\frac{\partial\phi^\dagger}{\partial t}\frac{\partial\phi}{\partial t} + \partial^i\phi^\dagger\partial^i\phi + m^2\phi^\dagger\phi \right] . \tag{2.100}$$

An integration by parts of the second term on the right-hand side leads, with the help of the Klein–Gordon equation (which ϕ obeys), to

$$H = -\int \mathrm{d}^3x\, \phi^\dagger \overleftrightarrow{\frac{\partial}{\partial t}} \left(\frac{\partial\phi}{\partial t}\right) . \tag{2.101}$$

A surface term has been dropped, which is justified by assuming ϕ to vanish on the integration surfaces. Now, using the expansion series (99) and the orthonormality relations (97), one can simplify the expression as follows, carefully keeping the relative order of all operators,

$$\begin{aligned} H &= \int \mathrm{d}^3x \sum_{p'} \left[a^\dagger_{\boldsymbol{p}'}\phi^{(+)*}_{\boldsymbol{p}'}(x) + b_{\boldsymbol{p}'}\phi^{(-)*}_{-\boldsymbol{p}'}(x) \right] \\ &\quad \times \mathrm{i} \overleftrightarrow{\frac{\partial}{\partial t}} \sum_p E_{\boldsymbol{p}} \left[a_{\boldsymbol{p}}\phi^{(+)}_{\boldsymbol{p}}(x) - b^\dagger_{\boldsymbol{p}}\phi^{(-)}_{-\boldsymbol{p}}(x) \right] \\ &= \sum_p E_{\boldsymbol{p}} \left(a^\dagger_{\boldsymbol{p}}a_{\boldsymbol{p}} + b_{\boldsymbol{p}}b^\dagger_{\boldsymbol{p}} \right) . \end{aligned} \tag{2.102}$$

The first term on the right-hand side, $E_{\boldsymbol{p}} a^\dagger_{\boldsymbol{p}} a_{\boldsymbol{p}} \equiv E_{\boldsymbol{p}} \mathcal{N}_{\boldsymbol{p}}$, can be interpreted as the energy density of the particle, and $\mathcal{N}_{\boldsymbol{p}} = a^\dagger_{\boldsymbol{p}} a_{\boldsymbol{p}}$, as the particle number operator. However, in the second term the order of the factors b and $b^\dagger$ must be reversed before their product can be similarly interpreted as an antiparticle number operator, $\overline{\mathcal{N}}_{\boldsymbol{p}} = b^\dagger_{\boldsymbol{p}} b_{\boldsymbol{p}}$.

The creation and annihilation operators and observables, such as H and $\mathcal{N}$, operate on an abstract space characterized by a vacuum state $|0\rangle$ defined by

$$a_{\boldsymbol{p}} \, |0\rangle \, = \, b_{\boldsymbol{p}} \, |0\rangle \, = \, 0 \, , \qquad \langle 0 \, | 0 \, \rangle \, = \, 1 \, . \tag{2.103}$$

The basis states of that space, called Fock space, are constructed by applying all positive powers of $a^\dagger$ and $b^\dagger$ on the vacuum, and Fock states are linear combinations of the basis states with definite numbers of excitation modes. In particular, $a^\dagger \, |0\rangle$ and $b^\dagger \, |0\rangle$ are one-particle states conventionally normalized:

$$\left\langle 0 \, \middle| \, a_{\boldsymbol{p}} a^\dagger_{\boldsymbol{p}} \, \middle| \, 0 \right\rangle = 1 \, , \qquad \left\langle 0 \, \middle| \, b_{\boldsymbol{p}} b^\dagger_{\boldsymbol{p}} \, \middle| \, 0 \right\rangle = 1 \, . \tag{2.104}$$

Therefore, when we reverse the order of b and $b^\dagger$ in (102), we may write

$$b_{\boldsymbol{p}} b^\dagger_{\boldsymbol{p}} = 1 \pm b^\dagger_{\boldsymbol{p}} b_{\boldsymbol{p}} \, ,$$

and obtain the expression for H in the form

$$H = \sum_p E_{\boldsymbol{p}} \, (a^\dagger_{\boldsymbol{p}} a_{\boldsymbol{p}} \, \pm \, b^\dagger_{\boldsymbol{p}} b_{\boldsymbol{p}}) \, + \sum_p E_{\boldsymbol{p}} \, . \tag{2.105}$$

The last term on the right-hand side is a c-number quantity which can be interpreted as the (infinite) vacuum energy. In general, it may be dropped (because it is not observable) by redefining the energy of the system relatively to the vacuum, such that the rescaled vacuum energy vanishes:

$$H \to H - \sum_p E_{\boldsymbol{p}} = \sum_p E_{\boldsymbol{p}} \, (a^\dagger_{\boldsymbol{p}} a_{\boldsymbol{p}} \, \pm \, b^\dagger_{\boldsymbol{p}} b_{\boldsymbol{p}}) \, . \tag{2.106}$$

Clearly, the Hamiltonian without the vacuum energy is non-negative, provided the contributions from antiparticles are non-negative, that is, provided that one adopts the $+$ sign, rather than the $-$ sign, in the above expressions. This leads to

$$H \, = \sum_p E_{\boldsymbol{p}} \, (a^\dagger_{\boldsymbol{p}} a_{\boldsymbol{p}} \, + \, b^\dagger_{\boldsymbol{p}} b_{\boldsymbol{p}}) = \sum_p E_{\boldsymbol{p}} \, (\mathcal{N}_{\boldsymbol{p}} + \overline{\mathcal{N}}_{\boldsymbol{p}}) \, . \tag{2.107}$$

Thus, the condition that the total energy of the system be non-negative implies that the operators a and b obey *commutation relations*. This is consistent with a natural generalization of the quantum rules of coordinates and momenta in a discrete system,

$$[q_i, \, p_j] = \mathrm{i}\delta_{ij}, \;\; [q_i, \, q_j] = [p_i, \, p_j] = 0 \, ,$$

to the quantization rules of a scalar field and its conjugate momentum density:

$$[\phi(t,\boldsymbol{x}),\,\pi(t,\boldsymbol{y})] = \mathrm{i}\delta^{(3)}(\boldsymbol{x}-\boldsymbol{y})\,,$$
$$[\phi(t,\boldsymbol{x}),\,\phi(t,\boldsymbol{y})] = [\pi(t,\boldsymbol{x}),\,\pi(t,\boldsymbol{y})] = 0\,, \tag{2.108}$$

where $\pi \equiv (\partial\phi^\dagger/\partial t)$. From this postulate, follow the commutation relations

$$[a_{\boldsymbol{p}'}, a^\dagger_{\boldsymbol{p}}] = \delta(\boldsymbol{p}'-\boldsymbol{p})\,; \quad [a_{\boldsymbol{p}'}, a_{\boldsymbol{p}}] = 0\,; \quad [a_{\boldsymbol{p}'}, b_{\boldsymbol{p}}] = 0\,;$$
$$[b_{\boldsymbol{p}'}, b^\dagger_{\boldsymbol{p}}] = \delta(\boldsymbol{p}'-\boldsymbol{p})\,; \quad [b_{\boldsymbol{p}'}, b_{\boldsymbol{p}}] = 0\,; \quad [a_{\boldsymbol{p}'}, b^\dagger_{\boldsymbol{p}}] = 0\,. \tag{2.109}$$

2.4.5 Physical Significance of the Fock Operators

The charge operator, which is given by

$$Q = \mathrm{i}\int \mathrm{d}^3x\, j^0(x) = \mathrm{i}\int \mathrm{d}^3x\, \phi^\dagger(x)\,\overleftrightarrow{\frac{\partial}{\partial t}}\,\phi(x)\,, \tag{2.110}$$

may be rewritten in the following form by retracing the same steps as in the calculation of H,

$$\begin{aligned} Q =& \int \mathrm{d}^3x \sum_{p'} \left[a^\dagger_{\boldsymbol{p}'}\phi^{(+)*}_{\boldsymbol{p}'}(x) + b_{\boldsymbol{p}'}\phi^{(-)*}_{-\boldsymbol{p}'}(x)\right] \\ & \times \mathrm{i}\,\overleftrightarrow{\frac{\partial}{\partial t}} \sum_{p} \left[a_{\boldsymbol{p}}\phi^{(+)}_{\boldsymbol{p}}(x) + b^\dagger_{\boldsymbol{p}}\phi^{(-)}_{-\boldsymbol{p}}(x)\right] \\ =& \sum_p (a^\dagger_{\boldsymbol{p}}a_{\boldsymbol{p}} - b_{\boldsymbol{p}}b^\dagger_{\boldsymbol{p}})\,. \end{aligned} \tag{2.111}$$

Using the commutation relations (109) and dropping the vacuum terms yields

$$Q = \sum_p (a^\dagger_{\boldsymbol{p}}a_{\boldsymbol{p}} - b^\dagger_{\boldsymbol{p}}b_{\boldsymbol{p}}) = \sum_p (\mathcal{N}_{\boldsymbol{p}} - \overline{\mathcal{N}}_{\boldsymbol{p}})\,. \tag{2.112}$$

Since $[\mathcal{N}_{\boldsymbol{p}}, \overline{\mathcal{N}}_{\boldsymbol{p}'}] = 0$, operators H and Q commute, which simply means that charge is conserved. Of course, when a particle is identical to its antiparticle, i.e. $\phi^\dagger = \phi$, its charge operator identically vanishes, $Q = 0$. A Hermitian field therefore represents a *neutral particle*.

As already mentioned, a, b, $a^\dagger$, $b^\dagger$, and the operators H and Q act on Fock states. For example, the Fock state of one particle of momentum p is

$$|p\rangle = \sqrt{(2\pi)^3 2E_p}\, a^\dagger_{\boldsymbol{p}}\,|0\rangle = C^{-1}_{\boldsymbol{p}}\, a^\dagger_{\boldsymbol{p}}\,|0\rangle\,, \tag{2.113}$$

normalized such that

$$\langle p'\,|p\rangle = (2\pi)^3\, 2E_{\boldsymbol{p}}\,\delta^3(\boldsymbol{p}'-\boldsymbol{p})\,. \tag{2.114}$$

With the help of (107), (109), and (112), one shows that it is an eigenstate of energy $E_{\boldsymbol{p}}$:

$$\begin{aligned} H\,|p\rangle \;&=\; C_{\boldsymbol{p}}^{-1} H a_{\boldsymbol{p}}^{\dagger}\,|0\rangle \;=\; C_{\boldsymbol{p}}^{-1}[H, a_{\boldsymbol{p}}^{\dagger}]\,|0\rangle \\ &=\; E_{\boldsymbol{p}}\;|p\rangle\;; \end{aligned} \tag{2.115}$$

and of one unit of charge:

$$\begin{aligned} Q\,|p\rangle \;&=\; C_{\boldsymbol{p}}^{-1} Q a_{\boldsymbol{p}}^{\dagger}\,|0\rangle \;=\; C_{\boldsymbol{p}}^{-1}[Q, a_{\boldsymbol{p}}^{\dagger}]\,|0\rangle \\ &=\; |p\rangle\;. \end{aligned} \tag{2.116}$$

Thus, the operator $a_{\boldsymbol{p}}^{\dagger}$ creates a discrete excitation of positive unit charge, of energy $E_{\boldsymbol{p}}$ related to momentum by the usual relation $E_{\boldsymbol{p}} = \sqrt{\boldsymbol{p}^2 + m^2}$. One may call such a state a *particle*. When applied on an arbitrary state describing the system, $a_{\boldsymbol{p}}^{\dagger}$ adds a particle to the system, increasing its energy by $E_{\boldsymbol{p}}$ and its charge by a unit of charge.

A similar analysis can be done for a one-antiparticle state, defined by

$$|\bar{p}\rangle \;=\; \sqrt{(2\pi)^3 2E_p}\, b_{\boldsymbol{p}}^{\dagger}\,|0\rangle \;=\; C_{\boldsymbol{p}}^{-1}\, b_{\boldsymbol{p}}^{\dagger}\,|0\rangle\;. \tag{2.117}$$

It will then be seen that $b_{\boldsymbol{p}}^{\dagger}$ adds an antiparticle to the system, thus increasing its energy by $E_{\boldsymbol{p}}$ and reducing its charge by a unit of charge:

$$[H,\, b_{\boldsymbol{p}}^{\dagger}] = E_{\boldsymbol{p}}\, b_{\boldsymbol{p}}^{\dagger}\,, \qquad [Q,\, b_{\boldsymbol{p}}^{\dagger}] = -b_{\boldsymbol{p}}^{\dagger}\,.$$

What is the connection between the classical wave function and the quantum field operators ϕ, $\phi^{\dagger}$? For a plane-wave free-particle state (99), it is given by the relation

$$\langle 0\,|\,\phi(x)\,|\,p\rangle \;=\; \mathrm{e}^{-\mathrm{i}p\cdot x}\,. \tag{2.118}$$

The matrix element $\langle 0\,|\,\phi(x)\,|\,p\rangle$, which represents the amplitude for the annihilation at point x of a particle carrying momentum p, is thus shown, with the help of (109) and (113), to be identical to the plane-wave function of a free particle having momentum p. Hence the following amplitudes describe the initial or final states of a particle or antiparticle:

$< 0\|\phi(x)\|p > = \mathrm{e}^{-\mathrm{i}p\cdot x}$	$\rightarrow\!\bullet$	annihilation of a particle,
$< p\|\phi^{\dagger}(x)\|0 > = \mathrm{e}^{\mathrm{i}p\cdot x}$	$\bullet\!\rightarrow$	creation of a particle,
$< \bar{p}\|\phi(x)\|0 > = \mathrm{e}^{\mathrm{i}p\cdot x}$	$\bullet\!\leftarrow$	creation of an antiparticle,
$< 0\|\phi^{\dagger}(x)\|\bar{p} > = \mathrm{e}^{-\mathrm{i}p\cdot x}$	$\leftarrow\!\bullet$	annihilation of an antiparticle.

2.5 Quantized Vector Fields

It was already shown that a vector field has intrinsic spin equal to 1. Quantum mechanics tells us that a spin-1 particle generally has three spinorial (polarization) states, eigenstates of the spin projection along some quantization axis, normally chosen to be the z axis. It is exactly the case of particles with nonvanishing masses, such as the vector mesons (ρ, ϕ, J/ψ) or the weak interaction bosons ($\mathrm{W}^{\pm}$, Z). However, the electromagnetic field does not have this property; it is described by a four-vector field which has just two independent components. The particle associated with it, the photon, is massless and has only two polarization states. We will describe in this section the nature of the vector fields, the differences between the two cases, and the role of gauge invariance and of relativistic invariance.

2.5.1 Massive Vector Fields

For a massive vector field in the absence of interactions, it is natural to assume in a relativistic theory that each of the four components of the field A^{μ} satisfies the Klein–Gordon equation

$$(\partial_{\nu}\partial^{\nu} + m^2)A^{\mu} = 0\,, \tag{2.119}$$

subject to the condition

$$\partial_{\nu}A^{\nu} = 0\,. \tag{2.120}$$

This manifestly Lorentz-invariant condition is needed to reduce the number of independent equations from four to three, which is precisely the number of the independent degrees of freedom of a spin-1 particle.

For a plane-wave solution

$$A^{\mu}(x) \approx a^{\mu}(k)\,\mathrm{e}^{-\mathrm{i}k\cdot x}\,, \tag{2.121}$$

the equation (119) implies a constraint on the four-momentum k:

$$k^2 \equiv (k^0)^2 - \boldsymbol{k}^2 = m^2\,, \qquad \text{or}$$
$$\pm k^0 = E_k = \sqrt{\boldsymbol{k}^2 + m^2}, \tag{2.122}$$

which shows that m is the mass of the particle associated with the field. On the other hand, (120) imposes a restriction on the vector $a^{\mu}(k)$ which describes the polarization states of the field:

$$k_{\mu}a^{\mu}(k) = 0\,. \tag{2.123}$$

It is not difficult to construct a set of three independent vectors satisfying this condition. For example, placing the z axis in the direction of the momentum $\boldsymbol{k}$, so that $k^{\mu} = (E_k, 0, 0, |\boldsymbol{k}|)$, a possible choice of polarization vectors is

$$e^{\mu}(k,1) = (0,1,0,0)\,,$$
$$e^{\mu}(k,2) = (0,0,1,0)\,,$$
$$e^{\mu}(k,3) = \frac{1}{m}\,(|\boldsymbol{k}\,|, 0, 0, E_k)\,. \tag{2.124}$$

In this choice, the basis vectors are *real*, spacelike, normalized to -1, and reduce to $\hat{\boldsymbol{x}}$, $\hat{\boldsymbol{y}}$, and $\hat{\boldsymbol{z}}$ when the particle comes to rest. Together with k^μ, they form a complete set in the sense that

$$\sum_{\lambda=1}^{3} e^\mu(k,\lambda)\, e^\nu(k,\lambda) = -g^{\mu\nu} + \frac{k^\mu k^\nu}{m^2}\,, \tag{2.125}$$

where $k^\mu \equiv (\sqrt{\boldsymbol{k}^2 + m^2},\, \boldsymbol{k})$. The vectors $e^\mu(k,1)$ and $e^\mu(k,2)$ are orthogonal to k^μ as well as to $\boldsymbol{k}$, and describe the *transverse polarizations* of the field, while the *longitudinal polarization* vector $e^\mu(k,3)$ has its space components parallel to $\boldsymbol{k}$ and its time component determined by (123). Of course, other choices are possible, leading generally to complex polarization vectors, which may be more convenient for some purposes.

The expansion series of the massive vector field in terms of plane waves and of complex polarization basis vectors reads

$$\begin{aligned} A^\mu(x) = \sum_k \sum_{\lambda=1}^{3} & \frac{1}{\sqrt{(2\pi)^3 2E_k}} \\ & \times \left[e^\mu(k,\lambda)\, a(\boldsymbol{k},\lambda)\, \mathrm{e}^{-\mathrm{i}k\cdot x} \;+\; e^{\mu\,*}(k,\lambda)\, b^\dagger(\boldsymbol{k},\lambda)\, \mathrm{e}^{\mathrm{i}k\cdot x} \right], \end{aligned} \tag{2.126}$$

where E_k and $e^\mu(k,\lambda)$ satisfy (122)–(123). Considered as a quantum field, A^μ is a superposition of eigenmodes of particles and antiparticles of momentum k^ν and polarization λ. These modes are described by the creation and annihilation operators $a^\dagger$, $b^\dagger$ and a, b, which obey the commutation relations

$$\begin{aligned} [a(\boldsymbol{k}',\lambda'),\, a(\boldsymbol{k},\lambda)] &= [b(\boldsymbol{k}',\lambda'),\, b(\boldsymbol{k},\lambda)] = 0\,, \\ [a(\boldsymbol{k}',\lambda'),\, b(\boldsymbol{k},\lambda)] &= [a^\dagger(\boldsymbol{k}',\lambda'),\, b^\dagger(\boldsymbol{k},\lambda)] = 0\,, \\ [a(\boldsymbol{k}',\lambda'),\, a^\dagger(\boldsymbol{k},\lambda)] &= \delta_{\lambda'\lambda}\, \delta(\boldsymbol{k}' - \boldsymbol{k})\,, \\ [b(\boldsymbol{k}',\lambda'),\, b^\dagger(\boldsymbol{k},\lambda)] &= \delta_{\lambda'\lambda}\, \delta(\boldsymbol{k}' - \boldsymbol{k})\,. \end{aligned} \tag{2.127}$$

2.5.2 The Maxwell Equations

A classical electromagnetic field is described by two three-vectors, the electric field $\boldsymbol{E}$ and the magnetic field $\boldsymbol{B}$, which obey a set of coupled differential equations discovered by James Clerk Maxwell in 1864. Equivalently, it can be represented by a four-vector that satisfies a Klein–Gordon equation for a particle of vanishing mass.

The Maxwell equations for a field in the presence of an electrical source ρ and a current $\boldsymbol{j}$ are

$$\begin{aligned} \nabla \cdot \boldsymbol{E} &= \rho\,, \\ \nabla \times \boldsymbol{B} - \frac{\partial \boldsymbol{E}}{\partial t} &= \boldsymbol{j}\,, \\ \nabla \cdot \boldsymbol{B} &= 0\,, \\ \nabla \times \boldsymbol{E} + \frac{\partial \boldsymbol{B}}{\partial t} &= 0\,. \end{aligned} \tag{2.128}$$

They are written in the rationalized Heaviside–Lorentz units, with $c = \hbar = 1$, so that the fine structure constant is defined by $\alpha = e^2/4\pi \approx 1/137$.

Experience tells us that electric and magnetic phenomena are not independent, except in static situations, and that a certain symmetry exists between $\boldsymbol{E}$ and $\boldsymbol{B}$ in spite of the absence of observable magnetic charges and currents. It is thus quite possible that these apparently distinct forces are actually different aspects of the same kind of interaction. Therefore, it would be of advantage to combine the components of $\boldsymbol{E}$ and $\boldsymbol{B}$ into a single object, which must be an antisymmetric tensor of second rank as it has the correct number of (six) independent components,

$$F^{i0} = -F^{0i} = E^i, \qquad F^{ij} = -\epsilon^{ijk}\, B^k\,,$$

or

$$F^{\mu\nu} = \begin{pmatrix} 0 & -E^1 & -E^2 & -E^3 \\ E^1 & 0 & -B^3 & B^2 \\ E^2 & B^3 & 0 & -B^1 \\ E^3 & -B^2 & B^1 & 0 \end{pmatrix}. \tag{2.129}$$

In terms of this field tensor, the two nonhomogeneous equations, which are driven by the charge and current densities, take on a really simple form

$$\partial_\mu\, F^{\mu\nu} = j^\nu\,, \tag{2.130}$$

where $j^\mu = (\rho, \boldsymbol{j})$. The current conservation is contained in this equation and is a direct consequence of the antisymmetry of $F^{\mu\nu}$ and of the commutativity of ∂_μ and ∂_ν:

$$\partial_\nu j^\nu = 0\,. \tag{2.131}$$

Now, if a *vector field* $A^\mu(x)$ is defined through

$$F^{\mu\nu} = \partial^\mu A^\nu \,-\, \partial^\nu A^\mu\,, \tag{2.132}$$

the homogeneous Maxwell equations are automatically satisfied because

$$\partial_\lambda F_{\mu\nu} + \partial_\mu F_{\nu\lambda} + \partial_\nu F_{\lambda\mu} = 0 \tag{2.133}$$

is an algebraic identity when considered in terms of A^μ, and reproduces the two homogeneous Maxwell equations when considered in terms of $F_{\mu\nu}$ with indices $(\lambda, \mu, \nu) = (i, j, k)$ and $(0, i, j)$. Given A^μ, the fields $\boldsymbol{E}$ and $\boldsymbol{B}$ may be calculated from (132):

$$\begin{aligned} \boldsymbol{E} &= -\frac{\partial \boldsymbol{A}}{\partial t} - \nabla A^0\,, \\ \boldsymbol{B} &= \nabla \times \boldsymbol{A}\,. \end{aligned} \tag{2.134}$$

To sum up, the Maxwell equations can be seen as the dynamic equations for the vector field $A^\mu(x)$:

$$\begin{aligned} F^{\mu\nu}(x) &= \partial^\mu A^\nu(x) \,-\, \partial^\nu A^\mu(x)\,, \\ \partial_\mu\, F^{\mu\nu}(x) &= j^\nu(x)\,. \end{aligned} \tag{2.135}$$

2.5.3 Quantization of the Electromagnetic Field

We now limit ourselves to a free field, in the absence of external sources. Then (135) reduces to

$$\partial_\mu F^{\mu\nu} = 0\,, \tag{2.136}$$

or, in terms of $A^\mu(x)$ now considered the fundamental variable, to

$$\Box A^\mu - \partial^\mu(\partial\cdot A) \;=\; 0\,. \tag{2.137}$$

This equation has a remarkable symmetry property, absent in the case of massive vector fields: it is invariant to transformations defined by

$$A^\mu(x) \to A'^\mu(x) = A^\mu(x) + \partial^\mu\Lambda(x)\,, \tag{2.138}$$

where $\Lambda(x)$ is an *arbitrary* function of x. Here the field argument x is unchanged, but the magnitude of each component is modified by an additive quantity (a 'gauge') that depends on the space-time coordinates of the field. Such a transformation is called a *local gauge transformation*. Invariance of the field tensor can easily be proved:

$$\begin{aligned} F'^{\mu\nu} &= \partial^\mu A'^\nu \;-\; \partial^\nu A'^\mu = \partial^\mu(A^\nu + \partial^\nu\Lambda) \;-\; \partial^\nu(A^\mu + \partial^\mu\Lambda) \\ &= \partial^\mu A^\nu \;-\; \partial^\nu A^\mu = F^{\mu\nu}\,, \end{aligned} \tag{2.139}$$

from which it follows that (137) is invariant. Saying that (137) is invariant to transformations (138) means that whenever $A^\mu(x)$ is a solution to (137), $A'^\mu(x)$ is also a solution. Thus, the equation for A^μ does not have a single solution but rather an infinite class of solutions related by (138). The situation in fact is quite similar to what one finds in general relativity where Einstein's field equations, which are invariant to general coordinate transformations, cannot determine by themselves the solution; to make the solution unique, one has to specify the coordinate system in which the problem is to be solved. In the present case, it corresponds to specifying the function $\Lambda(x)$, or as one says, to *choosing a gauge*. Since this choice is not unique, the intermediate steps in calculations may differ, depending on the gauge chosen. However, since the physics described by the theory is invariant to gauge transformations, it must not depend on the gauge chosen for the calculation and the final expression of any observable correctly calculated must be independent of the function Λ.

To solve (137), one can proceed in two steps: first, a constraint is imposed on the field to reduce the number of independent solutions from four to three, and next, an appropriate function $\Lambda(x)$ is chosen to fix the gauge, thus reducing the number of independent solutions from three to two.

In the first step, a possible constraint is inspired by the condition (120) for massive vector fields, which has the advantage of being covariant. With this constraint (the *Lorentz condition*), the problem is reduced to solving

$$\begin{aligned} &\Box A^\mu = 0 \qquad \text{(field equation)}, \\ &\partial_\mu A^\mu = 0 \qquad \text{(Lorentz condition)}\,. \end{aligned} \tag{2.140}$$

Even then, the field equation still remains invariant under a residual class of gauge transformations defined by

$$A^\mu(x) \to A'^\mu(x) = A^\mu(x) + \partial^\mu \Lambda(x)\,, \qquad \Box\Lambda = 0\,. \tag{2.141}$$

This means, if A_μ is solution to (140), then all $A'^\mu = A^\mu + \partial^\mu\Lambda$ such that $\Box\Lambda = 0$ form a class of admissible solutions. The solution becomes unique only by fixing $\Lambda(x)$.

A plane-wave solution to (140) has the form

$$A_\mu(x) \approx a_\mu(k)\,\mathrm{e}^{-\mathrm{i}k\cdot x}\,, \tag{2.142}$$

together with the constraints

$$k^2 = {k_0}^2 - \boldsymbol{k}^2 = 0 \qquad \to \qquad \pm k_0 = |\boldsymbol{k}| = \omega\,, \qquad k^\mu\, a_\mu(k) = 0\,. \tag{2.143}$$

The condition $k^2 = 0$ tells us that the quantum (photon) mass is zero, while the second, $k{\cdot}a = 0$, lets us determine one component of a^μ, say the time component, in terms of the others:

$$k^0 a_0(k) = \boldsymbol{k}{\cdot}\boldsymbol{a}(k) \;\to\; a_0(k) = \hat{\boldsymbol{k}}{\cdot}\boldsymbol{a}(k) \qquad \text{with } \hat{\boldsymbol{k}} \equiv \boldsymbol{k}/|\boldsymbol{k}|\,. \tag{2.144}$$

So $a_0(k)$ is equal to the length of the longitudinal component of $\boldsymbol{a}$:

$$\boldsymbol{a}_\parallel = (\boldsymbol{a}{\cdot}\hat{\boldsymbol{k}})\hat{\boldsymbol{k}}\,. \tag{2.145}$$

Suppose now that we have the solutions $A^\mu(x)$ of (140) and hence the coefficients $a^\mu(k)$. Let us consider a particular solution to $\Box\Lambda = 0$,

$$\Lambda(x) \approx \lambda(k)\,\mathrm{e}^{-\mathrm{i}k\cdot x}\,, \tag{2.146}$$

with the four-momentum k^μ satisfying the same constraint, $k^2 = 0$, as in $A_\mu(x)$. The residual gauge transformation (141) now appears as

$$a_\mu(k) \to a'_\mu(k) \;=\; a_\mu(k) \;-\; \mathrm{i}k_\mu\,\lambda(k)\,. \tag{2.147}$$

Now choose the gauge such that

$$\lambda(k) = a_0(k)/\mathrm{i}k_0\,, \tag{2.148}$$

which implies that the time component of the polarization vector, and hence also the longitudinal component, vanishes: $a'_0(k) = 0$, $\boldsymbol{a}'_\parallel = 0$. It follows that $a'_\mu(k)$ reduces to a spacelike vector lying in a plane perpendicular to $\boldsymbol{k}$;

that is, it reduces to a three-vector $\boldsymbol{a}' = \boldsymbol{a}'_\perp$ perpendicular to $\boldsymbol{k}$. Hence it can always be expressed in terms of two basis vectors spanning that plane:

$$\hat{\boldsymbol{k}}\cdot\boldsymbol{e}(k,\lambda) = 0, \qquad \lambda = 1,2\,. \tag{2.149}$$

The polarization vector $\boldsymbol{a}(k)$ (with its prime accent suppressed) can now be written in this basis as

$$\boldsymbol{a}(k) = \sum_{\lambda=1}^{2} a(k,\lambda)\,\boldsymbol{e}(k,\lambda)\,,$$

where $a(k,1)$ and $a(k,2)$ are scalar coefficients.

The three vectors $\boldsymbol{e}(k,1)$, $\boldsymbol{e}(k,2)$, and $\hat{\boldsymbol{k}}$ form a complete, orthonormal basis in three-dimensional space, so that

$$\sum_{\lambda=1}^{2} e^{i*}(k,\lambda)\,e^{j}(k,\lambda) = \delta_{ij} - \frac{k^i k^j}{\boldsymbol{k}^2}\,. \tag{2.150}$$

The simplest choice consists of two real unit vectors orthogonal to each other as well as to the propagation direction $\boldsymbol{k}$,

$$\begin{aligned}\boldsymbol{e}(k,\lambda)\cdot\boldsymbol{e}(k,\lambda') &= \delta_{\lambda\lambda'}\,,\\ \boldsymbol{e}(k,1)\times\boldsymbol{e}(k,2) &= \hat{\boldsymbol{k}}\,.\end{aligned} \tag{2.151}$$

These vectors represent transverse *linear polarizations*. If the z axis is chosen to coincide with $\boldsymbol{k}$, then $\boldsymbol{e}(k,1)$ and $\boldsymbol{e}(k,2)$ lie respectively along and in the same sense as $\hat{\boldsymbol{x}}$ and $\hat{\boldsymbol{y}}$ (see Fig. 2.2). In an infinitesimal rotation $(\hat{\boldsymbol{z}},\delta\theta)$, they transform according to (77), with $L^i = 0$ and $(S^i)_{ab} = -\mathrm{i}\,\epsilon^{iab}$,

$$\begin{aligned}\delta\boldsymbol{e}(k,1) &= -\epsilon_{12}\,\boldsymbol{e}(k,2) \;=\; \delta\theta\,\boldsymbol{e}(k,2)\,,\\ \delta\boldsymbol{e}(k,2) &= +\epsilon_{12}\,\boldsymbol{e}(k,1) \;=\; -\delta\theta\,\boldsymbol{e}(k,1)\,.\end{aligned} \tag{2.152}$$

Another possible basis may include the complex vectors

$$\boldsymbol{e}(k,\pm) = \mp\,\frac{1}{\sqrt{2}}\,[\boldsymbol{e}(k,1) \pm \mathrm{i}\boldsymbol{e}(k,2)]\,, \tag{2.153}$$

which transform without mixing as the spherical components of vectors:

$$\delta\boldsymbol{e}(k,\pm) \;=\; \mp\mathrm{i}\,\delta\theta\,\boldsymbol{e}(k,\pm)\,, \tag{2.154}$$

and therefore represent states of *circular polarizations*. They are associated with the $m = \pm 1$ spin components, with the quantization axis chosen to coincide with the propagation vector $\boldsymbol{k}$. Since the longitudinal polarization is absent, the $m = 0$ spin component is also absent. Thus, the massless

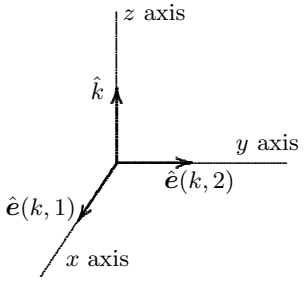

Fig. 2.2. Relative orientations of the linear polarization vectors and the propagation vector

photon has only two orientations, one parallel and the other antiparallel to the propagation vector.

The solution to (140) in the gauge thus chosen (called the *Coulomb gauge*),

$$A^0(x) \,=\, 0\,, \qquad \nabla{\cdot}\boldsymbol{A}(x) \,=\, 0\,, \tag{2.155}$$

is therefore uniquely determined. It reads, with $k = (\omega = |\boldsymbol{k}|,\, \boldsymbol{k})$,

$$\begin{aligned}\boldsymbol{A}(x) = &\sum_k \sum_{\lambda=1}^{2} \frac{1}{\sqrt{(2\pi)^3\, 2\omega}} \\ &\times \left[\boldsymbol{e}(k,\lambda)\, a(k,\lambda)\, \mathrm{e}^{-\mathrm{i}k\cdot x} + \boldsymbol{e}^*(k,\lambda)\, a^\dagger(k,\lambda)\, \mathrm{e}^{\mathrm{i}k\cdot x}\right]\,.\end{aligned} \tag{2.156}$$

As *quantum fields*, $A^i(x)$ and its conjugate momentum $\dot{A}^i(x)$ obey the canonical commutation relations at equal times

$$[A^i(t,\boldsymbol{x}),\, \dot{A}^j(t,\boldsymbol{y})] = \mathrm{i}\int \frac{\mathrm{d}^3k}{(2\pi)^3}\left(\delta_{ij} - \frac{k^i k^j}{\boldsymbol{k}^2}\right)\mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot(\boldsymbol{x}-\boldsymbol{y})}\,. \tag{2.157}$$

The transversality of the field explains the presence of a 'divergenceless' δ-function on the right-hand side. $A^i(x)$ is a superposition of particle modes of different momenta with two allowed polarizations represented by the creation and annihilation operators, $a(k,\lambda)$ and $a^\dagger(k,\lambda)$, which obey commutation relations (127). The field is Hermitian, $\boldsymbol{A}^\dagger = \boldsymbol{A}$, because the particle associated with it, the photon, has all generalized charges equal to 0.

By quantizing in Coulomb gauge (where A^0 is constrained), we have lost manifest Lorentz invariance; but since the Maxwell theory is Lorentz- and gauge-invariant, the final physical results will be both Lorentz-covariant and independent of gauge. Alternatively, we can quantize in the Lorentz gauge ($\partial_\nu A^\nu = 0$) by modifying the field equation, adding $\partial_\mu(\partial{\cdot}A)$, and treating A^0 as a dynamical variable. The commutation relations for the fields will appear in a more familiar form, $[A^\mu(t,\boldsymbol{x}),\, \dot{A}^\nu(t,\boldsymbol{y})] = -\mathrm{i}g^{\mu\nu}\delta(\boldsymbol{x}-\boldsymbol{y})$, but the extra minus sign in $[A^0,\, \dot{A}^0]$ indicates that the A^0 generates a space of indefinite metric. It turns out that quantization with constraints can be more easily done in a functional integral formulation, especially in the case of non-Abelian fields, as we will see in Chap. 15.

2.5.4 Field Energy and Momentum

As an application, we may calculate the energy and momentum of the electromagnetic field, given by the classical expressions

$$H = \tfrac{1}{2} \int \mathrm{d}^3x \, (\boldsymbol{E}^2 + \boldsymbol{B}^2) \,,$$
$$\boldsymbol{P} = \int \mathrm{d}^3x \, (\boldsymbol{E} \times \boldsymbol{B}) \,. \tag{2.158}$$

The same expressions remain valid in quantum mechanics, only with $\boldsymbol{E}$ and $\boldsymbol{B}$ interpreted as operators. In the gauge chosen, the fields are

$$\boldsymbol{B} = \nabla \times \boldsymbol{A} = \sum_{k,\lambda} C_{\boldsymbol{k}} (\mathrm{i}\boldsymbol{k} \times \boldsymbol{e}(k,\lambda)) \left[a(\boldsymbol{k},\lambda) \mathrm{e}^{-\mathrm{i}k.x} - a^\dagger(\boldsymbol{k},\lambda) \mathrm{e}^{\mathrm{i}k\cdot x} \right] ,$$
$$\boldsymbol{E} = \; -\frac{\partial \boldsymbol{A}}{\partial t} = \sum_{k,\lambda} C_{\boldsymbol{k}} \, \mathrm{i}\omega \, \boldsymbol{e}(k,\lambda) \left[a(\boldsymbol{k},\lambda) \mathrm{e}^{-\mathrm{i}k.x} - a^\dagger(\boldsymbol{k},\lambda) \mathrm{e}^{\mathrm{i}k\cdot x} \right] ,$$

with normalization constant $C_{\boldsymbol{k}} = \sqrt{1/(2\pi)^3 \, 2\omega}$ and energy $\omega \equiv |\boldsymbol{k}|$. The vectors $\boldsymbol{e}(k,\lambda)$ have been assumed, without loss in generality, to be real. The final results are

$$H \; = \sum_{k,\lambda} \omega \, a^\dagger(\boldsymbol{k},\lambda) a(\boldsymbol{k},\lambda) \,; \tag{2.159}$$

$$\boldsymbol{P} = \sum_{k,\lambda} \boldsymbol{k} \, a^\dagger(\boldsymbol{k},\lambda) a(\boldsymbol{k},\lambda) \,. \tag{2.160}$$

It follows from (159) and (160) that the energy and momentum of the one-photon state are

$$H a^\dagger(\boldsymbol{k},\lambda) \; |0\rangle = [H, a^\dagger(\boldsymbol{k},\lambda)] \; |0\rangle = \omega \, a^\dagger(\boldsymbol{k},\lambda) \; |0\rangle \;,$$
$$\boldsymbol{P} a^\dagger(\boldsymbol{k},\lambda) \; |0\rangle = [\boldsymbol{P}, a^\dagger(\boldsymbol{k},\lambda)] \; |0\rangle = \boldsymbol{k} \, a^\dagger(\boldsymbol{k},\lambda) \; |0\rangle \;.$$

Since $|\boldsymbol{k}|^2 = \omega^2$, the photon is massless: $(\text{mass})^2 = \omega^2 - |\boldsymbol{k}|^2 = 0$. The operator $a^\dagger(\boldsymbol{k},\lambda)$ creates a photon of energy ω, momentum $\boldsymbol{k}$, and polarization λ. The wave functions for a photon absorbed and a photon emitted at point x are given by

$$< 0|A_\mu(x)|k,\lambda > = e_\mu(k,\lambda) e^{-ik.x} \qquad \xrightarrow{k,\lambda} \; \sim\!\sim\!\sim\!\bigcirc \quad ,$$

$$< k,\lambda|A_\mu(x)|0 > = e^*_\mu(k,\lambda) e^{ik.x} \qquad \bigcirc\!\sim\!\sim\!\sim \xrightarrow{k,\lambda} \quad .$$

2.6 The Action

In particle physics, one of the most useful theoretical tools is the Lagrangian, more exactly the Lagrangian density. As it constitutes the complete definition of the model, from which the dynamic field equations are generated, one can incorporate into it from the beginning all the characteristics one wishes to see emerge from the model.

2.6.1 The Euler–Lagrange Equation

In the Hamiltonian formulation of classical mechanics, the equations of motion of a system of particles can be derived from a certain function of the generalized coordinates $q(t)$ and velocities $\dot{q}(t)$ of the particles, called the *action*,

$$S = \int_{t_1}^{t_2} \mathrm{d}t\, L(q(t), \dot{q}(t))\,, \tag{2.161}$$

where L is the Lagrange function. Hamilton's minimum action principle holds that, among all possible paths joining any two fixed points at times t_1 and t_2 ($t_2 > t_1$), the path for which S is minimum corresponds to the physical path that determines the actual motion of the particles. Thus, by varying $q \to q + \delta q$, subject to the constraint $\delta q(t_1) = \delta q(t_2) = 0$, one gets

$$S \to S + \delta S\,,$$

$$\delta S = \int_{t_1}^{t_2} \mathrm{d}t \left(\frac{\partial L}{\partial q} \delta q + \frac{\partial L}{\partial \dot{q}} \delta \dot{q} \right). \tag{2.162}$$

After an integration by part of the second term on the right-hand side, this becomes

$$\delta S = \int_{t_1}^{t_2} \mathrm{d}t \left(\frac{\partial L}{\partial q} - \frac{\mathrm{d}}{\mathrm{d}t} \frac{\partial L}{\partial \dot{q}} \right) \delta q + \int_{t_1}^{t_2} \mathrm{d}t \frac{\mathrm{d}}{\mathrm{d}t} \left(\frac{\partial L}{\partial \dot{q}}\, \delta q \right). \tag{2.163}$$

The last term on the right-hand side vanishes because $\delta q(t_1) = \delta q(t_2) = 0$. The variation δq being arbitrary, it follows that the minimization condition on δS implies the Euler–Lagrange equation of classical mechanics

$$\frac{\partial L}{\partial q} - \frac{\mathrm{d}}{\mathrm{d}t} \frac{\partial L}{\partial \dot{q}} = 0 \qquad \text{(Euler–Lagrange equation)}\,. \tag{2.164}$$

One may then introduce momentum p conjugate to coordinate q via

$$p = \partial L / \partial \dot{q}\,, \tag{2.165}$$

a relation that can be inverted to express $\dot{q}$ in terms of p and q. The Hamiltonian is now defined by a Lagrange transformation as a function of the dynamical variables p and q:

$$H = p\dot{q} - L\,, \tag{2.166}$$

When the situation is extended to a continuous distribution, as in field theories, each generalized coordinate $q(t)$ is replaced by a field, generically called $\varphi(t, \boldsymbol{x})$, that represents a continuum of particles in space-time, and each corresponding velocity $\dot{q}(t)$ by the four-derivative $\partial_\mu \varphi$. In a local field theory the Lagrangian can be written as a spatial integral of a functional of the field φ and its derivative $\partial_\mu \varphi$, called the *Lagrangian density*, $\mathcal{L}$,

$$L = \int \mathrm{d}^3x \, \mathcal{L}(\varphi, \partial_\mu \varphi) \,, \tag{2.167}$$

and the action assumes the form

$$S = \int_{\tau_1}^{\tau_2} \mathrm{d}^4x \, \mathcal{L}(\varphi, \partial_\mu \varphi) \,, \tag{2.168}$$

where τ_1 and τ_2 represent the limiting surfaces of integration. As in classical mechanics, L has the dimension of energy and S that of angular momentum. In natural units, $[S] = 1$ and $[\mathcal{L}] = [E]^4$. From now on we will refer to $\mathcal{L}$ simply as the 'Lagrangian', with little risk of confusing it with L since the latter function will scarcely reappear.

Now, in an arbitrary variation of the field

$$\varphi(x) \to \varphi'(x) = \varphi(x) + \delta_0 \varphi(x) \,, \tag{2.169}$$

such that $\delta_0 \varphi = 0$ at the integration limits, the action changes $S \to S + \delta S$ by an amount

$$\begin{aligned} \delta S &= \int_{\tau_1}^{\tau_2} \mathrm{d}^4x \left(\frac{\partial \mathcal{L}}{\partial \varphi} \delta_0 \varphi + \frac{\partial \mathcal{L}}{\partial(\partial_\mu \varphi)} \, \delta_0(\partial_\mu \varphi) \right) \\ &= \int_{\tau_1}^{\tau_2} \mathrm{d}^4x \left(\frac{\partial \mathcal{L}}{\partial \varphi} - \partial_\mu \frac{\partial \mathcal{L}}{\partial(\partial_\mu \varphi)} \right) \delta_0 \varphi + \int_{\tau_1}^{\tau_2} \mathrm{d}^4x \, \partial_\mu \Big(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \varphi)} \, \delta_0 \varphi \Big) . \end{aligned} \tag{2.170}$$

The second integral on the right-hand side vanishes because of the boundary condition, $\delta_0 \varphi = 0$ on the integration surfaces. Then the demand that S be stationary for such an arbitrary variation $\delta_0 \varphi$ immediately implies the Euler–Lagrange equation for the field φ :

$$\frac{\partial \mathcal{L}}{\partial \varphi} - \partial_\mu \frac{\partial \mathcal{L}}{\partial(\partial_\mu \varphi)} = 0 \qquad \text{(Euler–Lagrange equation).} \tag{2.171}$$

This equation can be readily generalized to cases with several fields of arbitrary tensor characters. For each independent field component (φ_i), there is an equation of the form (171). As a functional of the fields $\varphi_i(x)$ and their conjugate momentum densities, $\pi_i(x)$, the Hamiltonian density is given by

$$\mathcal{H}(\pi_i, \varphi_i) = \sum_i \left(\pi_i \, \frac{\partial \varphi_i}{\partial t} \right) - \mathcal{L} \,, \quad \text{where} \quad \pi_i = \frac{\partial \mathcal{L}}{\partial \left(\partial \varphi_i / \partial t \right)} \,. \tag{2.172}$$

How to find $\mathcal{L}$? If the equations of motion for a system are known, it is always possible to find the corresponding Lagrangian. Let us consider the three cases we have studied so far, and for each let us write down the dynamic equations and the associated Lagrangian.

(a) *Real scalar field:*

$$(\Box + m^2)\phi = -\frac{\lambda}{6}\phi^3\,;$$
$$\mathcal{L} = \frac{1}{2}\partial_\mu\phi\,\partial^\mu\phi - \frac{1}{2}m^2\phi^2 - \frac{\lambda}{4!}\,\phi^4\,. \tag{2.173}$$

The expression for the Lagrangian contains, in order, a kinematic term, a mass term (for particle of mass m), and a self-interaction term (of coupling constant λ). It can be checked that the Euler–Lagrange equation derived from $\mathcal{L}$ coincides precisely with the dynamic equation given. As $\mathcal{L}$ has the dimension of $[E]^4$, the field ϕ has dimension $[E]$, and the coupling constant λ is dimensionless. The momentum density conjugate to ϕ is $\pi \equiv \partial\phi/\partial t$.

(b) *Complex scalar field:*

$$(\Box + m^2)\phi = -2\lambda\phi(\phi^*\phi)\,, \qquad (\Box + m^2)\phi^* = -2\lambda\phi^*(\phi^*\phi)\,; \tag{2.174}$$

$$\mathcal{L} = \partial_\mu\phi^*\partial^\mu\phi - m^2\phi^*\phi - \lambda(\phi^*\phi)^2\,. \tag{2.175}$$

Note the absence of factors $1/2$ in the present case; but with $\phi = \frac{1}{\sqrt{2}}(\phi_1 + \mathrm{i}\phi_2)$, (175) will coincide with the Lagrangian for two real fields ϕ_1 and ϕ_2. The momentum density conjugate to ϕ is $\pi \equiv \partial\phi^*/\partial t$.

(c) *Real vector field:*

$$\partial_\mu F^{\mu\nu} + m^2 A^\nu = j^\nu\,, \tag{2.176}$$

$$\mathcal{L} = -\tfrac{1}{4}F_{\mu\nu}F^{\mu\nu} + \tfrac{1}{2}\,m^2 A_\mu A^\mu - j_\mu A^\mu\,. \tag{2.177}$$

The overall sign in $\mathcal{L}$ is conventional, and the sign of the mass term for the physical degrees of freedom (spacelike components A^i) agrees with that found in the scalar field case. The vector field has dimension $[A_\mu] = [E]$. For each component A_μ considered as an independent field, one gets

$$\frac{\partial\mathcal{L}}{\partial A_\nu} = m^2 A^\nu - j^\nu\,,$$

and then

$$\begin{aligned}\partial_\mu\frac{\partial\mathcal{L}}{\partial(\partial_\mu A_\nu)} &= -\frac{1}{4}\partial_\mu\frac{\partial}{\partial(\partial_\mu A_\nu)}[(\partial_\rho A_\sigma - \partial_\sigma A_\rho)(\partial^\rho A^\sigma - \partial^\sigma A^\rho)] \\ &= -\frac{1}{4}\partial_\mu\frac{\partial}{\partial(\partial_\mu A_\nu)}(2\partial_\rho A_\sigma\partial^\rho A^\sigma - 2\partial_\rho A_\sigma\partial^\sigma A^\rho) \\ &= -\partial_\mu(\partial^\mu A^\nu - \partial^\nu A^\mu) \;=\; -\partial_\mu F^{\mu\nu}\,.\end{aligned}$$

Hence the Euler–Lagrange equations for A^ν correctly reproduce the equations of motion of the vector field.

On the other hand, if the dynamic equations are unknown, which is the case most often encountered in practice and where the Lagrangian approach turns out to be the most fruitful, one can construct the Lagrangian for the model by imposing certain conditions, general enough for the model to be useful, yet restrictive enough to make it well defined, even unique. The most important conditions one may consider are

(a) $\mathcal{L}$ must be Hermitian, so that the Hamiltonian itself is Hermitian;
(b) it must be relativistically invariant, that is, invariant to space-time translations and Lorentz transformations;
(c) it should be invariant to other universal symmetries, but may break certain symmetries in some definite manner, as required by experiments;
(d) it could be limited so that it contains no space-time derivatives of fields higher than the first (so that the field equations are at most of second order), and so that it contains only local couplings built up from field quantities evaluated at the same space-time point.

2.6.2 Conserved Current

The advantages of using a formalism based on the Lagrangian density $\mathcal{L}$ become apparent when we want to study the implications of the symmetries that the model described by $\mathcal{L}$ may have. In particular, it can be shown that there is a close relationship between the *invariance of the action* in an arbitrary *continuous global transformation* and the existence of a *conserved current*. This important result is referred to as *Noether's theorem*.

The symmetries of a model may be uncovered by studying the changes in the Lagrangian or the action function that defines the model following transformations on space-time coordinates or on internal variables. Leaving the second class of symmetries for later chapters, we limit our discussion for now to coordinate transformations. To begin, consider the following infinitesimal coordinate transformations and their subsequent effects on the typical field:

$$\begin{aligned} x^\mu \to x'^\mu \quad &= x^\mu + \delta x^\mu \,, \\ \varphi(x) \to \varphi'(x') &= \varphi(x) + \delta\varphi(x) \,. \end{aligned} \tag{2.178}$$

For example, for a translation, $\delta x^\mu = -a^\mu$, while for a Lorentz transformation, $\delta x^\mu = \epsilon^\mu{}_\nu \, x^\nu$. Generally, δx^μ may depend on coordinates even if the transformation parameters are themselves constant.

The total variation of the field, $\delta\varphi(x)$, may be determined by a series expansion of $\varphi'(x')$:

$$\begin{aligned} \delta\varphi(x) &= \varphi'(x') - \varphi(x) \\ &= (\varphi'(x) - \varphi(x)) + (\varphi'(x') - \varphi'(x)) \\ &\approx \delta_0\varphi(x) + \delta x^\mu \, \partial_\mu \varphi(x) \,. \end{aligned} \tag{2.179}$$

It consists in general of a variation of the field due to its coordinate dependence and of a variation specific to the functional structure of the field, $\delta_0\varphi$. This variation is left arbitrary and, in particular, need *not* vanish on the integration surfaces, in contrast to the situation where one wants to derive the Euler–Lagrange equation. The action function, defined for an arbitrary volume V, varies by an amount

$$\delta S = \int_V \delta(\mathrm{d}^4x)\mathcal{L} + \int_V \mathrm{d}^4x\,\delta\mathcal{L}\,, \tag{2.180}$$

in which the two terms arise from the variations of the integration volume element and from the variations of the Lagrangian itself.

To calculate the variation $\delta(\mathrm{d}^4x)$, consider the element of volume in the transformed coordinates

$$\begin{aligned}\mathrm{d}^4x' &= \left|\det\left(\frac{\partial x'^\mu}{\partial x^\nu}\right)\right| \mathrm{d}^4x = |\det[\delta^\mu{}_\nu + \partial_\nu(\delta x^\mu)]|\ \mathrm{d}^4x \\ &\approx [1 + \partial_\mu(\delta x^\mu)]\,\mathrm{d}^4x\,,\end{aligned} \tag{2.181}$$

which implies

$$\delta(\mathrm{d}^4x) = (\partial_\mu(\delta x^\mu))\,\mathrm{d}^4x\,. \tag{2.182}$$

The first term in (180) then becomes

$$\int_V \delta(\mathrm{d}^4x)\mathcal{L} = \int_V \mathrm{d}^4x\,(\partial_\mu(\delta x^\mu))\,\mathcal{L}\,. \tag{2.183}$$

In the second term of (180), the variation $\delta\mathcal{L}$ arises from both δx and $\delta\varphi$, the latter coming from δx and $\delta_0\varphi$. The x variations of $\mathcal{L}$ as function of x yield the quantity $\int \mathrm{d}^4x\,(\partial\mathcal{L}/\partial x^\mu)\delta x^\mu$, while contributions from $\delta_0\varphi$ are given by the same expression as in (170), which includes two terms. If φ satisfies the equation of motion (171), which is being assumed, the first term vanishes, leaving only the second, nonzero term (which vanished there because of the imposed boundary conditions). Therefore,

$$\int_V \mathrm{d}^4x\,\delta\mathcal{L} = \int_V \mathrm{d}^4x\,\left[(\partial_\mu\mathcal{L})\delta x^\mu + \partial_\mu\left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\varphi)}\,\delta_0\varphi\right)\right]\,. \tag{2.184}$$

Adding together both terms (183) and (184), we get for the total variation of the action

$$\delta S = \int_V \mathrm{d}^4x\,\partial_\mu\left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\varphi)}\,\delta_0\varphi + \mathcal{L}\delta x^\mu\right)\,. \tag{2.185}$$

We are interested in *global transformations*, that is, those with constant parameters. The corresponding infinitesimal transformations are defined by

the small constant parameters $\delta\omega_a$, with the values of the subscripts left undefined for now. Then invariance of S to a symmetry transformation defined by $\delta\omega_a$ means that its variations in an arbitrary finite volume vanish:

$$\frac{\partial S}{\partial \omega_a} \equiv \int_V \mathrm{d}^4x\, \partial_\mu j_a^\mu = 0\,, \tag{2.186}$$

for every value of a. Since this is true for an arbitrary volume, the integrand itself vanishes, and so does the divergence:

$$\partial_\mu j_a^\mu = 0\,. \tag{2.187}$$

The four-vector current introduced is shorthand for

$$j_a^\mu = \frac{\partial \mathcal{L}}{\partial(\partial_\mu \varphi)}\,\frac{\delta_0 \varphi}{\delta \omega_a} + \mathcal{L}\frac{\delta x^\mu}{\delta \omega_a}\,. \tag{2.188}$$

Note that contributions from variations of the coordinates and of the field are well separated in the two terms. When $\mathcal{L}$ contains several fields φ_i for $i = 1, 2, \ldots$, all fields contribute to the current

$$j_a^\mu = \frac{\partial \mathcal{L}}{\partial(\partial_\mu \varphi_i)}\,\frac{\delta_0 \varphi_i}{\delta \omega_a} + \mathcal{L}\,\frac{\delta x^\mu}{\delta \omega_a}\,. \tag{2.189}$$

The current conservation in (187) implies that the associated 'charge' defined by the space integral of the current time component

$$Q_a \equiv \int_V \mathrm{d}^3x\, j_a^0(t, \boldsymbol{x}) \tag{2.190}$$

is a *constant of the motion*, since

$$\begin{aligned}\frac{dQ_a}{\mathrm{d}t} &= \int_V \mathrm{d}^3x\, \partial_0\, j_a^0 \;= \int_V \mathrm{d}^3x\, (\partial_\mu j_a^\mu - \nabla\cdot \boldsymbol{j}_a) \\ &= -\int_S \mathrm{d}S\cdot \boldsymbol{j}_a \;= 0\,,\end{aligned} \tag{2.191}$$

where use has been made of (187) and of Gauss's theorem. In the last step it is assumed that $\boldsymbol{j} = 0$ on the surface of integration.

In summary, when the action function is invariant to a continuous symmetry transformation of the coordinates and of the fields involved in the model, a locally conserved density current and a conserved (constant in time) charge associated with the symmetry can be defined. In other words, if the fields or the different components of each of the tensor fields describing the model can be transformed among themselves without changing the physical content of the fields and their interactions, the system possesses a symmetry, and this symmetry implies well-defined conserved quantities. This result, due to

Emmy Noether, makes mathematically precise the relationship between the symmetries of a system and the existence of conserved quantum numbers, and therefore, between dynamics and conservation laws. In the examples that follow, we consider, for simplicity, free fields. But it must be kept in mind that Noether's theorem may be applied to general situations, classical or quantized, with or without interactions.

Example 2.4 Translation of a Generic Field
Consider the translation $x^\mu \to x'^\mu = x^\mu - \delta a^\mu$. It implies the following variations:

$$\delta x^\mu = -\delta a^\mu\,, \quad \delta_0\varphi = \delta a^\mu\,\partial_\mu\varphi\,;$$
$$\delta\omega_a \equiv \delta a^\mu\,,$$

which then give

$$\frac{\delta x^\mu}{\delta a^\nu} = -\delta^\mu{}_\nu, \quad \frac{\delta_0\varphi}{\delta a^\nu} = \partial_\nu\varphi\,.$$

If the action for a generic field $\varphi(x)$ is invariant to space-time translations, the conserved current and charge associated with this invariance are

$$j^\mu{}_\nu \equiv \mathcal{T}^\mu{}_\nu = \left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\varphi)}\right)\partial_\nu\varphi - \delta^\mu{}_\nu\,\mathcal{L}\,, \tag{2.192}$$

$$P_\nu \equiv T^0_\nu \quad = \int \mathrm{d}^3x\,\mathcal{T}^0{}_\nu\,. \tag{2.193}$$

It turns out that the current, $\mathcal{T}^\mu{}_\nu$, is the energy-momentum tensor of the system and the charge, P_ν, is the energy-momentum vector. In particular, the energy of the field can be identified with P_0 :

$$P_0 \;=\; H \;=\; \int \mathrm{d}^3x\,(\pi\dot\varphi - \mathcal{L})\,, \tag{2.194}$$

where the velocity and momentum conjugates to the field are

$$\dot\varphi = \partial_0\varphi \quad \text{and} \quad \pi = \partial\mathcal{L}/\partial\dot\varphi\,,$$

in agreement with the classical relation (166).

Example 2.5 Lorentz Transformation of a Real Scalar Field
In this case, the variations are

$$\delta x^\mu = \epsilon^\mu{}_\nu\,x^\nu = \epsilon^{\mu\nu}\,x_\nu\,; \qquad \delta_0\phi = -\frac{\mathrm{i}}{2}\,\epsilon^{\mu\nu}\,L_{\mu\nu}\,\phi\,.$$

Letting $\delta\omega_a \equiv \epsilon^{\rho\sigma}$, we obtain the derivatives

$$\frac{\delta x^\mu}{\delta\epsilon^{\rho\sigma}} = \delta^\mu{}_\rho\,x_\sigma - \delta^\mu{}_\sigma\,x_\rho\,,$$
$$\frac{\delta_0\phi}{\delta\epsilon^{\rho\sigma}} = -\mathrm{i}L_{\rho\sigma}\phi = (x_\rho\,\partial_\sigma - x_\sigma\,\partial_\rho)\,\phi\,.$$

The Lagrangian for a scalar field being given Lorentz-invariant, we can immediately deduce the associated current (called the angular momentum density tensor)

$$j^{\mu}{}_{\rho\sigma} \equiv \mathcal{M}^{\mu}{}_{\rho\sigma} = x_{\rho}\,\mathcal{T}^{\mu}{}_{\sigma} - x_{\sigma}\,\mathcal{T}^{\mu}{}_{\rho}\,, \tag{2.195}$$

and the corresponding conserved charge (or angular momentum)

$$Q_{\rho\sigma} \equiv M_{\rho\sigma} = \int \mathrm{d}^3x\,\mathcal{M}^{0}{}_{\rho\sigma}\,. \tag{2.196}$$

The space components of $Q_{\rho\sigma}$ give the familiar angular momentum

$$M_{ij} = \int \mathrm{d}^3x\,(x_i\mathcal{P}_j - x_j\mathcal{P}_i)\,, \tag{2.197}$$

where $\mathcal{P}_i = \mathcal{T}^{0}{}_{i}$ is the momentum density of the field.

Example 2.6 Internal Transformation of a Complex Scalar Field
We consider here an example of transformation on internal space. In such a transformation, space-time coordinates are not affected, $\delta x^{\mu} = 0$, only the fields vary, $\delta\phi = \delta_0\phi \neq 0$. As an example of physical interest, consider the Lagrangian for a free complex scalar field (175). It is clearly invariant to the *phase transformations*

$$\begin{aligned}\phi &\to \phi' = \mathrm{e}^{-\mathrm{i}\alpha}\phi \approx \phi - \mathrm{i}\alpha\,\phi\,,\\ \phi^* &\to \phi'^* = \mathrm{e}^{\mathrm{i}\alpha}\phi^* \approx \phi^* + \mathrm{i}\alpha\phi^*\,,\end{aligned} \tag{2.198}$$

where α is a real constant. Posing $\delta\omega \equiv \delta\alpha$, one gets the derivatives

$$\frac{\delta_0\phi}{\delta\alpha} = -\mathrm{i}\phi\,, \qquad \frac{\delta_0\phi^*}{\delta\alpha} = \mathrm{i}\phi^*\,, \tag{2.199}$$

and subsequently the associated current and charge

$$j^{\mu} = \mathrm{i}\left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi^*)}\,\phi^* - \frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)}\,\phi\right)\,, \tag{2.200}$$

$$Q = \mathrm{i}\int \mathrm{d}^3x\,\left(\frac{\partial\mathcal{L}}{\partial(\partial_0\phi^*)}\,\phi^* - \frac{\partial\mathcal{L}}{\partial(\partial_0\phi)}\,\phi\right)\,. \tag{2.201}$$

Thus, the Noether current associated with the phase invariance is identical to the charge current density postulated in Sect. 2.4.2. ■

Problems

2.1 Spin of π^0 meson. (a) Show that two real photons in the reference frame of their center-of-mass (where the total momentum is zero) cannot be in a state of angular momentum 1. Your proof will be based on rotational invariance, Bose statistics, the transversality of photons, and the superposition principle (which says that the state of two photons is a homogeneous linear function of their polarization vectors).
(b) The π^0 meson decays mainly through the channel $\pi^0 \to 2\gamma$. Show that if angular momentum is conserved, the π^0 spin cannot be 1 (it is known now that its spin is 0).
(c) Given that the π^0 spin is 0, show that in the rest frame of π^0, the two photons emitted in the decay have the same polarization.

2.2 Spin of K^0 meson. The mode $K^0 \to 2\pi^0$ accounts for 31% of all K^0 meson decays. Given that π^0 spin is 0, show that the K^0 spin is an even number (in $\hbar$ units). We now know that its spin is 0.

2.3 Dilation. Consider the Lagrangian for a real scalar field $\phi(x)$ in four-space-time,

$$\mathcal{L} = \frac{1}{2}(\partial\phi)^2 - \frac{1}{2}m^2\phi^2 - \frac{1}{4!}g\phi^4 .$$

Under a scale transformation, $x' = \lambda^{-1}x$ and $\phi'(x) = \exp[D\ln\lambda]\phi(\lambda x)$, (with $\lambda > 0$). Show that the action is invariant when $D = 1$ and $m = 0$, with the field satisfying the equation of motion.

2.4 Angular momentum of the electromagnetic field. Consider an electromagnetic field in a region where the sources are absent. In an infinitesimal Lorentz transformation (cf. Sect. 2.3) the variations are defined by

$$\delta x^\mu = \epsilon^\mu{}_\nu\, x^\nu ,$$
$$\delta_0 A^\alpha(x) = -\frac{\mathrm{i}}{2}\,\epsilon^{\mu\nu}\,(L_{\mu\nu} + \Sigma_{\mu\nu})\,A^\alpha(x) .$$

(a) Show that the density of the energy-momentum tensor is given by

$$\mathcal{T}^\mu{}_\nu \;=\; -F^{\mu\lambda}\partial_\nu A_\lambda + \delta^\mu{}_\nu\frac{1}{4}F_{\rho\sigma}F^{\rho\sigma},$$

and that the density of the angular momentum tensor is given by

$$\mathcal{M}^\mu_{\rho\sigma} \;=\; x_\rho\mathcal{T}^\mu{}_\sigma - x_\sigma\mathcal{T}^\mu{}_\rho \;-\; F^{\mu\lambda}(g_{\lambda\rho}A_\sigma - g_{\lambda\sigma}A_\rho) .$$

(b) The intrinsic spin of the field is defined by

$$S^i = \frac{1}{2}\epsilon^{ijk}S_{jk} ,$$

where

$$S_{jk} = \int \mathrm{d}^3x\, \mathcal{M}^0_{jk} \,=\, -\int \mathrm{d}^3x\, (\dot{A}_j A_k \,-\, \dot{A}_k A_j)\,.$$

Prove that

$$\begin{aligned}\boldsymbol{S} &= -\mathrm{i}\sum_k \hat{\boldsymbol{k}}\,[a^\dagger(\boldsymbol{k},1)a(\boldsymbol{k},2) \,-\, a^\dagger(\boldsymbol{k},2)a(\boldsymbol{k},1)] \\ &= \sum_k \hat{\boldsymbol{k}}\,[a^\dagger(\boldsymbol{k},+)a(\boldsymbol{k},+) \,-\, a^\dagger(\boldsymbol{k},-)a(\boldsymbol{k},-)]\end{aligned}$$

and hence, with $S_\parallel \equiv \boldsymbol{S}\cdot\hat{\boldsymbol{k}}$,

$$[S_\parallel,\, a^\dagger(\boldsymbol{k},\pm)] \,=\, \pm a^\dagger(\boldsymbol{k},\pm)\,,$$

where $a(\boldsymbol{k},\lambda)$ and $a^\dagger(\boldsymbol{k},\lambda)$ are defined in Sect. 2.5.3.

Suggestions for Further Reading

For their historical value:
Gordon, W., Z. Phys. **40** (1926) 117, 121
Klein, O., Z. Phys. **37** (1926) 895
Pauli, W. and Weisskopf, V. F., Helv. Phys. Acta **7** (1934) 709
Proca, A., J. Phys. Radium **9** (1938) 61
Detailed study of the Klein–Gordon equation:
Feshbach, H. and Villars, F., Rev. Mod. Phys. **30** (1958) 24
Discussion on Noether's theorem:
Hill, E. L., Rev. Mod. Phys. **23** (1957) 253
Noether, E., Nachr. K. Geo. Wiss. Göttingen **37** (1918) 235