11 Canonical quantization of classical fields

In Chapters 6-9 relativistic quantum field theory was formulated adopting the view that the underlying physical system (the basic "ontology") are particles. Therefore the natural starting point was the relativistic (in the sense of existence in the Hilbert space of the Poincaré group generators P_0^μ and $J_0^{\mu\nu}$) theory of free particles which could be built by exploiting the formalism of second quantization of Chapter 5. Making strong assumptions about the relation between the spectra and the eigenvectors of the full and free Hamiltonians, $H = H_0 + V_{\rm int}$ and H_0, it was possible to show that interaction operators $V_{\rm int}$ constructed according to a certain set of rules can lead to relativistic (again, in the sense of the possibility to construct the Poincaré group generators P^μ and $J^{\mu\nu}$) theories of interacting particles and a perturbative scheme of computing S-matrix elements (strongly relying on these assuptions) was developed.

Here we present another approach to the formulation of the theory of interactions of relativistic particles. It is based on the canonical quantization of relativistic fields. Particles emerge in this formalism as states ("excitations") of quantized fields. The basic entities - at least as far as bosonic degrees of freedom are concerned - are, however, fluctuating fields. The advantage of the field theory approach is twofold: firstly, it provides a well defined prescription for constructing the full Poincaré group generators (the structure of which in terms of the creation and annihilation operators in the approach of Chapter 7 could only be guessed at) and, secondly, allows for an easy discussion of various internal symmetries, especially gauge ones (providing through the Noether theorem a concrete prescription for constructing the associated conserved charges) and their spontaneous breaking. Moreover, in the case of fields "the quanta" of which[1] are bosons (integer spin particles) the picture of fluctuating field values at different space-time points (complementary to the description in terms of particles) seems more fundamental; it allows for deeper understanding of global aspects of the theory (such as the role of topologically nontrivial classical field configurations, symmetry breaking etc.) and proves extremely useful in diverse modern applications of quantum field theory (e.g. in cosmology).

The picture of fluctuating fields seems however not naturally applicable to quantum fields, the "quanta" of which are fermions (half-integer spin particles). While it is possible to take as a starting point of quantum theories of fermions the Lagrangian formalism which has formal aspects analogous to that of classical fields (Section 11.8), the basic entities in this case are anticommuting generators of an abstract Grassmann (or Bieriezin) algebra which have no classical counterpart and can hardly be considered physical. The difference between bosonic and fermionic fields[2] becomes particularly clear in the path integral

[1]We put the quotation marks because, as will become clear in Chapter 13, the correspondence between the fields entering the Lagrangian and the particle states predicted by a given theory need not be direct, nor one-to-one.

[2]The reason for this sharp difference between bosonic and fermionic fields seems to be that in the classical, $\hbar \to 0$, limit bosonic quantum fields go over into true, measurable (at least in principle) classical fields of forces (like the electromagnetic forces, and various potential forces), whereas fermionic fields get

approach to field quantization (see Chapter 16). While quantizing in this way bosonic fields one integrates functionally over (that is performs summation of contributions of) all possible classically realizable (and measurable) field configurations, quantizing fermionic fields one uses for this purposes an operation (formally defined for Grassmann algebra generators) only formal aspects of which resemble the integration. Thus no "classical fermionic fields" can be measurable and the only physical picture that in quantum field theory can be associated with fermions seems to be that given in Chapters 5 and 6. Quantization of classical fermion fields is only a mathematical trick allowing to easily reproduce the results of the formalism presented there. In particular it cannot be viewed (as is sometimes presented in older textbooks) as "quantization of the wave function satisfying the Dirac equation".

In this chapter quantum theories of fields are constructed by using the operator method, that is by applying to their classical counterparts the so-called canonical quantization procedure. It consists of choosing a set of field variables in terms of which the Lagrangian is written, identifying the associated set of canonical momenta and constructing the corresponding Hamilton's formalism, imposing on operators which are going to represent the canonical variables the canonical commutation rules (anticommutation rules in the case of Bierezin algebra valued fields) and constructing a Hilbert space (and selecting in it a proper Fock space) in which these operators act. Therefore, in the first section a brief review of the classical theory of relativistic fields and their symmetries is given. As the next step we show, that quantum theory of noninteracting fields, Lagrangians of which consist of terms at most bilinear in field variables, leads to the interpretation of the corresponding Hamiltonian eigenstates in terms of (noninteracting) particles. In principle, quantization of classical fields is very similar to the quantization of the system of many coupled harmonic oscillators discussed in Section 5.6: the field value at a given point $\mathbf{x}$ of the space can be treated as an oscillator coupled to other oscillators in neighbouring points (in particular, the common features of both systems: of quantized coupled anharmonic oscillators and of interacting quantized fields, is the nonconservation of the numbers of phonons and particles, respectively). However, despite these deep similarities, quantization of some classical fields (e.g. the electromagnetic one) require special treatment because direct construction of the Hamilton's formalism cannot be carried out. The proper operator quantization of such fields requires developing the formalism of constrained system which will be presented in Section 11.6. The formalism allowing to treat fermionic fields on formally equal footing with bosonic ones is outlined in Section 11.8. After quantizing various types of classical field theories we apply to the theories of interacting fields the so-called transition to the interaction picture which allows to compute the S-matrix elements using the methods developed in Chapters 7 and 9.

---

replaced by matter particles the behaviour of which is first, at low energies, captured in terms of wave functions satisfying the Schrödinger (or Pauli) equation (as in Chapter 3) and ultimately, in the $\hbar \to 0$ limit, become paricles of classical mechanics.

11.1 Action and the Noether theorem

Dynamics of a system of classical fields[3] ϕ_i is determined by the action $I[\phi]$ which is a functional over all differentiable space-time field configurations. Actions considered in connection with quantum field theories are given in terms of Lagrangians L obtained as space integrals of local Lagrangian densities

$$I[\phi] = \int_{t_0}^{t_1} dt\, L = \int_{t_0}^{t_1} dt \int_V d^3\mathbf{x}\, \mathcal{L}(\phi, \partial_\mu \phi) \equiv \int_\Omega d^4x\, \mathcal{L}(\phi, \partial_\mu \phi)\,. \tag{11.1}$$

The Lagrangian density $\mathcal{L}$ is a function of fields ϕ_i, $i = 1, \ldots, N$ and of their space-time derivatives. In the following we consider Lagrangian densities depending on first order derivatives only. The volume V may be finite or infinite; in the latter case one assumes that all the fields ϕ_i vanish sufficiently fast at spatial infinity; in the finite volume V some spatial boundary conditions must be specified. The action $I[\phi]$ must be a real and - if the field theory is to be relativistic (which is possible only if the volume is infinite) - Poincaré invariant quantity. In general the fields ϕ_i transform nontrivially (as some, in general reducible, representation of the $SO(1,3)$ group or, in the case of fermionic fields to be discussed in Section 11.8, of its universal covering $SL(2,C)$) under changes of the inertial frame and can also transform nontrivially under some internal symmetries.

Equations of motion of a system of fields are obtained by requiring that the true field configurations $\phi_i(x)$ are stationary points of the action functional (11.1), i.e. that $I[\phi]$ does not change to first order in $\delta\phi_i(x)$ when the substitution $\phi_i(x) \to \phi_i(x) + \delta\phi_i(x)$ is made, provided the variations $\delta\phi_i(x)$ of the fields configuration around the stationary configuration $\phi_i(x)$ are bound to vanish at $t = t_1$ and $t = t_2$ as well as in the limit $|\mathbf{x}| \to \infty$ (if the volume V in (11.1) is finite, $\delta\phi_i(x)$ are assumed to vanish at the boundaries). Concisely this requirement is written as $\delta I[\phi] = 0$. If the action (11.1) depends only on fields and their first derivatives, its variation δI due to any variations $\delta\phi_i(x)$ of fields (not necessarily subject to some specific boundary conditions) reads[4]

$$\delta I = \int_\Omega d^4x \left[\frac{\partial \mathcal{L}}{\partial \phi_i}\, \delta\phi_i + \frac{\partial \mathcal{L}}{\partial(\partial_\nu \phi_i)}\, \delta(\partial_\nu \phi_i) \right] \tag{11.2}$$

(summation over the index i is understood). In the following we will consider such variations $\phi_i(x) \to \phi_i(x) + \delta\phi_i(x)$ that $\delta(\partial_\nu \phi_i) = \partial_\nu(\delta\phi_i)$. The variation δI of the action can

[3] The formalism presented in this section carries over unmodified to fermionic fields discussed in Section 11.8 provided all derivatives of the Lagrangian with respect to field variables are treated as right derivatives and the variations stand to the right of these derivatives.

[4] Generalization to Lagrangian densities depending on higher derivatives of field is straightforward but requires imposing appropriate boundary conditions also on derivatives of field variations. However the Hamilton's formalism presented here, which is crucial in formulating the transition to the quantum theory (without appealing to functional integrals which will be introduced in Chapter 16), is adapted to Lagrangian densities depending on first derivatives only. It can be extended to theories Lagrangians of which depend on higher derivatives using the formalism of constraints presented in Section 11.6.

be then rewritten in the form

$$\delta I = \int_\Omega d^4x \left\{ \left[\frac{\partial \mathcal{L}}{\partial \phi_i} - \partial_\nu \frac{\partial \mathcal{L}}{\partial(\partial_\nu \phi_i)} \right] \delta\phi_i + \partial_\nu \left[\frac{\partial \mathcal{L}}{\partial(\partial_\nu \phi_i)} \, \delta\phi_i \right] \right\} . \tag{11.3}$$

Due to the boundary conditions imposed on the variations $\delta\phi_i$ considered in connecttion with the determination of the true field configurations, the second, surface term in (11.3) vanishes and the condition $\delta I[\phi] = 0$ leads to the Euler-Lagrange equations of motions:

$$\partial_\nu \, \frac{\partial \mathcal{L}(\phi, \partial\phi)}{\partial(\partial_\nu \phi_i)} - \frac{\partial \mathcal{L}(\phi, \partial\phi)}{\partial \phi_i} = 0\,. \tag{11.4}$$

For further discussion it is important to observe that two Lagrangian densities which differ from one another by a total derivative of an arbitrary function of fields[5]

$$\mathcal{L}' = \mathcal{L} + \partial_\mu \mathcal{X}^\mu(\phi)\,, \tag{11.5}$$

give the same classical equations of motion (due to the assumed vanishing of the field variations on the boundary of the domain Ω).

Canonical quantization of a system of fields requires[6] going over to the Hamilton's formalism. To this end one defines the momenta $\Pi_i(t, \mathbf{x})$ canonically conjugated to the field variables $\phi_i(t, \mathbf{x})$ by

$$\Pi_i(t, \mathbf{x}) = \frac{\delta L}{\delta \dot{\phi}_i(t, \mathbf{x})} \equiv \frac{\partial \mathcal{L}}{\partial(\partial_0 \phi_i(t, \mathbf{x}))}\,, \tag{11.6}$$

and forms the Hamiltonian density

$$\mathcal{H}(x) = \sum_i \Pi_i(x) \, \dot{\phi}_i(x) - \mathcal{L}(x)\,, \tag{11.7}$$

and the Hamiltonian

$$H = \int d^3\mathbf{x} \, \mathcal{H}(t, \mathbf{x})\,. \tag{11.8}$$

We have assumed here that the momenta Π_i and the field variables ϕ_i are not subject to any constraints and that the relations (11.6) can be inverted to give

$$\dot{\phi}_i(x) = \dot{\phi}_i(\Pi(x), \phi(x))\,, \tag{11.9}$$

[5] $\mathcal{X}^\mu[\phi]$ depending only on fields but not their derivatives (as is required to get $\mathcal{L}'$ independent of higher field derivatives) can be constructed only in the presence of fermionic (the four-vector index μ can be then carried by a gamma or a sigma matrix) or four-vector fields; admiting $\mathcal{X}^\mu$ depending also on field derivatives is possible provided the allowed class of variations $\delta\phi_i(x)$ is restricted to those with derivatives vanish at the boundaries of Ω.

[6] In the alternative approach based on path integrals (see Chapter 16) fields are quantized using directly the action $I[\phi]$, but in fact a proper justification of thus approach also requires the Hamilton's formalism.

where the dependence on ϕ means possibly also a dependence on $\partial_i\phi(x)$. As we will see, in physically interesting cases this assumption is not always true and we will have to consider appropriate modifications of the canonical formalism. If no such problems arise, the equations of motions in the Hamilton's formalism read

$$\begin{aligned}\dot{\phi}_i(t,\mathbf{x}) &= \{\phi_i(t,\mathbf{x}),\ H\}_{\mathrm{PB}}\ ,\\ \dot{\Pi}_i(t,\mathbf{x}) &= \{\Pi_i(t,\mathbf{x}),\ H\}_{\mathrm{PB}}\ ,\end{aligned} \tag{11.10}$$

where the Poisson bracket (PB) of any two functionals $F[\phi,\Pi]$ and $G[\phi,\Pi]$ is defined as

$$\Big\{F[\phi,\Pi],\ G[\phi,\Pi]\Big\}_{\mathrm{PB}} \equiv \sum_i \int d^3\mathbf{x}\left(\frac{\delta F[\phi,\Pi]}{\delta\phi_i(t,\mathbf{x})}\frac{\delta G[\phi,\Pi]}{\delta\Pi_i(t,\mathbf{x})} - \frac{\delta G[\phi,\Pi]}{\delta\phi_i(t,\mathbf{x})}\frac{\delta F[\phi,\Pi]}{\delta\Pi_i(t,\mathbf{x})}\right), \tag{11.11}$$

so that

$$\begin{aligned}\{\phi_i(t,\mathbf{x}),\ \Pi_j(t,\mathbf{y})\}_{\mathrm{PB}} &= \delta_{ij}\,\delta^{(3)}(\mathbf{x}-\mathbf{y})\,,\\ \{\phi_i(t,\mathbf{x}),\ \phi_j(t,\mathbf{y})\}_{\mathrm{PB}} &= \{\Pi_i(t,\mathbf{x}),\ \Pi_j(t,\mathbf{y})\}_{\mathrm{PB}} = 0\,.\end{aligned} \tag{11.12}$$

If the canonical variables ϕ_i and Π_i are not subject to any constraints, the canonical equations of motion (11.10) are exactly equivalent to the original Euler-Lagrange equations (11.4).

Formulation of the field dynamics in terms of the action I allows to easily identify symmetries of the field equations of motion and, via the Noether theorem, of the corresponding conserved quantities. We begin with the case of symmetries which do not affect space-time coordinates (these can be ordinary global symmetries or, in the case of theories involving fermionic fields, "rigid" supersymmetries). We consider first a general change of the field variables which can always be written in the form

$$\phi_i(x) \to \phi'_i(x) = \phi'_i(\phi(x))\,. \tag{11.13}$$

The dynamics in the new field variables $\phi'_i(x)$ is detemined by a new Lagrangian density $\mathcal{L}'$ which depends on the fields ϕ'_i, and which can be chosen so that[7]

$$\mathcal{L}'(\phi',\partial_\mu\phi') = \mathcal{L}(\phi,\partial_\mu\phi)\,. \tag{11.14}$$

[7]This is just as in classical mechanics in which one is allowed to use any set of dynamical variables, $q_i(t)$ or $q'_i(t) = q'_i(q(t),t)$, to characterize the state of motion of a given system. The equations of motion in the new variables follow then from the new Lagrangian L'

$$\frac{d}{dt}\frac{\partial L'(q',\dot{q}',t)}{\partial\dot{q}'_i} = \frac{\partial L'(q',\dot{q}',t)}{\partial q'_i}\,.$$

Because the Lagrangian has a well defined physical interpretation (e.g. $L = T - V$ in nonrelativistic mechanics), the new Lagrangian $L'(q',\dot{q}',t)$ is obtained just by inverting the relations $q'_i = q'_i(q(t),t)$ and inserting $q_i(t) = q_i(q'(t),t)$ into the original Lagrangian: $L'(q'(q),\,\dot{q}'(q),\,t) = L(q(q'(t),t),\,\dot{q}(q'(t),\dot{q}'(t),t),\,t)$ where $\dot{q}(q'(t),\dot{q}'(t),t) \equiv (\partial q_i/\partial q'_j)\,\dot{q}'_j + \partial q_i/\partial t$.

In this case

$$I' - I \equiv \int d^4x\, \mathcal{L}'(\phi', \partial_\mu\phi') - \int d^4x\, \mathcal{L}(\phi, \partial_\mu\phi) \equiv 0\,, \tag{11.15}$$

i.e. $I' = I$ provided the field configurations $\phi_i(x)$ and $\phi_i'(x)$ are related by (11.13). Consequently, if the configuration $\phi_i(x)$ of the original fields is a stationary point of I, that is, if $\phi_i(x)$ satisfy the equations (11.4), then the configuration $\phi_i'(x)$ related to $\phi_i(x)$ by (11.13) is a stationary point of I', which implies that the fields $\phi_i'(x)$ satisfiy the equations

$$\partial_\nu \frac{\partial\mathcal{L}'(\phi', \partial\phi')}{\partial(\partial_\nu\phi_i')} - \frac{\partial\mathcal{L}'(\phi', \partial\phi')}{\partial\phi_i'} = 0\,, \tag{11.16}$$

which are in general of different form than the equations (11.4) because $\mathcal{L}'(\cdot\,,\,\cdot)$ is in general a different function of its arguments than $\mathcal{L}(\cdot\,,\,\cdot)$.

The choice of $\mathcal{L}'$ satisfying the condition (11.14) is not the only possibility: any $\mathcal{L}'$ such that

$$\mathcal{L}'(\phi', \partial_\mu\phi') = \mathcal{L}(\phi, \partial_\mu\phi) + \partial_\mu\mathcal{X}^\mu(\phi)\,, \tag{11.17}$$

is equally good. In this case, the relation (11.15), gets replaced by

$$I' = I + \int d^4x\, \partial_\mu\mathcal{X}^\mu(\phi) = I + \int d\sigma^\mu\, \mathcal{X}_\mu(\phi)\,, \tag{11.18}$$

but still, if a configuration $\phi_i(x)$ is the stationary point of $I[\cdot]$, then $\phi_i'(x)$ related to $\phi_i(x)$ by (11.13) is the stationary point of $I'[\cdot]$ because $\delta\int d^4x\, \partial_\mu\mathcal{X}^\mu(\phi) = 0$ for variations $\delta\phi_i$ vanishing at t_1 and at t_2 and in the limit $|\mathbf{x}| \to \infty$.

It is important to realize, that we are not yet speaking of a symmetry of the field equations of motion, but rather of the well known fact that in the Lagrangian formalism states of a given system can be specified by an arbitrarily chosen set of dynamical variables. A change of variables[8] $\phi_i \to \phi_i'(\phi)$ is a *symmetry* of the equations of motion if $\mathcal{L}'(\cdot\,,\,\cdot)$ leading to (11.17) is such that for some choice of $\mathcal{X}_\mu'(\cdot)$

$$\mathcal{L}'(\cdot\,,\,\cdot) - \partial^\mu\mathcal{X}_\mu'(\cdot) = \mathcal{L}(\cdot\,,\,\cdot)\,, \tag{11.19}$$

or - using the freedom to redefine $\mathcal{L}'$ by subtracting from it a total four-divergence - if for $\mathcal{L}'(\cdot\,,\,\cdot)$ leading to (11.17) one can just take $\mathcal{L}(\cdot\,,\,\cdot)$, because then the equations of

[8]Mathematically speaking, the fields $\phi_i(x)$, $i = 1, \dots, N$ define mappings from the space-time into some N-dimensional "target" space $T^{(N)}$, called by physicists the internal space. Therefore $\phi_i(x)$ should be viewed as coordinates of a point of a manifold $T^{(N)}$ onto which the space-time point (that is a point of the so-called base manifold), characterized by the coordinates x^μ is mapped. A change of variables $\phi_i \to \phi_i'$ can be due to changing the coordinate system on $T^{(N)}$, in which case we have to do with a passive transformation, or due to considering a transformed system (active transformation) - cf. Section 4.1.

motion (11.4) satisfied by the fields configuration $\phi_i(x)$ which is a stationary point of $I[\cdot]$ have *the same form* as the equations of motion (11.16) satisfied by the configuration $\phi'_i(x)$ which are stationary points of $I'[\cdot]$. In other words, $\phi_i(x)$ and $\phi'_i(x)$ are both solutions of the same Euler-Lagrange equations. Thus, the condition that the change $\phi_i \to \phi'_i(\phi)$ is a symmetry reads

$$\mathcal{L}(\phi', \partial\phi') = \mathcal{L}(\phi, \partial\phi) + \partial_\mu \mathcal{X}^\mu(\phi) \,. \tag{11.20}$$

Restricting now the discussion to transformations depending on some parameters θ_a, $a = 1, \ldots, n$, which can be continously deformed to the identity transformation (see Section 4.2),[9] we consider an infinitesimal symmetry (in the sense specified above) transformation[10]

$$\phi_i(x) \to \phi'_i(x) = \phi_i(x) + \delta\theta_a F^a_i(\phi) \equiv \phi_i(x) + \delta_0\phi_i(x) \,, \tag{11.21}$$

we write

$$\mathcal{L}(\phi', \partial\phi') = \mathcal{L}(\phi + \delta_0\phi,\, \partial\phi + \delta_0\partial\phi) = \mathcal{L}(\phi, \partial\phi) + \delta\mathcal{L}(\phi, \partial\phi) \,, \tag{11.22}$$

with $\delta\mathcal{L}$ being of first order in $\delta_0\phi_i$ (i.e. in $\delta\theta_a$). Using then (11.3) as well as the condition (11.20) combined with (11.18) we can write the equality (correspondingly to the infinitesimal character of the considered field transformations we write $\delta\mathcal{X}^\mu$ instead of $\mathcal{X}^\mu$)[11]

$$\begin{aligned} 0 &\equiv \int_\Omega d^4x \left[\mathcal{L}(\phi', \partial\phi') - \mathcal{L}(\phi, \partial\phi) - \partial_\mu \delta\mathcal{X}^\mu(\phi)\right] \qquad (11.23) \\ &= \int_\Omega d^4x \left\{ \left[\frac{\partial\mathcal{L}}{\partial\phi_i} - \partial_\mu \frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_i)}\right] \delta_0\phi_i + \partial_\mu \left[\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_i)}\, \delta_0\phi_i - \delta\mathcal{X}^\mu(\phi)\right] + \ldots \right\} . \end{aligned}$$

The first bracket in the above formula vanishes for the fields ϕ_i satisfying the equations of motion (11.4). Thus, for such a fields configuration $\phi_i(x)$ the quantity

$$j_\mu(x) = \frac{\partial\mathcal{L}}{\partial(\partial^\mu\phi_i)}\, \delta_0\phi_i - \delta\mathcal{X}_\mu(\phi) \,, \tag{11.24}$$

is conserved, i.e. satisfies the equation

$$\partial^\mu j_\mu(x) = 0 \,, \tag{11.25}$$

[9]In theories with fields taking values in Grassmann-Bieriezin algebras - see Section 11.8 - one can also consider transformations of the supersymmetric type the parameters θ_a of which are anticommuting Grassmann variables; although in such a case the notion of continuous parameter changes is only formal, the results obtained below still apply.

[10]The reasoning can be straightforwardly generalized to local transformations with $\delta_0\phi_i(x) = \delta\theta_a(x)F^a_i(\phi)$.

[11]This reasoning can easily be generalized to Lagrangian densities depending on higher field derivatives.

because if (11.21) is a symmetry transformation, (11.23) is identically equal to zero, and no special requirements on the values at $t = t_1$ and $t = t_2$ of the field changes $\delta_0\phi_i(x)$ related to a symmetry transformation are imposed[12] (of course, $\delta_0\phi_i(x)$ must vanish at spatial infinity or, in the finite volume, should not change the boundary conditions satisfied by the fields). If the transformations depend on n independent parameters, there are, therefore, n conserved *Noether currents*

$$j^a_\mu = \frac{\partial\mathcal{L}}{\partial(\partial^\mu\phi_i)}\, F^a_i(\phi) - \mathcal{X}^a_\mu(\phi)\,, \qquad a = 1, \ldots, n, \tag{11.26}$$

(since $\delta\mathcal{X}^\mu$ must be of first order in the transformation parameters $\delta\theta_a$, it must take the form $\delta\mathcal{X}_\mu(\phi) = \delta\theta_a\mathcal{X}^a_\mu(\phi)$). The Noether charges Q^a given by

$$Q^a = \int d^3\mathbf{x}\, j^a_0(t, \mathbf{x})\,, \qquad a = 1, \ldots, n, \tag{11.27}$$

are then time independent (conserved) quantities provided the fields fall off sufficiently rapidly at spatial infinity (or, in a finite volume, satisfy the appropriate boundary conditions). This can be seen directly from (11.23): upon using the Stokes theorem this equality reduces, for fields $\phi_i(x)$ satisfying the equations of motion, to

$$0 = \int_{\partial\Omega} d\sigma^\mu\, j_\mu(x)\,. \tag{11.28}$$

which, if Ω is the part of the space-time bounded by two hyperplanes $t = t_1$ and $t = t_2$, means precisely that[13]

$$\int d^3\mathbf{x}\, j_0(t_1, \mathbf{x}) = \int d^3\mathbf{x}\, j_0(t_2, \mathbf{x})\,. \tag{11.29}$$

In most cases the transformations (11.21) are linear[14] in the fields ϕ_i

$$\delta\theta_a F^a_i(\phi) = -i\delta\theta_a T^a_{ij}\phi_j\,, \tag{11.30}$$

(summations over a and j are understood) with T^a being a set of Hermitian matrices - a matrix representation of the generators of a symmetry group G - forming a basis of a

[12]For this reason the second term on the right hand side of (11.18), if present, is nonvanishing - it receives contributions from the $t = t_1$ and $t = t_2$ hypersufaces.

[13]Notice that the situation is quite different in the case of theories formulated in the Euclidean space with coordinates $\bar{x}^\mu$: because in this case fields and, therefore, also the symmetry changes $\delta_0\phi_i$ are bound to vanish for $|\bar{x}| \to \infty$ in *all* directions, one cannot infer from (11.23) the existence of conserved currents and, hence, there are no conserved charges (11.27).

[14]However, symmetries realized on fields nonlinearly also play an important role, mainly in effective quantum field theories. E.g. the so-called chiral Lagrangians of effective theories of strong interactions of low energy lightest mesons are invariant with respect to nonlinear transformations of fields.

representation of the dimension n Lie algebra of G (see Chapter 4). The matrix generators T^a satisfy then the commutation rule

$$\left[T^a,\ T^b\right] = iT^c f_c{}^{ab}\,, \tag{11.31}$$

with some structure constants $f_c{}^{ab}$ and will be assumed to be normalized by the condition $\mathrm{tr}(T^aT^b) = \frac{1}{2}\delta^{ab}$. The transformations (11.13) are then infinitesimal forms of finite symmetry transformations of fields. The characteristic feature of such *internal symmetries*[15] is that $\mathcal{X}^a_\mu \equiv 0$ and it is then easy to see that in the canonical Hamilton's formalism the Noether charges Q^a generate via Poisson brackets the infinitesimal symmetry transformations (11.30) of field variables. Indeed, the time component of the current is then simply

$$j^a_0(t,\mathbf{x}) = -i\Pi_i(t,\mathbf{x})\,T^a_{ij}\,\phi_j(t,\mathbf{x})\,, \tag{11.32}$$

and using (11.12) one gets

$$\delta\theta_a\,\{Q^a,\ \phi_i(t,\mathbf{x})\}_{\mathrm{PB}} = i\delta\theta_a T^a_{ij}\,\phi_j(t,\mathbf{x})\,, \tag{11.33}$$

so that $\phi^\prime_i(x) \approx \phi_i(x) - \delta\theta_a\,\{Q^a,\ \phi_i(x)\}_{\mathrm{PB}}$.

In order to discuss space-time transformations (like translations, rotations or Lorentz boosts) the formalism has to be generalized. An infinitesimal such transformation $x^\mu \to x^{\prime\mu}$

$$x^{\prime\mu} = x^\mu + \delta x^\mu(x) = x^\mu + \delta\theta_a f^{a\mu}(x)\,, \tag{11.34}$$

should be accompanied by a corresponding infinitesimal transformation $\phi_i(x) \to \phi^\prime_i(x^\prime)$ of the fields:[16]

$$\phi^\prime_i(x^\prime) = \phi_i(x) + \delta\phi_i(x) = \phi_i(x) + \delta\theta_a F^a_i(\phi(x))\,. \tag{11.35}$$

For example, in the case of the infinitesimal transformation $x^\mu \to x^{\prime\mu}$ with

$$x^{\prime\mu} = x^\mu + \delta\omega^\mu{}_\nu x^\nu - \delta\epsilon^\mu\,, \tag{11.36}$$

corresponding to a change of the reference frame, the fields $\phi_i(x)$ transform under some regular (in general reducible) matrix representation $(\mathcal{J}^{\mu\nu})_{ij}$ of the Lorentz group[17]

$$\phi^\prime_i(x^\prime) = \phi_i(x) - \frac{i}{2}\,\delta\omega_{\mu\nu}\,(\mathcal{J}^{\mu\nu})_{ij}\,\phi_j(x)\,. \tag{11.37}$$

[15] More generally, whenever $\mathcal{X}^\mu(\phi)$ cannot be removed by a suitable redefinition of $\mathcal{L}$ by a total four-divergence, we have to do with transformations having a space-time character (for example, $\mathcal{X}^\mu(\phi)$ cannot be removed in the case of supersymmetric transformations); if $\mathcal{X}^a_\mu \equiv 0$ one speaks of genuinely internal symmetries.

[16] This means that the change of the space-time (base manifold) coordinate system entails a related coordinate change of the internal (target) space.

[17] The simplest nontrivial is the vector representation with $\left(\mathcal{J}^{\lambda\nu}_{\mathrm{vec}}\right)^\mu{}_\kappa$ given by (D.3) acting on vector fields $\phi_i \equiv V^\kappa$, but higher rank tensor or spinorial representations can also be considered.

One can also consider other space-time transformations $x^\mu \to x'^\mu$ (and the associated transformations of fiels) such as e.g. conformal transformations etc.

The system is in the new space-time coordinates x'^μ represented by the new field variables $\phi'_i(x')$ if there is such a new Lagrangian density $\mathcal{L}'$ that when any two field configurations $\phi_i(x)$ and $\phi'_i(x')$ are related to each other by (11.34) and (11.35)

$$I' - I \equiv \int_{\Omega'} d^4x'\, \mathcal{L}'(\phi'(x'), \partial'_\mu\phi'(x')) - \int_\Omega d^4x\, \mathcal{L}(\phi(x), \partial_\mu\phi(x)) = 0\,, \tag{11.38}$$

where ∂'_μ denotes the derivative with respect to x'^μ and Ω' is the image of the integration domain Ω under the change of variables (11.34). (11.38) is a sufficient condition for the new fields $\phi'_i(x')$ obtained via (11.35) from the solutions $\phi_i(x)$ of the equations of motion (11.4) following from $\mathcal{L}$ to be solutions of the equations of motion following from $\mathcal{L}'$. However as previously, the same conclusion concerning $\phi'_i(x')$ obtains if $\mathcal{L}'(\phi'(x'), \partial'_\mu\phi'(x'))$ is chosen so that

$$d^4x'\mathcal{L}'(\phi'(x'), \partial'_\mu\phi'(x')) = d^4x\, [\mathcal{L}(\phi(x), \partial_\mu\phi(x)) + \partial_\mu\mathcal{X}^\mu(\phi(x))]\,, \tag{11.39}$$

with some four-vector function $\mathcal{X}^\mu(\phi)$.

Again, one speaks of a space-time symmetry, if (by an appropriate choice of the factor $\mathcal{X}'_\mu(\phi')$) for the new Lagrangian density $\mathcal{L}'(\cdot,\cdot)$ leading to (11.39) one can take the original $\mathcal{L}(\cdot,\cdot)$:

$$\mathcal{L}'(\phi'(x'), \partial'_\mu\phi'(x')) = \mathcal{L}(\phi'(x'), \partial'_\mu\phi'(x'))\,, \tag{11.40}$$

because then the equations of motion satisfied by $\phi'_i(x')$ (in the space-time coordinates x'^μ) have the same form as the equations of motion (in x^μ) satisfied by $\phi_i(x)$. Thus, if (11.34) combined with (11.35) are symmetry transformations of the theory, the condition (11.39) implies the following identity

$$d^4x'\mathcal{L}(\phi'(x'), \partial'_\mu\phi'(x')) = d^4x\, [\mathcal{L}(\phi(x), \partial_\mu\phi(x)) + \partial_\mu\mathcal{X}^\mu(\phi(x))]\,, \tag{11.41}$$

in which one understands that x'^μ in the left hand side is expressed in terms of x^μ using (11.34), so that in (11.38) $\Omega' \to \Omega$ (for a general change (11.34) of the coordinates $d^4x' \neq d^4x$, i.e. $\det(\partial x'/\partial x) \neq 1$).

The transformations (11.36) accompanied by (11.37) should be symmetries in the above sense of the action I of a relativistic field theory. In this case $\det(\partial x'/\partial x) = 1$ and the condition (11.41) takes the form similar to (11.20), except for different space-time coordinates on both sides.

Conserved quantities corresponding to space-time symmetry transformations can be found using the condition (11.39) (or (11.41)). To derive them it is technically convenient to split the total change $\delta\phi_i(x)$ of the fields as follows[18]

$$\begin{aligned}\delta\phi_i(x) &\equiv \phi'_i(x') - \phi_i(x)\\ &= \phi'_i(x') - \phi'_i(x) + \phi'_i(x) - \phi_i(x) \equiv \delta x^\mu\partial_\mu\phi_i(x) + \delta_0\phi_i(x)\,,\end{aligned}$$

[18]It is also possible to arrive at the final result (11.51) without this splitting.

(to the first order in the transformation parameters $\delta\theta_a$ the fields ϕ'_i have been in the term with δx^μ replaced by ϕ_i) and to work with the functional changes of fields $\delta_0\phi_i(x) = \phi'_i(x) - \phi_i(x)$ which have the property

$$\delta_0(\partial_\mu\phi_i(x)) = \partial_\mu(\delta_0\phi_i(x))\,. \tag{11.42}$$

Thus, to the first order in the infinitesimal parameters $\delta\theta_a$, we write the transformed fields and their derivatives as[19]

$$\begin{aligned}
\phi'_i(x') &= \phi_i(x) + \delta_0\phi_i(x) + \delta x^\lambda\partial_\lambda\phi_i(x) + \mathcal{O}(\delta\theta^2)\,,\\
\partial'_\mu\phi'_i(x') &= \partial_\mu\phi_i(x) + \partial_\mu(\delta_0\phi_i(x)) + \delta x^\lambda\partial_\lambda\partial_\mu\phi_i(x) + \mathcal{O}(\delta\theta^2)\,,\\
\partial'_\nu\partial'_\mu\phi'_i(x') &= \partial_\nu\partial_\mu\phi_i(x) + \partial_\nu\partial_\mu(\delta_0\phi_i(x)) + \delta x^\lambda\partial_\lambda\partial_\nu\partial_\mu\phi_i(x) + \mathcal{O}(\delta\theta^2)\,,
\end{aligned} \tag{11.43}$$

etc. This allows to represent $\mathcal{L}(\phi'(x'), \partial'_\mu\phi'(x'))$ in the identity (11.41) in the form[20]

$$\mathcal{L}(\phi'(x'), \partial'_\mu\phi'(x')) = \mathcal{L}(\phi(x), \partial_\mu\phi(x)) + \delta\mathcal{L}\,, \tag{11.44}$$

in which

$$\delta\mathcal{L} = \frac{\partial\mathcal{L}}{\partial\phi_i}\,\delta_0\phi_i + \frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_i)}\,\partial_\mu\delta_0\phi_i + \delta x^\mu\partial_\mu\mathcal{L} + \mathcal{O}(\delta\theta^2)\,, \tag{11.45}$$

where in turn

$$\delta x^\mu\partial_\mu\mathcal{L} \equiv \delta x^\mu\left[\frac{\partial\mathcal{L}}{\partial\phi_i}\,\partial_\mu\phi_i(x) + \frac{\partial\mathcal{L}}{\partial(\partial_\lambda\phi_i)}\,\partial_\mu\partial_\lambda\phi_i(x)\right]. \tag{11.46}$$

The Lagrangian density change (11.45) can also be rewritten as

$$\delta\mathcal{L} = \delta x^\mu\partial_\mu\mathcal{L} + \left[\frac{\partial\mathcal{L}}{\partial\phi_i} - \partial_\mu\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_i)}\right]\delta_0\phi_i + \partial_\mu\left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_i)}\,\delta_0\phi_i\right). \tag{11.47}$$

Using then the Jacobian[21]

$$\det\left(\frac{\partial x'_\mu}{\partial x_\nu}\right) = \det\left(\delta^\nu_\mu + \frac{\partial\delta x_\mu(x)}{\partial x_\nu}\right) \approx 1 + \mathrm{tr}\left(\frac{\partial\delta x_\mu}{\partial x_\nu}\right) = 1 + \partial_\mu\delta x^\mu\,, \tag{11.48}$$

to explicitly express d^4x' through d^4x, one writes the left hand side of (11.41) to the first order in $\delta\theta_a$ in the form

$$\begin{aligned}
d^4x'\,\mathcal{L}(\phi'(x'), \partial'_\mu\phi'(x')) &= d^4x'\,[\mathcal{L}(\phi(x), \partial_\mu\phi(x)) + \delta\mathcal{L}]\\
&= d^4x\,\mathcal{L}(\phi(x), \partial_\mu\phi(x)) + d^4x\,[\delta\mathcal{L} + \mathcal{L}\,\partial_\mu\delta x^\mu]\,.
\end{aligned} \tag{11.49}$$

[19]From (11.34) it follows that $\partial x'^\lambda/\partial x^\kappa = \delta^\lambda_\kappa + \partial(\delta x^\lambda)/\partial x^\kappa$, and $\partial x^\lambda/\partial x'^\mu = \delta^\lambda_\mu - \partial(\delta x^\lambda)/\partial x^\mu$. Thus, $\partial'_\mu = \partial_\mu - (\partial_\mu(\delta x^\lambda))\partial_\lambda$ and the terms with derivatives of δx^λ cancel out in (11.43).

[20]We again assume, that $\mathcal{L}$ depends only on fields and their first derivatives. Conserved Noether charges corresponding to more complicated Lagrangian densities can be derived using similar methods.

[21]We use the relation $\det(1 + A) = \exp\{\mathrm{tr}\ln(1 + A)\} \approx 1 + \mathrm{tr}(A)$.

Thus, for field configurations $\phi_i(x)$ satisfying the equations of motion (11.4) the relation (11.39) leads to the identity (again $\delta\mathcal{X}^\mu$ denotes the term of first order in $\delta\theta$ in $\mathcal{X}^\mu$):

$$\begin{aligned}
&\int_\Omega d^4x \left[\mathcal{L}\,\partial_\mu\delta x^\mu + \delta x^\mu\partial_\mu\mathcal{L} + \partial_\mu\left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_i)}\,\delta_0\phi_i\right) - \partial_\mu\delta\mathcal{X}^\mu(\phi)\right] + \mathcal{O}(\delta\theta^2)\\
&\qquad = \int_\Omega d^4x\,\partial_\mu\left[\delta x^\mu\,\mathcal{L} + \frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_i)}\,\delta_0\phi_i - \delta\mathcal{X}^\mu(\phi)\right] + \mathcal{O}(\delta\theta^2) = 0\,.
\end{aligned} \tag{11.50}$$

Taking into account the arbitrariness of Ω and reexpressing $\delta_0\phi_i$ back in terms of $\delta\phi_i \equiv \phi_i'(x') - \phi_i(x) = \delta_0\phi_i + \delta x^\mu\partial_\mu\phi_i$ we see that the quantity

$$J^\mu = \left[\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_i)}\,\partial_\rho\phi_i - g^\mu{}_\rho\,\mathcal{L}\right]\delta x^\rho - \frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_i)}\,\delta\phi_i + \delta\mathcal{X}^\mu(\phi)\,. \tag{11.51}$$

evaluated on field configuration $\phi_i(x)$ satusfying the equation (11.4) is conserved, that is, $\partial_\mu J^\mu(x) = 0$. It plays, therefore, the role of the Noether symmetry current of space-time transformations.

We now consider the Noether currents (11.51) associated with the Poincaré transformations of the form (11.36). Translations correspond to $\delta\omega_{\mu\nu} = 0$, $\delta\epsilon^\mu \neq 0$ and $\phi_i'(x') = \phi_i(x)$, so that in this case $\delta\phi_i = 0$. The corresponding Noether current which (assuming that $\delta\mathcal{X}^\mu = 0$, which is usually the case) reads

$$T^{\mu\nu}_{\text{can}} = \frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_i)}\,\partial^\nu\phi_i - g^{\mu\nu}\mathcal{L}\,, \tag{11.52}$$

is called the *canonical energy-momentum tensor*. It is by construction conserved:

$$\partial_\mu T^{\mu\nu}_{\text{can}}(x) = 0\,. \tag{11.53}$$

The four constants of motion (i.e. time independent quantities)

$$P^\mu = \int d^3\mathbf{x}\,T^{0\mu}_{\text{can}}(t,\mathbf{x})\,, \tag{11.54}$$

play the role of the total energy P^0 of the system of fields and of its total momentum vector P^i. It can be shown that if $T^{\nu\mu}_{\text{can}}$ is conserved (and only then!), P^μ given by (11.54) transforms as a true four-vector when the reference frame is changed.

In the case of Lagrangian densities which are of the general form

$$\mathcal{L}(\phi_i, \partial_\mu\phi_i) = \frac{1}{2}\partial_\mu\phi_i\partial^\mu\phi_i - V(\phi)\,, \tag{11.55}$$

with a function $V(\phi)$ called the *field potential*, going over to the Hamilton's formalism, one finds

$$\begin{aligned}
T^{00}_{\text{can}} &= \frac{1}{2}\Pi_i\Pi_i + \frac{1}{2}\boldsymbol{\nabla}\phi_i\cdot\boldsymbol{\nabla}\phi_i + V(\phi) \equiv \mathcal{H}\,,\\
T^{0k}_{\text{can}} &= \Pi_i\partial^k\phi_i\,.
\end{aligned} \tag{11.56}$$

In this case it is easy to check that the transformations of the field $\phi_i(x)$ corresponding to spacetime translations are generated by the Poisson brackets:

$$\{P^\mu,\ \phi_i(x)\}_{\rm PB} = -\,\partial^\mu\phi_i(x)\,, \tag{11.57}$$

so that $\phi'_i(x) = \phi_i(x) - \{P^\mu,\ \phi_i(x)\}_{\rm PB}\,\delta\epsilon_\mu$.

The canonical energy-momentum tensor (11.52) is not always symmetric in its indices $\mu\nu$. However, instead of $T^{\mu\nu}_{\rm can}$ one can always use a modified tensor

$$T^{\mu\nu}(x) = T^{\mu\nu}_{\rm can}(x) + H^{\mu\nu}(x)\,, \tag{11.58}$$

In which $H^{\mu\nu}(x)$ is a tensor which is conserved, $\partial_\mu H^{\mu\nu}(x) = 0$, and such that

$$\int d^3\mathbf{x}\, H^{0\nu}(x) = 0\,. \tag{11.59}$$

The associated conserved charges P^μ obtained from the modified tensor $T^{\mu\nu}$ are then (if the fields vanish sufficiently fast at spatial infinity) the same as the ones obtained from $T^{\mu\nu}_{\rm can}$. These conditions are met if

$$H^{\mu\nu}(x) = \partial_\rho H^{\rho\mu\nu}(x)\,, \tag{11.60}$$

where $H^{\rho\mu\nu}(x)$ is antisymmetric in its $\rho\mu$ indices. Using this freedom one can always replace $T^{\mu\nu}_{\rm can}$ by a symmetric tensor $T^{\mu\nu}_{\rm symm} = T^{\nu\mu}_{\rm symm}$. In the case of fields which transform nontrivially under the Lorentz group (as in (11.37)) of particular interest is the Belinfante symmetric tensor obtained by taking

$$H^{\rho\mu\nu} = \frac{1}{2}\left[\frac{\partial\mathcal{L}}{\partial(\partial_\rho\phi_i)}\left(-i\mathcal{J}^{\mu\nu}\right)_{ij}\phi_j - \frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_i)}\left(-i\mathcal{J}^{\rho\nu}\right)_{ij}\phi_j - \frac{\partial\mathcal{L}}{\partial(\partial_\nu\phi_i)}\left(-i\mathcal{J}^{\rho\mu}\right)_{ij}\phi_j\right]. \tag{11.61}$$

It is the Belinfante symmetric energy-momentum tensor which appears as the right hand side of the Einstein's equations of General Relativity.[22]

Conserved currents associated with the Lorentz transformations ($\delta\epsilon^\mu = 0$, $\delta\omega^\mu{}_\nu \neq 0$ in (11.36)) are derived in a similar way. Using $\delta x^\mu = \delta\omega^\mu{}_\nu x^\nu$ and $\delta\phi_i = -\frac{i}{2}\delta\omega_{\mu\nu}\left(\mathcal{J}^{\mu\nu}\right)_{ij}\phi_j(x)$ following from (11.37) we get (as previously asuming that $\delta\mathcal{X}^\mu = 0$)

$$M^{\mu\nu\kappa}_{\rm can}(x) = x^\nu\, T^{\mu\kappa}_{\rm can} - x^\kappa\, T^{\mu\nu}_{\rm can} + \frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_i)}\left(-i\mathcal{J}^{\nu\kappa}\right)_{ij}\phi_j\,. \tag{11.62}$$

[22] That is, it coincides with the energy-momentum tensor defined as the variational derivative with respect to the metric tensor $g_{\mu\nu}(x)$ of the action $I[\phi]$ written in the generally covariant form ($g \equiv -\det(g_{\mu\nu})$)

$$T^{\mu\nu}_{\rm symm}(x) = \frac{\delta I[\phi]}{\delta g_{\mu\nu}(x)} = \frac{\delta}{\delta g_{\mu\nu}(x)}\int d^4x\,\sqrt{g}\,\mathcal{L}\,.$$

The conserved (i.e. time independent) charges

$$J^{\nu\kappa} = \int d^3\mathbf{x}\, M^{0\nu\kappa}_{\text{can}}(t, \mathbf{x})\,, \tag{11.63}$$

are antisymmetric $J^{\nu\kappa} = J^{\kappa\nu}$. Again, it can be shown that if $M^{\mu\nu\kappa}_{\text{can}}$ is conserved, $J^{\nu\kappa}$ transforms as a true four-dimensional second rank tensor. The spatial components J^{ij} of (11.63) play the role of the total angular momentum of the considered system of fields. It is straightforward to check that if the Lagrangian density has the form (11.55), the tensor $J^{\nu\kappa}$ generates, through the Poisson brackets, Lorentz transformations of the fields $\phi_i(x)$:

$$\{J^{\mu\nu}, \phi_i(x)\}_{\text{PB}} = -\left[(x^\mu\partial^\nu - x^\nu\partial^\mu)\delta_{ij} + (-i\mathcal{J}^{\mu\nu})_{ij}\right]\phi_j(x)\,, \tag{11.64}$$

so that $\phi_i'(x) = \phi_i(x) - \frac{1}{2}\,\delta\omega_{\mu\nu}\,\{J^{\mu\nu},\ \phi_i(x)\}_{\text{PB}}$. It can also be shown, that the tensor (11.62) differs by a total four-divergence from the tensor[23]

$$M^{\mu\nu\kappa} = x^\nu\, T^{\mu\kappa}_{\text{symm}} - x^\kappa\, T^{\mu\nu}_{\text{symm}}\,, \tag{11.65}$$

in which $T^{\mu\nu}_{\text{symm}}$ is the Belinfante symmetric energy-momentum tensor obtained from (11.61), which gives therefore the same conserved Noether charges $J^{\mu\nu}$ (for field configurations satisfying the equations of motion (11.4) and vanishing sufficiently fast at spatial infinity).

11.2 Canonical quantization of a real scalar field

In this section we discuss in details quantization of the simplest example of a relativistic field - the real scalar field $\varphi(x)$. No special difficulties, beyond those inherent in treating systems of infinitely many degrees of freedom, arise in this case. More complicated cases of vector (in particular of the electromagnetic field) and spinor fields will be discussed sections 11.7 and 11.8, respectively. We consider first quantization of the noninteracting field $\varphi(x)$ and then outline modifications introduced by interactions. At the end of this section these results are generalized to systems of many interacting fields.

The simplest dynamics of a real field transforming as a scalar when the inertial reference frame is changed is given (in units in which $c = \hbar = 1$) by the Poincaré invariant Lagrangian density

$$\mathcal{L}_0 = \frac{1}{2}\partial_\mu\varphi\partial^\mu\varphi - \frac{1}{2}M^2\varphi^2\,. \tag{11.66}$$

By a rescaling of the field variable φ (i.e. by a simple canonical transformation) the first term quadratic in the first derivatives, called the *kinetic term*, can always be brought into the canonical form as above. In the absence in the Lagrangian density of powers of φ

[23] The tensor (11.65) is conserved due to the symmetry and conservation of $T^{\mu\kappa}_{\text{symm}}$.

higher than the second (a feature which makes this theory solvable both classically and quantum mechanically) the negative sign of the second term is indispensable to ensure, as will be seen, the boundedness from below of the spectrum of the resulting quantum Hamiltonian and the relativistic relation $E^2(\mathbf{k}) = \mathbf{k}^2 + M^2$ between three-momentum $\mathbf{k}$ and energy E of the Hamiltonian eigenstates identified with one-particle states. To stress the similarity to the system of coupled oscillators discussed in Section 5.6 we write the corresponding Lagrangian in the more general, a priori spatially nonlocal, form

$$L_0 = \frac{1}{2}\int d^3\mathbf{x}\,\dot{\varphi}^2(t,\mathbf{x}) - \frac{1}{2}\int d^3\mathbf{x}\int d^3\mathbf{y}\,\varphi(t,\mathbf{x})K(\mathbf{x},\mathbf{y})\,\varphi(t,\mathbf{y})\,, \tag{11.67}$$

with a real and symmetric, $K(\mathbf{x},\mathbf{y}) = K(\mathbf{y},\mathbf{x})$, kernel. Formally, the value of φ at every point $\mathbf{x}$ can be treated as an independent canonical variable $Q_{\mathbf{x}}(t) = \varphi(t,\mathbf{x})$. Accordingly, the Euler-Lagrange equation of motion of the system defined by (11.67) can be written in the form

$$\frac{d}{dt}\frac{\delta L_0}{\delta\dot{\varphi}(t,\mathbf{x})} - \frac{\delta L_0}{\delta\varphi(t,\mathbf{x})} = 0\,, \tag{11.68}$$

(the derivatives with respect to φ and $\dot{\varphi}$ are functional derivatives). Applied to (11.67) this gives the equation

$$\frac{d^2}{dt^2}\,\varphi(t,\mathbf{x}) + \int d^3\mathbf{y}\,K(\mathbf{x},\mathbf{y})\,\varphi(t,\mathbf{y}) = 0\,. \tag{11.69}$$

The kernel $K(\mathbf{x},\mathbf{y})$ of the local Lagrangian density (11.66) takes (restoring the constants c and $\hbar$ for decoration) the form

$$K(\mathbf{x}-\mathbf{y}) = -c^2\boldsymbol{\nabla}^2_{(\mathbf{x})}\delta^{(3)}(\mathbf{x}-\mathbf{y}) + \frac{M^2c^4}{\hbar^2}\,\delta^{(3)}(\mathbf{x}-\mathbf{y})\,, \tag{11.70}$$

so that[24]

$$L_0 = \frac{1}{2}\int d^3\mathbf{x}\left[\dot{\varphi}^2(t,\mathbf{x}) - c^2(\boldsymbol{\nabla}\varphi(t,\mathbf{x}))^2 - \frac{M^2c^4}{\hbar^2}\varphi^2(t,\mathbf{x})\right]. \tag{11.71}$$

The Euler-Lagrange equation (11.68) is in this case equivalent to the field equation (11.4) and reads

$$\left(\frac{\partial^2}{\partial t^2} - c^2\boldsymbol{\nabla}^2 + \frac{M^2c^4}{\hbar^2}\right)\varphi(t,\mathbf{x}) = 0\,. \tag{11.72}$$

[24]In units $[M]$, $[T]$ and $[L]$ (mass, time and length) the action $I = \int dt L$ has dimension of $\hbar$, that is, $[M][L]^2[T]^{-1}$. It follows from (11.71) that the field φ has dimension $[M]^{1/2}[L]^{-1/2}$. Furthermore, since $\Pi = \dot{\varphi}$, it has dimension $[M]^{1/2}[L]^{-1/2}[T]^{-1}$ and the Hamiltonian (11.75) has the right dimension $[M][L]^2[T]^{-2}$ of energy. See also Appendix I.

To quantize the system one has first to set up the canonical (Hamilton's) formalism. To this end one defines the canonical momenta $P_{\mathbf{x}}(t) \equiv \Pi(t, \mathbf{x})$ conjugated to the canonical variables $Q_{\mathbf{x}}(t) \equiv \varphi(t, \mathbf{x})$:

$$\Pi(t, \mathbf{x}) = \frac{\delta L_0}{\delta \dot{\varphi}(t, \mathbf{x})} = \dot{\varphi}(t, \mathbf{x})\,, \tag{11.73}$$

and constructs the Hamiltonian

$$H_0 = \int d^3\mathbf{x}\,\Pi(t, \mathbf{x})\dot{\varphi}(t, \mathbf{x}) - L_0\,, \tag{11.74}$$

in which $\dot{\varphi}(t, \mathbf{x})$ has to be expressed in terms of $\Pi(t, \mathbf{x})$ and $\varphi(t, \mathbf{x})$:

$$H_0 = \frac{1}{2}\int d^3\mathbf{x}\,\Pi^2(t, \mathbf{x}) + \frac{1}{2}\int d^3\mathbf{x}\int d^3\mathbf{y}\,\varphi(t, \mathbf{x})K(\mathbf{x} - \mathbf{y})\,\varphi(t, \mathbf{y})\,. \tag{11.75}$$

It is easy to check, that the classical Hamilton's equations

$$\begin{aligned}
\frac{d}{dt}\varphi(t, \mathbf{x}) &= \{\varphi(t, \mathbf{x}),\ H\}_{\rm PB} = \frac{\delta H}{\delta \Pi(t, \mathbf{x})}\,,\\
\frac{d}{dt}\Pi(t, \mathbf{x}) &= \{\Pi(t, \mathbf{x}),\ H\}_{\rm PB} = -\frac{\delta H}{\delta \varphi(t, \mathbf{x})}\,,
\end{aligned} \tag{11.76}$$

are fully equivalent to the Euler-Lagrange equation (11.4), i.e. to (11.72).

Quantization in the Schrödinger picture of a classical system means promoting (a set of) its real canonical variables Q^i and P_j taken at *one particular instant*, usually $t = 0$, to time independent Hermitian operators $\hat{Q}^i$ and $\hat{P}_j$ satisfying the canonical commutation rules[25]

$$[\hat{Q}^i,\ \hat{P}_j] = i\hbar\left\{Q^i,\ P_j\right\}_{\rm PB} = i\hbar\,\delta^i{}_j\,, \tag{11.77}$$

and $[\hat{Q}^i, \hat{Q}^j] = [\hat{P}_i, \hat{P}_j] = 0$ and representing the resulting algebra of operators in some Hilbert space of states. Following this prescription, in the considered case one imposes on the operators $\hat{\varphi}$ and $\hat{\Pi}$ the conditions

$$\begin{aligned}
&[\hat{\varphi}(\mathbf{x}),\ \hat{\Pi}(\mathbf{y})] = i\hbar\,\delta^{(3)}(\mathbf{x} - \mathbf{y})\,,\\
&[\hat{\varphi}(\mathbf{x}),\ \hat{\varphi}(\mathbf{y})] = [\hat{\Pi}(\mathbf{x}),\ \hat{\Pi}(\mathbf{y})] = 0\,.
\end{aligned} \tag{11.78}$$

[25] A stronger requirement, that for any pair of classical observables $F(Q, P)$ and $G(Q, P)$ such that $\{F(Q, P),\ G(Q, P)\}_{\rm PB} = A(Q, P)$ the corresponding operators in the quantum theory satisfy the relation

$$\left[F(\hat{Q}, \hat{P}),\ G(\hat{Q}, \hat{P})\right] = i\hbar A(\hat{Q}, \hat{P})\,,$$

cannot in general be imposed because of problems with ordering of operators; it can hold only for observables $F(Q, P)$ and $G(Q, P)$ which are at most linear in the canonical variables Q and P.

Upon quantization the Hamiltonian (11.75) also becomes a Hermitian (owing to the Hermiticity of $\hat{\varphi}$ and $\hat{\Pi}$) operator:

$$H_0 = \frac{1}{2}\int d^3\mathbf{x}\left[\hat{\Pi}^2(\mathbf{x}) + (\boldsymbol{\nabla}\hat{\varphi}(\mathbf{x}))^2 + M^2\hat{\varphi}^2(\mathbf{x})\right]. \tag{11.79}$$

In principle, in full analogy with the ordinary quantum mechanics of a system having n degrees of freedom formulated in the position space, i.e. in $L_2(\mathbb{R}^n)$ as the Hilbert space, in which the system's states are represented by wave functions $\psi(Q^1, Q^2, \dots, t)$ (the probability amplitude of finding the system in the classical state characterized by the values Q^i of its canonical variables) on which $\hat{Q}^i$'s act by multiplication by Q^i and $\hat{P}_j$'s act as $-i\hbar\partial/\partial Q^j$, one can represent the algebra (11.78) of the operators $\hat{\varphi}(\mathbf{x})$ and $\hat{\Pi}(\mathbf{x})$ in the space of all functionals $\Psi[\varphi]$ defined on (classical) field configurations $\varphi(\mathbf{x})$ vanishing at $|\mathbf{x}| \to \infty$. In this nonseparable Hilbert space $\hat{\varphi}(\mathbf{x})$ and $\hat{\Pi}(\mathbf{x})$ act through

$$\hat{\varphi}(\mathbf{x})\Psi[\varphi, t] = \varphi(\mathbf{x})\Psi[\varphi, t]\ , \qquad \hat{\Pi}(\mathbf{x})\Psi[\varphi, t] = -i\hbar\,\frac{\delta\Psi[\varphi, t]}{\delta\varphi(\mathbf{x})}\,, \tag{11.80}$$

and the "wave functionals" $\Psi[\varphi, t] \equiv \langle\varphi(\mathbf{x})|\Psi(t)\rangle$ can heuristically be treated as representing the probability amplitude that the field takes on at time t the configuration $\varphi(\mathbf{x})$. While this approach can offer a useful insight into the global structure of the theory's Hilbert space (which proves indispensable when e.g. topological aspects of gauge theories are to be investigated), it does not immediately lead to the interpretation of field states in terms of particles (in the nonseparable Hilbert space the fundamental commutation relations (11.78) are not realized irreducibly). Such an interpretation emerges naturally (at least in the case of free field theories) if one finds a representation of the $\hat{\varphi}(\mathbf{x})$ and $\hat{\Pi}(\mathbf{x})$ operators in terms of some other operators the commutation relations of which (their algebra) can be represented in some Fock space (cf. Chapter 5) - a separable subspace of the "big" Hilbert space $\mathcal{H}$ (the one specified above) of all possible state-vectors. In the case of field theories defined on the flat Minkowski space-time, Hamiltonians of which are time independent, this Fock space should be chosen in $\mathcal{H}$ in such a way as to include one of the lowest energy eigenvectors (some field theories, like e.g. supersymmetric ones or gauge theories can have in the nonseparable Hilbert space multiple such vectors) of the Hamiltonian operator of the theory. In free field theories it is relatively easy to chose the Fock space so that the Hamilonian lowest energy eigenvector $|\Omega_0\rangle$ is just the vector $|0_{\rm Fock}\rangle$ out of which other Fock state-vectors are built by the action of the creation operators. In theories of interacting fields the lowest energy eigenvector of the full Hamiltonian H, the true vacuum $|\Omega\rangle$, is at best a complicated formal superposition of vectors belonging to the the chosen Fock space to which the vector $|\Omega_0\rangle$ belongs (and usually is identical with the vector $|0_{\rm Fock}\rangle$), but usually, from the orthodox mathematical point of view, it does not belong to this Fock space. (Renormalization consisting essentially of working with some cutoffs making effectively the number of degrees of freedom finite can be viewed also as a way of going around this difficulty). In most cases[26] representing the algebra of the

[26]The exception is e.g. the Hamiltonian of the electromagnetic field quantized using the Gupta-Bleuler

field operators in the suitable Fock space has also the welcome effect of giving the free part H_0 of the complete Hamiltonian the form similar to the one of uncoupled harmonic oscillators (thereby making its spectrum explicit).

To illustrate the programme oulined above and to show how the Fock space vectors are related to the states of the quantized field we consider first quantization of the classical real scalar field $\varphi(\mathbf{x})$ satisfying periodic boundary conditions in the box of finite volume[27] $V = L^3$. To implement these conditions every field configuration is written in the form[28]

$$\varphi(\mathbf{x}) = \frac{1}{\sqrt{V}} \sum_{\mathbf{k}} \varphi_{\mathbf{k}}\, e^{i\mathbf{k}\cdot\mathbf{x}}\,, \tag{11.81}$$

with the wave vectors $\mathbf{k}$ forming a countably infinite (i.e. discrete) set: $\mathbf{k} = (2\pi/L)\mathbf{n}$, where $\mathbf{n} = (n_x, n_y, n_z)$, $n_i \in \mathbb{Z}$. Since the field $\varphi(\mathbf{x})$ is real, $\varphi_{-\mathbf{k}} = \varphi^*_{\mathbf{k}}$. Expressed in terms of the coefficients $\varphi_{\mathbf{k}}$, which play now the role of new variables, the lagrangian L_0 corresponding to the Lagrangian density (11.66), which sets the dynamics of the field $\varphi(\mathbf{x})$, takes the form[29]

$$\begin{aligned} L_0 = \int_V d^3\mathbf{x}\, \mathcal{L}_0 &= \frac{1}{2} \sum_{\mathbf{k}} \left(\dot\varphi_{\mathbf{k}}\dot\varphi^*_{\mathbf{k}} - \omega^2_{\mathbf{k}}\varphi_{\mathbf{k}}\varphi^*_{\mathbf{k}}\right) \\ &= \frac{1}{2}\left(\dot\varphi^2_{\mathbf{0}} - \omega^2_{\mathbf{0}}\varphi^2_{\mathbf{0}}\right) + \sum_{\mathbf{k}>0} \left(\dot\varphi_{\mathbf{k}}\dot\varphi^*_{\mathbf{k}} - \omega^2_{\mathbf{k}}\varphi_{\mathbf{k}}\varphi^*_{\mathbf{k}}\right). \end{aligned} \tag{11.82}$$

Here $\omega^2_{\mathbf{k}} \equiv \mathbf{k}^2 + M^2$. In the second line the sum over discrete wave vectors $\mathbf{k}$ has been split into the term with $\mathbf{k} = \mathbf{0}$ and the sum over only half of nonzero vectors $\mathbf{k}$ (this is somewhat heuristically denoted by $\mathbf{k} > 0$) accounting for the fact that $\dot\varphi_{\mathbf{k}}\dot\varphi^*_{\mathbf{k}} \equiv \dot\varphi_{\mathbf{k}}\dot\varphi_{-\mathbf{k}} = \dot\varphi_{-\mathbf{k}}\dot\varphi_{\mathbf{k}}$, etc. To completely reduce the considered field to a system characterized by an infinite set of ordinary real canonical variables one writes

$$\varphi_{\mathbf{0}} \equiv q_{\mathbf{0}}\,, \qquad \varphi_{\mathbf{k}} \equiv \frac{1}{\sqrt{2}}\left(q_{\mathbf{k}} + i\bar q_{\mathbf{k}}\right) \qquad \text{for } \mathbf{k} > 0\,, \tag{11.83}$$

so that the Lagrangian takes the form

$$L_0 = \frac{1}{2}(\dot q^2_{\mathbf{0}} - \omega^2_{\mathbf{0}} q^2_{\mathbf{0}}) + \frac{1}{2}\sum_{\mathbf{k}>0}\left(\dot q^2_{\mathbf{k}} - \omega^2_{\mathbf{k}} q^2_{\mathbf{k}} + \dot{\bar q}^2_{\mathbf{k}} - \omega^2_{\mathbf{k}}\bar q^2_{\mathbf{k}}\right). \tag{11.84}$$

---

method (Section 11.11) with the gauge parameter $\xi \neq 1$ or the nonabelian Yang-Mills fields (Section 20.3) quantized in an analogous gauge.

[27] Do not confuse this L with the Lagrangian.

[28] The reasoning presented here is essentially identical with the one used in Section 3.8 in quantizing the free electromagnetic field (in the gauge $\varphi = 0$, $\boldsymbol{\nabla}\cdot\mathbf{A} = 0$.

[29] The necessary orthogonality and completeness relations read

$$\int_V d^3\mathbf{x}\, e^{i\mathbf{x}\cdot(\mathbf{k}-\mathbf{k}')} = V\,\delta_{\mathbf{k},\mathbf{k}'}\,, \qquad \sum_{\mathbf{k}} e^{i\mathbf{k}\cdot(\mathbf{x}-\mathbf{y})} = V\,\delta^{(3)}(\mathbf{x}-\mathbf{y})\,.$$

Going over to the Hamiltonian is then straightforward:

$$H_0 = \frac{1}{2}(p_{\mathbf{0}}^2 + \omega_{\mathbf{0}}^2 q_{\mathbf{0}}^2) + \frac{1}{2}\sum_{\mathbf{k}>0}\left(p_{\mathbf{k}}^2 + \omega_{\mathbf{k}}^2 q_{\mathbf{k}}^2 + \bar{p}_{\mathbf{k}}^2 + \omega_{\mathbf{k}}^2 \bar{q}_{\mathbf{k}}^2\right), \tag{11.85}$$

and the quantization just means promoting the new canonical variables $q_{\mathbf{0}}$, $p_{\mathbf{0}}$, $q_{\mathbf{k}}$, $p_{\mathbf{k}}$ and $\bar{q}_{\mathbf{k}}$, $\bar{p}_{\mathbf{k}}$ (with $\mathbf{k} > 0$) to Hermitian operators satisfying the standard commutation rules

$$[q_{\mathbf{0}},\, p_{\mathbf{0}}] = i\hbar, \qquad [q_{\mathbf{k}},\, p_{\mathbf{k}'}] = [\bar{q}_{\mathbf{k}},\, \bar{p}_{\mathbf{k}'}] = i\hbar\delta_{\mathbf{k}\mathbf{k}'}\,. \tag{11.86}$$

etc. One can then construct the field operators $\hat{\varphi}(\mathbf{x})$ and $\hat{\Pi}(\mathbf{x})$:

$$\begin{aligned}
\hat{\varphi}(\mathbf{x}) &\equiv \frac{1}{\sqrt{V}}\left(q_{\mathbf{0}} + \sum_{\mathbf{k}>0}\frac{q_{\mathbf{k}} + i\bar{q}_{\mathbf{k}}}{\sqrt{2}}\, e^{i\mathbf{k}\cdot\mathbf{x}} + \sum_{\mathbf{k}>0}\frac{q_{\mathbf{k}} - i\bar{q}_{\mathbf{k}}}{\sqrt{2}}\, e^{-i\mathbf{k}\cdot\mathbf{x}}\right), \\
\hat{\Pi}(\mathbf{x}) &\equiv \frac{1}{\sqrt{V}}\left(p_{\mathbf{0}} + \sum_{\mathbf{k}>0}\frac{p_{\mathbf{k}} + i\bar{p}_{\mathbf{k}}}{\sqrt{2}}\, e^{i\mathbf{k}\cdot\mathbf{x}} + \sum_{\mathbf{k}>0}\frac{p_{\mathbf{k}} - i\bar{p}_{\mathbf{k}}}{\sqrt{2}}\, e^{-i\mathbf{k}\cdot\mathbf{x}}\right),
\end{aligned} \tag{11.87}$$

Writing them in the form (thereby defining the operators $\hat{\varphi}_{\mathbf{k}}$ and $\hat{\Pi}_{\mathbf{k}}$)

$$\hat{\varphi}(\mathbf{x}) = \frac{1}{\sqrt{V}}\sum_{\mathbf{k}}\hat{\varphi}_{\mathbf{k}}\, e^{i\mathbf{k}\cdot\mathbf{x}}, \qquad \hat{\Pi}(\mathbf{x}) = \frac{1}{\sqrt{V}}\sum_{\mathbf{k}}\hat{\Pi}_{\mathbf{k}}\, e^{i\mathbf{k}\cdot\mathbf{x}}, \tag{11.88}$$

(with the sums extending now to all $\mathbf{k}$) one finds, using the rules (11.86), that

$$[\hat{\varphi}_{\mathbf{k}},\ \hat{\Pi}_{\mathbf{k}'}] = i\hbar\,\delta_{-\mathbf{k},\mathbf{k}'}\,, \qquad [\hat{\varphi}_{\mathbf{k}},\ \hat{\varphi}_{\mathbf{k}'}] = [\hat{\Pi}_{\mathbf{k}},\ \hat{\Pi}_{\mathbf{k}'}] = 0\,. \tag{11.89}$$

This ensures that the canonical commutation relations (11.78) are satisfied[30] by the operators $\hat{\varphi}(\mathbf{x})$ and $\hat{\Pi}(\mathbf{x})$.

The "big" Hilbert space $\mathcal{H}$, the vectors of which represent all possible states of the field as the quantum system, consists of all functions $\Phi(q_{\mathbf{0}},\, q_{\mathbf{k}_1},\, \bar{q}_{\mathbf{k}_1}, \dots)$ of the countably infinite set of variables $q_{\mathbf{0}}$, $q_{\mathbf{k}}$ and $\bar{q}_{\mathbf{k}}$. The natural scalar product in $\mathcal{H}$ is given by

$$(\Phi_2|\Phi_1) = \int dq_{\mathbf{0}}\prod_{\mathbf{k}>0} dq_{\mathbf{k}} d\bar{q}_{\mathbf{k}}\, \Phi_2^*(q_{\mathbf{0}},\, q_{\mathbf{k}_1},\, \bar{q}_{\mathbf{k}_1}, \dots)\Phi_1(q_{\mathbf{0}},\, q_{\mathbf{k}_1},\, \bar{q}_{\mathbf{k}_1}, \dots)\,. \tag{11.90}$$

(Due to the infinite number of variables, finiteness of such scalar products and normalizability of the "wave functions" Φ is, of course, a delicate question). Since each set of values of $q_{\mathbf{0}}$, $q_{\mathbf{k}}$ and $\bar{q}_{\mathbf{k}}$ uniquely specifies a configuration of the classical field $\varphi(\mathbf{x})$, the "wave function" $\Phi(q_{\mathbf{0}},\, q_{\mathbf{k}_1},\, \bar{q}_{\mathbf{k}_1}, \dots)$ of the system (the field) can be given the standard probablistic interpretation: the quantity

$$|\Phi(q_{\mathbf{0}},\, q_{\mathbf{k}_1},\, \bar{q}_{\mathbf{k}_1}, \dots)|^2 dq_{\mathbf{0}}\prod_{\mathbf{k}>0} dq_{\mathbf{k}} d\bar{q}_{\mathbf{k}}\,,$$

[30] One can also take the viewpoint that this is the proper justification of these relations.

can be treated (provided the product $\prod_{\mathbf{k}>0} dq_{\mathbf{k}} d\bar{q}_{\mathbf{k}}$ of the differentials is finite), as the probablity that in the field quantum state represented by Φ the classical configuration of the field will be found between $\varphi(\mathbf{x})$ and $\varphi(\mathbf{x}) + \delta\varphi(\mathbf{x})$, where $\delta\varphi(\mathbf{x})$ is related to $dq_{\mathbf{k}}$ and $d\bar{q}_{\mathbf{k}}$.

The peculiar feature of the "big" Hilbert space $\mathcal{H}$ introduced above is its nonseparability. As a result, a choice of a set of functions which would constitute the basis of the whole Hilbert space $\mathcal{H}$ if the number of variables were finite, specifies only a particular subspace of $\mathcal{H}$; two different choices of such sets can specify subspaces which are mutually orthogonal in the scalar product (11.90). As such a set of functions one can take the products

$$\Psi_{n_{\mathbf{0}}\, n_{\mathbf{k}_1}\bar{n}_{\mathbf{k}_1}\dots}(q_{\mathbf{0}},\, q_{\mathbf{k}_1},\, \bar{q}_{\mathbf{k}_1}, \dots) = \psi_{n_{\mathbf{0}}}(q_{\mathbf{0}})\, \psi_{n_{\mathbf{k}_1}}(q_{\mathbf{k}_1})\, \psi_{\bar{n}_{\mathbf{k}_1}}(\bar{q}_{\mathbf{k}_1}) \dots\,, \tag{11.91}$$

of functions $\psi_n(q)$, $n = 0, 1, 2, \dots$ of one variable forming a complete discrete set of normalizable functions of $L_2(\mathbb{R})$. Such complete sets of functions can be different for different variables $q_{\mathbf{0}}$, $q_{\mathbf{k}_1}$, $\bar{q}_{\mathbf{k}_1}$, ... For instance, as $\psi_n(q)$ one can take the sets of harmonic oscillator wave functions corresponding to frequencies Ω which can arbitrarily change from one variable to another (i.e. $\Omega_{\mathbf{0}}$, $\Omega_{\mathbf{k}}$ and $\bar{\Omega}_{\mathbf{k}}$ can be arbitrary functions of $\mathbf{k}$); it should be clear that a priori this dependence needs not be correlated in any particular way with $\omega_{\mathbf{k}} = \sqrt{\mathbf{k}^2 + M^2}$. State vectors (11.91) can be denoted $|n_{\mathbf{0}}, n_{\mathbf{k}_1}, \bar{n}_{\mathbf{k}_1}, \dots\rangle$. Singling out the "Fock-vacuum" vector[31] $|0, 0, \dots\rangle \equiv |0_{\mathrm{Fock}}\rangle$, introducing the operators $A_{\mathbf{0}}$, $A^\dagger_{\mathbf{0}}$, $A_{\mathbf{k}}$, $A^\dagger_{\mathbf{k}}$ and $\bar{A}_{\mathbf{k}}$, $\bar{A}^\dagger_{\mathbf{k}}$ ($\mathbf{k} > 0$) acting on the vectors $|n_{\mathbf{0}}, n_{\mathbf{k}_1}, \bar{n}_{\mathbf{k}_1}, \dots\rangle$ as do the bosonic creation and annihilation operators in the occupation number representation (see Section 5.2) that is, so that

$$\begin{aligned}
&A_{\mathbf{0}}|0_{\mathrm{Fock}}\rangle = A_{\mathbf{k}}|0_{\mathrm{Fock}}\rangle = \bar{A}_{\mathbf{k}}|0_{\mathrm{Fock}}\rangle = 0\,,\\
&A^\dagger_{\mathbf{k}_i}|n_{\mathbf{0}}, \dots, n_{\mathbf{k}_i}, \bar{n}_{\mathbf{k}_i}, \dots\rangle = \sqrt{n_{\mathbf{k}_i} + 1}\,|n_{\mathbf{0}}, \dots, n_{\mathbf{k}_i} + 1, \bar{n}_{\mathbf{k}_i}, \dots\rangle\,,
\end{aligned}$$

etc., and chosing as the basis the vectors $|n_{\mathbf{0}}, n_{\mathbf{k}_1}, \bar{n}_{\mathbf{k}_1}, \dots\rangle$ with[32] $n_{\mathbf{0}} + n_{\mathbf{k}_1} + \bar{n}_{\mathbf{k}_1} + \dots < \infty$ (and taking the Cauchy completion of the set of such vectors) one constructs the separable Fock space. Since the operators $q_{\mathbf{0}}$, $p_{\mathbf{0}}$, $q_{\mathbf{k}}$, $p_{\mathbf{k}}$, $\bar{q}_{\mathbf{k}}$, $\bar{p}_{\mathbf{k}}$, can be expressed through $A_{\mathbf{0}}$, $A^\dagger_{\mathbf{0}}$, $A_{\mathbf{k}}$, $A^\dagger_{\mathbf{k}}$ and $\bar{A}_{\mathbf{k}}$, $\bar{A}^\dagger_{\mathbf{k}}$, their algebra (11.86) and therefore also the algebra (11.78) in this way gets represented (irreducibly) in a separable Fock space. If the oscillator wave functions are used in the products (11.91), the operators $q_{\mathbf{0}}$, $p_{\mathbf{0}}$, $q_{\mathbf{k}}$, $p_{\mathbf{k}}$, $\bar{q}_{\mathbf{k}}$, $\bar{p}_{\mathbf{k}}$, are related to the operators $A_{\mathbf{0}}$, $A^\dagger_{\mathbf{0}}$, $A_{\mathbf{k}}$, $A^\dagger_{\mathbf{k}}$, $\bar{A}_{\mathbf{k}}$, $\bar{A}^\dagger_{\mathbf{k}}$ and the frequencies $\Omega_{\mathbf{0}}$, $\Omega_{\mathbf{k}}$, $\bar{\Omega}_{\mathbf{k}}$ in the standard way.

To explain why different choices of the sets of functions (11.91) can select orthogonal subspaces in the "big" Hilbert space $\mathcal{H}$, suppose $\Psi_{n_{\mathbf{0}}n_{\mathbf{k}_1}\bar{n}_{\mathbf{k}_1}\dots}$ and $\Phi_{n_{\mathbf{0}}n_{\mathbf{k}_1}\bar{n}_{\mathbf{k}_1}\dots}$ are two such

[31] An "ontological" difference behind the formal similarity to the $|\mathrm{void}\rangle$ vector introduced in Section 5.1 and the vector $|0_{\mathrm{Fock}}\rangle$ is perhaps worth noting: while in the second quantization formulation of many particle quantum mechanics the vector $|\mathrm{void}\rangle$ represents an artificial state of no particles (the literal "void"), here $|0_{\mathrm{Fock}}\rangle$ is a real state of a fluctuating quantum field.

[32] Recall that the set of all vectors $|n_{\mathbf{0}}, n_{\mathbf{k}_1}, \bar{n}_{\mathbf{k}_1}, \dots\rangle$ without any restiction on the sum of n's is still uncountably infinite.

sets constructed using the oscillator wave functions with different assignments of the frequencies $\Omega_{\mathbf{0}}$, $\Omega_{\mathbf{k}}$, $\bar{\Omega}_{\mathbf{k}}$ to the complete sets of functions of variables $q_{\mathbf{0}}$, $q_{\mathbf{k}}$ and $\bar{q}_{\mathbf{k}}$ (which may also be centered at different values $q^{(0)}_{\mathbf{k}}$, $\bar{q}^{(0)}_{\mathbf{k}}$). Since the individual integrals in the scalar product (11.90) are then certainly such that

$$\left|\int_{-\infty}^{\infty} dq_{\mathbf{k}}\, \phi^*_{m_{\mathbf{k}}}(q_{\mathbf{k}})\, \psi_{n_{\mathbf{k}}}(q_{\mathbf{k}})\right| < 1\,,$$

for any $n_{\mathbf{k}}$ and $m_{\mathbf{k}}$ (also if $n_{\mathbf{k}} = m_{\mathbf{k}}$), it is easy to figure out that typically

$$(\Phi_{m_{\mathbf{0}}m_{\mathbf{k}_1}\bar{m}_{\mathbf{k}_1}\dots}|\Psi_{n_{\mathbf{0}}n_{\mathbf{k}_1}\bar{n}_{\mathbf{k}_1}\dots}) = 0\,,$$

for any choice of the quantum numbers $m_{\mathbf{0}}m_{\mathbf{k}_1}\bar{m}_{\mathbf{k}_1}\dots$ and $n_{\mathbf{0}}n_{\mathbf{k}_1}\bar{n}_{\mathbf{k}_1}\dots$ This means that vectors of $\mathcal{H}$ constructed as normalizable superpositions of the basis vectors $\Psi_{n_{\mathbf{0}}n_{\mathbf{k}_1}\bar{n}_{\mathbf{k}_1}\dots}$ spanning the first Fock space cannot be obtained as superpositions of the vectors $\Phi_{m_{\mathbf{0}}m_{\mathbf{k}_1}\bar{m}_{\mathbf{k}_1}\dots}$ spanning another Fock space and vice versa. Since the algebra of the operators $q_{\mathbf{0}}$, $p_{\mathbf{0}}$, $q_{\mathbf{k}}$, $p_{\mathbf{k}}$, $\bar{q}_{\mathbf{k}}$, $\bar{p}_{\mathbf{k}}$ can be represented (through the respective annihilation and creation operators) on any of the sets of functions, $\Psi_{n_{\mathbf{0}}n_{\mathbf{k}_1}\bar{n}_{\mathbf{k}_1}\dots}$, or $\Phi_{n_{\mathbf{0}}n_{\mathbf{k}_1}\bar{n}_{\mathbf{k}_1}\dots}$, it follows that the big Hilbert space $\mathcal{H}$ furnishes a reducible representation of the algebra (11.86). Thus, in principle selecting (by chosing the right Fock space) a subspace of $\mathcal{H}$ in which the algebra of the operators $q_{\mathbf{0}}$, $p_{\mathbf{0}}$, $q_{\mathbf{k}}$, $p_{\mathbf{k}}$, $\bar{q}_{\mathbf{k}}$, $\bar{p}_{\mathbf{k}}$ is realized irreducibly is an important part of constructing the quantum theory of any field. Note also that the existence of many possible Fock spaces in which the algebra of the operators can be represented is not related to the particular form of the function $\omega_{\mathbf{k}}$ in H_0, that is to the Hamiltonian of the free field. It is also clear that if the Hamiltonian is not quadratic in field variables (an interaction term is added to H_0 and the field is no longer free) the construction of the Fock spaces presented here does not change because the form of the Hamiltonian has nowhere been used.

If the field is quantized in the flat Minkowski space-time, in which the time-independent Hamiltonian operator plays a distinguished role, physically motivated is the choice of the separable $\mathcal{H}$ subspace containing the lowest energy eigenvector of the theory Hamiltonian. In a curved space-time, in which the notion of the Hamiltonian is more delicate, or if the Hamiltonian is explicitly time-dependent (the field is coupled to some varying in time external agents), the choice of the subspace of $\mathcal{H}$ is more problematic and one tries to develop methods allowing, at least in the case of free fields defined on curved space-times, to extract out of the theory a physical information without making a concrete choice of the Fock space. In the physical (as opposed to mathematical) practice, however, one is forced to accept that in all field theories one has to impose some sort of an UV cutoff on lengts of the considered wave vectors $\mathbf{k}$ and, therefore, at least in the finite volume V, as long as there is an UV cutoff, any choice of the Fock space is equally good - any "wave function" $\Phi(q_{\mathbf{0}}, q_{\mathbf{k}_1}, \bar{q}_{\mathbf{k}_1}, \dots)$ depending on a finite (due to the cutoff) number of variables can be expressed as a superposition of the vectors of a basis of an arbitrarily chosen Fock space.

In the case of the free Hamiltonian (11.85) which is simply a sum of an infinite number of independent harmonic oscillator Hamiltonians, the choice of the proper separable subspace of $\mathcal{H}$ is rather trivial: it is the one spanned by the vectors $\Psi_{n_{\mathbf{0}} n_{\mathbf{k}_1} \bar{n}_{\mathbf{k}_1} \cdots}$ constructed out of the harmonic oscillator functions centered at $q_{\mathbf{k}} = 0$ and corresponding to frequencies $\Omega_{\mathbf{k}} = \bar{\Omega}_{\mathbf{k}} = \omega_{\mathbf{k}}$. The eigenvector $|\Omega_0\rangle$ of H_0 corresponding to the lowest energy is then the Fock space "vacuum" vector $|0, 0, 0, \ldots\rangle \equiv |0_{\text{Fock}}\rangle$. To see this it suffices to introduce the operators ($\mathbf{k} > 0$)

$$
\begin{aligned}
a_{\mathbf{k}} &= \frac{A_{\mathbf{k}} + i\bar{A}_{\mathbf{k}}}{\sqrt{2}}, \qquad & a^\dagger_{\mathbf{k}} &= \frac{A^\dagger_{\mathbf{k}} - i\bar{A}^\dagger_{\mathbf{k}}}{\sqrt{2}}, \\
a_{-\mathbf{k}} &= \frac{A_{\mathbf{k}} - i\bar{A}_{\mathbf{k}}}{\sqrt{2}}, \qquad & a^\dagger_{-\mathbf{k}} &= \frac{A^\dagger_{\mathbf{k}} + i\bar{A}^\dagger_{\mathbf{k}}}{\sqrt{2}},
\end{aligned} \tag{11.92}
$$

and $a_{\mathbf{0}} = A_{\mathbf{0}}$, $a^\dagger_{\mathbf{0}} = A^\dagger_{\mathbf{0}}$, satisfying the standard rules

$$
[a_{\mathbf{k}},\ a^\dagger_{\mathbf{k}'}] = \delta_{\mathbf{k},\mathbf{k}'}\,, \qquad [a_{\mathbf{k}},\ a_{\mathbf{k}'}] = [a^\dagger_{\mathbf{k}},\ a^\dagger_{\mathbf{k}'}] = 0\,, \tag{11.93}
$$

and form the basis of the same Fock space out of the state-vectors ($\mathbf{k}_0 \equiv \mathbf{0}$)

$$
\begin{aligned}
|n_{\mathbf{k}_0}, n_{\mathbf{k}_1}, \ldots\rangle &\equiv \frac{1}{\sqrt{n_{\mathbf{k}_0}!\, n_{\mathbf{k}_1}! \ldots}}\, |\mathbf{k}_0, \ldots, \mathbf{k}_0, \mathbf{k}_1, \ldots, \mathbf{k}_1, \ldots\rangle \\
&= \frac{\left(a^\dagger_{\mathbf{k}_0}\right)^{n_{\mathbf{k}_0}}}{\sqrt{n_{\mathbf{k}_0}!}} \frac{\left(a^\dagger_{\mathbf{k}_1}\right)^{n_{\mathbf{k}_1}}}{\sqrt{n_{\mathbf{k}_1}!}} \ldots |0_{\text{Fock}}\rangle\,,
\end{aligned} \tag{11.94}
$$

in which now the vectors $\mathbf{k}_i$ are not restricted to $\mathbf{k}_i \geq 0$. These vectors can be interpreted (see below) as representing $n_{\mathbf{k}_0}$ bosons having momentum $\mathbf{k}_0$, $n_{\mathbf{k}_1}$ bosons having momentum $\mathbf{k}_1$, etc. Using the standard relations $q_{\mathbf{k}} = (A_{\mathbf{k}} + A^\dagger_{\mathbf{k}})/\sqrt{2\omega_{\mathbf{k}}}$ etc. and (11.83) one then finds that for all $\mathbf{k}$

$$
\hat{\varphi}_{\mathbf{k}} = \frac{1}{\sqrt{2\omega_{\mathbf{k}}}} \left(a_{\mathbf{k}} + a^\dagger_{-\mathbf{k}}\right), \qquad \hat{\Pi}_{\mathbf{k}} = \frac{1}{i}\sqrt{\frac{\omega_{\mathbf{k}}}{2}} \left(a_{\mathbf{k}} - a^\dagger_{-\mathbf{k}}\right), \tag{11.95}
$$

so that

$$
\begin{aligned}
\hat{\varphi}(\mathbf{x}) &= \sum_{\mathbf{k}} \sqrt{\frac{\hbar}{2V\omega_{\mathbf{k}}}} \left(a_{\mathbf{k}}\, e^{i\mathbf{k}\cdot\mathbf{x}} + a^\dagger_{\mathbf{k}}\, e^{-i\mathbf{k}\cdot\mathbf{x}}\right), \\
\hat{\Pi}(\mathbf{x}) &= \frac{1}{i} \sum_{\mathbf{k}} \sqrt{\frac{\hbar\omega_{\mathbf{k}}}{2V}} \left(a_{\mathbf{k}}\, e^{i\mathbf{k}\cdot\mathbf{x}} - a^\dagger_{\mathbf{k}}\, e^{-i\mathbf{k}\cdot\mathbf{x}}\right).
\end{aligned} \tag{11.96}
$$

The Hamiltonian (11.85) takes the form

$$
\hat{H}_0 = \frac{1}{2} \sum_{\mathbf{k}} \left(\hat{\Pi}_{\mathbf{k}} \hat{\Pi}_{-\mathbf{k}} + \omega^2_{\mathbf{k}}\, \hat{\varphi}_{\mathbf{k}} \hat{\varphi}_{-\mathbf{k}}\right) = \sum_{\mathbf{k}} \hbar\omega_{\mathbf{k}} \left(a^\dagger_{\mathbf{k}} a_{\mathbf{k}} + \frac{1}{2}\right), \tag{11.97}
$$

and the basis vectors (11.94) are (here by construction) its (normalizable) eigenvectors. In particular it is clear that[33] $|\Omega_0\rangle = |0_{\rm Fock}\rangle$.

The interpretation of the $\hat{H}_0$ eigenvectors in terms of noninteracting particles (bosons) suggested above[34] is supported by the statistical properties of the quantized (free) field: To show this let us consider the quantum field (still enclosed in the box of volume $V = L^3$) in contact with a heat bath of temperature T and compute, using the Gibbs Canonical Ensemble, the statistical sum $Z_{\rm stat} = e^{-\beta F}$, where[35] $\beta = 1/k_{\rm B}T$ and $F(T,V)$ is the Helmholtz free energy. This reduces to

$$\begin{aligned} Z_{\rm stat} = {\rm Tr}\, e^{-\beta H_0} &= \sum_{n_{\mathbf{k}_0}=0}^{\infty} \sum_{n_{\mathbf{k}_1}=0}^{\infty} \dots \ \langle n_{\mathbf{k}_0}, n_{\mathbf{k}_1}, \dots | e^{-\beta \hat{H}_0} | n_{\mathbf{k}_0}, n_{\mathbf{k}_1}, \dots \rangle \\ &= \prod_{\mathbf{k}} \left[e^{-\hbar\omega_{\mathbf{k}}/2k_{\rm B}T} \sum_{n_{\mathbf{k}}=0}^{\infty} \left(e^{-\hbar\omega_{\mathbf{k}}/k_{\rm B}T} \right)^{n_{\mathbf{k}}} \right]. \end{aligned} \tag{11.98}$$

The geometric series can easily be summed and one finds

$$F(T,V) = -k_{\rm B}T \ln Z_{\rm stat} = \frac{1}{2}\sum_{\mathbf{k}} \hbar\omega_{\mathbf{k}} + k_{\rm B}T \sum_{\mathbf{k}} \ln\left(1 - e^{-\hbar\omega_{\mathbf{k}}/k_{\rm B}T}\right), \tag{11.99}$$

or, going over, with the help of the prescription (5.48), to the continuous normalization,

$$F(T,V) = V\int \frac{d^3\mathbf{k}}{(2\pi)^3} \left[\frac{1}{2}\hbar\omega_{\mathbf{k}} + k_{\rm B}T \ln\left(1 - e^{-\hbar\omega_{\mathbf{k}}/k_{\rm B}T}\right) \right], \tag{11.100}$$

with $\omega_{\mathbf{k}} = \sqrt{c^2\mathbf{k}^2 + c^4M^2/\hbar^2}$. Apart from the first term which represents the contribution of the zero point oscillations (which could have been subtracted from the beginning by redefining $\hat{H}_0$) the temperature dependent part of F is precisely the free energy of a system of nointeracting relativistic bosons which do not carry any conserved quantum number and the total number of which - similarly to the total number of photons - is therefore determined solely by the condition of thermal equilibrium. Thus, one of the arguments that allow us to interpret states of quantized fields as particles is essentially identical to the statistical argument used by Einstein in 1905 to argue that electromagnetic radiation behaves as a collection of noninteracting particles (photons).

Calculation of the partition function $Z_{\rm stat}$ of an interacting quantum field is most easily formulated in the path integral approach. Still, presenting that approach in Chapter 16 we

[33] One can go further and show that the "wave function" of the ground state has the form

$$\langle q_{\mathbf{0}}, q_{\mathbf{k}_1}, \bar{q}_{\mathbf{k}_1}, \dots | \Omega_0 \rangle \equiv \Psi_{\Omega_0}[\varphi] = \prod_{\mathbf{k}} \left(\frac{\omega_{\mathbf{k}}}{\pi\hbar} \right)^{1/4} \exp\left(-\frac{1}{2\hbar} \sum_{\mathbf{k}} \omega_{\mathbf{k}} \varphi_{\mathbf{k}} \varphi^*_{\mathbf{k}} \right).$$

[34] Notice that having quantized the field in a finite box one cannot appeal to the Poincaré transformation properties of these state-vectors.

[35] $k_{\rm B} = 8.617343 \times 10^{-5}$ eV/K is the Boltzmann constant.

will need the result (11.100) to fix an additive constant in $F(T,V)$ which in the functional approach is difficult to compute.

It is clear that the choice of the proper Fock space, fairly obvious in the case of a free field, ceases to be such when the Hamiltonian of the field involves an interaction term. In connection with this it is perhaps instructive to consider a slightly less trivial example of the field $\varphi(\mathbf{x})$ satisfying as previously periodic boundary conditions in the box of volume $V = L^3$ and the dynamics of which is set by the Lagrangian density

$$\mathcal{L} = \frac{1}{2}\partial_\mu\varphi\partial^\mu\varphi - \frac{1}{2}M^2\varphi^2 - g f(\mathbf{x})\varphi(t,\mathbf{x}) \equiv \mathcal{L}_0 + \mathcal{L}_{\text{int}}\,, \tag{11.101}$$

in which g is the coupling constant and $f(\mathbf{x})$ is a given real function also satisfying the periodic boundary conditions. It can therefore be written as

$$f(\mathbf{x}) = \frac{1}{\sqrt{V}}\sum_{\mathbf{k}} f_{\mathbf{k}}\, e^{i\mathbf{k}\cdot\mathbf{x}}\,, \qquad \text{with} \quad f^*_{\mathbf{k}} = f_{-\mathbf{k}}$$

Introducing as previously the variables $\varphi_{\mathbf{k}} = (q_{\mathbf{k}} + i\bar{q}_{\mathbf{k}})/\sqrt{2}$ for $\mathbf{k} > 0$ etc. one can represent the algebra of the operators $q_{\mathbf{0}}$, $p_{\mathbf{0}}$, $q_{\mathbf{k}}$, $p_{\mathbf{k}}$, $\bar{q}_{\mathbf{k}}$ and $\bar{p}_{\mathbf{k}}$ or, alternatively, of the operators $a_{\mathbf{k}}$ and $a^\dagger_{\mathbf{k}}$ defined in (11.92) in the Fock space spanned by the vectors (11.94). The Hamiltonian takes in this case the form $\hat{H} = \hat{H}_0 + \hat{V}_{\text{int}}$ with $\hat{H}_0$ given by (11.97) and

$$V_{\text{int}} = g\left\{\frac{1}{\sqrt{2\omega_{\mathbf{0}}}}\,(a_{\mathbf{0}} + a^\dagger_{\mathbf{0}})f_{\mathbf{0}} + \sum_{\mathbf{k}>0}\frac{1}{\sqrt{2\omega_{\mathbf{k}}}}\left[(a_{\mathbf{k}} + a^\dagger_{-\mathbf{k}})f_{-\mathbf{k}} + (a_{-\mathbf{k}} + a^\dagger_{\mathbf{k}})f_{\mathbf{k}}\right]\right\}.$$

It is clear that now the basis vectors (11.94) are not eigenvectors of H and that $|0_{\text{Fock}}\rangle$ is not the lowest energy eigenvector of H. To find the eigenvectors of H one can notice that if one introduced the new operators[36]

$$\tilde{a}_{\mathbf{k}} \equiv a_{\mathbf{k}} + c_{\mathbf{k}}\,, \qquad \tilde{a}^\dagger_{\mathbf{k}} \equiv a^\dagger_{\mathbf{k}} + c^*_{\mathbf{k}}\,, \tag{11.102}$$

(satisfying the same commutation relations as do the operators $a_{\mathbf{k}}$ and $a^\dagger_{\mathbf{k}}$, because the factors $c_{\mathbf{k}}$'s are c-numbers) with

$$c_{\mathbf{k}} = \frac{g f_{\mathbf{k}}}{\sqrt{2\omega^3_{\mathbf{k}}}}\,, \qquad c^*_{\mathbf{k}} = \frac{g f_{-\mathbf{k}}}{\sqrt{2\omega^3_{\mathbf{k}}}}\,,$$

the Hamiltonian would take the form ($\hbar = 1$ again)

$$H = \sum_{\mathbf{k}}\omega_{\mathbf{k}}\left(\tilde{a}^\dagger_{\mathbf{k}}\tilde{a}_{\mathbf{k}} + \frac{1}{2}\right) - \sum_{\mathbf{k}}\frac{g^2 f_{\mathbf{k}} f_{-\mathbf{k}}}{2\omega^2_{\mathbf{k}}}\,.$$

[36]The steps performed below are essentially the ones done in Section 1.3 in solving the problem of the harmonic oscillator subject to the action of an external force.

It follows that the vectors $|\tilde{n}_{\mathbf{k}_0}, \tilde{n}_{\mathbf{k}_1}, \ldots\rangle$ built out of the vector $|\tilde{0}, \tilde{0}, \ldots\rangle \equiv |\tilde{0}_{\rm Fock}\rangle$ annihilated by all $\tilde{a}_{\mathbf{k}}$ are also the eigenvectors of $\hat{H}$ and therefore $|\tilde{0}_{\rm Fock}\rangle = |\Omega\rangle$. These vectors are given by

$$|\tilde{n}_{\mathbf{k}_0}, \tilde{n}_{\mathbf{k}_1}, \ldots\rangle = U(c)|n_{\mathbf{k}_0}, n_{\mathbf{k}_1}, \ldots\rangle \,,$$

where the formally unitary operator $U(c)$ is given by

$$U(c) = \exp\left(-\sum_{\mathbf{k}} c_{\mathbf{k}} a^\dagger_{\mathbf{k}} + \sum_{\mathbf{k}} c^*_{\mathbf{k}} a_{\mathbf{k}}\right) . \tag{11.103}$$

Indeed, because[37]

$$U^{-1}(c)\, a_{\mathbf{k}} U(c) = a_{\mathbf{k}} - c_{\mathbf{k}} \,, \qquad U^{-1}(c)\, a^\dagger_{\mathbf{k}} U(c) = a^\dagger_{\mathbf{k}} - c^*_{\mathbf{k}} \,, \tag{11.104}$$

$$\begin{aligned}
\hat{H}|\tilde{n}_{\mathbf{k}_0}, \tilde{n}_{\mathbf{k}_1}, \ldots\rangle &= U(c)\, U^{-1}(c)\, \hat{H}\, U(c)|n_{\mathbf{k}_0}, n_{\mathbf{k}_1}, \ldots\rangle \\
&= U(c) \sum_{\mathbf{k}} \left[\omega_{\mathbf{k}}\left(a^\dagger_{\mathbf{k}} a_{\mathbf{k}} + \frac{1}{2}\right) - \frac{g^2|f_{\mathbf{k}}|^2}{2\omega^3_{\mathbf{k}}}\right] |n_{\mathbf{k}_0}, n_{\mathbf{k}_1}, \ldots\rangle \\
&= \left\{\sum_{\mathbf{k}} \left[\omega_{\mathbf{k}}\left(n_{\mathbf{k}} + \frac{1}{2}\right) - \frac{g^2|f_{\mathbf{k}}|^2}{2\omega^3_{\mathbf{k}}}\right]\right\} |\tilde{n}_{\mathbf{k}_0}, \tilde{n}_{\mathbf{k}_1}, \ldots\rangle \,.
\end{aligned}$$

Since $U(c)$ can be written also in the form[38]

$$U(c) = \exp\left(-\frac{1}{2}\sum_{\mathbf{k}} \frac{g^2|f_{\mathbf{k}}|^2}{2\omega^3_{\mathbf{k}}}\right) \exp\left(-\sum_{\mathbf{k}} c_{\mathbf{k}} a^\dagger_{\mathbf{k}}\right) \exp\left(\sum_{\mathbf{k}} c^*_{\mathbf{k}} a_{\mathbf{k}}\right) ,$$

it is clear that all scalar products

$$\langle m_{\mathbf{k}_0}, m_{\mathbf{k}_1}, \ldots |\tilde{n}_{\mathbf{k}_0}, \tilde{n}_{\mathbf{k}_1}, \ldots\rangle = \langle m_{\mathbf{k}_0}, m_{\mathbf{k}_1}, \ldots |U(c)|n_{\mathbf{k}_0}, n_{\mathbf{k}_1}, \ldots\rangle \,,$$

are proportional to $\exp\left(-\frac{1}{2}\sum_{\mathbf{k}} \frac{g^2|f_{\mathbf{k}}|^2}{2\omega^3_{\mathbf{k}}}\right)$ and vanish either if the sum over $\mathbf{k}$ diverges or in the infinite volume limit $V \to \infty$ in which the sum over $\mathbf{k}$ is replaced according to the rule (5.48) by the integral. In particular this is the case when $f(\mathbf{x}) \equiv 1$, so that $f_{\mathbf{k}} = \sqrt{V}\delta_{\mathbf{0},\mathbf{k}}$. In such cases the true ground state-vector $|\Omega\rangle$ of the system and all other $\hat{H}$ eigenvectors inhabit, if $V = \infty$, a subspace of $\mathcal{H}$ orthogonal to the Fock space (11.94) in which the algebra of the operators has originally been realized. It is also easy to see that while $\langle 0_{\rm Fock}|\hat{\varphi}(\mathbf{x})|0_{\rm Fock}\rangle = 0$,

$$\begin{aligned}
\langle\Omega|\hat{\varphi}(\mathbf{x})|\Omega\rangle \equiv \langle\tilde{0}_{\rm Fock}|\hat{\varphi}(\mathbf{x})|\tilde{0}_{\rm Fock}\rangle &= \langle 0_{\rm Fock}|U^{-1}(c)\, \hat{\varphi}(\mathbf{x}) U(c)|0_{\rm Fock}\rangle \\
&= \frac{1}{\sqrt{V}} \sum_{\mathbf{k}} \frac{1}{\sqrt{2\omega_{\mathbf{k}}}} \left(c_{\mathbf{k}}\, e^{i\mathbf{k}\cdot\mathbf{x}} + c^*_{\mathbf{k}}\, e^{-i\mathbf{k}\cdot\mathbf{x}}\right) \equiv h(\mathbf{x}) \,,
\end{aligned} \tag{11.105}$$

[37]Notice also that $U(c)\, a_{\mathbf{k}} U^{-1}(c) = \tilde{a}_{\mathbf{k}}$, $U(c)\, a^\dagger_{\mathbf{k}} U^{-1}(c) = \tilde{a}^\dagger_{\mathbf{k}}$. Therefore, on the vectors $|\tilde{n}_{\mathbf{k}_0}, \tilde{n}_{\mathbf{k}_1}, \ldots\rangle$ the operators $\tilde{a}_{\mathbf{k}}$ and $\tilde{a}^\dagger_{\mathbf{k}}$ act the same way as do the operators $a_{\mathbf{k}}$ and $a^\dagger_{\mathbf{k}}$ on the states $|n_{\mathbf{k}_0}, n_{\mathbf{k}_1}, \ldots\rangle$.

[38]Recall, that $e^{A+B} = e^A e^B e^{-\frac{1}{2}[A,B]}$ if, as here, $[A,\,[A,\,B]] = [B,\,[A,\,B]] = 0$.

- the expectation value of the field operator $\hat{\varphi}(\mathbf{x})$ in the H ground state is nonvanishing.

Of course to build the Fock space in which the algebra of the operators $q_{\mathbf{0}}$, $p_{\mathbf{0}}$, $q_{\mathbf{k}}$, $p_{\mathbf{k}}$, $\bar{q}_{\mathbf{k}}$, $\bar{p}_{\mathbf{k}}$, is realized one could use the wave functions of displaced harmonic oscillators (i.e. of harmonic oscillator the origins of which are at $q^{(0)}_{\mathbf{k}} \neq 0$) with the displacements appropriately correlated with the factors $f_{\mathbf{k}}$. The basis vectors of such a Fock space would then precisely be the vectors $|\tilde{n}_{\mathbf{k}_0}, \tilde{n}_{\mathbf{k}_1}, \ldots\rangle$ and the corresponding creation and annihilation operators would be $\tilde{a}_{\mathbf{k}}$ and $\tilde{a}^{\dagger}_{\mathbf{k}}$ defined in (11.102), while the relation (11.95) would in this case read

$$\begin{aligned}\hat{\varphi}_{\mathbf{k}} &= \frac{1}{\sqrt{2\omega_{\mathbf{k}}}}\left(c_{\mathbf{k}} + c^{*}_{-\mathbf{k}}\right) + \frac{1}{\sqrt{2\omega_{\mathbf{k}}}}\left(\tilde{a}_{\mathbf{k}} + \tilde{a}^{\dagger}_{-\mathbf{k}}\right),\\ \hat{\Pi}_{\mathbf{k}} &= \frac{1}{i}\sqrt{\frac{\omega_{\mathbf{k}}}{2}}\left(\tilde{a}_{\mathbf{k}} - \tilde{a}^{\dagger}_{-\mathbf{k}}\right),\end{aligned} \tag{11.106}$$

leading to $\langle\tilde{0}_{\text{Fock}}|\hat{\varphi}(\mathbf{x})|\tilde{0}_{\text{Fock}}\rangle = h(\mathbf{x})$, where $|\tilde{0}_{\text{Fock}}\rangle$ is the Fock vacuum vector of the "displaced" Fock space which is now the lowest energy eigenvector $|\Omega\rangle$ of $\hat{H}$. One can say that the algebra of the field operators $\hat{\varphi}(\mathbf{x})$ and $\hat{\Pi}(\mathbf{x})$ is in this case realized in another, in general unitarily inequvalent, Fock space.

Yet another way of quantizing the theory defined by the classical Lagrangian density (11.101) consists of introducing first (before quantization) another canonical field variable $\chi(t,\mathbf{x})$ related to $\varphi(t,\mathbf{x})$ by $\varphi(t,\mathbf{x}) = \chi(t,\mathbf{x}) + h(\mathbf{x})$, with the function $h(\mathbf{x})$ satisfying (in addition to periodic boundary conditions in the volume V) the differential equation

$$(\partial_i\partial_i - M^2)h(\mathbf{x}) = gf(\mathbf{x})\,.$$

Obviously,

$$h(\mathbf{x}) = \frac{1}{\sqrt{V}}\sum_{\mathbf{k}}\left(\frac{-gf_{\mathbf{k}}}{\omega^2_{\mathbf{k}}}\right)e^{i\mathbf{k}\cdot\mathbf{x}}\,.$$

is precisely the same function as in (11.105). The Lagrangian density (11.101) expressed in terms of $\chi(t,\mathbf{x})$ is (after integrating by parts) equivalent to

$$\mathcal{L} = \frac{1}{2}\partial_\mu\chi\partial^\mu\chi - \frac{1}{2}M^2\chi^2 - \frac{1}{2}ghf\,.$$

Upon quantization the resulting algebra of the operators $\hat{\chi}(\mathbf{x})$ and $\hat{\Pi}_\chi(\mathbf{x})$ would be then from the beginning realized in the Fock space spanned by the vectors $|\tilde{n}_{\mathbf{k}_0}, \tilde{n}_{\mathbf{k}_1}, \ldots\rangle$ (which are directly related to probabilities of finding a given classical configuration of the field $\chi(\mathbf{x})$), because the creation and annihilation operators entering $\hat{\chi}(\mathbf{x})$ and $\hat{\Pi}_\chi(\mathbf{x})$ would be precisely the operators $\tilde{a}_{\mathbf{k}}$ and $\tilde{a}^{\dagger}_{\mathbf{k}}$ defined above. Thus quantization of systems characterized by infinite numbers of degrees of freedom using different sets of their canonical variables (at the classical level related to each other by decent canonical transformations) may lead to realizations of the respective operator algebras in orthogonal Fock spaces (orthogonal subspaces of the same "big" Hilbert space).

Of course, in the case of quantum theories of truly interacting fields (for example if in (11.101) $\mathcal{L}_{\rm int} \propto \varphi^4$) finding the right Fock space is practically impossible and one contents oneself by realizing the algebra of field operators in the Fock space built on the lowest energy eigenvector of some appropriately chosen free hamiltonian $\hat{H}_0$ after making an educated guess at which classical canonical variables are the most appropriate for quantization in the given case.

Quantizing in the infinite space the real scalar field $\varphi(\mathbf{x})$, one could try to follow closely the approach adopted in the finite volume: taking a complete set of real functions $f_l(\mathbf{x})$ vanishing for $|\mathbf{x}| \to \infty$ and orthonormal,[39] i.e. such that

$$\int d^3\mathbf{x}\, f_{l'}(\mathbf{x})\, f_l(\mathbf{x}) = \delta_{l'l}\,, \qquad \sum_l f_l(\mathbf{x})\, f_l(\mathbf{x}') = \delta^{(3)}(\mathbf{x}-\mathbf{x}')\,,$$

one could write every field configuration $\varphi(\mathbf{x})$ in the form

$$\varphi(\mathbf{x}) = \sum_l q_l f_l(\mathbf{x})\,, \tag{11.107}$$

introducing thereby a countably infinite set of canonical variables q_l. If the field is free its Lagrangian (11.66) would then take the form

$$L_0 = \int d^3\mathbf{x}\, \mathcal{L}_0 = \frac{1}{2}\sum_l \dot{q}_l^2 - \frac{1}{2}\sum_{l'l} V_{l'l}\, q_{l'} q_l\,, \tag{11.108}$$

with

$$V_{l'l} = M^2 \delta_{l'l} + \int d^3\mathbf{x}\, \partial_i f_{l'}(\mathbf{x})\, \partial_i f_l(\mathbf{x})\,. \tag{11.109}$$

The free Hamiltonian would then read

$$H_0 = \frac{1}{2}\sum_l p_l^2 + \frac{1}{2}\sum_{l'l} V_{l'l}\, q_{l'} q_l\,. \tag{11.110}$$

If the field is not free and its Lagrangian and Hamiltonian is not quadratic in the canonical field variables, nothing changes except that the Hamiltonian would have in such a case additional terms expressed through the variables q_l. Quantization would now just mean

[39] One assumes here that all configurations of the fluctuating quantum fields vanish at spatial infinity; in some cases this may be too strong a requirement, especially when the gauge fields are considered because classically they are not by themselves observable.

promoting the canonical variables q_l and p_l to Hermitian operators $\hat{q}_l$ and $\hat{p}_l$ and imposing the standard commutation relations $[\hat{q}_l,\, \hat{p}_{l'}] = i\hbar\,\delta_{ll'}$. The Hermitian field operators

$$\hat{\varphi}(\mathbf{x}) = \sum_l \hat{q}_l f_l(\mathbf{x})\,,$$
$$\hat{\Pi}(\mathbf{x}) = \sum_l \hat{p}_l f_l(\mathbf{x})\,,$$

woud then satisfy the canonical commutation relations[40] (11.78). Introducing for each variable q_l a set of basis functions $\psi_{n_l}(q_l)$ of $L_2(\mathbb{R})$ one could then form the Fock space, a separable subspace of the "big" Hilbert space of all states of the system, spanned by the vectors $|n_1, n_2, \ldots\rangle$ with the Fock "vacuum" state $|0_{\text{Fock}}\rangle \equiv |0, 0, 0, \ldots\rangle$ in which the algebra of the operators $\hat{\varphi}(\mathbf{x})$ and $\hat{\Pi}(\mathbf{x})$ would be realized. (Of course, different choices of the basis functions $f_l(\mathbf{x})$ select different and in general mutually orthogonal Fock spaces in the big Hilbert space of all states of the quantum field). However, the vectors $|n_1, n_2, \ldots\rangle$, although having direct interpretation in terms of probabilities of finding various classical field configurations would not be eigenvectors of $\hat{H}_0$. Therefore in this case it is better to follow a slightly different approach which is equivalent to taking instead of the set of normalizable functions $f_l(\mathbf{x})$, a set of nonnormalizable plane waves; as a result one is essentially constructing a basis of $\mathcal{H}^*$, the dual of $\mathcal{H}$ which admits nonnormalizable state-vectors. To this end one first defines the new operators (from now on we omit "hats") $\tilde{\varphi}(\mathbf{k})$ and $\tilde{\Pi}(\mathbf{k})$ by

$$\tilde{\varphi}(\mathbf{k}) = \int d^3\mathbf{x}\,\varphi(\mathbf{x})\,e^{-i\mathbf{k}\cdot\mathbf{x}}\,, \qquad \tilde{\Pi}(\mathbf{k}) = \int d^3\mathbf{x}\,\Pi(\mathbf{x})\,e^{-i\mathbf{k}\cdot\mathbf{x}}\,. \tag{11.111}$$

which (because of Hermiticity of $\varphi(\mathbf{x})$ and $\Pi(\mathbf{x})$) satisfy the relations

$$\tilde{\varphi}^\dagger(\mathbf{k}) = \tilde{\varphi}(-\mathbf{k})\,, \qquad \tilde{\Pi}^\dagger(\mathbf{k}) = \tilde{\Pi}(-\mathbf{k})\,. \tag{11.112}$$

From the commutation relations (11.78) it then follows that

$$\begin{aligned} [\tilde{\varphi}(\mathbf{k}),\ \tilde{\Pi}(\mathbf{k}')] &= i\hbar\,(2\pi)^3\delta^{(3)}(\mathbf{k}+\mathbf{k}')\,,\\ [\tilde{\varphi}(\mathbf{k}),\ \tilde{\varphi}(\mathbf{k}')] &= \left[\tilde{\Pi}(\mathbf{k}),\ \tilde{\Pi}(\mathbf{k}')\right] = 0\,. \end{aligned} \tag{11.113}$$

Expressing the Hamiltonian (11.75) with $K(\mathbf{x}-\mathbf{y})$ given by (11.70) in terms of $\tilde{\varphi}(\mathbf{k})$, $\tilde{\Pi}(\mathbf{k})$ and $\omega^2(\mathbf{k})$ given by

$$\omega^2(\mathbf{k}) = \int d^3\mathbf{x}\,K(\mathbf{x})\,e^{-i\mathbf{k}\cdot\mathbf{x}} = c^2\mathbf{k}^2 + \frac{M^2c^4}{\hbar^2}\,, \tag{11.114}$$

one obtains

$$H_0 = \frac{1}{2}\int\frac{d^3\mathbf{k}}{(2\pi)^3}\left[\tilde{\Pi}(\mathbf{k})\tilde{\Pi}(-\mathbf{k}) + \omega^2(\mathbf{k})\,\tilde{\varphi}(\mathbf{k})\tilde{\varphi}(-\mathbf{k})\right]. \tag{11.115}$$

[40] Again, this can be viewed as a justification of these relations.

By analogy with the harmonic oscillator case we define now the operators

$$\begin{aligned} a(\mathbf{k}) &= \sqrt{\frac{\omega(\mathbf{k})}{2\hbar}} \left(\tilde{\varphi}(\mathbf{k}) + \frac{i}{\omega(\mathbf{k})} \tilde{\Pi}(\mathbf{k}) \right), \\ a^\dagger(\mathbf{k}) &= \sqrt{\frac{\omega(\mathbf{k})}{2\hbar}} \left(\tilde{\varphi}(-\mathbf{k}) - \frac{i}{\omega(\mathbf{k})} \tilde{\Pi}(-\mathbf{k}) \right), \end{aligned} \tag{11.116}$$

the commutation rules of which:

$$\begin{aligned} \left[a(\mathbf{k}),\ a^\dagger(\mathbf{k}')\right] &= (2\pi)^3 \delta^{(3)}(\mathbf{k} - \mathbf{k}'), \\ \left[a(\mathbf{k}),\ a(\mathbf{k}')\right] &= \left[a^\dagger(\mathbf{k}),\ a^\dagger(\mathbf{k}')\right] = 0 . \end{aligned} \tag{11.117}$$

follow directly from (11.113). In terms of $a(\mathbf{k})$ and $a^\dagger(\mathbf{k})$ the operators (11.111) are given by

$$\tilde{\varphi}(\mathbf{k}) = \sqrt{\frac{\hbar}{2\omega(\mathbf{k})}} \left[a(\mathbf{k}) + a^\dagger(-\mathbf{k})\right], \qquad \tilde{\Pi}(\mathbf{k}) = \frac{1}{i}\sqrt{\frac{\hbar\omega(\mathbf{k})}{2}} \left[a(\mathbf{k}) - a^\dagger(-\mathbf{k})\right].$$

Inserting these expressions in (11.115) we get the Hamiltonian in the form

$$\begin{aligned} H_0 &= \frac{1}{2} \int \frac{d^3\mathbf{k}}{(2\pi)^3}\ \hbar\omega(\mathbf{k}) \left[a^\dagger(\mathbf{k}) a(\mathbf{k}) + a(\mathbf{k}) a^\dagger(\mathbf{k})\right] \\ &= \int \frac{d^3\mathbf{k}}{(2\pi)^3}\ E(\mathbf{k}) \left[a^\dagger(\mathbf{k}) a(\mathbf{k}) + \frac{1}{2}(2\pi)^3 \delta^{(3)}(\mathbf{0})\right], \end{aligned} \tag{11.118}$$

with $E(\mathbf{k}) = \hbar\omega(\mathbf{k})$. The delta function $(2\pi)^3\delta^{(3)}(\mathbf{0})$ in the second term should be interpreted as the (infinite) volume factor. The term

$$\frac{1}{2} \int \frac{d^3\mathbf{k}}{(2\pi)^3} E(\mathbf{k})\, (2\pi)^3 \delta^{(3)}(\mathbf{0}) \to V \times \frac{1}{2} \int \frac{d^3\mathbf{k}}{(2\pi)^3}\ E(\mathbf{k}), \tag{11.119}$$

can be therefore identified with the contribution to the field energy contained in the volume V of the zero point oscillations of infinitely many field modes (numbered by the wave vectors $\mathbf{k}$). This infinite constant contribution can be discarded so long as the effect of the quantized field on gravity is not considered and so long as one does not compare energies of vacuum states of quantized fields subject to different boundary conditions (see Section 11.3).

Expressed in terms of the operators $a^\dagger(\mathbf{k})$ and $a(\mathbf{k})$, the Schrödinger picture (time independent) field operators $\varphi(\mathbf{x})$ and $\Pi(\mathbf{x})$ take the form

$$\varphi(\mathbf{x}) = \int \frac{d^3\mathbf{k}}{(2\pi)^3} \sqrt{\frac{\hbar}{2\omega(\mathbf{k})}} \left(a(\mathbf{k})\, e^{i\mathbf{k}\cdot\mathbf{x}} + a^\dagger(\mathbf{k})\, e^{-i\mathbf{k}\cdot\mathbf{x}}\right), \tag{11.120}$$

$$\Pi(\mathbf{x}) = \frac{1}{i}\int\frac{d^3\mathbf{k}}{(2\pi)^3}\sqrt{\frac{\hbar\omega(\mathbf{k})}{2}}\left(a(\mathbf{k})\,e^{i\mathbf{k}\cdot\mathbf{x}} - a^\dagger(\mathbf{k})\,e^{-i\mathbf{k}\cdot\mathbf{x}}\right), \tag{11.121}$$

It is then easy to check, that with (11.117) the basic commutation rules (11.78) are satisfied.

One can now construct a separable Fock space in which the algebra (11.117) of the operators can be naturally represented. To this end one can take an arbitrary complete countable set of normalizable functions $f_l(\mathbf{k})$, $l = 1, 2, \ldots, \infty$ such that

$$\int\frac{d^3\mathbf{k}}{(2\pi)^3}\,f^*_{l'}(\mathbf{k})f_l(\mathbf{k}) = \delta_{l'l}\,, \qquad \sum_l f_l(\mathbf{k})f^*_l(\mathbf{k}') = (2\pi)^3\delta^{(3)}(\mathbf{k}-\mathbf{k}')\,,$$

(for instance, $f_l(\mathbf{k})$ can be the momentum space three-dimensional harmonic oscillator functions) and define the new operators[41]

$$a_l = \int\frac{d^3\mathbf{k}}{(2\pi)^3}\,f^*_l(\mathbf{k})\,a(\mathbf{k})\,, \qquad a^\dagger_l = \int\frac{d^3\mathbf{k}}{(2\pi)^3}\,f_l(\mathbf{k})\,a^\dagger(\mathbf{k})\,, \tag{11.122}$$

satisfying the familiar rules

$$[a_{l'},\ a^\dagger_l] = \delta_{l'l}\,, \qquad [a_{l'},\ a_l] = [a^\dagger_{l'},\ a^\dagger_l] = 0\,.$$

The inverse relations read

$$a(\mathbf{k}) = \sum_l a_l\,f_l(\mathbf{k})\,, \qquad a^\dagger(\mathbf{k}) = \sum_l a^\dagger_l\,f^*_l(\mathbf{k})\,. \tag{11.123}$$

It is then possible to use the same argument as for the ordinary harmonic oscillator (see the first footnote in section 1.3) that in the space in which a_l and $a^\dagger_l$ act there must exist a vector $|0,\ldots,0,\ldots\rangle \equiv |0_{\rm Fock}\rangle$ annihilated by all a_l's and to construct the states

$$|n_1, n_2, \ldots, n_l, \ldots\rangle = \frac{(a^\dagger_1)^{n_1}}{\sqrt{n_1!}}\frac{(a^\dagger_2)^{n_2}}{\sqrt{n_2!}}\ldots\frac{(a^\dagger_l)^{n_l}}{\sqrt{n_l!}}\ldots|0_{\rm Fock}\rangle\,, \tag{11.124}$$

with $n_1 + n_2 + \ldots < \infty$, which span the Fock space.[42] In terms of these creation and annihilation operators the Hamiltonian (11.118) takes the form

$$\begin{aligned} H_0 &= \sum_{l'l}\left[\int\frac{d^3\mathbf{k}}{(2\pi)^3}\,\frac{\omega(\mathbf{k})}{2}\,f^*_{l'}(\mathbf{k})f_l(\mathbf{k})\right]\left(a^\dagger_{l'}a_l + a_l a^\dagger_{l'}\right) \\ &= \sum_{l'l}\left[\int\frac{d^3\mathbf{k}}{(2\pi)^3}\,\omega(\mathbf{k})\,f^*_{l'}(\mathbf{k})f_l(\mathbf{k})\right]\left(a^\dagger_{l'}a_l + \frac{1}{2}\,\delta_{l'l}\right). \end{aligned} \tag{11.125}$$

[41] The operators $a^\dagger(\mathbf{k})$ and $a(\mathbf{k})$, and therefore also $\varphi(\mathbf{x})$ and $\Pi(\mathbf{x})$, are operator-valued distributions which acting on normalizable vectors of the proper Fock space throw them out of it.

[42] As discussed in Chapter 5, the Hilbert space spanned by all vectors $|n_1, n_2, \ldots, n_l, \ldots\rangle$ with no restriction on the sum of the numbers n_l is not separable and the abstract algebra (11.117) can be represented in infinitely many unitarily inequivalent ways.

It is then clear that the Fock space "vacuum" vector $|0_{\rm Fock}\rangle \equiv |0,0,0,\dots\rangle$ is the normalizable eigenvector $|\Omega_0\rangle$ of H_0 with the infinite eigenvalue given by (11.119). The remaining eigenvectors can be easily constructed but are not normalizable: it is straightforward to check (using the formulae (11.123) and the completeness relation) that the vectors

$$\begin{aligned}|\mathbf{k}\rangle &= a^\dagger(\mathbf{k})|\Omega_0\rangle\,,\\ |\mathbf{k}_1,\mathbf{k}_2\rangle &= a^\dagger(\mathbf{k}_1)\,a^\dagger(\mathbf{k}_2)|\Omega_0\rangle\,,\\ |\mathbf{k}_1,\mathbf{k}_2,\mathbf{k}_3\rangle &= a^\dagger(\mathbf{k}_1)\,a^\dagger(\mathbf{k}_2)\,a^\dagger(\mathbf{k}_3)|\Omega_0\rangle\,,\\ &\dots\end{aligned}\tag{11.126}$$

where all momenta $\mathbf{k}$, $\mathbf{k}_1$, $\mathbf{k}_2$, etc. are arbitrary are eigenvectors of H_0. (These state-vectors are also not properly normalized - in the generalized sense - when two or more momenta coincide; this has to be taken care of in the completeness relation as it was done in (5.18)). Although non-normalizable (if we assume that the state $|\Omega_0\rangle$ has the norm equal to 1) the vectors (11.126) are true (generalized) eigenvectors of the Hamiltonian H_0 (11.118). Their particle interpretation follows now from their Poincaré transformation properties which we discuss at the end of this subsection. The normalizable basis state-vectors (11.124) of the Fock space represent quantum excitations of the field φ, which can also be given a particle interpretation: the state $|n_1,n_2,\dots\rangle$ is interpreted as the state in which n_1 particles are in the first one-particle state characterized by the momentum space wave function $f_1(\mathbf{k})$, n_2 particles are in the second one-particle state characterized by $f_2(\mathbf{k})$, etc. The particles can occupy infinitely many one-particle states (in the case of N coupled oscillators there were only N types of phonons corresponding to N possible 1-phonon states). Although the state-vectors (11.124) are not eigenvectors of the Hamiltonian (11.118), they can, with the appropriately chosen functions $f_l(\mathbf{k})$ - essentially modeling the delta functions $\delta^{(3)}(\mathbf{k}-\mathbf{k}_l)$ - be made arbitrarily close to the true generalized eigenvectors (11.126) of H_0.

If the Lagrangian defining the theory is, unlike (11.66), not quadratic in the field φ, but instead takes the more general form (we set $c=1$ but for decoration keep $\hbar$ in some formulae below)

$$\mathcal{L} = \frac{1}{2}\partial_\mu\varphi\partial^\mu\varphi - V(\varphi) = \frac{1}{2}\partial_\mu\varphi\partial^\mu\varphi - \frac{1}{2}M^2\varphi^2 - \mathcal{H}_{\rm int}(\varphi)\,,\tag{11.127}$$

in which $\mathcal{H}_{\rm int}(\varphi)$, and therefore also the whole *field potential* $V(\varphi)$, is usually a polynomial of fields,[43] e.g. $\mathcal{H}_{\rm int}(\varphi) = (\lambda/4!)\varphi^4$, the quantization in the Schrödinger picture proceeds in the same way as described above: the commutation rules (11.78) of the operators φ and Π remain unchanged. The Hamiltonian takes then the form

$$\begin{aligned}H &= H_0 + V_{\rm int} = H_0 + \int d^3\mathbf{x}\,\mathcal{H}_{\rm int}(\varphi)\,,\\ H_0 &= \int d^3\mathbf{x}\left[\frac{1}{2}\Pi^2 + \frac{1}{2}(\boldsymbol{\nabla}\varphi)^2 + \frac{1}{2}M^2\varphi^2\right].\end{aligned}\tag{11.128}$$

[43] We will see that in the presence of $\mathcal{H}_{\rm int}(\varphi)$ the sign of the term of (11.127) quadratic in φ can also be positive (negative M^2).

The form (11.128) of H remains valid also if $\mathcal{H}_{\rm int}(\varphi)$, or even $V(\varphi)$, explicitly depends on time (e.g. if the quantized field φ interacts with some external agents). As in the case of the free field, one can now expand the Schrödinger picture operators $\varphi(\mathbf{x})$ and $\Pi(\mathbf{x})$ as in (11.120) and (11.121) into the operators $a(\mathbf{k})$ and $a^\dagger(\mathbf{k})$ with the same commutation rules as previously but the Hamiltonian (11.128) will now contain terms $(a(\mathbf{k}))^4$, $(a^\dagger(\mathbf{k}))^4$, etc. and the generalized state-vectors (11.126) will not be its eigenvectors.[44] Still, if the coefficient M^2 does not depend on time, they are eigenvectors of the free part H_0 of the Hamiltonian (11.128) and will play an important role in the formulation of the S-matrix approach to the scattering theory. This will be discussed in Section 11.9.

In older formulations one used to *define* the quantum theory by taking for the Hamiltonian operator not the expression (11.128), but its counterpart $\hat{H} = \ :H(11.128):$ ordered normally with respect to the vector $|\Omega_0\rangle = |0_{\rm Fock}\rangle$. Operators ordered in this way have all (except for the unit operator) zero expectation value in the Fock state $|0_{\rm Fock}\rangle$ annihilated (by definition) by all the annihilation operators. This prescription for $\hat{H}$ removes some of the infinities encountered in practical calculations, in particular it removes the additive infinite part (11.119) in the expectation value of the free field Hamiltonian (11.118) in the state $|\Omega_0\rangle$. Normal ordering, ubiquitous on older approaches to quantum field theory, has been now largely abandoned. First of all, it does not remove all divergences which must be renormalized anyway (see Chapter 14) and there is no point to invoke two different prescription for removing divergences having a common ultraviolet origin. Furthermore, normal ordering of operators defined in terms of the creation and annihilation operators diagonalizing H_0 (i.e. normal ordering with respect to the Fock state $|0_{\rm Fock}\rangle$) does not imply that $\langle\Omega|\ :O_S:\ |\Omega\rangle = 0$, where $|\Omega\rangle$ is the lowest energy eigenvector of the Hamiltonian - the example (11.105) clearly shows that matrix elements of operators ordered with respect to the Fock space "vacuum" can have nonzero expectation values in the true vacuum.[45] Finally, and probably most importantly, in the modern functional approach (see Chapter 16) quantum field theory is viewed as a formalism allowing to take into account real quantum fluctuations of fields; from this point of view[46] every quantum field theory is defined with Fourier momenta bounded by an ultraviolet cutoff Λ which (as we expect) should have a real physical meaning. All contributions to amplitudes, expectation values, etc. one computes using the quantum field theory formalism are then equally physical and should not be subtracted using an arbitrarily defined prescription. The divergences (arising in the limit $\Lambda \to \infty$) disappear however if the computed quantities are expressed in terms of some other measurable quantities; subtracting divergences is then merely done for computational convenience and has no fundamental meaning; hence defining the Hamiltonian as a normal ordered operator becomes then a completely useless

[44]Of course, if $\mathcal{H}_{\rm int}$ depends explicitly on time, the Hamiltonian H does not, strictly speaking, possess time-independent eigenstates.

[45]In nontrivial models of quantum field theory this may result as a consequence of interactions which cannot be treated by perturbative methods based on the expansion exploiting the Gell-Mann - Low construction of Section 1.2.

[46]This point of view is possible also in gauge theories although it entails in this case serious technical complications.

and obsolete prescription.[47]

The classical real scalar field φ has been quantized here in the Schrödinger picture in which the operators $\varphi(\mathbf{x})$ and $\Pi(\mathbf{x})$ satisfying the rules (11.78) do not depend on time. The canonical commutation rules (11.78), the Hamiltonian expressed in terms of the operators $\varphi(\mathbf{x})$ and $\Pi(\mathbf{x})$ constitute, together with the choice of the appropriate Fock space, the complete formulation of the theory. It can be then used to investigate various problems using the standard methods of quantum mechanics. Since in the case of relativistic theories one is interested primarily in scattering processes and in preserving manifest covariance it is, however, convenient to go over to the Heisenberg picture, in which the operators depend on time whereas the the state-vectors do not. In general (see Section 1.1), this is achieved by choosing the moment t_0 at which the two pictures coincide and defining Heisenberg picture operators $O_H(t,\mathbf{x})$ corresponding to Schrödinger picture operators $O_S(t,\mathbf{x}) = O(t,\varphi(\mathbf{x}),\Pi(\mathbf{x}))$ by the formula

$$O_H(t,\mathbf{x}) = U^\dagger(t,t_0)O_S(t,\mathbf{x})U(t,t_0) = O(t,\varphi_H(t,\mathbf{x}),\Pi_H(t,\mathbf{x}))\,, \tag{11.129}$$

in which

$$U(t,t_0) = T\exp\left(-\frac{i}{\hbar}\int_{t_0}^{t} dt'\,H(t')\right), \tag{11.130}$$

is the Schrödinger picture evolution operator. By construction the operators $O_H(t,\mathbf{x})$ satisfy then the Heisenberg equation (see Section 1.1)

$$\begin{aligned}\frac{d}{dt}\,Q_H(t,\mathbf{x}) &= \frac{1}{i\hbar}\,[O_H(t,\mathbf{x}),\ H_H(t)] + \left(\frac{\partial O}{\partial t}\right)_H \\ &\equiv \frac{1}{i\hbar}\,U^\dagger(t,t_0)\,[O_S(t,\mathbf{x}),\ H(t_0)]\,U(t,t_0) + \left(\frac{\partial O}{\partial t}\right)_H\,,\end{aligned} \tag{11.131}$$

in which $H_H(t) \equiv U^\dagger(t,t_0)H(t)U(t,t_0)$. (Most of the Schrödinger picture operators O_S do not depend on time and the second term in (11.131) is absent; one notable exception is the boost generators given by the space integrals (11.63) of M^{00i} given by (11.62)).

It is easy to check that the Heisenberg picture operators $\varphi_H(t,\mathbf{x})$ and $\Pi_H(t,\mathbf{x})$, corresponding to the canonical variables and obtained from the prescription (11.129), satisfy the *canonical equal time commutation rules*

$$\begin{gathered}[\varphi_H(t,\mathbf{x}),\ \Pi_H(t,\mathbf{y})] = i\delta^{(3)}(\mathbf{x}-\mathbf{y})\,, \\ [\varphi_H(t,\mathbf{x}),\ \varphi_H(t,\mathbf{y})] = [\Pi_H(t,\mathbf{x}),\ \Pi_H(t,\mathbf{y})] = 0\,,\end{gathered} \tag{11.132}$$

for arbitrary times t. These rules together with the Hamiltonian H (following from the action $I[\phi]$) are in fact the basic relations defining the quantum version of the classical field

[47]Normal ordering remains relevant (as a technical tool) for the Wick theorem - see Section (5.9) - used to set the perturbative expansion.

theory model. Furthermore, if H is of the form (11.128) (even if $V(\varphi)$ in the Lagrangian density (11.127) depends explicitly on time, i.e. even if the system interacts with external agents) the Heisenberg equations (11.131) satisfied by $\varphi_H(t,\mathbf{x})$ and $\Pi_H(t,\mathbf{x})$,

$$\begin{aligned}\frac{d}{dt}\,\varphi_H(t,\mathbf{x}) &= \frac{1}{i\hbar}\left[\varphi_H(t,\mathbf{x}),\ H_H(t)\right],\\ \frac{d}{dt}\,\Pi_H(t,\mathbf{x}) &= \frac{1}{i\hbar}\left[\Pi_H(t,\mathbf{x}),\ H_H(t)\right],\end{aligned} \tag{11.133}$$

yield $\dot{\varphi}_H = \Pi_H$ and $\dot{\Pi}_H = c^2\boldsymbol{\nabla}^2\varphi_H - M^2\varphi_H - \mathcal{H}'_{\rm int}(\varphi_H)$. As a result, the Heisenberg operator $\varphi_H(t,\mathbf{x})$ satisfies the "classical" field equation of motion

$$\left(\frac{\partial^2}{\partial t^2} - \boldsymbol{\nabla}^2 + M^2\right)\varphi_H(t,\mathbf{x}) = -\mathcal{H}'_{\rm int}(\varphi_H(t,\mathbf{x}))\,, \tag{11.134}$$

while the operator $\Pi_H(t,\mathbf{x})$ satisfies the equation

$$\left(\frac{\partial^2}{\partial t^2} - \boldsymbol{\nabla}^2 + M^2 + \mathcal{H}''_{\rm int}(\varphi_H(t,\mathbf{x}))\right)\Pi_H(t,\mathbf{x}) = 0\,. \tag{11.135}$$

Together with the equal time commutation rules (11.132) these two equations can also constitute the complete specification of the theory.

Of course, in the case of closed systems, which the system of fields usually are, H is independent of time. The evolution operator $U(t,t_0)$ (11.130) reduces then to the ordinary exponent, $H_H(t) = H$, so that in the considered theory (11.127) it takes the form

$$H = \int d^3\mathbf{x}\left[\frac{1}{2}\Pi^2_H(t,\mathbf{x}) + \frac{1}{2}(\boldsymbol{\nabla}\varphi_H)^2(t,\mathbf{x}) + V(\varphi_H(t,\mathbf{x}))\right], \tag{11.136}$$

formally the same as in the classical theory, and choosing $t_0 = 0$ as it is customary, one has

$$\begin{aligned}\varphi_H(t,\mathbf{x}) &= e^{iHt/\hbar}\,\varphi(\mathbf{x})\,e^{-iHt/\hbar}\,,\\ \Pi_H(t,\mathbf{x}) &= e^{iHt/\hbar}\,\Pi(\mathbf{x})\,e^{-iHt/\hbar}\,.\end{aligned} \tag{11.137}$$

In the case of the free scalar field the Hamiltonian H_0 (11.75) of which can be represented in the form (11.118), the equations (11.133) can be easily solved. $\varphi_H(t,\mathbf{x})$ and $\Pi_H(t,\mathbf{x})$ are then given by the expressions (11.120) and (11.121) but with $a(\mathbf{k})$ and $a^\dagger(\mathbf{k})$ replaced by $a_H(t,\mathbf{k})$ and $a^\dagger_H(t,\mathbf{k})$ satisfying the equations

$$\begin{aligned}\frac{d}{dt}\,a_H(t,\mathbf{k}) &= \frac{1}{i\hbar}\left[a_H(t,\mathbf{k}),\ H_0\right],\\ \frac{d}{dt}\,a^\dagger_H(t,\mathbf{k}) &= \frac{1}{i\hbar}\left[a^\dagger_H(t,\mathbf{k}),\ H_0\right],\end{aligned} \tag{11.138}$$

with the initial conditions $a_H(0,\mathbf{k}) = a(\mathbf{k})$, $a_H^\dagger(0,\mathbf{k}) = a^\dagger(\mathbf{k})$. One then finds $[a_H(t,\mathbf{k}),\ H_0] = \hbar\omega(\mathbf{k})a_H(t,\mathbf{k})$ and $[a_H^\dagger(t,\mathbf{k}),\ H_0] = -\hbar\omega(\mathbf{k})a_H^\dagger(t,\mathbf{k})$. The equations (11.138) can be then easily integrated[48] to give

$$a_H(t,\mathbf{k}) = e^{-i\omega(\mathbf{k})t}\,a(\mathbf{k})\,, \qquad a_H^\dagger(t,\mathbf{k}) = e^{i\omega(\mathbf{k})t}\,a^\dagger(\mathbf{k})\,, \tag{11.139}$$

so that

$$\varphi_H(t,\mathbf{x}) = \int\frac{d^3\mathbf{k}}{(2\pi)^3}\sqrt{\frac{\hbar}{2\omega(\mathbf{k})}}\left[a(\mathbf{k})\,e^{-i\omega(\mathbf{k})t+i\mathbf{k}\cdot\mathbf{x}} + a^\dagger(\mathbf{k})\,e^{i\omega(\mathbf{k})t-i\mathbf{k}\cdot\mathbf{x}}\right], \tag{11.140}$$

$$\Pi_H(t,\mathbf{x}) = \frac{1}{i}\int\frac{d^3\mathbf{k}}{(2\pi)^3}\sqrt{\frac{\hbar\omega(\mathbf{k})}{2}}\left[a(\mathbf{k})\,e^{-i\omega(\mathbf{k})t+i\mathbf{k}\cdot\mathbf{x}} - a^\dagger(\mathbf{k})\,e^{i\omega(\mathbf{k})t-i\mathbf{k}\cdot\mathbf{x}}\right], \tag{11.141}$$

that is, the exponents in the Heisenberg operators depend on the Lorentz invariant products $x^\mu k_\mu$, where $x^\mu = (ct,\mathbf{x})$ and $k^\mu = (\omega(\mathbf{k})/c,\mathbf{k})$. Both field operators, φ_H and Π_H, satisfy in this case the free Klein-Gordon equation.

It should be said, that in certain cases (the most notable being the one of the free electromagnetic field quantized using the approach of Gupta and Bleuler (discussed in Section 11.11) with the gauge fixing parameter $\xi \neq 1$) eventhough the free Hamiltonian is quadratic in field variables and conjugated momenta, their expansions analogous to (11.120), (11.121) are not easy to find and do not render the free Hamiltonian diagonal; the time dependent operators (11.137) are then not simply given by replacing $e^{\pm i\mathbf{k}\cdot\mathbf{x}}$ in (11.120) and (11.121) by $e^{\mp ikx}$ as in (11.140) and (11.141). Still, the explicit form of the time dependent field operators and their associated canonical momenta operators can be found by directly solving (though not so easily) the canonical equations (11.133).

At this point, in order to simplify the notation, we set also $\hbar = 1$ and drop the subscript H on Heisenberg picture operators. In addition, it is convenient to change the normalization of the the creation and annihilation operators $a(\mathbf{k}) \to a(\mathbf{k})/\sqrt{2E(\mathbf{k})}$, $a^\dagger(\mathbf{k}) \to a^\dagger(\mathbf{k})/\sqrt{2E(\mathbf{k})}$, where $E(\mathbf{k}) = \hbar\omega_\mathbf{k}$, so that their commutator becomes

$$\left[a(\mathbf{k}),\ a^\dagger(\mathbf{k}')\right] = (2\pi)^3 2E(\mathbf{k})\delta^{(3)}(\mathbf{k}-\mathbf{k}') \equiv \delta_\Gamma(\mathbf{k}-\mathbf{k}'), \tag{11.142}$$

and the (Heisenberg picture) free field operators (11.140), (11.141) take the simple form

$$\varphi(x) = \int d\Gamma_\mathbf{k}\left[a(\mathbf{k})\,e^{-ik\cdot x} + a^\dagger(\mathbf{k})\,e^{ik\cdot x}\right], \tag{11.143}$$

$$\Pi(x) = \frac{1}{i}\int d\Gamma_\mathbf{k} E_\mathbf{k}\left[a(\mathbf{k})\,e^{-ik\cdot x} - a^\dagger(\mathbf{k})\,e^{ik\cdot x}\right], \tag{11.144}$$

[48]Notice, that if the Lagrangian contains terms with φ in powers higher than the second one, the Heisenberg equations resulting from (11.138) are more complicated and their solution, the operators $a_H(t,\mathbf{k})$ and $a_H^\dagger(t,\mathbf{k})$, cannot be found in a closed form. Hence, the Heisenberg picture operators $\varphi_H(x)$ and $\Pi_H(x)$ corresponding to the interacting field cannot be written in the forms (11.140) and (11.141); only at $t=0$ $\varphi_H(0,\mathbf{x}) \equiv \varphi(\mathbf{x})$ and $\Pi_H(0,\mathbf{x}) \equiv \Pi(\mathbf{x})$ (recall, we have chosen to equate the Schrödinger and Heisenberg pictures at $t_0 = 0$) can be written as in (11.120) and (11.121).

where the measure $d\Gamma_{\mathbf{k}} \equiv d^3\mathbf{k}/(2\pi)^3 2E(\mathbf{k})$ is separately Lorentz invariant.

In theories of interacting fields (when in the full Hamiltonian there are terms with powers of fields higher than the second one) the time dependent operators termed the *interaction picture* operators obtained by solving the equations (11.133) but with time independent H_0 replacing the full $H_H(t)$ and hence given in closed form by (11.137) again with H_0 instead of H, play an important role in formulating the perturbation expansion of S-matrix elements. This will be discussed in Section 11.9.

We can now discuss the Poincaré transformations. Using the rules given in Section 11.1 one finds the canonical energy-momentum tensor corresponding to the Lagrangian density (11.127)

$$T^{\mu\nu}_{\rm can} = \partial^\mu\varphi\partial^\nu\varphi - g^{\mu\nu}\left[\frac{1}{2}\partial^\lambda\varphi\partial_\lambda\varphi - \frac{1}{2}M^2\varphi^2 - \mathcal{H}_{\rm int}(\varphi)\right]. \tag{11.145}$$

It is in this case automatically symmetric.[49] In the quantum theory $\partial^0\varphi$ must of course be expressed in terms of the canonical variables φ and Π. One gets in this way (we drop the subscript H)

$$\begin{aligned} T^{00}_{\rm can}(t,\mathbf{x}) &= \frac{1}{2}\left[\Pi^2(t,\mathbf{x}) + \boldsymbol{\nabla}\varphi(t,\mathbf{x})\cdot\boldsymbol{\nabla}\varphi(t,\mathbf{x}) + M^2\varphi^2(t,\mathbf{x})\right] + \mathcal{H}_{\rm int}(\varphi(t,\mathbf{x}))\,, \\ T^{0i}_{\rm can}(t,\mathbf{x}) &= \Pi(t,\mathbf{x})\,\partial^i\varphi(t,\mathbf{x})\,. \end{aligned} \tag{11.146}$$

P^0 given by the integral of $T^{00}_{\rm can}$ over $d^3\mathbf{x}$ is just the Hamiltonian (11.128) written in terms of the Heisenberg field operators whereas the momentum operator is given by

$$P^i = \int d^3\mathbf{x}\, T^{0i}_{\rm can}(t,\mathbf{x}) = \int d^3\mathbf{x}\, \Pi(t,\mathbf{x})\partial^i\varphi(t,\mathbf{x})\,. \tag{11.147}$$

From $T^{\mu\nu}_{\rm can}$ the generators of the Lorentz transformations are obtained according to the formulae (11.62) and (11.63):

$$J^{\mu\nu} = \int d^3\mathbf{x}\left(x^\mu\, T^{0\nu}_{\rm can} - x^\nu\, T^{0\mu}_{\rm can}\right). \tag{11.148}$$

The Poincaré symmetry generators P^i, H, J^i and K^i are by construction independent of time. This can be checked by using the Heisenberg equation (11.131). As such they can be computed for any time t. A particularly convenient is the choice of $t = 0$, because then the Heisenberg picture operators φ_H and Π_H can be expanded into the creation and annihilation operators and can be shown, using the canonical commutation rules (11.78) (or 11.142)) to satisfy the commutation rules of the Poincaré algebra (6.19) or (6.21).

[49]Still, as it turns out, even in this simple case $T^{\mu\nu}_{\rm can}$ must be modified by adding to it a tensor $H^{\mu\nu} = \partial_\rho H^{\rho\mu\nu}$ as in (11.58), in order to ensure finiteness of matrix elements of the energy-momentum tensor (treated as an operator) in the theory of the interacting field.

Moreover, exploiting their time-independence, P^μ and $J^{\mu\nu}$ can be easily seen to generate transformations of the field operator $\varphi(x)$ related to changes of the reference frame:

$$\begin{aligned} e^{ia_\mu P^\mu}\varphi(x)\, e^{-ia_\mu P^\mu} &= \varphi(x+a)\,, \\ e^{\frac{i}{2}\omega_{\mu\nu}J^{\mu\nu}}\varphi(x)\, e^{-\frac{i}{2}\omega_{\mu\nu}J^{\mu\nu}} &= \varphi(\Lambda^{-1}(\omega)\cdot x)\,, \end{aligned} \tag{11.149}$$

or, in the infinitesimal form,

$$\begin{aligned} i\,[P^\mu,\ \varphi(x)] &= \partial^\mu\varphi(x)\,, \\ i\,[J^{\mu\nu},\ \varphi(x)] &= (x^\mu\partial^\nu - x^\nu\partial^\mu)\varphi(x)\,. \end{aligned} \tag{11.150}$$

It is important to stress that the operators P^i, H, J^i and K^i (the components of $J^{\mu\nu}$) satisfy the commutation rules (6.21) of the Poincaré algebra and generate the field transformations as in (11.149) solely by virtue of the canonical commutation relations (11.132), independently of the form of $\mathcal{H}_{\rm int}(\varphi)$ in (11.127), i.e. also in the quantum theory of the interacting field φ.

If $V(\varphi) = \frac{1}{2}M^2\varphi^2$, i.e. in the quantum theory of the free field (with $\mathcal{H}_{\rm int} = 0$), H takes in terms of the creation and annihilation operators the form (11.118) and the generators P^i and J^{ij} and $K^i = J^{0i}$ are given by

$$\begin{aligned} P^i &= \int d\Gamma_{\mathbf{k}}\, k^i\, a^\dagger(\mathbf{k})a(\mathbf{k})\,, \\ J^{ij} &= i\int d\Gamma_{\mathbf{k}}\, a^\dagger(\mathbf{k})\left(k^i\frac{\partial}{\partial k_j} - k_j\frac{\partial}{\partial k_i}\right)a(\mathbf{k})\,, \\ K^i &= i\int d\Gamma_{\mathbf{k}}\, a^\dagger(\mathbf{k})\, E_{\mathbf{k}}\frac{\partial}{\partial k_i}\, a(\mathbf{k})\,. \end{aligned} \tag{11.151}$$

In this case, the vectors (11.126) are the generalized eigenvectors of H and $\mathbf{P}$ with the eigenvalues $(E_{\mathbf{k}}, \hbar\mathbf{k})$, $(E_{\mathbf{k}_1} + E_{\mathbf{k}_2}, \hbar\mathbf{k}_1 + \hbar\mathbf{k}_2)$, etc., where $E_{\mathbf{k}}$ is given by the relativistic formula $E_{\mathbf{k}} = \sqrt{\hbar^2\mathbf{k}^2c^2 + M^2c^4} = \hbar\omega(\mathbf{k})$. One can also check that the one particle state-vectors $|\mathbf{k}\rangle$ transform properly under the Lorentz transformations,

$$U(\Lambda)|\mathbf{k}\rangle = U(\Lambda)\, a^\dagger(\mathbf{k})\, U^{-1}(\Lambda)\, U(\Lambda)|\Omega_0\rangle = a^\dagger(\mathbf{k}_\Lambda)|\Omega_0\rangle\,, \tag{11.152}$$

i.e. that they transform in the way appropriate for states of a spin 0 particle (see Section 6.2) provided the vacuum state-vector $|\Omega_0\rangle$ is Lorentz invariant $U(\Lambda)|\Omega_0\rangle = |\Omega_0\rangle$. In the interacting field case, if $\varphi_H(0,\mathbf{x})$ and $\Pi_H(0,\mathbf{x})$ are expanded into creation and annihilation operators, the operators P^i and J^{ij} obtained from (11.147) and (11.148) take usually[50] the same form as in (11.151), whereas the boost generators K^i are, as argued in Section 7.5, modified by the interaction.

Quantization of a single real scalar field discussed above can immediately be generalized to the case of many real scalar fields $\varphi_i(x)$ where $i = 1, \ldots, N$, the classical dynamics

[50] Exception are theories quantized around a nontrivial classical background $\varphi_{\rm cl}(\mathbf{x})$ in which the field operator takes the form $\hat\varphi_H(t,\mathbf{x}) = \varphi_{\rm cl}(\mathbf{x})\,\hat{1} + \hat\chi_H(t,\mathbf{x})$.

of which is governed by the Lagrangian density (11.55). Furthermore, a system of N classical complex fields can always be represented as a system of $2N$ real fields and quantized using the same prescriptions.

A Lagrangian density depending on several fields φ_i may be invariant (or invariant up to a total divergence - see the formula (11.20)) under the action of a group of continuous transformations which in the infinitesimal form can be written as $\varphi_i \to \varphi'_i = \varphi_i + \delta\theta_a F^a_i(\varphi)$. In this case, upon replacing classical fields φ_i and their derivatives by the corresponding field operators (as $\partial^0\varphi_i$ one has of course to use the expressions $\dot{\varphi}_i = \dot{\varphi}_i(\Pi, \varphi)$ given by the canonical formalism), the classical expressions for Noether currents j^a_μ become current operators and the quantities

$$\hat{Q}^a = \int d^3\mathbf{x}\, j^a_0(t, \mathbf{x})\,, \tag{11.153}$$

become the symmetry generators acting in the Hilbert space.[51] By using the canonical commutation rules (11.132) one can show that independently of the precise form of $\mathcal{L}$ (provided it is symmetric) the *time-like* components of the Noether currents satisfy the commutation relations

$$\left[j^a_0(t, \mathbf{y}),\ j^b_0(t, \mathbf{x})\right] = j^c_0(t, \mathbf{x})\, i f_c{}^{ab}\, \delta^{(3)}(\mathbf{x} - \mathbf{y})\,, \tag{11.154}$$

which ensure that the Noether charges (11.153) satisfy the symmetry algebra relations

$$[\hat{Q}^a,\ \hat{Q}^b] = \hat{Q}^c\, i f_c{}^{ab}\,. \tag{11.155}$$

Moreover, by virtue of the canonical commutation rules (11.78) the Noether charges $\hat{Q}^a$ generate symmetry transformations of the field operators

$$i\delta\theta_a\, [\hat{Q}^a,\ \varphi_i(x)] = \delta\theta_a F^a_i(\varphi(x))\,. \tag{11.156}$$

In contrast, space-like components of the Noether currents usually can satisfy the commutation rules analogous to (11.154) only in the theory of noninteracting fermionic fields (to be discussed in Section 11.8). In the general case, it is only possible to infer (using the Lorentz covariance) that

$$\left[j^a_0(t, \mathbf{y}),\ j^b_i(t, \mathbf{x})\right] = j^c_i(t, \mathbf{x})\, i f_c{}^{ab}\, \delta^{(3)}(\mathbf{x} - \mathbf{y}) + S^{ab}_{ik}(t, \mathbf{x})\, \partial^k_{(\mathbf{x})} \delta^{(3)}(\mathbf{x} - \mathbf{y})\,,$$

where $S^{ik}_{ab}(t, \mathbf{x})$ are the so-called Schwinger terms. Even in theories (like electrodynamics of spin 1/2 particles) in which the canonical (anti)commutation relations formally imply

[51] Of course, the operators $\hat{Q}^a$ in (11.153) are well defined if their matrix elements between physical states: $\langle\Phi|\hat{Q}^a|\Psi\rangle = \int d^3\mathbf{x}\, \langle\Phi|j^0_a(t, \mathbf{x})|\Psi\rangle$ are well defined; it may happen that the right hand side is not integrable due to the presence in the H spectrum of massless particles which mediate long range forces. This is so whenever a classical continuous symmetry of the Lagrangian is spontaneously broken by the vacuum state (see Chapter 22).

the absence of Schwinger terms,[52] it is possible to prove, using arguments based on general principles of quantum mechanics, that they cannot vanish. The contradiction with the reasoning based on the canonical (anti)commutation relations stems from the fact that Noether currents are composite operators (products of elementary field operators taken at the same space-time point) which are in general singular objects - their matrix elements require some regularization; Schwinger terms can therefore depend on the adopted regularization prescription, and a regularization can, a priori, also induce nontrivial Schwinger terms in the commutators (11.154) of the time components of the currents. If a given symmetry can be recovered after regularization in the renormalized theory, Schwinger terms in the commutators like (11.154) are absent (although are generally nonvanishing in the commutators involving spatial components of the Noether currents). In fact, anomalies discussed in Chapter 23 can be understood as manifestation of nontrivial Schwinger terms induced by the necessity of regularization in the commutators like (11.154) of the time components of the currents.

Finally, if the Hamiltonian commutes with the Noether charges $\hat{Q}^a$ and the symmetry is not spontaneously broken by the vacuum state, i.e. if $\hat{Q}^a|\Omega\rangle = 0$ for all a (see Chapter 22), the Hamiltonian eigenvectors (the *in* and *out* state-vectors) form multiplets of the symmetry algebra. Usually also the H_0 part quadratic in field operators of the full Hamiltonian H commutes with $\hat{Q}^a$'s separately, and the H_0 eigenvectors, i.e. particles created and annihilated by the Hermitian free-field (interaction picture) operators $\varphi_i(x)$ out of $|\Omega_0\rangle$ also form multiplets of the symmetry algebra. As all members of the symmetry multiplets have the same mass (the $P^\mu P_\mu$ operator eigenvalues), one then forms linear combinations of the particle states (both *in*, *out* and the free-particle ones) which diagonalize the Noether charges $\hat{Q}^a$ forming the Cartan subalgebra (see Chapter 4) of the full symmetry algebra. It is then convenient to form also the appropriate (complex in general) linear combinations of the free-field operators $\varphi_i(x)$ creating and destroying free one-particle eigenstates of the the Cartan subalgebra generators.[53] In the case of the symmetry multiplets transforming as complex representations, such that each particle (except for essentially neutral ones) finds its antiparticle within the same multiplet, this corresponds precisely to forming (non-Hermitian) field operators creating a particle and annihilating its antiparticle as described in Section 8.2.

Existence of conserved charges has also important consequences for statistical properties of the system of quantum fields. If the Hamiltonian commutes with Noether charges $\hat{Q}^a$, the quantum numbers corresponding to the Cartan subalgebra generators can have simultaneously definite values (because all these generators commute with one another) which are constants of motion. This would have to be taken into account in the Gibbs

[52] In theories in which symmetries are realized on scalar fields, the presence of nontrivial (regularization independent) Schwinger terms is usually revealed already by using the canonical commutation relations.

[53] The corresponding Heisenberg picture operators obtained from the free-field ones as in (11.137) (or, more generally, as in (11.129)) have then "diagonal" matrix elements between the one-particle states of the full Hamiltonian H and the vacuum: $\langle\Omega|\phi^i_H|\mathbf{p},\sigma,j\rangle \propto \delta^{ij}$ with the same proportionality constant for all ϕ^i_H and all states $|\mathbf{p},\sigma,j\rangle$ belonging to the same symmetry multiplet.

Canonical Ensemble statistical sum $Z_{\rm stat}(T,V)$, making it essentially intractable analytically (similarly as in the case of a system consisting of a fixed number N of particles) even for noninteracting fields. One then passes to the Grand Canonical Ensemble introducing the chemical potentials μ_a for each Cartan subalgebra generator. Choosing in the Hilbert space the basis in which the Cartan subalgebra generators are diagonal facilitates then computation of the statistical sum $\Xi(T,V,\mu_a)$.

For example, the Lagrangian

$$\mathcal{L}=\frac{1}{2}\sum_{l=1}^{2}\left(\partial_\mu\varphi_l\partial^\mu\varphi_l-M^2\varphi_l^2\right), \tag{11.157}$$

of two noninteracting scalar fields φ_1 and φ_2 has the $O(2)\simeq U(1)$ symmetry the conserved Noether current of which is $j_\mu=\varphi_2\partial_\mu\varphi_1-\varphi_1\partial_\mu\varphi_2$. After quantization it is the operators $a^\dagger(\mathbf{k})$ and $a^{c\dagger}(\mathbf{k})$ formed as appropriate linear combinations of $a_l^\dagger(\mathbf{k})$, $l=1,2$ which, acting on the vacuum $|\Omega_0\rangle$, generate the state-vectors $|n_{\mathbf{k}},\dots,n^c_{\mathbf{k}},\dots\rangle$ of the Hilbert space basis in which the Hamiltonian

$$H=\sum_{\mathbf{k}}\hbar\omega_{\mathbf{k}}\left(a^\dagger_{\mathbf{k}}a_{\mathbf{k}}+a^{c\dagger}_{\mathbf{k}}a^c_{\mathbf{k}}+1\right), \tag{11.158}$$

where $\omega_{\mathbf{k}}=\sqrt{c^2\mathbf{k}^2+c^4M^2/\hbar^2}$, and the Noether charge

$$\hat{Q}=\sum_{\mathbf{k}}(a^\dagger_{\mathbf{k}}a_{\mathbf{k}}-a^{c\dagger}_{\mathbf{k}}a^c_{\mathbf{k}}), \tag{11.159}$$

are simultaneously diagonal. Obtaining the potential $\Omega(T,V,\mu)=-k_{\rm B}T\ln\Xi$ by computing the statistical sum

$$\Xi=e^{-\beta\Omega}={\rm Tr}\,e^{-\beta(H-\mu\hat{Q})}, \tag{11.160}$$

is then straightforward:

$$\begin{aligned}\Omega(T,V,\mu)=V\int\frac{d^3\mathbf{k}}{(2\pi)^3}\,&\Big[\hbar\omega_{\mathbf{k}}+k_{\rm B}T\ln\left(1-e^{-(\hbar\omega_{\mathbf{k}}-\mu)/k_{\rm B}T}\right)\\&+k_{\rm B}T\ln\left(1-e^{-(\hbar\omega_{\mathbf{k}}+\mu)/k_{\rm B}T}\right)\Big].\end{aligned} \tag{11.161}$$

We have passed already to the continuum (thermodynamical) limit as in (11.100). The value of the chemical potential μ determines the mean (in the sense of the ensemble) total charge Q of the field in the box V. Convergence of the summations leading to (11.161) imposes the constraint $|\mu|<M={\rm min}_{\mathbf{k}}\omega_{\mathbf{k}}$, where M is the mass of the particles. Of course, for $\mu=0$ the system is neutral ($Q=0$ on average), and $|\mu|$ grows as the (mean) charge $|Q|$ of the system grows. Eventually, if $|Q|$ becomes so large that $|\mu|\to M$, the occupancy of the zero momentum ($\mathbf{k}=\mathbf{0}$) state becomes macroscopic (it has to be extracted before the transition to the continuous normalization) and the usual Bose-Einstein condensation occurs.

The states created by the operators $a_v^\dagger(\mathbf{k})$ related to the mode functions $v(\eta, \mathbf{k})$ and by the operators $a_u^\dagger(\mathbf{k})$ related to the mode functions $u(\eta, \mathbf{k})$ can be then interpreted as normal particles in the far past and in the far future, respectively. Using the Bogolyubov coefficients one can compute then probabilities of creating some number of particles by a variable gravitational field.

11.4 Lagrangian of the electromagnetic field

Canonical quantization of fields transforming under changes of the reference frame as nontrivial representations of the Lorentz group requires special treatment because the assumption that all generalized velocities (time derivatives of canonical variables) can be expressed as in (11.9) through the conjugated momenta is usually not fulfilled. This is, in particular, the case of the electromagnetic and Proca fields the elementary excitations of which, should, after quantization, be respectively massless and massive spin 1 particles. One method of dealing with this difficulty consists of eliminating some of the canonical variables (thereby reducing the number of independent ones and therefore also the number of the conjugated momenta). This can be achieved either directly, before quantization (as it is possible in the case of the Proca field - see Section 11.5) or in the course of quantization, through the use of the Dirac's quantization formalism adapted to system with constraints. This very important formalism, most useful when direct elimination of redundant variables is either difficult or not possible, will be presented in Section 11.6.

We first discuss possible forms of the action I (11.1) setting classical dynamics of vector fields. We begin with the familiar case of the electromagnetic field. The Maxwell equations (in the Gauss' system of units, in which the fields $\mathbf{E}$ and $\mathbf{B}$ have the same dimension - the relation of electromagnetic quantities in the Gauss system to their counterparts in the "official" SI (Système des Idiots) of units is recalled in Appendix I) read

$$\boldsymbol{\nabla}\times\mathbf{E} + \frac{1}{c}\frac{\partial \mathbf{B}}{\partial t} = \mathbf{0}\,, \qquad \boldsymbol{\nabla}\cdot\mathbf{B} = 0\,, \tag{11.241}$$

$$\boldsymbol{\nabla}\times\mathbf{B} - \frac{1}{c}\frac{\partial \mathbf{E}}{\partial t} = \frac{4\pi}{c}\,\mathbf{j}\,, \qquad \boldsymbol{\nabla}\cdot\mathbf{E} = 4\pi\rho\,. \tag{11.242}$$

The first two are automatically satisfied if the scalar and vector potentials, forming together a four-vector $A^\mu = (\varphi, \mathbf{A})$, are introduced, in terms of which

$$\mathbf{B} = \boldsymbol{\nabla}\times\mathbf{A}\,, \qquad \mathbf{E} = -\boldsymbol{\nabla}\varphi - \frac{1}{c}\frac{\partial \mathbf{A}}{\partial t}\,. \tag{11.243}$$

The remaining two equations, (11.242), follow from the Lagrangian density[65]

$$\mathcal{L}_{\text{EM}} = -\frac{1}{16\pi}\, f_{\mu\nu} f^{\mu\nu} - \frac{1}{c}\, e J_\mu A^\mu\,. \tag{11.244}$$

[65]In the context of the quantum theory it is more natural to factorize $e > 0$ - the fundamental coupling constant - out of the four-current J^ν.

The electromagnetic antisymmetric field strength tensor $f_{\mu\nu}$ is defined as

$$f_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu \,. \tag{11.245}$$

As it is easy to find,

$$\frac{\partial \mathcal{L}_{\rm EM}}{\partial(\partial_\mu A_\nu)} = -\frac{1}{8\pi}\left[\frac{\partial}{\partial(\partial_\mu A_\nu)}\, f_{\lambda\rho}\right] f^{\lambda\rho} = -\frac{1}{4\pi}\, f^{\mu\nu} \,. \tag{11.246}$$

The Euler-Lagrange equations (11.4) which in the case of a four-vector field have the general form

$$\partial_\mu \frac{\partial \mathcal{L}_{\rm EM}}{\partial(\partial_\mu A_\nu)} = \frac{\partial \mathcal{L}_{\rm EM}}{\partial A_\nu} \,, \tag{11.247}$$

therefore read

$$\partial_\mu f^{\mu\nu} = \frac{4\pi}{c}\, eJ^\nu \,. \tag{11.248}$$

Since the four-divergence of the left hand side of (11.248) vanishes by antisymmetry of the field strength tensor $f^{\mu\nu}$ (11.245), the current J^ν, to which the electromagnetic field couples, must be conserved $\partial_\mu J^\mu = 0$ (otherwise the equations (11.248) are inconsistent). It is instructive to examine the content of (11.248). Setting in (11.248) $\nu = 0$, since $f^{00} \equiv 0$ by antisymmetry, on the left hand side one gets[66]

$$\partial_i(\partial^i A^0 - \partial^0 A^i) = \partial_i(-\partial_i A^0 - \partial_0 A^i) = \boldsymbol{\nabla}\cdot\!\left(-\boldsymbol{\nabla}\varphi - \frac{1}{c}\frac{\partial \mathbf{A}}{\partial t}\right) . \tag{11.249}$$

The second Maxwell equation (11.242) is then recovered if

$$e\, J^0 = c\rho \,. \tag{11.250}$$

For $\nu = k$ (11.248) yields the equation

$$\partial_0(\partial_0 A^k + \partial_k A^0) + \partial_i f^{ik} = \frac{4\pi}{c}\, eJ^k \,. \tag{11.251}$$

Furthermore, it is easy to check that

$$f^{ij} \equiv -\partial_i A^j + \partial_j A^i = -\epsilon^{ijk} B^k \,,$$

so that (11.251) is equivalent to the equation

$$\frac{1}{c}\frac{\partial}{\partial t}\left(\partial_k \varphi + \frac{1}{c}\frac{\partial}{\partial t} A^k\right) - \epsilon^{ikl}\partial_i B^l = \frac{4\pi}{c}\, eJ^k \,, \tag{11.252}$$

[66] Since we want to keep contact with the ordinary three-dimensional notation, we convert all expressions to a form in which A^0 and A^i have always upper indices whereas the derivatives ∂_0 and ∂_i have always lower indices, so that $\partial_i = \partial/\partial \mathbf{x}^i$ is the ordinary gradient in the contravariant coordinates $\mathbf{x}^i$.

which, if the second relation (11.243) is taken into account, is just the first equation (11.242), provided one identifies $e\mathbf{J}$ with $\mathbf{j}$. Thus $eJ^\mu = (c\rho, \mathbf{j})$.

It is sometimes useful to have explicit forms of the tensors $f_{\mu\nu}$ and $f^{\mu\nu}$ in terms of the Cartesian components of the ordinary three-vectors $\mathbf{E}$ and $\mathbf{B}$:

$$f_{\mu\nu} = \begin{pmatrix} 0 & E^x & E^y & E^z \\ -E^x & 0 & -B^z & B^y \\ -E^y & B^z & 0 & -B^x \\ -E^z & -B^y & B^x & 0 \end{pmatrix}, \qquad f^{\mu\nu} = \begin{pmatrix} 0 & -E^x & -E^y & -E^z \\ E^x & 0 & -B^z & B^y \\ E^y & B^z & 0 & -B^x \\ E^z & -B^y & B^x & 0 \end{pmatrix}.$$

Having the explicit formulae it is easy to see that (11.244) can equivalently be written as

$$\mathcal{L}_{\rm EM} = \frac{1}{8\pi}\left(\mathbf{E}^2 - \mathbf{B}^2\right) - \frac{1}{c}\, eJ^\mu A_\mu\,. \tag{11.253}$$

There exist yet another term bilinear in the field A_μ, which, being a scalar with respect to proper ortochroneous Lorentz transformations,[67] could be added to the Lagrangian density (11.244): it is the term

$$\Delta\mathcal{L} \propto \tilde{f}^{\mu\nu} f_{\mu\nu}\,, \tag{11.254}$$

in which (recall that we use $\epsilon^{0123} = +1$)

$$\tilde{f}^{\mu\nu} \equiv \frac{1}{2}\,\epsilon^{\mu\nu\lambda\rho} f_{\lambda\rho} = \begin{pmatrix} 0 & -B^x & -B^y & -B^z \\ B^x & 0 & E^z & -E^y \\ B^y & -E^z & 0 & E^x \\ B^z & E^y & -E^x & 0 \end{pmatrix}. \tag{11.255}$$

$\Delta\mathcal{L}$ is, however, a total four-divergence: $\tilde{f}^{\mu\nu} f_{\mu\nu} = \partial_\mu(2\,\epsilon^{\mu\nu\lambda\rho} A_\nu \partial_\lambda A_\rho)$ and, as such, it does not modify the Euler-Lagrange (i.e. Maxwell) equations. It neither has any impact on the dynamics of the quantized electromagnetic fields, because no topologically nontrivial configurations of the electromagnetic field exist classically (this statement cannot be given a justification here). Nonabelian gauge fields (see Chapter 20) can, however, form topologically nontrivial configurations and the nonabelian analog of $\Delta\mathcal{L}$ is not innocuous in the quantum theory of such fields.[68]

The Maxwell equations (11.241), (11.242) are written in the Gauss system of units, in which the Coulomb potential looks simple (it does not have the $1/4\pi$ factor nor the ε_0 factor) and the fine structure constant (defined by the Thomson limit of the Compton scattering cross section) is $\alpha_{\rm EM} = e^2/\hbar c = 1/137.035999679(94) \approx 1/137$. This system of

[67] As a result of the parity and time reversal transformations $\Delta\mathcal{L}$ changes sign; it is therefore neither P- nor T- (and, hence, neither CP-) invariant.

[68] The reasons for complete absence or an unnaturally smal value of the (effective) coefficient of a term analogous to (11.254) in the Lagrangian of quantum chromodynamics is a still unsolved problem, called the strong CP problem - that is the problem why CP symmetry is not violated in strong interactions.

units is not very convenient in the quantum theory of the electromagnetic field because of the factor $\sqrt{4\pi}$ which would appear in many places (e.g. in the free field operators - see Section 3.8 and e.g. the textbook *Quantum Electrodynamics* by W.B. Bierestecki, J.M. Lifshitz, L.P. Pitaevski). It is much more practical to go over to the Heaviside-Lorentz system of units in which electromagnetic quantities have the same physical dimensions as in the Gauss' system but are rescaled by the appropriate factors of $\sqrt{4\pi}$:

$$A_\mu \to \sqrt{4\pi}A_\mu, \qquad eJ^\mu \to eJ^\mu/\sqrt{4\pi}\,, \tag{11.256}$$

i.e. $A_\mu^{\rm Gauss} = \sqrt{4\pi}A_\mu^{\rm Heaviside-Lorentz}$ and $e_{\rm Gauss} = e_{\rm Heaviside-Lorentz}/\sqrt{4\pi}$. The rescaling of A_μ amounts to the rescalings $\mathbf{E} \to \sqrt{4\pi}\,\mathbf{E}$ and $\mathbf{B} \to \sqrt{4\pi}\,\mathbf{B}$, while as a result of the charge rescaling, the fine structure constant $\alpha_{\rm EM}$ expressed through the charge in the Heaviside-Lorentz units reads

$$\alpha_{\rm EM} = \frac{e^2}{4\pi\hbar c}\,. \tag{11.257}$$

The current J^μ, from which the elementary charge has been factorized out, remains, of course, unaffected by the rescalings (11.256). In the rescaled variables the Lagrangian density (11.244) takes, therefore, the form

$$\mathcal{L}_{\rm EM} = -\frac{1}{4}\,f_{\mu\nu}f^{\mu\nu} - \frac{1}{c}\,eJ_\mu A^\mu\,. \tag{11.258}$$

In the Lagrangian density (11.258) the current J^μ couples to A_μ linearly. The task of quantizing the electromagnetic field coupled to a given external classical current would be therefore similar to the second example discussed in Section 11.3, were it not for the complications which will be discussed below.

If the electromagnetic field interacts with "matter" (particles or other fields) and both, the electromagnetic field and the matter, are treated as a single dynamical quantum system, the current $\mathbf{j}$ and the charge density ρ, that is J^μ, are given in terms of the dynamical variables representing matter. In such a case the current J^μ can, if the electromagnetic field is coupled e.g. to a system of (complex) scalar fields, even depend on the variable A_μ (but not on its derivatives if the canonical quantization is to be carried along the lines described below); A_μ is then coupled to "matter" nonlinearly. In all these cases quantization of the system as a whole in the Schrödinger picture, as in (11.2), should proceed essentially as if the electromagnetic field was free - the interactions do not modify the momenta conjugated to the canonical variables of the electromagnetic field itself (although can modify canonical momenta of the "matter"). The canonical quantization of systems involving the electromagnetic field encounters, however, immediately an obstacle. If all four components A^μ are taken for independent dynamical variables, one finds that the momentum canonically conjugated to A^0 vanishes identically:

$$\Pi_\nu \equiv \frac{\partial\mathcal{L}_{\rm EM}}{\partial\dot{A}^\nu} = \frac{\partial\mathcal{L}_{\rm EM}}{\partial(c\,\partial^0A^\nu)} = -\frac{1}{c}\,f_{0\nu}\,, \tag{11.259}$$

which implies that $\Pi_0 \propto f_{00} \equiv 0$. Systems coupled to the electromagnetic field are therefore necessarily systems subject to constraints.

In fact, the peculiarity of the electromagnetic field is twofold: one peculiarity is $\Pi_0 = 0$. The other one is *gauge invariance* - the symmetry of the Lagrangian density (11.258) under (local) transformations[69]

$$A_\mu(x) \to A'_\mu(x) = A_\mu(x) + \frac{1}{e}\,\partial_\mu\theta(x)\,, \tag{11.260}$$

in which θ is an arbitrary differentiable function of the space-time coordinates x^μ. Invariance of the $f^{\mu\nu}f_{\mu\nu}$ term of (11.258) is obvious as under the transformations (11.260) the tensor $f_{\mu\nu}$ does not change. Assuming that J_μ itself does not change when the transformation (11.260) is made (i.e. that it is gauge invariant), the $eJ_\mu A^\mu$ term of (11.258) transforms into $eJ^\mu A_\mu + J^\mu\partial_\mu\theta$; the term $J^\mu\partial_\mu\theta = \theta\partial_\mu J^\mu - \partial_\mu(\theta J^\mu)$ does not contribute to the action $I = \int d^4x\,\mathcal{L}_{\rm EM}$ if the current is conserved, which is anyway the necessary condition for consistency of the field equations (11.248).

Invariance of the current J_μ with respect to the transformations (11.260) is obvious and its conservation, $\partial_\mu J^\mu = 0$, is easy to check, if it depends, as in the case of nonrelativistic charged particles coupled to the electromagnetic field, discussed in Section 11.7, only on dynamical variables of the "matter" (particles or fields, as opposed to the "radiation" i.e. the electromagnetic field). In more complicated cases, for example in the one of scalar fields representing "matter" coupled to the electromagnetic field, the basic assumption is the gauge invariance of the complete action

$$I[A_\mu, \text{"matter"}] \equiv -\frac{1}{4}\int d^4x\, f_{\mu\nu}f^{\mu\nu} + I_{\rm rest}[A_\mu, \text{"matter"}]\,, \tag{11.261}$$

by which one understands its invariance with respect to the transformation (11.260) of the electromagnetic field potentials supplemented with appropriate, dependent on $\theta(x)$, that is *local*, transformations of other canonical variables (of the "matter"). For θ independent of x, these transformations must form a $U(1)$ group of *global* symmetries (which were discussed in Section 11.1) of the action $I_{\rm rest}$. For x-dependent θ this group becomes the *local* $U(1)$ symmetry group of the complete action $I[A_\mu, \text{"matter"}]$. Gauge invariance of the complete classical action (11.261) may be viewed as a *fundamental principle*, which must always be respected. This principle, generalized to nonabelian symmetry groups (see Chapter 20), constitutes in fact the cornerstone of the whole modern development of quantum field theory.[70] Within this generalized formulation of (classical) theories in

[69]The same two problems: vanishing of time components of the canonical momenta and gauge invariance with respect to transformations slightly more complicated than (11.260) are also characteristic of systems coupled to nonabelian gauge (Yang - Mills) fields discussed in Chapter 20.

[70]It will be seen, however, (Section 20.3) that to properly formulate quantum theories of gauge fields corresponding to nonabelian symmetry groups it is advantageous to take a more general approach and to start with an extended action $I = I_{\rm gauge\ inv} + I_{\rm non-gauge\ inv}$, which as a whole is not gauge invariant but possesses a continuous *global* symmetry of the supersymmetric type, called the BRST symmetry (Section

which the electromagnetic field couples to matter or other fields, gauge invariance ensures that the generalized electromagnetic current J^μ, which enters the right hand side of the Euler-Lagrange equation of the electromagnetic field

$$\partial_\mu f^{\mu\nu} = \frac{e}{c}\, J^\nu\,,$$

which is, therefore, defined as[71]

$$-eJ_\mu(x) \equiv \frac{\delta}{\delta A^\mu(x)}\, I_{\rm rest}[A_\mu, \text{"matter"}]\,, \tag{11.262}$$

even if it depends on A_μ, it is gauge invariant and conserved. To see this let us consider $I_{\rm rest}$ of the general form

$$I_{\rm rest}[A_\mu, \phi_i] = \int d^4x\, \mathcal{L}_{\rm rest}(A_\mu, \phi_i, \partial\phi_i)\,, \tag{11.263}$$

depending on A_μ and a set of fields ϕ_i, which under the changes (11.260) of the gauge transform according to the rule

$$\phi_i(x) \to \phi_i'(x) = \left(e^{-i\theta(x)T}\right)_{ij} \phi_j(x) \approx \phi_i(x) - i\theta(x) T_{ij}\phi_j(x)\,, \tag{11.264}$$

in which T is the matrix of the $U(1)$ group generator represented on the real fields ϕ_i. The assumed gauge invariance of $I_{\rm rest}$ means that $I_{\rm rest}[A_\mu', \phi_i'] - I_{\rm rest}[A_\mu, \phi_i] = 0$. Applied to an infinitesimal gauge transformations (11.260) and (11.264) this gives the identity

$$\begin{aligned}&\int d^4x\, \frac{1}{e}\, \frac{\delta I_{\rm rest}}{\delta A_\mu(x)}\, \partial_\mu\theta(x)\\ &-i\int d^4x \left\{\frac{\partial\mathcal{L}_{\rm rest}}{\partial\phi_i(x)}\, \theta(x)T_{ij}\phi_j(x) + \frac{\partial\mathcal{L}_{\rm rest}}{\partial(\partial_\mu\phi_i(x))}\, \partial_\mu\left[\theta(x)T_{ij}\phi_j\right]\right\} = 0\,.\end{aligned}$$

The last line is zero, as can be seen by integrating by parts and using the Euler-Lagrange equations of motion[72] (11.4) of the fields ϕ_i. In view of the arbitrariness of $\theta(x)$ one gets (integrating by parts also in the first term) therefore the condition of the conservation[73] of the current (11.262). Gauge invariance of the current (11.262) is of course guaranteed by the gauge invariance of the action (11.263). But it is precisely the gauge invariance of the complete action $I[A_\mu, \text{"matter"}]$ of the electromagnetic field coupled to "matter fields", that is responsible for additional difficulties which one will have to overcome in quantizing the electromagnetic field in Section 11.7.

20.2). Quantization with the help of the Dirac method of Section 11.6 is then fairly easy and the BRST invariance of the quantum theory ensures that all the implications of gauge invariance are recovered in the properly defined physical subspace of the extended Hilbert space. Moreover, in some special cases, e.g. of gauge theories coupled to chiral femions, or when an explicit momentum cutoff is implemented in the Lagrangian, even the BRST invariance of the classical action must be abadoned if the resulting quantum gauge theory is to be consistent.

[71] In most cases this current coincides with the Noether current of the global $U(1)$ symmetry.

[72] The reasoning remains valid also if the action $I_{\rm rest}$ depends on higher derivatives of the fields ϕ_i, provided one uses the appropriate generalization of the Euler-Lagrange equations of motion.

[73] In the case of nonabelian gauge fields infinitesimal transformations of which analogous to (11.260) involve the covariant derivative instead of the ordinary one, this reasoning leads to the covariant conservation of the current to which the gauge fields couple.

11.5 The Proca field and its quantization

Before dealing with the difficulties of the electromagnetic field, it is instructive to quantize first a simpler system with constraints, the so-called Proca field[74] V^μ (a vector field with a nonzero mass) the classical dynamics of which is governed by the Lagrangian density

$$\mathcal{L}_{\rm Proca} = -\frac{1}{4}\,V_{\mu\nu}V^{\mu\nu} + \frac{1}{2}M^2 V_\mu V^\mu - g\,V_\mu J^\mu\,, \tag{11.265}$$

in which $V_{\mu\nu} \equiv \partial_\mu V_\nu - \partial_\nu V_\mu$ and g is the coupling constant (we have set $\hbar = c = 1$). In the following we assume that the current J^μ is independent of V_μ. Conservation of the current J^μ is not assumed here (it is not necessary for consistency of the field equations).

The transformation (11.260) is not a symmetry of the the Lagrangian (11.265) due to the mass term $\frac{1}{2}M^2V_\mu V^\mu$ by which it differs from (11.258). Still (11.265) leads to the similar (at first sight) obstacle in the canonical quantization as (11.258): $\Pi_0 \equiv 0$.

Before disscussing this problem we comment on the form of the Lagrangian density (11.265). Since it is not invariant under the gauge transformations (11.260), there is a priori no reason for which its derivative part should take the restricted form (11.265). In principle, one could write down the most general Lagrangian density, at most quadratic in V_μ and with at most two derivatives in the form

$$\mathcal{L}_{\rm Proca} = -\frac{1}{2}\,\partial_\mu V_\nu \partial^\mu V^\nu - \frac{1}{2}\,\kappa\,\partial_\mu V_\nu \partial^\nu V^\mu + \frac{1}{2}M^2 V_\mu V^\mu - g\,V_\mu J^\mu\,, \tag{11.266}$$

with an arbitrary real κ (any other Lagrangian density satisfying our requirements differ from this one by a total derivative; an arbitrary real negative constant, which could multiply the first term, can always be made equal to $-1/2$ by appropriately rescalings of V_μ, κ, M^2 and $g\,J_\mu$). However, taking the four-divergence of both sides of the Euler-Lagrange equation

$$\partial_\mu\partial^\mu V^\nu + \kappa\,\partial^\nu(\partial_\mu V^\mu) + M^2 V^\nu = g\,J^\nu\,, \tag{11.267}$$

resulting from (11.266), we discover that the four-divergence $\partial_\nu V^\nu \equiv \phi$ satisfies the independent equation

$$(1+\kappa)\partial_\mu\partial^\mu\phi + M^2\phi = g\,\tilde{J}\,, \tag{11.268}$$

in which $\tilde{J} \equiv \partial_\mu J^\mu$. Therefore, the Lagrangian density (11.266) gives rise, among other things, also to the independent propagation of a scalar field[75] which couples to the scalar current $\tilde{J}$. To remove the propagation of this scalar field ϕ (that is, to remove the

[74] We denote this field V_μ, instead of A_μ, to distinguish it from the electromagnetic field.

[75] That an unconstrained vector field V_μ contains in it in addition a massive spin zero particle was also revealed in Section 8.4 in the course of constructing the free-field operator transforming under changes of the reference frame as a vector representation of the Lorentz group.

homogeneous part of the classical solution for ϕ, which would remain even if the current J^μ was conserved, leading to $\tilde{J} \equiv 0$), we set $\kappa = -1$; taking then the four-divergence of both sides of the Euler-Lagrange equation (11.267) we find the relation

$$M^2 \partial_\mu V^\mu = g\, \partial_\mu J^\mu \,, \tag{11.269}$$

which shows that now ϕ can be expressed algebraically in terms of the external current (which depends on dynamical variables of other parts of the system). In particular, now $\phi \equiv 0$ if $\partial_\mu J^\mu = 0$. With $\kappa = -1$ the Lagrangian density (11.266) is just the Proca Lagrangian density (11.265).

We now quantize the Proca field. From its Lagrangian density (11.265) one obtains the canonical momenta Π_ν conjugate to the variables V^μ:

$$\Pi_i = \frac{\partial \mathcal{L}_{\rm Proca}}{\partial \dot{V}^i} = -V_{0i} = \dot{V}^i + \partial_i V^0 \,, \qquad \Pi_0 = 0 \,. \tag{11.270}$$

As has been said, $\Pi_0 \equiv 0$ is an obstacle in the canonical quantization. The problem lies in the fact that the naive Hamilton's formalism constructed ignoring the difficulty would not lead to equations equivalent to the original Euler-Lagrange equations

$$\partial_\mu V^{\mu\nu} = -M^2 V^\nu + g\, J^\nu \,, \tag{11.271}$$

following from the Proca Lagrangian density (11.265), that is to the equations

$$\partial_i V^{i0} = -M^2 V^0 + g\, J^0 \,, \tag{11.272}$$

$$\partial_0(\partial_0 V^k + \partial_k V^0) + \partial_i(-\partial_i V^k + \partial_k V^i) = -M^2 V^k + g\, J^k \,, \tag{11.273}$$

obtained by setting in (11.271) $\nu = 0$ and $\nu = k$, respectively. Trying to construct the Hamiltonian corresponding to the Proca Lagrangian and following the standard recipe one would write:

$$\begin{aligned}\mathcal{H} &= \Pi_0 \dot{V}^0 + \Pi_i \dot{V}^i - \mathcal{L}_{\rm Proca} \\ &= \Pi_0 \dot{V}^0 + \Pi_i \dot{V}^i - \frac{1}{2}(\dot{V}^i + \partial_i V^0)(\dot{V}^i + \partial_i V^0) + \frac{1}{2}\partial_i V^j(\partial_i V^j - \partial_j V^i) \\ &\qquad + \frac{1}{2}M^2 V^i V^i - \frac{1}{2}M^2 V^0 V^0 - g\, V^i J^i + g\, V^0 J^0 \,. \end{aligned} \tag{11.274}$$

Eliminating $\dot{V}^i$ by using (11.270), i.e. substituting $\dot{V}^i = \Pi_i - \partial_i V^0$, one would then get

$$\begin{aligned}\mathcal{H} &= \frac{1}{2}\Pi_i \Pi_i + \Pi_0 \dot{V}^0(V^\lambda, \Pi_\lambda) - \Pi_i \partial_i V^0 + \frac{1}{2}(\boldsymbol{\nabla} \times \mathbf{V})^2 \\ &\qquad + \frac{1}{2}M^2 V^i V^i - \frac{1}{2}M^2 V^0 V^0 - g\, V^i J^i + g\, V^0 J^0 \,. \end{aligned} \tag{11.275}$$

Of course, since $\Pi_0 \equiv 0$, there is no way to express the generalized velocity $\dot{V}^0$ through the canonical variables V^λ and Π_λ. Let us therefore see, what one gets, setting naively $\Pi_0 = 0$

in the Hamiltonian (11.275). The canonical equations following from the Hamiltonian $H = \int d^3\mathbf{x}\,\mathcal{H}$ obtained in this way would read

$$\dot{V}^i = \left\{V^i,\ H\right\}_{\rm PB} = \frac{\partial\mathcal{H}}{\partial\Pi_i}\,, \qquad \dot{\Pi}_i = \{\Pi_i,\ H\}_{\rm PB} = -\frac{\partial\mathcal{H}}{\partial V^i}\,, \tag{11.276}$$
$$\dot{V}^0 = \left\{V^0,\ H\right\}_{\rm PB} = \frac{\partial\mathcal{H}}{\partial\Pi_0} \equiv 0\,, \qquad \dot{\Pi}_0 = \{\Pi_0,\ H\}_{\rm PB} = -\frac{\partial\mathcal{H}}{\partial V^0} \neq 0\,,$$

and would evidently be incompatible with the Euler-Lagrange equation (11.271): the latter imply firstly that $\Pi_0 = 0$ for any time t, whereas here, even if one sets $\Pi_0 = 0$ for $t = 0$, a nonzero Π_0 would be generated for $t \neq 0$ due to a nonzero derivative $\dot{\Pi}_0$. Secondly, from (11.272) expressed in terms of Π_i it follows that the time evolution of V^0 is fully determined in terms of J^0 and $\partial_i\Pi_i$:

$$V^0 = \frac{1}{M^2}\left(\partial_i\Pi_i + g\,J^0\right). \tag{11.277}$$

In contrast, the third canonical equation (11.276) would imply that V^0 is constant. Thus, the dynamics of V^μ following from the Hamiltonian (11.275) would not be the same as the one generated by the Euler-Lagrange equations (11.272) and (11.273).

The equivalence is restored if in the Hamiltonian (11.275), in addition to setting $\Pi_0 = 0$, one substitutes for V^0 the expression (11.277), that is, if one "algebraically" eliminates one of the canonical variables (in other words, if one expresses its value at any t by a combination of other canonical variables taken at the same instant t). One then obtains (after integrating by parts) the "physical" Hamiltonian:

$$\mathcal{H}^{\rm ph} = \frac{1}{2}\Pi_i\Pi_i + \frac{1}{2M^2}\left(\partial_i\Pi_i\right)^2 + \frac{1}{2}(\boldsymbol{\nabla}\times\mathbf{V})^2$$
$$+\frac{1}{2}M^2V^iV^i - g\,V^iJ^i + \frac{g^2}{2M^2}J^0J^0 + \frac{g}{M^2}J^0\partial_i\Pi_i\,. \tag{11.278}$$

As the canonical equations one now gets

$$\dot{V}^i = \frac{\partial\mathcal{H}^{\rm ph}}{\partial\Pi_i} = \Pi_i - \frac{1}{M^2}\,\partial_i\left(\partial_j\Pi_j\right) - \frac{g}{M^2}\,\partial_iJ^0\ , \tag{11.279}$$
$$\dot{\Pi}_i = -\frac{\partial\mathcal{H}^{\rm ph}}{\partial V^i} = -\epsilon^{ijk}\epsilon^{klm}\partial_j\partial_lV^m - M^2V^i + g\,J^i\,. \tag{11.280}$$

Their equivalence with the Euler-Lagrange (11.271) equations can be established as follows: If we *define* an auxiliary variable V^0 as in (11.277); the equation (11.272) is then automatically satisfied. Next, the equation (11.279) can be written in the form $\Pi_i = \dot{V}^i + \partial_iV^0$. Differentiating this relation with respect to time and inserting $\dot{\Pi}_i$ obtained in this way into (11.280) transforms the latter equation into (11.273). Thus, by eliminating "algebraically"one of the variables, we have constructed the Hamilton's formalism which is equivalent to the field equations following from the original Lagrangian (11.265).

Quantization is now straightforward: one promotes $V^i(\mathbf{x},t)$ and $\Pi_i(\mathbf{x},t)$ taken at $t=0$ to operators in the Schrödinger picture and imposes the standard canonical commutation relations

$$\begin{aligned}[\hat{V}^i(\mathbf{x}),\ \hat{\Pi}_j(\mathbf{y})] = i\left\{V^i(\mathbf{x}),\ \Pi_j(\mathbf{y})\right\}_{\rm PB} = i\delta^i_j\delta(\mathbf{x}-\mathbf{y})\,,\\ [\hat{V}^i(\mathbf{x}),\ \hat{V}^j(\mathbf{y})] = [\hat{\Pi}_i(\mathbf{x}),\ \hat{\Pi}_j(\mathbf{y})] = 0\,.\end{aligned}\tag{11.281}$$

on the operators representing the independent canonical variables. In this way canonically quantized get only independent physical degrees of freedom of the system of fields. If the current J^μ depends on field variables ϕ_a and Π_a other than V^i and Π_i themselves, the states $|\Psi\rangle$ of the system can be represented as wave functionals $\Psi[V^i(\mathbf{x}),\phi_a(\mathbf{x}),t] = \langle\,[V^i],[\phi_a]|\Psi(t)\rangle$ on which the operators $\hat{V}^i(\mathbf{x})$ and $\hat{\Pi}_i(\mathbf{x})$ act as

$$\hat{V}^i(\mathbf{x}) = V^i(\mathbf{x})\,,\qquad \hat{\Pi}_i(\mathbf{x}) = -i\,\frac{\delta}{\delta V^i(\mathbf{x})}\,.\tag{11.282}$$

Of course, the Hilbert space of all functionals $\Psi[V^i(\mathbf{x}),\phi_a(\mathbf{x})]$ is nonseparable and one seeks an appropriate separable Fock space in which the algebra of operators can be represented irreducibly. As usually one choses the Fock space in which the free Hamiltonian has its lowest energy eigenvector.

In order to realize the algebra (11.281) of field operators in this separable Fock space it is necessary to find a representation of the operators[76] $V^i(\mathbf{x})$ and $\Pi_i(\mathbf{x})$ in terms of the creation and annihilation operators (having the standard commutation rules) To simplify this task we can apply the method outlined in Section 11.3 allowing to easily quantize the free field (i.e. for $J^\mu=0$) in the Heisenberg picture; the free-field Heisenberg picture operators $V^i_H(t,\mathbf{x})$, $\Pi^H_i(t,\mathbf{x})$ taken at $t=0$ will then provide the sought representations of the Schrödinger picture operators of the interacting (i.e. coupled to $J^\mu\neq 0$) Proca field.

To this end first write down the the most general solution to the classical field equations (11.267) and (11.269) following from the free (with $J^\mu=0$) Proca Lagrangian which read

$$(\partial_\nu\partial^\nu + M^2)V^\mu(t,\mathbf{x}) = 0\,,\qquad \partial_\mu V^\mu(t,\mathbf{x}) = 0\,.\tag{11.283}$$

The most general classical solution of the first equation can be written in the form (11.164) but with the coefficients $a(\mathbf{k})$ and $a^*(\mathbf{k})$ replaced by some four-vector coefficients $a_\mu(\mathbf{k})$ and $a^*_\mu(\mathbf{k})$. In order to satisfy the second of these equations we write them in the form

$$a^\mu(\mathbf{k}) = \sum_{\lambda=0,\pm1} a(\mathbf{k},\lambda)\,\epsilon^\mu(\mathbf{k},\lambda)\,,\tag{11.284}$$

with three four-vectors $\epsilon^\mu(\mathbf{k},\lambda)$ satisfying the condition

$$k_\mu\epsilon^\mu(\mathbf{k},\lambda) = 0\,.\tag{11.285}$$

[76] We suppress hats on operators from now on.

Since $k^2 = M^2 \neq 0$, the are three linearly independent such four-vectors. Two of them, corresponding to $\lambda = \pm 1$, can be chosen to have only space-like components (in the plane perpendicular to $\mathbf{k}$) and will be normalized so that

$$\epsilon^\mu(\mathbf{k}, \lambda)\, \epsilon^*_\mu(\mathbf{k}, \lambda) = -1\,. \tag{11.286}$$

(i.e. $\boldsymbol{\epsilon}(\mathbf{k}, \lambda) \cdot \boldsymbol{\epsilon}^*(\mathbf{k}, \lambda) = 1$). As the third one, corresponding to $\lambda = 0$, we take the four-vector

$$\epsilon^\mu(\mathbf{k}, \lambda = 0) = \left(\frac{|\mathbf{k}|}{M}\,,\ \frac{\mathbf{k}}{|\mathbf{k}|}\frac{E(\mathbf{k})}{M}\right), \tag{11.287}$$

which also satisfies the conditions (11.285) and (11.286). The vectors $\epsilon^\mu(\mathbf{k}, \lambda)$ satisfy then the sum rule

$$\sum_{\lambda=0,\pm1} \epsilon_\mu(\mathbf{k}, \lambda)\, \epsilon^*_\nu(\mathbf{k}, \lambda) = -g_{\mu\nu} + \frac{k_\mu k_\nu}{M^2}\,. \tag{11.288}$$

Thus, when $J^\mu = 0$, the most general solution of the Proca field equations of motion has the form

$$V^\mu(t, \mathbf{x}) = \int d\Gamma_{\mathbf{k}} \sum_{\lambda=0,\pm1} \left[a(\mathbf{k}, \lambda)\, \epsilon^\mu(\mathbf{k}, \lambda)\, e^{-iEt+i\mathbf{k}\cdot\mathbf{x}} + \text{h.c.}\right], \tag{11.289}$$

where, as usually, $d\Gamma_{\mathbf{k}} = d^3\mathbf{k}/(2\pi)^3 2E$ and $E = \sqrt{\mathbf{k}^2 + M^2}$. Since there are only three independent vectors $\epsilon^\mu(\mathbf{k}, \lambda)$, out of the four components of V^μ only three, for which one can take V^i's, are independent. To explicitly write down their canonical momenta $\Pi_i(t, \mathbf{x})$ which are the corresponding solutions to the classical canonical equations, one recalls that in the classical Proca theory $\Pi_i(t, \mathbf{x}) = \dot{V}^i(t, \mathbf{x}) + \partial_i V^0(t, \mathbf{x})$. In agreement with this relation one postulates that

$$\Pi_i(t, \mathbf{x}) = \frac{1}{i} \int d\Gamma_{\mathbf{k}}\, E(\mathbf{k}) \sum_{\lambda=0,\pm1} \left[a(\mathbf{k}, \lambda)\, \tilde{\epsilon}^i(\mathbf{k}, \lambda)\, e^{-iEt+i\mathbf{k}\cdot\mathbf{x}} - \text{h.c.}\right], \tag{11.290}$$

where

$$\tilde{\epsilon}^i(\mathbf{k}, \lambda) \equiv \epsilon^i(\mathbf{k}, \lambda) - \frac{k^i}{E}\, \epsilon^0(\mathbf{k}, \lambda)\,. \tag{11.291}$$

Obviously, $\tilde{\epsilon}^i(\mathbf{k}, \lambda) = \epsilon^i(\mathbf{k}, \lambda)$ for $\lambda = \pm 1$, while for $\lambda = 0$ one finds $\tilde{\epsilon}^i(\mathbf{k}, \lambda = 0) = (k^i/|\mathbf{k}|)(M/E)$. It is then easy to verify that $V^0(t, \mathbf{x})$ as given by (11.289) is correctly reproduced by (11.277) with $J^\mu = 0$.

One can then check, that the time dependent Heisenberg picture operators $V^i_H(t, \mathbf{x})$ and $\Pi^H_i(t, \mathbf{x})$ of the free Proca theory are obtained by simply promoting the coefficients $a(\mathbf{k}, \lambda)$ and $a^\lambda(\mathbf{k}, \lambda)$ to operators satisfying the

$$\begin{aligned} \left[a(\mathbf{k}, \lambda),\ a^\dagger(\mathbf{k}', \lambda')\right] &= (2\pi)^3 2E(\mathbf{k}) \delta_{\lambda\lambda'} \delta^{(3)}(\mathbf{k} - \mathbf{k}')\,, \\ \left[a(\mathbf{k}, \lambda),\ a(\mathbf{k}', \lambda')\right] &= \left[a^\dagger(\mathbf{k}, \lambda),\ a^\dagger(\mathbf{k}', \lambda')\right] = 0\,, \end{aligned} \tag{11.292}$$

because the canonical equal-time commutation rules (11.281) and all necessary operator equations are then satisfied.

It follows that for the Schrödinger picture operators $V^i(\mathbf{x})$ and $\Pi_i(\mathbf{x})$, which at $t=0$ will be equal to the Heisenberg picture operators of the interacting Proca field one can take (11.289) and (11.290) with t set to zero and the coefficients a and a^* replaced by the operators $a(\mathbf{k},\lambda)$ and $a^\dagger(\mathbf{k},\lambda)$. (Of course, at $t\neq 0$ the Heisenberg picture operators $V^i_H(t,\mathbf{x})$ and $\Pi_H(t,\mathbf{x})$ of the Proca field coupled to a nonvanishing current J^μ operator will, unlike the Heisenberg picture operators of the free Proca field constructed above, not have the simple forms (11.289) and (11.290) with a and a^* replaced by the creation and annihilation operators!)

In terms of the creation and annihilation operators introduced in this way, the free part H_0 of the Hamiltonian (11.278) takes the form

$$H_0=\int d\Gamma_{\mathbf{k}}\,E(\mathbf{k})\sum_{\lambda=0,\pm1}a^\dagger(\mathbf{k},\lambda)\,a(\mathbf{k},\lambda)\,, \tag{11.293}$$

plus an infinite constant which we discard. Thus, the free part H_0 of the Hamiltonian H (that is, its part independent of the current J^μ) becomes diagonal. It is also easy to see that the operators obtained by substituting into the classical solutions (11.289) and (11.290) (for $t\neq 0$) the operators $a(\mathbf{k},\lambda)$ and $a^\dagger(\mathbf{k},\lambda)$, will in the theory of the interacting Proca field be the interaction picture operators $V^i_I(x)$ and $\Pi^I_i(x)$ (see Section 11.9), because they are related to the Schrödinger picture ones by the standard rules

$$V^i_I(t,\mathbf{x})=e^{iH_0t}\,V^i(\mathbf{x})\,e^{-iH_0t}\,,\qquad \Pi^I_i(t,\mathbf{x})=e^{iH_0t}\,\Pi_i(\mathbf{x})\,e^{-iH_0t}\,.$$

Finally, one can compute the commutators of $V^0(\mathbf{x})$ which, if the Proca field V^μ is coupled to some "matter", is given by (11.277). Assuming that J^0 depends on the canonical variables of systems coupled to the Proca field but not on the canonical variables V^i or Π_i of the vector field itself, one finds

$$\begin{aligned}\left[V^i(\mathbf{x}),\ V^0(\mathbf{y})\right]=\frac{1}{M^2}\left[V^i(\mathbf{x}),\ \partial^{\mathbf{y}}_j\Pi_j(\mathbf{y})\right]=-\frac{i}{M^2}\,\partial^{\mathbf{x}}_i\delta^{(3)}(\mathbf{x}-\mathbf{y})\,,\\ \left[\Pi_i(\mathbf{x}),\ V^0(\mathbf{y})\right]=0\,.\end{aligned} \tag{11.294}$$

Thus, the operator representing the time component of V^μ does not commute with the operators representing the remaining components, contrary to what could naively be expected.

11.6 Systems subject to constraints

In Section 11.5 the Proca vector field, which is an example of a system subject to constraints which in the Dirac terminology are second class constraints, has been quantized by expressing "algebraically"one of its components (V^0) in terms of the remaining independent canonical variables (Π_i and the variables out of which J^0 is built). The form of the canonical commutation relations imposed on the independent canonical variables (that is, the right quantization prescription) was then obvious and they determined also the commutation relations satisfied by the operator V^0. Explicit solutions for dependent variables (like V^0) in terms of the remaining variables, chosen as independent, even if possible in principle, may not always be easy, especially in systems composed of several mutually interacting subsystems (in field theories expressing "algebraically" may mean solving differential equations in the space variable $\mathbf{x}$ - the term "algebraic" refers only to the fact that no time derivatives are involved). Moreover, simple "algebraic" elimination of dependent variables is not directly applicable to the electromagnetic field which is also an example of a system subject to constraints. Therefore in this section we present a systematic and general method proposed by Dirac, which allows to "hamiltonize" systems subject to constraints and to quantize them canonically. In the case of systems subject to constraints, called second class, similar to the ones encountered in the Proca theory of the vector field, the Dirac method allows to find, without solving explicitly for dependent variables, the right commutation relations which must be imposed directly on the original set of canonical variables along with a set of identities which must be satisfied by operators representing these variables, in order to properly quantize the system.[77] The same method proves useful also in quantizing systems subject, like the electromagnetic field, to constraints of the first class. Such constraints reflect invariance of the (classical) physical state of the system with respect to changes of "gauge" - i.e. with respect to appropriate transformations of the canonical variables (which are therefore not uniquely specified by the physical state of the system). Such systems can be quantized either by fixing the gauge, that is by introducing additional constraints which convert them into systems subject to second class constraints only, or by treating the first class constraints as conditions which select in the Hilbert space (or in the selected Fock space) a subset of vectors (rays) which represent physical states of the system. The Dirac method allows, among other things, for an easy canonical quantization of theories of non-Abelian Yang-Mills fields which in the BRST formulation become systems subject to second class constraints (see Section 20.3).

Canonical quantization consists of the identification of the canonical variables q^i, $i = 1, \dots, n$ and the conjugated momenta p_i, their subsequent promotion to the Schrödinger picture operators $\hat{q}^i$ and $\hat{p}_i$ satisfying the commutation rule

$$\left[\hat{q}^i,\ \hat{p}_j\right] = i\hbar\left\{q^i,\ p_j\right\}_{\mathrm{PB}} = i\hbar\,\delta^i{}_j\,, \tag{11.295}$$

and, finally, realization of the resulting algebra of operators in some Hilbert (or Fock) space. The method in this simple form is applicable if the canonical variables and their

[77]The method itself, however, does not give any clues, how to satisfy these commutation relations and identities, that is how to obtain a representation of the resulting algebra of operators in the Hilbert or a Fock space - in the latter case by expanding field operators taken at $t = 0$ into creation and annihilation operators.

momenta are all independent. Classical systems may, however, be subject to constraints as a result of which not all their canonical variables are independent. The *primary* constraints $\Phi_M(q,p)=0$, $M=1,\dots,n-r$, follow from the structure of the Lagrangian and reflect the impossibility to solve the equations[78]

$$p_j = \frac{\partial}{\partial \dot{q}^j}\, L(q,\dot{q})\,, \tag{11.296}$$

for $n-r$ velocities $\dot{q}^j$, $j=r+1,\dots,n$. In such a case the original Euler-Lagrange equations of motion are equivalent (see Appendix J) to the set of canonical equations which follow from the *total Hamiltonian* $H_{\rm T}$ (Dirac's terminology) of the form

$$\begin{aligned} H_{\rm T}(q,p) &= \left(\sum_{i=1}^{r} \frac{\partial L}{\partial \dot{q}^i}\,\dot{q}^i - L(q,\dot{q})\right)_{\dot{q}^i=\dot{q}^i(q,p,u)} + \sum_{M=1}^{n-r} \Phi_M\, u^M \\ &= H(q,p) + \sum_{M=1}^{n-r} \Phi_M(q,p)\, u^M\,, \end{aligned} \tag{11.297}$$

combined with the set of (primary) constraints $\Phi_M(q,p)=0$. The quantitities u^M with $M=1,\dots,n-r$, are the $n-r$ generalized velocities $\dot{q}^{r+M}$ with respect to which the equations (11.296) cannot be solved and $\pi_M\equiv p_{r+M}$ are the canonical momenta associated with the variables q^{r+M}. The subscript $\dot{q}^i=\dot{q}^i(q,p,u)$ on the bracket in the first term of (11.297) means that r of the generalized velocities $\dot{q}^i$ (those with respect to which the relations (11.296) can be solved) are expressed in terms of the canonical variables q, the conjugate momenta p_i, $i=1,\dots,r$ and the remaining velocities $u^M=\dot{q}^{r+M}$. In Appendix J it is shown that the Hamiltonian $H(q,p)$ in the second line of (11.297), is independent of the $n-r$ velocities u^M. The sum in the second term runs over the set of $n-r$ primary constraint functions $\Phi_M(q,p)$ - the equality $\Phi_M=0$ is not used at this stage.[79] By construction (see Appendix J) the primary constraints $\Phi_M(q,p)=0$, $M=1,\dots,n-r$ following directly from the the Lagrangian have the structure $\pi_M - f_M(q,p_1,\dots,p_r)=0$. The mentioned equivalence is to be understood in the sense that the constraints $\Phi_M(q,p)=0$ differentiated with respect to time, combined with the evolution equations $\dot{q}^i=\{q^i,H_{\rm T}\}_{\rm PB}$ and $\dot{p}_i=\{p_i,H_{\rm T}\}_{\rm PB}$, in which the Poisson brackets with the total Hamiltonian $H_{\rm T}$ (11.297) are computed as if all canonical variables q^i and p_i were independent (i.e. as if there were no constraints at all), yield second order differential equations for q^i's which are equivalent to the original Euler-Lagrange equations. This

[78]As aready said, in the case of gauge systems primary constraints can be also imposed on the system's variables from the outside.

[79]The factors $u^M(t)$ in $H_{\rm T}$ can also be interpreted as Lagrange multipliers allowing to take into account the constraints $\Phi_M(q,p)=0$ in the standard variational formulation

$$\delta\int_{t_1}^{t_2} dt\,[p_i\dot{q}^i - H(q,p)] = 0\,, \qquad \delta q^i(t_1)=\delta q^i(t_2)=0\,,$$

of the Hamilton's equations of motion.

classical dynamics can be cast into a fully Hamiltonian form by giving the factors u^M an appropriate dependence on p_i and q^i. The prescription for achieving this, formulated by Dirac, is as follows.

As in the Lagrangian formalism the primary constraints $\Phi_M(q,p) = 0$ are identities relating the canonical variables at any instant t, one should ensure that they are also preserved by the dynamics generated by the Hamiltonian (11.297) - in this case it will suffice to impose them on the initial data only. This means that $\dot{\Phi}_N(q,p)$, that is the Poisson brackets $\{\Phi_N,\ H_{\rm T}\}_{\rm PB}$, must be made vanishing. Investigating these Poisson brackets one can encounter different situations. Barring the case of obvious contradictions, like $1 = 0$ (which would mean that the Lagrangian itself leads to inconsistent Euler-Lagrange equations), the first possibility is that the Poisson bracket $\{\Phi_N,\ H_{\rm T}\}_{\rm PB}$ vanishes when the already identified constraints are imposed *after* computing it. This is written as

$$\{\Phi_N,\ H_{\rm T}\}_{\rm PB} \simeq 0\,,$$

where the symbol $\simeq 0$ means "vanishes weakly" (Dirac's terminology again). Also in this case the Poisson bracket should be computed as if the variables were unconstrained and independent; the constraints $\Phi_M = 0$ are imposed only afterwards. Thus, in general, the symbol $\simeq 0$ is equivalent to the equality to a combination $\alpha^M(q,p)\Phi_M(q,p)$ of the constraints. The second possibility is that the investigated Poisson bracket of the constraint Φ_N with the Hamiltonian (11.297) does not vanish, even weakly, and the result does not depend on the (yet) unknown coefficients u^N. This means that the new constraint $\Phi_K(q,p) \equiv \{\Phi_N,\ H_{\rm T}\}_{\rm PB} = 0$ has to be added to the list of constraints. Constraints identified in this way are called *secondary.* The obvious next step is to investigate the Poisson bracket of the newly identified constraint Φ_K with the Hamiltonian (11.297) which can lead to a yet new constraint and so on. Proceeding in this way, one eventually identifies all the constraints $\Phi_M = 0$, $M = 1, \dots, k$, where $n - r \leq k < 2n$, the investigated system is subjected to. Finally, there are Poisson brackets of Φ_N with the Hamiltonian $H_{\rm T}$ which do not vanish weakly but depend on the coefficients u^M. Requiring that all such Poisson brackets, of the primary as well as of all secondary constraint (weakly) vanish one gets the system of linear equations

$$\{\Phi_N,\ H\}_{\rm PB} + \sum_{M=1}^{n-r} \{\Phi_N,\ \Phi_M\}_{\rm PB}\, u^M \simeq 0\,, \qquad N = 1, \dots, k\,, \tag{11.298}$$

in which the sum runs over the $n - r$ primary constraints. The solution of this system of linear inhomogeneous equations takes the general form[80]

$$u^M = c^M(q,p) + \sum_{a=1}^{p} v^M_{(a)}(q,p)\, s^a\,, \tag{11.299}$$

[80]Since the number $n - r$ of unknowns u^M is smaller than the number k of the equations, the existence of a solution is a nontrivial fact; it follows from the assumption that the original Euler-Lagrange equations are not inconsistent.

in which $c^M(q,p)$ form a particular solution of the inhomogeneous system (11.298), while $v^M_{(a)}(q,p)$, $a=1,\dots,p$, are all linearly independent solutions of the homogeneous system

$$\sum_{M=1}^{n-r}\{\Phi_N,\ \Phi_M\}_{\rm PB}\, v^M_{(a)}(q,p)\simeq 0\,. \tag{11.300}$$

Plugging the solution (11.299) into the total Hamiltonian (11.297) completes the hamiltonization of the classical system: once the constraints $\Phi_M(q,p)=0$, $M=1,\dots,k$, are imposed on the initial data, they are automatically preserved by the time evolution.

An important feature of the total Hamiltonian (11.297) is its possible dependence, after inserting in it the solution (11.299), on the coefficients s^a, which may be arbitrary functions of time. They are present whenever there exist nontrivial linear combinations

$$\Phi_a(q,p)\equiv\Phi_M(q,p)\, v^M_{(a)}(q,p)\,, \tag{11.301}$$

of the primary constraints which, as follows form (11.300), have (weakly) vanishing Poisson brackets with all other constraints (the primary and secondary ones). Generally, one calls a quantity (a function of q and p) a first class quantity if its Poisson brackets with *all* the constraints $\Phi_M(q,p)$ weakly vanish. Correspondingly, linear combinations $\Phi_a(q,p)$ such that

$$\{\Phi_N,\ \Phi_a\}_{\rm PB}\simeq 0 \qquad \text{for all} \quad N \tag{11.302}$$

are called *first class* constraints. The remaining ones, which have at least one nonzero Poisson bracket with the other constraints, are called *second class*. It can be noted at this point that the total Hamiltonian

$$\begin{aligned} H_{\rm T} = H &+ \sum_{M\in\text{primary}}\Phi_M(q,p)\, c^M(q,p) + \sum_{a\in\text{primary}}\Phi_a(q,p)\, s^a(t)\\ &\equiv H'(q,p)+\sum_{a\in\text{primary}}\Phi_a(q,p)\, s^a(t)\,, \end{aligned} \tag{11.303}$$

obtained using the Dirac procedure outlined above is the sum of two terms, which are separately first class. It is also straightforward to prove, using the Jacobi identity, that the Poisson bracket of two first class quantities is also a first class quantity.

Due to the presence of the arbitrary functions $s^a(t)$, which is always the case if there are primary first class constraints, the time evolution of the canonical variables $q^i(t)$ and $p_i(t)$ generated by the Hamiltonian (11.297) is not unique - the initial data $q^i(t_0)$ and $p_i(t_0)$ set at some $t=t_0$ do not specify uniquely the values of $q^i(t)$ and $p_i(t)$ at other times t. This can only make sense if the physical state of the system does not uniquely determine the variables q^i and p_i. In other words, characterization of the system in terms of canonical variables exhibits some kind of "gauge invariance" by which term one means the situation in which the (classical) physical state of a system does not change when the variables used to characterize it are transformed in some specific way. In such a case infinitely many different values of *the same* canonical variables correspond to the same (classical) physical state; owing to this the time evolution generated by the Hamiltonian (11.303) (and by the Euler-Lagrange equations, to which it is equivalent) can lead, despite the presence of the arbitrary functions $s^a(t)$, to a unique evolution of physical states.

Let us identify possible gauge transformations of the variables q^i and p_i. One obvious class of such transformations is of the form

$$\begin{aligned} \delta q^i &= \delta\theta^a \{q^i,\ \Phi_a\}_{\rm PB}\,, \\ \delta p_i &= \delta\theta^a \{p_i,\ \Phi_a\}_{\rm PB}\,, \end{aligned} \tag{11.304}$$

where Φ_a are the system's first class primary constraints and $\delta\theta^a$ are arbitrary parameters. Indeed, starting at $t = t_0$ from a given point in the phase space and using two infinitesimally different Hamiltonians (11.303), one with $s^a(t)$ and the other one with $s^a(t) + \delta s^a(t)$, one reaches, after an infinitesimal time Δt two different phase space points the coordinates of which are related by (11.304) with $\delta\theta^a = \Delta t\, \delta s^a(t_0)$. Thus, variables connected by the transformations (11.304) generated through the Poisson brackets by the primary first class constraints must define the same physical state. Furthermore, taking the difference of the results of two successive transformations (11.304) performed in two different orders one concludes that Poisson brackets of two primary first class constraints must also generate admissible gauge transformations. Finally, considering changes of the variables generated by applying in two different orders: a transformation (11.304) and an infinitesimal time evolution generated by $H_{\rm T}$, one concludes that the Poisson brackets of H' (a first class quantity) defined in (11.303) with the primary first class constraints are also generators of admissible gauge transformations. Since, as has been noted above, Poisson brackets of first class quantities are also first class quantities, that is,

$$\begin{aligned} \{\Phi_a,\ \Phi_{a'}\}_{\rm PB} &= \alpha_{aa'}{}^{b}\Phi_b\,, \\ \{H',\ \Phi_{a'}\}_{\rm PB} &= \beta_{a'}{}^{b}\Phi_b\,, \end{aligned}$$

where the indices b run over all first class constraints (not necessarily the primary ones), it follows that some gauge transformations can be also generated by first class constraints which are not primary. The only difference at this point between gauge transformations generated by the first class secondary and primary constraints is that only the latter class of gauge transformations contributes to the arbitrariness of the time evolution generated by $H_{\rm T}$ (and by the corresponding Euler-Lagrange equations). Although it is possible to construct examples (nonrelativistic) in which not all secondary first class constraints generate transformations not affecting (in the classical theory) the physical state of the system, one assumes that in all physically sensible cases *all* first class constraints (including all secondary ones) are generators of gauge transformations which can be iteratively reached from the identity transformation.[81] It also seems that adopting this assumption is

[81]This very important fact is crucial in discussing topological properties of quantized nonabelian Yang Mills theories.

necessary for consistent quantization. Accordingly, one defines the *extended Hamiltonian* H_{E} (Dirac's terminology again) by including in the second term of (11.303) all the first class constraints. The classical dynamics generated by H_{E} has more arbitrariness than does the one generated by H_{T} and that is inherent in the time evolution determined by the underlying Euler-Lagrange equations. Such an extension of the evolution, possible in the Hamilton's formulation is, however, fully admissible since it still leads to a unique evolution of the physical state of the system.[82]

The final remark is that the secondary second class constraints also could have been included in the sum in (11.297) - in the resulting set of equations (11.298) there would be then more factors u^N to determine (their number would in this case match the number of equations), but it can be shown that the solutions for additional u^N's would vanish weakly, i.e. would be proportional to a linear combination of constraints (and therefore the corresponding extra terms in (11.297) would not change the canonical equations for q^i and p_i, just because $\{f(q,p),\ \Phi_M\Phi_N\}_{\mathrm{PB}} = \Phi_M\{f(q,p),\ \Phi_N\}_{\mathrm{PB}} + \{f(q,p),\ \Phi_M\}_{\mathrm{PB}}\Phi_N \simeq 0$). Thus, all constraints, primary and secondary of both classes can be from the beginning included in (11.297) and treated on the same footing.

Hamiltonization is the necessary first step towards the canonical (operator) quantization. We will consider first systems subject to first class constraints only. There are two major methods of quantizing such systems. One consists of promoting all the canonical variables q^i and p_i to independent operators satisfying the standard commutation rules (11.295) and representing their algebra in a Hilbert (or in a Fock) space of states $|\Psi\rangle$. The classical constraints $\Phi_a = 0$ are in this method *not* imposed as operator relations but instead become conditions selecting in the full Hilbert (or Fock) space vectors (rays) $|\Psi_{\mathrm{phys}}\rangle$, forming a subspace, which represent physical states. Physical are then only the states satisfying the conditions

$$\Phi_a(\hat{q},\hat{p})|\Psi_{\mathrm{phys}}\rangle = 0\,, \qquad \text{all } \ a\,. \tag{11.305}$$

As the first class constraints generate gauge transformations, this means that physical are those states which are invariant with respect to such transformations.[83] This is possible provided the operators $\Phi_a(\hat{q},\hat{p})$ can be ordered in such a way as to ensure the operator relations $[\hat{\Phi}_a,\ \hat{\Phi}_b] = i\hbar\, c_{ab}{}^{d}(\hat{q},\hat{p})\,\hat{\Phi}_d$ (necessary for consistency of the conditions (11.305)) and $[\hat{H}',\ \hat{\Phi}_a] = i\hbar\, b_a{}^{d}(\hat{q},\hat{p})\,\hat{\Phi}_d$, where H' is defined in (11.303) (necessary to ensure that the time evolution does not map physical states into unphysical ones or the other way around).

[82] The fact that the Hamiltonian formulation allows for a more general time evolution of gauge system's variables should not surprise: the Hamilton's formulation allows also for a wider class of transformations of the canonical variables than does the Lagrangian one.

[83] More precisely, the conditions (11.305) ensure invariance of physical states with respect to those gauge transformations (called "small") which can be reached iteratively from the identity transformation. The requirement that the states $|\Psi_{\mathrm{phys}}\rangle$ are invariant also with respect to so-called "large" gauge transformations, if imposed, is an extra assumption not following from the consistency of the quantization prescription.

A variant of this approach, called the Dirac-Fock quantization is employed when, as it is the case in quantum field theories, the first class constraints Φ_a expressed in terms of the canonical variables promoted to Schrödinger picture operators the algebra of which is realized in some Fock space naturally split into $\Phi_a^{(+)} + \Phi_a^{(-)}$, where $\Phi_a^{(+)}$ ($\Phi_a^{(-)}$) involve only the annihilation (creation) operators defined with respect to the vector $|0_{\rm Fock}\rangle$ of the chosen Fock space and imposing the conditions (11.305) would either lead to inconsistency due to commutators of Φ_a with some other operators or would just leave in the Fock space no physical state-vectors at all (or both). In this case one identifies physical states by the weaker conditions,

$$\Phi_a^{(+)}|\Psi_{\rm phys}\rangle = 0\,, \qquad \text{all}\ \ a\,, \tag{11.306}$$

(so that still $\langle\Psi'_{\rm phys}|\Phi_a|\Psi_{\rm phys}\rangle = 0$). A characteristic feature of this approach is the presence in the Fock space of state-vectors of negative or zero norm. If the system is properly quantized in this way, all state-vectors of negative norm are manifestly unphysical (in the sense of the condition (11.306)), whereas the zero norm ones, while being classified as physical, have zero scalar products among themselves and with all physical, positive norm states. Arbitrary linear combinations of "physical" zero norm vectors can be then added to a positive norm physical state-vector without changing its norm or scalar products (transition rates). A physical state of the system is then in the Fock space represented not by a ray but by a larger equivalence class of state-vectors differnig one from another by a zero norm vector; the zero norm vectors represent in this approach a residual gauge invariance not removed by the conditions (11.306). This approach is one of possible methods of the electromagnetic field quantization (free or interacting with other fields) and leads to the same structure of the theory as the Gupta-Bleuler quantization in the covariant Lorentz gauge $\partial_\mu A^\mu = 0$ outlined in Section 11.11. It is also used for quantization of the relativistic string (see the old review by C. Rebbi).

The second method of quantizing systems subject to first class constraints, which will be used in section 11.7 to quantize the electromagnetic field in the Coulomb gauge, is to fix the gauge completely. This is done by imposing additional constraint(s), called *gauge fixing conditions*, $\Phi(q,p) = 0$ from the outside,[84] so that the entire system of constraints (determined as described at the beginning of this section, treating the gauge fixing conditions as all other primary constraints and including them together with all second class constraints multiplied by the corresponding u^M factors in the Hamiltonian H_T (11.297)) becomes second class and the time evolution of the canonical variables becomes uniquely determined by the equations of motion and the initial data. The resulting classical systems subject to second class constraints can be quantized with the help of the Dirac prescription described below. The method works outside the perturbative expansion provided the gauge can be fixed globally, that is in such a way that the canonical variables satisfying the chosen gauge condition are uniquely determined. This is not always the case.

[84] This is possible because, as explained above, the physical state of the system does not fix the values of the canonical variables uniquely; the added constraints can be then made consistent with the dynamics generated by $H_{\rm T}$ from which the arbitrary functions s_a are now absent.

For example, it is known that in the case of non-Abelian Yang-Mills theories (Chapter 20), there are multiple solutions to the Coulomb gauge condition (this is known under the name of Gribov ambiguity). In this case the validity of the method is restricted to the perturbative expansion only. In field theory a consequence of imposing extra gauge conditions may be the spatial nonlocality of the resulting Hamiltonian.

To quantize a theory in which first class constraints are absent from the beginning or have been eliminated by imposing additional constraints from the outside, one forms the matrix C_{NM}:

$$C_{NM} \equiv \{\Phi_N,\ \Phi_M\}_{\rm PB}\ . \tag{11.307}$$

Since the matrix C_{NM} is antisymmetric (the basic property of the Poisson brackets defined for commuting variables, to which we restrict ourselves in this section) and $\det(C_{NM}) \neq 0$ (vanishing, even weak, of this determinant would mean that one can form at least one more linear combination of the constraints which has zero Poisson brackets with all others and is, hence, a first class constraint which we have assumed to be already eliminated), it must be of even dimension. The Dirac prescription for quantizing systems subject to second class constraints then reads

$$\begin{aligned}&[\hat{q},\ \hat{p}] = i\hbar\,\{q,\ p\}_{\rm D}\big|_{q\to\hat{q},\,p\to\hat{p}}\ ,\\ &\Phi_N(\hat{q},\hat{p}) \equiv 0\,, \qquad \text{for all } N\,,\end{aligned} \tag{11.308}$$

that is, the commutators (or anticommutators in the case of fermionic fields - see section 11.8) are determined by the *Dirac brackets* (instead of Poisson brackets) and all the constraints are realized as *strong operator identities* $\hat{\Phi}_M \equiv \Phi_M(\hat{q},\hat{p}) = 0$. The Dirac bracket $\{A(p,q),\ B(p,q)\}_{\rm D}$ of two functions of the canonical variables is defined as follows:

$$\{A,\ B\}_{\rm D} \equiv \{A,\ B\}_{\rm PB} - \sum_{N,M} \{A,\ \Phi_N\}_{\rm PB} \left(C^{-1}\right)^{NM} \{\Phi_M,\ B\}_{\rm PB}\ , \tag{11.309}$$

(again the Poisson brackets have to be computed as if there were no constraints). The Dirac bracket shares with the Poisson bracket all the algebraic properties: it is bilinear in canonical variables, antisymmetric and satisfies the Jacobi identity. It is also a matter of a simple algebra to check that the definition (11.309) is invariant with respect to linear changes

$$\Phi_M(q,p) = O^N_M(q,p)\,\tilde{\Phi}_N(q,p)\,,$$

of the basis of constraints, provided the matrix $O^N_M(q,p)$ is nonsingular (invertible) on the surface of constraints, that is provided its determinant does not vanish weakly. Finally, the Dirac bracket has the (easy to verify) property

$$\{A,\ \Phi_N\}_{\rm D} = 0\,, \qquad \text{for all } N\,, \tag{11.310}$$

which ensures compatibility of the commutation relation (11.308) with the constraints $\Phi_N = 0$ realized as strong operator identities. This also means that the Hamiltonian operator is obtained from the H part of the total classical Hamiltonian $H_{\rm T}$ (11.297): the sum of the primary constraints (which in the classical theory must be kept in order to ensure the compatibility of constraints with the dynamics) is in the quantum theory just a zero operator. Obviously, the entire algebra (11.308) must be realized by operators acting in a Hilbert or a Fock space and finding this realization may turn out to be the most important difficulty in carrying out the quantization à la Dirac, especially when not all right hand sides of the first set of relations in the prescription (11.308) are c-numbers.

The Dirac prescription can be justified as follows. First of all, if the first class constraints are absent, the second term, depending on arbitrary functions s^a in the solution (11.299) of the consistency conditions (11.298) is absent and the factors u^M entering the Hamiltonian $H_{\rm T}$ (11.297) in which all second class constraints have been included multiplied by the corresponding u^M factors are given by[85]

$$u^M(q,p) = c^M(q,p) = -(C^{-1})^{MN}\{\Phi_N,\, H\}_{\rm PB}\,.$$

The classical canonical equations of motion $\dot{q}^i = \{q^i,\, H_{\rm T}\}_{\rm PB}$ and $\dot{p}_i = \{p_i,\, H_{\rm T}\}_{\rm PB}$ take then just the form

$$\dot{q}^i = \{q^i,\ H_{\rm T}\}_{\rm PB} = \{q^i,\ H\}_{\rm PB} + \sum_{M=1}^{2k}\{q^i,\, \Phi_M\}_{\rm PB}\, c^M(q,p) \equiv \{q^i,\ H\}_{\rm D}\,,$$

and similarly for $\dot{p}_i$. Since in the quantum theory the equations of motion satisfied by the Heisenberg picture operators are given by their commutators with the Hamiltonian, the above form of the classical equations of motion strongly supports the Dirac quantization rule. More formally, exploiting the fact that the number of the second class constraints Φ_M, $M = 1, 2, \ldots, 2k$ is always even, it can be shown that there always exists a classical canonical transformation $(q_i,\ p_i) \to (Q_i(q,p),\ P_i(q,p))$, $i = 1, 2, \ldots, n$, such that in the new canonical variables the constraints Φ_M take the simple form

$$Q_i = 0\,, \qquad P_i = 0 \qquad \text{for} \quad i = 1, \ldots, k. \tag{11.311}$$

One can therefore quantize the system in the remaining new variables Q_i, and P_i with $i = k+1, \ldots, n$ forgetting about the first $k + k = 2k$ variables which are zero. The commutators of the quantum operators $\hat{q}_i$, and $\hat{p}_j$ when derived from the commutators $[\hat{Q}_i,\ \hat{P}_j] = i\hbar\,\delta_{ij}$ with $i, j > k$ turns out to be given by the Dirac bracket (times $i\hbar$). The advantage of the Dirac method for finding the right commutation relations which must be satisfied by the canonical variables is that one does not have to find the canonical transformation $(q_i,\ p_i) \to (Q_i(q,p),\ P_i(q,p))$ explicitly nor does one have to solve the constraints for some variables in terms of others.

[85]It is assumed now that all second class constraints, primary and secondary, have been included in $H_{\rm T}$ in agreement with the remark made earlier.

Quantization of a system subject to both, first and second class constraints can be done either by fixing the gauge and turning them into systems with second class constraints only or by handling the second class constraints with the help of the Dirac prescription (11.308), while imposing the first class ones as subsidiary conditions selecting those vectors of the Hilbert space which represent quantum states of the physical system. As has been remarked above, practical implementation of the Dirac prescription for treating second class constraints may in more complicated cases be very difficult because realization of the operator algebra defined by (11.308), especially if some of the right hand sides of the commutation relations (determined by the Dirac bracket) turn out to be operators, and not c-numbers as in the case of the Proca theory or the electromagnetic field. It is then helpful to know that in principle any system subject to second class constraints can be turned into a (gauge invariant) system subject to first class constraints only by appropriately enlarging the number of its canonical variables; this opens the possibility of realizing all constraints as subsidiary conditions selecting physical states in the whole Hilbert of Fock space and not as complicated operator identities.

We end this section by applying the Dirac prescription for quantizing systems subject to second class constraints to the Proca theory of the vector field. The primary constraint[86]

$$\Phi_1 \equiv \Pi_0(t, \mathbf{x}) = 0\,, \tag{11.312}$$

follow in this case from the structure of the Lagrangian (11.265). The total Hamiltonian takes in this case the form

$$\begin{aligned}\mathcal{H}_{\mathrm{T}} = \frac{1}{2}\Pi_i\Pi_i - \Pi_i\partial_i V^0 + \frac{1}{2}(\boldsymbol{\nabla}\times\mathbf{V})^2 + \frac{1}{2}M^2 V^i V^i \\ -\frac{1}{2}M^2 V^0 V^0 + \Pi_0 u^1 - g\, V^i J^i + g\, V^0 J^0\,,\end{aligned} \tag{11.313}$$

with an initially unknown function u^1. Since $\dot{\Phi}_1 = \{\Phi_1,\, H_{\mathrm{T}}\}_{\mathrm{PB}}$ does not vanish (even weakly), the secondary constraint

$$\Phi_2 \equiv M^2 V^0(t, \mathbf{x}) - \partial_i \Pi_i(t, \mathbf{x}) - g\, J^0(t, \mathbf{x}) = 0\,, \tag{11.314}$$

must be imposed as the consistency condition. There no more constraints in this case - vanishing of $\dot{\Phi}_2 = \{\Phi_2,\, H_{\mathrm{T}}\}_{\mathrm{PB}}$ can be ensured by adjusting the function u^1. This completes the Hamiltonization of the classical dynamics of the Proca field.

In the quantum theory $\Phi_1 = 0$ and $\Phi_2 = 0$ become strong operator identities; therefore, in the Hilbert space Π_0 must be represented by the zero operator and through (11.314) the operator V^0 becomes completely determined by Π_i and J^0. The canonical commutation relations of the V^i, Π_i, V^0 and Π_0 operators are given by the Dirac prescription (11.308)

[86] In field theory the index i of q^i in the formulae (11.296)-(11.311) includes also the space variable $\mathbf{x}$. Hence, each Φ_M stands for an infinite set of constraints $\Phi_{M\mathbf{x}} \equiv \Phi_M(\mathbf{x})$.

which ensures their compatibility with the operator relations $\Phi_1 = 0$ and $\Phi_2 = 0$. The antisymmetric matrix C_{NM} (11.307) has in this case the form

$$C_{1\mathbf{x},2\mathbf{y}} = \{\Phi_{1\mathbf{x}},\ \Phi_{2\mathbf{y}}\}_{\mathrm{PB}} = \{\Phi_1(\mathbf{x}),\ \Phi_2(\mathbf{y})\}_{\mathrm{PB}} = -M^2\delta^{(3)}(\mathbf{x}-\mathbf{y})\,, \tag{11.315}$$

and its inverse reads

$$C^{-1} = \begin{pmatrix} 0 & \frac{1}{M^2}\,\delta^{(3)}(\mathbf{x}-\mathbf{y}) \\ -\frac{1}{M^2}\,\delta^{(3)}(\mathbf{x}-\mathbf{y}) & 0 \end{pmatrix}. \tag{11.316}$$

The Dirac bracket of any two functions $A(\mathbf{x})$ and $B(\mathbf{y})$ of the canonical variables V^μ and Π_μ therefore reads

$$\begin{aligned}\{A(\mathbf{x}),\ B(\mathbf{y})\}_{\mathrm{D}} &= \{A(\mathbf{x}),\ B(\mathbf{y})\}_{\mathrm{PB}} \\ &-\frac{1}{M^2}\int d^3\mathbf{z}\,\{A(\mathbf{x}),\ \Pi_0(\mathbf{z})\}_{\mathrm{PB}}\left\{M^2V^0(\mathbf{z}) - \partial_i\Pi_i(\mathbf{z}) - gJ^0(\mathbf{z}),\ B(\mathbf{y})\right\}_{\mathrm{PB}} \\ &+\frac{1}{M^2}\int d^3\mathbf{z}\left\{A(\mathbf{x}),\ M^2V^0(\mathbf{z}) - \partial_i\Pi_i(\mathbf{z}) - gJ^0(\mathbf{z})\right\}_{\mathrm{PB}}\{\Pi_0(\mathbf{z}),\ B(\mathbf{y})\}_{\mathrm{PB}}\,.\end{aligned}$$

Computing using this formula the commutator

$$\left[V^i(\mathbf{x}),\ V^0(\mathbf{y})\right] = i\hbar\left\{V^i(\mathbf{x}),\ V^0(\mathbf{y})\right\}_{\mathrm{D}}\,, \tag{11.317}$$

one finds the same result as was obtained in Section 11.5 by treating the V^0 operator as constructed out of Π_i and J^0 in agreement with the constraint $\Phi_2 = 0$. It is also easy to check that the Dirac prescription gives all the remaining commutators in their standard forms. The rest of the quantization procedure is then unmodified: the expansions (11.289) with $\mu = i$ and (11.290) give the Schrödinger picture operators $V^i(\mathbf{x})$ and $\Pi_i(\mathbf{x})$ satisfying the commutation relations following from the Dirac quantization procedure, $\Pi_0(\mathbf{x})$ is just the zero operator and $V^0(\mathbf{x})$ is a sum of (11.289) with $\mu = 0$ (and t set to zero) and a of term depending on the matter variables.

11.7 Quantization of the electromagnetic field

We now depart a little bit from the main logic of this chapter (devoted mainly to quantization of systems of relativistic fields) and consider quantization of the system consisting of the classical electromagnetic field coupled to N classical nonrelativistic charged particles. The reason for this departure is that in this way we will obtain a working theory, called nonrelativistic quantum electrodynamics (NRQED), which is applicable to a wide range of physical phenomena related to interaction of light with nonrelativistic matter, so long as spin effects can be neglected. In this way we complete here the quantum theory of radiation presented in Section 3.8. The resulting theory constitutes also the starting point of modern calculations of properties of bound states in the fully relativistic quantum electrodynamics.

The Lagrangian which determines classical dynamics of such a system is given by the sum of three terms: $L = L_{\rm EM} + L_{\rm matt} + L_{\rm int}$:

$$L_{\rm EM} = -\frac{1}{4}\int d^3\mathbf{x}\, f_{\mu\nu} f^{\mu\nu}, \tag{11.318}$$

$$L_{\rm matt} = \frac{1}{2}\sum_{i=1}^{N} m_i \mathbf{v}_i^2 - \sum_{i=1}^{N} V(\mathbf{r}_i) - \frac{1}{2}\sum_{i\neq j}^{N} V(\mathbf{r}_i - \mathbf{r}_j)\,, \tag{11.319}$$

$$L_{\rm int} = -e\sum_{i=1}^{N}\left[q_i\phi(\mathbf{r}_i) - \frac{q_i}{c}\mathbf{A}(\mathbf{r}_i)\cdot\mathbf{v}_i\right] \equiv -\frac{e}{c}\int d^3\mathbf{x}\, J_\mu(\mathbf{x}) A^\mu(\mathbf{x})\,, \tag{11.320}$$

where the current in the last formula has the form (q_i are particle charges in units of $e > 0$)

$$\begin{aligned} J^0(t,\mathbf{x}) &= \sum_{i=1}^{N} q_i c\, \delta^{(3)}(\mathbf{x} - \mathbf{r}_i(t))\,, \\ \mathbf{J}(t,\mathbf{x}) &= \sum_{i=1}^{N} q_i \mathbf{v}_i\, \delta^{(3)}(\mathbf{x} - \mathbf{r}_i(t))\,. \end{aligned} \tag{11.321}$$

It can be checked that the current J^μ defined by (11.321) is conserved: $\partial_\mu J^\mu = 0$. For greater generality we have allowed in $L_{\rm matt}$ also for 1-particle interactions $V(\mathbf{r}_i)$ with external potentials (which can be also of electromagnetic origin - e.g. the electrostatic interactions binding electrons in atoms, which we may treat separately from the dynamical electromagnetic field) and two-particle interactions $V(\mathbf{r}_i - \mathbf{r}_j)$ of *non*-electromagnetic origin.

To build the quantum theory of this system we first try to take for its canonical variables and conjugated momenta $\mathbf{r}_i(t)$, $\mathbf{P}_i(t)$, $A^\mu(t,\mathbf{x})$ and $\Pi_\mu(t,\mathbf{x})$. In the matter part everything goes in the standard way:

$$\mathbf{P}_i(t) = \frac{\partial L}{\partial \mathbf{v}_i(t)} = m_i\mathbf{v}_i(t) + \frac{q_i}{c}\, e\mathbf{A}(t,\mathbf{r}_i(t))\,. \tag{11.322}$$

Expressing $\mathbf{v}_i(t)$ in the matter part of the Hamiltonian through $\mathbf{P}_i(t)$ we get

$$\begin{aligned} H = H_0^{\rm EM} &+ \sum_{i=1}^{N} \frac{1}{2m_i}\left(\mathbf{P}_i - \frac{q_i}{c}\, e\mathbf{A}(\mathbf{r}_i)\right)^2 \\ &+ e\sum_{i=1}^{N} q_i\phi(\mathbf{r}_i) + \sum_{i=1}^{N} V(\mathbf{r}_i) + \frac{1}{2}\sum_{i\neq j}^{N} V(\mathbf{r}_i - \mathbf{r}_j)\,, \end{aligned} \tag{11.323}$$

where $H_0^{\rm EM}$ is the current-independent (here "current-independent" *before* expressing ϕ in terms of J^0!) part of the (total) electromagnetic field Hamiltonian (11.348) which we derive below. Quantization of the matter part is standard: we promote $\mathbf{r}_i(t)$ and $\mathbf{P}_i(t)$ taken at $t=0$ to Schrödinger picture operators satisfying the commutation rules $[\hat{\mathbf{r}}_i,\ \hat{\mathbf{P}}_j]=i\hbar\delta_{ij}$. We can then work either in the position representation in which $\hat{\mathbf{P}}_i=-i\hbar\partial/\partial\mathbf{r}_i$, or we can go over to the second quantization formalism (presented in Chapter 5), which is especially convenient if the matter particles of the system are numerous and indistinguishable.

The electromagnetic part of the system is more troublesome. Written in terms of the potentials the Lagrangian density $\mathcal{L}$ (11.258) of the electromagnetic field coupled to the current J^μ reads (c is kept for decoration, $\phi\equiv A^0$):

$$\mathcal{L}^{\rm EM}=\frac{1}{2}\left(\frac{1}{c}\frac{\partial}{\partial t}\mathbf{A}+\boldsymbol{\nabla}\phi\right)^2-\frac{1}{2}\left(\boldsymbol{\nabla}\times\mathbf{A}\right)^2-\frac{1}{c}\,eJ^0\phi+\frac{1}{c}\,e\mathbf{J}\cdot\mathbf{A}\,. \tag{11.324}$$

Of course, $\int d^3\mathbf{x}\,\mathcal{L}_{\rm EM}=L_{\rm EM}+L_{\rm int}$ in the notation of (11.318-11.320). As already checked in Section 11.4, the Euler-Lagrange equations following from this Lagrangian density are ($\rho\equiv eJ^0$)

$$\boldsymbol{\nabla}\times\mathbf{B}-\frac{1}{c}\,\frac{\partial\mathbf{E}}{\partial t}=\frac{1}{c}\,e\mathbf{J}\,,\qquad \boldsymbol{\nabla}\cdot\mathbf{E}=\rho\,. \tag{11.325}$$

As the canonical momenta Π_i we get

$$\Pi_i=\frac{1}{c^2}\,\frac{\partial A^i}{\partial t}+\frac{1}{c}\,\partial_i\phi=-\frac{1}{c}\,E^i\,, \tag{11.326}$$

and $\Pi_0=0$, similarly as in the case of the Proca vector field. We have therefore the primary constraint $\Phi_0\equiv\Pi_0$ and we have to check whether it is compatible with the Hamilton's canonical equations. To this end we form the total Hamiltonian density $\mathcal{H}_{\rm T}\equiv\mathcal{H}_{\rm T}^{\rm EM}+\mathcal{H}^{\rm matt}$ (the interaction of the field with matter is now included in $\mathcal{H}_{\rm T}^{\rm EM}$) according to the Dirac method expressing $\dot{A}^i$ through Π_i by using the relation (11.326) and adding the primary constraint with an unknown coefficient u^0. This gives:[87]

$$\begin{aligned}\mathcal{H}_{\rm T}^{\rm EM}&=\Pi_i\,\dot{A}^i+u^0\Phi_0-\mathcal{L}^{\rm EM}\\&=\frac{c^2}{2}\,\Pi_i\Pi_i-c\,\Pi_i\partial_i\phi+\frac{1}{2}\left(\boldsymbol{\nabla}\times\mathbf{A}\right)^2+\Pi_0u^0+\frac{1}{c}\,eJ^0\phi-\frac{1}{c}\,e\mathbf{J}\cdot\mathbf{A}\\&=\frac{1}{2}\left(\mathbf{E}^2+\mathbf{B}^2\right)+\mathbf{E}\cdot\boldsymbol{\nabla}\phi+\Pi_0u^0+\rho\,\phi-\frac{1}{c}\,e\mathbf{J}\cdot\mathbf{A}\,,\end{aligned} \tag{11.327}$$

[87]Note, that $\mathbf{J}$ depends now on $\hat{\mathbf{r}}_i$ and $\hat{\mathbf{P}}_i$.

Using this Hamiltonian density we find[88]

$$\dot{\Phi}_0 \equiv \left\{\Pi_0,\ \int d^3\mathbf{x}\,\mathcal{H}_{\rm T}\right\}_{\rm PB} = -\frac{\partial\mathcal{H}^{\rm EM}_{\rm T}}{\partial\phi} = -\Pi_0\,\frac{\partial u^0}{\partial A^0} - c\,\partial_i\Pi_i - \frac{1}{c}\,eJ^0$$
$$\simeq -c\,\partial_i\Pi_i - \frac{1}{c}\,eJ^0\,, \qquad (11.328)$$

(the constraint $\Pi_0 = 0$ has been imposed after computing the Poisson bracket). As this is not zero (even weakly), the secondary constraint $\Phi_2 \equiv c^2\partial_i\Pi_i + eJ^0$ has to be imposed. This is obviously the Gauss law, which is the second of the Euler-Lagrange equations (11.325), but is lost in the Hamilton's formalism because of vanishing of Π_0. The constraint Φ_2 is already compatible with the dynamics: $\{\Phi_2,\ H_{\rm T}\}_{\rm PB} = c\,\partial_\mu(eJ^\mu)$ vanishes provided the current is conserved[89] - as remarked, the current (11.321) doe satisfy this requirement.

The system is therefore subject to two constraints, Φ_0 and Φ_2, but they turn out to be the first class:

$$\{\Phi_0,\ \Phi_2\}_{\rm PB} = 0\,,$$

and the Dirac prescription for quantization cannot be applied because the (classical) state of the electromagnetic field, fully characterized by $\mathbf{E}$ and $\mathbf{B}$, does not determine uniquely the potentials A^μ. The mathematical reason for this is the gauge invariance of the Lagrangian (11.324). The Euler-Lagrange equations do not determine $A^\mu(t,\mathbf{x})$ uniquely at $t \neq 0$ from initial conditions $A^\mu(\mathbf{x})$ and $\dot{A}^\mu(\mathbf{x})$ specified at $t = 0$: one can always imagine a function $\theta(t,\mathbf{x})$ such that $\partial^\mu\theta(t,\mathbf{x}) = 0$ and $\partial^\mu\dot{\theta}(t,\mathbf{x}) = 0$ at $t = 0$, but nonvanishing at $t \neq 0$. If $A^\mu(t,\mathbf{x})$ solves the Euler-Lagrange equations with the given initial conditions, then so does $A^\mu(t,\mathbf{x}) + \partial^\mu\theta(t,\mathbf{x})$ with the same initial conditions at $t = 0$. Since the constraints $\Phi_0 = 0$ and $\Phi_2 = 0$ essentially make the canonical Hamilton's equations equivalent to the Euler-Lagrange equations, it is clear that they cannot help to determine $A^\mu(t,\mathbf{x})$ uniquely at all t as it was possible in the case of the Proca vector field.

Unique determination of the time evolution of $A^\mu(t,\mathbf{x})$ becomes possible if one chooses a gauge. In a fixed gauge the Euler-Lagrange equations can determine $A^\mu(t,\mathbf{x})$ for all t unambiguously and it should be possible to find an equivalent Hamilton's canonical formulation too. One possibility is the Coulomb gauge

$$\boldsymbol{\nabla}\cdot\mathbf{A} = 0\,. \qquad (11.329)$$

[88]In computing Poisson brackets all variables and momenta have to be treated as independent; therefore, for any two functionals G and H we have

$$\left\{G[A^\mu,\Pi_\mu],\ H[A^\mu,\Pi_\mu]\right\}_{\rm PB} \equiv \int d^3\mathbf{x}\left(\frac{\delta G}{\delta A^\lambda(\mathbf{x})}\frac{\delta H}{\delta\Pi_\lambda(\mathbf{x})} - \frac{\delta H}{\delta A^\lambda(\mathbf{x})}\frac{\delta G}{\delta\Pi_\lambda(\mathbf{x})}\right).$$

[89]To derive the relation $\{\Phi_2,\ H_{\rm T}\}_{\rm PB} = c\,\partial_\mu(eJ^\mu)$ crucial is taking into account the full Hamiltonian $\mathcal{H}_{\rm T} = \mathcal{H}^{\rm EM}_{\rm T} + \mathcal{H}_{\rm matt}$ which in the case considered here includes, in addition to the terms displayed in (11.327), also the term $H_{\rm matt}$ dynamics of the nonrelativistic particles; one then finds that $\{J^0,\ H^{\rm matt}\}_{\rm PB} = c\,\partial_0 J^0$.

Another, convenient in some applications, is the temporal gauge $A^0 = 0$. However, while the latter makes the time evolution unique, it does not fix uniquely the initial data: gauge transformations with time independent function $\theta(\mathbf{x})$ are still possible. As a result, the constraint $\Phi_2 = 0$ (the Gauss law) remains in this approach first class and has to be imposed as a subsidiary condition selecting physical state-vectors of the Hilbert space. For this reason we here choose to work in the Coulomb gauge (11.329) which classically has a clear physical interpretation: it separates electrostatic effects from those due to radiation.

We restart, therefore, the whole procedure with the same Lagrangian density (11.324) and one primary constraint

$$\Phi_1 \equiv \boldsymbol{\nabla}\cdot\mathbf{A} \equiv \partial_i A^i\,, \tag{11.330}$$

imposed from outside. As previously one finds $\Phi_0 \equiv \Pi_0$ as the primary constraint and has to consider the extended Hamiltonian which now involves two initially unspecified functions u^0 and u^1:

$$\mathcal{H}^{\rm EM}_{\rm T} = \frac{c^2}{2}\,\Pi_i\Pi_i - c\,\Pi_i\partial_i\phi + \frac{1}{2}\,(\boldsymbol{\nabla}\times\mathbf{A})^2 + \Phi_0 u^0 + \Phi_1 u^1 + \frac{1}{c}\,eJ^0\phi - \frac{1}{c}\,e\mathbf{J}\cdot\mathbf{A}\,.$$

The consistency condition $\{\Phi_0,\ H_{\rm T}\}_{\rm PB} \simeq 0$, leads again to the secondary constraint (the Gauss law)

$$\Phi_2 \equiv c^2\partial_i\Pi_i + eJ^0\,. \tag{11.331}$$

In turn, the requirement that $\{\Phi_1,\ H^{\rm EM}_{\rm T}\}_{\rm PB} \simeq 0$ leads to yet another constraint

$$\Phi_3 \equiv c^2\partial_i\Pi_i - c\,\partial_i\partial_i\phi\,. \tag{11.332}$$

There are no more constraints: conservation of the current J^μ implies vanishing of $\dot{\Phi}_2$, provided the function u^1 is set to zero, while u^0 can be adjusted to ensure $\dot{\Phi}_3 = 0$.

It is convenient to replace the constraint Φ_3 (11.332) by the linear combination $\Phi_{3'}$ of (11.332) and (11.331):

$$\Phi_{3'} \equiv c\,\partial_i\partial_i\phi + eJ^0 = 0\,. \tag{11.333}$$

This shows that in the quantum theory the operator $A^0(t,\mathbf{x}) \equiv \phi(t,\mathbf{x})$ is related to the operator J^0 by the identity:[90]

$$\phi(t,\mathbf{x}) = \frac{e}{4\pi c}\int d^3\mathbf{y}\,\frac{J^0(t,\mathbf{y})}{|\mathbf{x}-\mathbf{y}|}\,. \tag{11.334}$$

Thus, the operator $A^0(t,\mathbf{x}) \equiv \phi(t,\mathbf{x})$ becomes completely determined in terms of operators representing the matter canonical variables only (note, that J^0 given by (11.321) does not

[90]Recall that $\boldsymbol{\nabla}^2\frac{1}{|\mathbf{x}|} = -4\pi\,\delta^{(3)}(\mathbf{x})$.

depend on $\mathbf{v}_i$ and, hence, J^0, unlike $\mathbf{J}$, when expressed through $\hat{\mathbf{r}}_i$ and $\hat{\mathbf{P}}_i$ does not depend on $\mathbf{A}^i$). Since $\Phi_0 = 0$, the canonical momentum Π_0 must be represented by the zero operator, while the constraint $\Phi_{3'} = 0$ realized as an operator identity effectively eliminates the variable $\phi = A^0$ as an independent operator from the quantum theory.

The complete system of constraints Φ_0, Φ_1, Φ_2 and $\Phi_{3'}$ is of second class. The corresponding 4×4 matrix $C_{MN} = -C_{NM}$ is nonsingular - it has nonzero elements on its antidiagonal:

$$\begin{aligned} C_{1\mathbf{x},2\mathbf{y}} &= \{\Phi_1(\mathbf{x}),\ \Phi_2(\mathbf{y})\}_{\rm PB} \\ &= c^2\partial_i^{(\mathbf{x})}\partial_j^{(\mathbf{y})}\left\{A^i(\mathbf{x}),\ \Pi_j(\mathbf{y})\right\}_{\rm PB} = -c^2\partial_i^{(\mathbf{x})}\partial_i^{(\mathbf{x})}\delta^{(3)}(\mathbf{x}-\mathbf{y})\,, \end{aligned} \tag{11.335}$$

and

$$C_{0\mathbf{x},3'\mathbf{y}} = \{\Phi_0(\mathbf{x}),\ \Phi_{3'}(\mathbf{y})\}_{\rm PB} = \frac{1}{c}\,C_{1\mathbf{x},2\mathbf{y}} = -c\,\partial_k^{(\mathbf{x})}\partial_k^{(\mathbf{x})}\delta^{(3)}(\mathbf{x}-\mathbf{y})\,. \tag{11.336}$$

The canonical commutation relation which must be imposed on the operators A^i, Π_i, A^0 and Π_0 are determined by the Dirac prescription (11.308). The inverse $(C^{-1})^{N\mathbf{x},M\mathbf{y}}$ of the C_{MN} matrix has also nonzero elements only on its antidiagonal. These are

$$\left(C^{-1}\right)^{1\mathbf{x},2\mathbf{y}} = -\frac{1}{4\pi c^2|\mathbf{x}-\mathbf{y}|} = \frac{1}{c}\left(C^{-1}\right)^{0\mathbf{x},3'\mathbf{y}}\,, \tag{11.337}$$

Using this matrix one determines the basic commutator:

$$\begin{aligned} \frac{1}{i\hbar}\left[A^i(\mathbf{x}),\ \Pi_j(\mathbf{y})\right] &= \left\{A^i(\mathbf{x}),\ \Pi_j(\mathbf{y})\right\}_{\rm D} \equiv \left\{A^i(\mathbf{x}),\ \Pi_j(\mathbf{y})\right\}_{\rm PB} \\ &- \int d^3\mathbf{z}\int d^3\mathbf{w}\left\{A^i(\mathbf{x}),\ \Phi_2(\mathbf{z})\right\}_{\rm PB}\left(C^{-1}\right)^{2\mathbf{z},1\mathbf{w}}\{\Phi_1(\mathbf{w}),\ \Pi_j(\mathbf{y})\}_{\rm PB} \end{aligned} \tag{11.338}$$

(terms with the remaining pairs of constraints give all zero). Using the Poisson brackets

$$\begin{aligned} \left\{A^i(\mathbf{x}),\ \Phi_2(\mathbf{z})\right\}_{\rm PB} &= -c^2\partial_i^{(\mathbf{x})}\delta^{(3)}(\mathbf{x}-\mathbf{z})\,, \\ \{\Phi_1(\mathbf{w}),\ \Pi_j(\mathbf{y})\}_{\rm PB} &= \partial_j^{(\mathbf{w})}\delta^{(3)}(\mathbf{w}-\mathbf{y})\,, \end{aligned}$$

it is easy to find that

$$\left\{A^i(\mathbf{x}),\ \Pi_j(\mathbf{y})\right\}_{\rm D} = \delta^i_j\delta^{(3)}(\mathbf{x}-\mathbf{y}) + \partial_i^{(\mathbf{x})}\partial_j^{(\mathbf{x})}\frac{1}{4\pi|\mathbf{x}-\mathbf{y}|}\,. \tag{11.339}$$

Proceeding in the similar way, one can check that all the remaining Dirac brackets among the system of variables A^i, Π_j, A^0 and Π_0 vanish (vanishing of the Dirac brackets of $\Pi_0 \equiv \Phi_0$ with all functions of the canonical variables of the entire theory is ensured by the property (11.310)). In particular,

$$\left\{A^i(\mathbf{x}),\ A^j(\mathbf{y})\right\}_{\rm D} = \{\Pi_i(\mathbf{x}),\ \Pi_j(\mathbf{y})\}_{\rm D} = 0\,. \tag{11.340}$$

Therefore, the basic quantization conditions read

$$\begin{aligned} \left[A^i(\mathbf{x}),\ \Pi_j(\mathbf{y})\right] &= i\hbar\,\delta^i_j\,\delta^{(3)}(\mathbf{x}-\mathbf{y}) + i\hbar\,\partial_i^{(\mathbf{x})}\partial_j^{(\mathbf{x})}\frac{1}{4\pi|\mathbf{x}-\mathbf{y}|}\,, \\ \left[A^i(\mathbf{x}),\ A^j(\mathbf{y})\right] &= \left[\Pi_i(\mathbf{x}),\ \Pi_j(\mathbf{y})\right] = 0\,. \end{aligned} \tag{11.341}$$

The function on the right hand side of the first commutation rule (11.341) is called the *transverse delta function*. It has the following momentum space representation

$$\begin{aligned} &\delta^{ij}\delta^{(3)}(\mathbf{x}-\mathbf{y}) + \partial^i_{(\mathbf{x})}\partial^j_{(\mathbf{x})}\frac{1}{4\pi|\mathbf{x}-\mathbf{y}|} \\ &\qquad \equiv \delta^{ij}_{\mathrm{tr}}(\mathbf{x}-\mathbf{y}) = \int\frac{d^3\mathbf{k}}{(2\pi)^3}\,e^{i\mathbf{k}\cdot(\mathbf{x}-\mathbf{y})}\left(\delta^{ij} - \frac{k^ik^j}{|\mathbf{k}|^2}\right), \end{aligned} \tag{11.342}$$

which makes it explicit that

$$\partial_i^{\mathbf{x}}\left(\delta^i_j\delta^{(3)}(\mathbf{x}-\mathbf{y}) + \partial_i^{(\mathbf{x})}\partial_j^{(\mathbf{x})}\frac{1}{4\pi|\mathbf{x}-\mathbf{y}|}\right) = 0\,, \tag{11.343}$$

and, therefore, that the commutation rules (11.341) are compatible with the constraint (11.330).

The Hamiltonian density operator

$$\mathcal{H}^{\mathrm{EM}} = \frac{c^2}{2}\,\Pi_i\Pi_i + \frac{1}{2}\,(\boldsymbol{\nabla}\times\mathbf{A})^2 - c\,\Pi_i\partial_i\phi + \frac{1}{c}\,eJ^0\phi - \frac{1}{c}\,e\mathbf{J}\cdot\mathbf{A}\,, \tag{11.344}$$

in which ϕ is now a shorthand for the operator (11.334) and Φ_0 as well as Φ_1 have been set to zero, together with the quantization rules (11.341) specify in principle the dynamics of the quantized electromagnetic field. There is, however, one technical problem that the operator $\Pi_i(\mathbf{x})$ does not commute with matter sector canonical variables. This is because the Dirac brackets of $\Pi_i(\mathbf{x})$ with functions F depending on the matter sector canonical variables do not necessarily vanish.[91] Indeed, according to the Dirac prescription such commutators are given by $i\hbar$ times the Dirac brackets

$$\begin{aligned} &\{F(\mathrm{mat}),\ \Pi_i(\mathbf{x})\}_{\mathrm{D}} = \{F(\mathrm{mat}),\ \Pi_i(\mathbf{x})\}_{\mathrm{PB}} \\ &\qquad - \int d^3\mathbf{z}\int d^3\mathbf{w}\,\{F(\mathrm{mat}),\ \Phi_M(\mathbf{z})\}_{\mathrm{PB}}\left(C^{-1}\right)^{M\mathbf{z},N\mathbf{w}}\{\Phi_N(\mathbf{w}),\ \Pi_i(\mathbf{x})\}_{\mathrm{PB}}\,. \end{aligned}$$

The Poisson bracket in the first line is of course zero. To the second line contributes only the term with $MN = 21$. This line therefore gives

$$\begin{aligned} &-\int d^3\mathbf{z}\int d^3\mathbf{w}\,\{F(\mathrm{mat}),\ \Phi_2(\mathbf{z})\}_{\mathrm{PB}}\,\frac{1}{4\pi c^2|\mathbf{z}-\mathbf{w}|}\,\partial_i^{(\mathbf{w})}\delta^{(3)}(\mathbf{w}-\mathbf{x}) \\ &\qquad\qquad = \int d^3\mathbf{z}\,\left\{F(\mathrm{mat}),\ eJ^0(\mathbf{z})\right\}_{\mathrm{PB}}\,\partial_i^{(\mathbf{x})}\frac{1}{4\pi c^2|\mathbf{z}-\mathbf{x}|}\,. \end{aligned}$$

[91] It is easy to check that $\{F(\mathrm{mat}),\ A^i(\mathbf{x})\}_{\mathrm{D}} = 0$.

This is evidently nonzero, if F is a function of $\mathbf{P}_j$ because $J^0(\mathbf{z})$ given by (11.321) depends on $\mathbf{r}_i$, and $\{\mathbf{r}_i,\ \mathbf{P}_j\}_{\mathrm{PB}} = \delta_{ij}$. Dealing with such an operator would be difficult. Therefore, one defines another operator

$$\Pi_i^{\mathrm{T}} = \Pi_i - \frac{1}{c}\,\partial_i\phi\,, \tag{11.345}$$

which in fact is just $(1/c^2)\dot{A}^i$ (cf. (11.326)), i.e. the transverse (the divergenceless) part of the electric field, which has better properties. Firstly,

$$\begin{aligned}
-\frac{1}{c}\,\partial_i^{(\mathbf{x})}&\{F(\mathrm{mat}),\ \phi(\mathbf{x})\}_{\mathrm{D}}\\
&= \frac{1}{c}\,\partial_i^{(\mathbf{x})}\int d^3\mathbf{z}\int d^3\mathbf{w}\,\{F(\mathrm{mat}),\ \Phi_M(\mathbf{z})\}_{\mathrm{PB}}\left(C^{-1}\right)^{M\mathbf{z},N\mathbf{w}}\{\Phi_N(\mathbf{w}),\ \phi(\mathbf{x})\}_{\mathrm{PB}}\\
&= -\frac{1}{c}\,\partial_i^{(\mathbf{x})}\int d^3\mathbf{z}\left\{F(\mathrm{mat}),\ eJ^0(\mathbf{z})\right\}_{\mathrm{PB}}\frac{1}{4\pi c|\mathbf{z}-\mathbf{x}|}
\end{aligned}$$

(only the element $(C^{-1})^{3'\mathbf{z},0\mathbf{w}}$ contributes to the sum over constraints; the Poisson bracket $\{F(\mathrm{mat}),\ \phi(\mathbf{x})\}_{\mathrm{PB}}$ obviously vanishes), so this precisely cancels out the unwanted term in the Dirac bracket of Π_i with $F(\mathrm{mat})$. Hence, the operator Π_i^{T} does commute with the matter variables. Secondly, since

$$-\frac{1}{c}\,\partial_i^{(\mathbf{x})}\left\{A^i(\mathbf{x}),\ \phi(\mathbf{y})\right\}_{\mathrm{D}} = 0$$

(the Poisson bracket of $A^i(\mathbf{x})$ with $\phi(\mathbf{y})$ vanishes, A^i has nonzero Poisson bracket only with Φ_2, while ϕ only with Φ_0 and $(C^{-1})^{2\mathbf{z},0\mathbf{w}} = 0$, so the second term in the Dirac bracket vanishes too), it follows that

$$\left[A^i(\mathbf{x}),\ \Pi_j^{\mathrm{T}}(\mathbf{y})\right] = i\hbar\,\delta^i_j\,\delta^{(3)}(\mathbf{x}-\mathbf{y}) + i\hbar\,\partial_i^{(\mathbf{x})}\partial_j^{(\mathbf{x})}\frac{1}{4\pi|\mathbf{x}-\mathbf{y}|}\,, \tag{11.346}$$

i.e. the commutator of Π_j^{T} with A^i is the same as that of $\Pi_i(\mathbf{y})$. Finally, one has to check the Dirac bracket

$$\begin{aligned}
\left\{\Pi_i^{\mathrm{T}}(\mathbf{x}),\ \Pi_j^{\mathrm{T}}(\mathbf{y})\right\}_{\mathrm{D}} &= \{\Pi_i(\mathbf{x}),\ \Pi_j(\mathbf{y})\}_{\mathrm{D}} + \frac{1}{c^2}\,\partial_i^{(\mathbf{x})}\partial_j^{(\mathbf{y})}\,\{\phi(\mathbf{x}),\ \phi(\mathbf{y})\}_{\mathrm{D}}\\
&\quad -\frac{1}{c}\,\partial_i^{(\mathbf{x})}\{\phi(\mathbf{x}),\ \Pi_j(\mathbf{y})\}_{\mathrm{D}} - \frac{1}{c}\,\partial_j^{(\mathbf{y})}\{\Pi_i(\mathbf{x}),\ \phi(\mathbf{y})\}_{\mathrm{D}}\ .
\end{aligned}$$

As already has been established in (11.340), the first term on the right hand side is zero. To show that the remaining three terms also vanish, we note that

$$\begin{aligned}
\{\phi(\mathbf{x}),\ \phi(\mathbf{y})\}_{\mathrm{D}} &= \{\phi(\mathbf{x}),\ \phi(\mathbf{y})\}_{\mathrm{PB}}\\
&\quad -\int d^3\mathbf{z}\int d^3\mathbf{w}\,\{\phi(\mathbf{x}),\ \Phi_M(\mathbf{z})\}_{\mathrm{PB}}\left(C^{-1}\right)^{M\mathbf{z},N\mathbf{w}}\{\Phi_N(\mathbf{w}),\ \phi(\mathbf{y})\}_{\mathrm{PB}} = 0\,:
\end{aligned}$$

the Poisson bracket of two ϕ's vanishes and the second term vanishes because ϕ's have nonzero Poisson bracket only with Π_0 and $(C^{-1})^{0\mathbf{z},0\mathbf{w}} = 0$. To complete the proof, we consider $\left\{\Pi_i(\mathbf{x}),\ \partial_j^{\mathbf{y}}\phi(\mathbf{y})\right\}_{\mathrm{D}}$: the corresponding Poisson bracket is zero and

$$-\int d^3\mathbf{z}\int d^3\mathbf{w}\,\{\Pi_i(\mathbf{x}),\ \Phi_M(\mathbf{z})\}_{\mathrm{PB}}\left(C^{-1}\right)^{M\mathbf{z},N\mathbf{w}}\left\{\Phi_N(\mathbf{w}),\ \partial_j^{\mathbf{y}}\phi(\mathbf{y})\right\}_{\mathrm{PB}}$$

also vanishes because Π_i has a nonzero Poisson bracket only with Φ_1, ϕ only with Φ_0 and $(C^{-1})^{1\mathbf{z},0\mathbf{w}} = 0$. We have thus shown, that the new operator Π_i^{T} has all the properties required from the canonical momentum operator but in contrast to Π_i, it commutes with operators representing the matter sector variables. In addition, Π_i^{T} satisfies a simpler constraint than $\Phi_2 = 0$:

$$\Phi_{2'} \equiv \partial_i\Pi_i^{\mathrm{T}} = 0\,. \tag{11.347}$$

Expressing now the Hamiltonian density operator (11.344) through Π_i^{T} we obtain

$$\begin{aligned}\mathcal{H}^{\mathrm{EM}} &= \frac{c^2}{2}\,\Pi_i^{\mathrm{T}}\Pi_i^{\mathrm{T}} + \frac{1}{2}\left(\boldsymbol{\nabla}\times\mathbf{A}\right)^2 - \frac{1}{2}\partial_i\phi\partial_i\phi + \frac{1}{c}\,e\,J^0\phi - \frac{1}{c}\,e\mathbf{J}\cdot\mathbf{A}\\ &= \frac{c^2}{2}\Pi_i^{\mathrm{T}}\Pi_i^{\mathrm{T}} + \frac{1}{2}\left(\boldsymbol{\nabla}\times\mathbf{A}\right)^2 + \frac{1}{2c}\,eJ^0\phi - \frac{1}{c}\,e\,\mathbf{J}\cdot\mathbf{A}\,.\end{aligned} \tag{11.348}$$

The noncovariantly looking term $eJ^0\phi/2c$ (obtained after integrating by parts and using the constraint $\Phi_{3'} = 0$ (11.333)) produces in the Hamiltonian the term representing energy of the ordinary electrostatic Coulomb interaction

$$H^{\mathrm{EM}} \supset \frac{1}{2}\int d^3\mathbf{x}\int d^3\mathbf{y}\,\frac{e^2}{4\pi c^2}\,\frac{J^0(\mathbf{x})J^0(\mathbf{y})}{|\mathbf{x}-\mathbf{y}|}\,. \tag{11.349}$$

It remains to find the representation of the Schrödinger picture operators A^i and Π_i^{T} in terms of the creation and annihilation operators satisfying simple commutation rules and diagonalizing (if it is possible) the Hamiltonian $H_0^{\mathrm{EM}} = \int d^3\mathbf{x}\,\mathcal{H}_0^{\mathrm{EM}}$ obtained from (11.348) by setting $J^\mu = 0$. In addition, these representations should automatically ensure that the constraints $\Phi_1 = \Phi_{2'} = 0$ hold as operator identities. It is easy to guess that

$$A^i(\mathbf{x}) = \int\frac{d^3\mathbf{k}}{(2\pi)^3}\sqrt{\frac{\hbar c^2}{2\omega_{\mathbf{k}}}}\sum_{\lambda=\pm1}\left[\epsilon^i(\mathbf{k},\lambda)a_\lambda(\mathbf{k})e^{i\mathbf{k}\cdot\mathbf{x}} + \epsilon^{i*}(\mathbf{k},\lambda)a_\lambda^\dagger(\mathbf{k})e^{-i\mathbf{k}\cdot\mathbf{x}}\right], \tag{11.350}$$

where $\omega_{\mathbf{k}} = c|\mathbf{k}|$ and $\epsilon^i(\mathbf{k},\lambda)$ are two polarization vectors such that

$$\mathbf{k}\cdot\boldsymbol{\epsilon}(\mathbf{k},\lambda) = 0 \tag{11.351}$$

(so that the constraint $\Phi_1 \equiv \boldsymbol{\nabla}\cdot\mathbf{A} = 0$ is satisfied) and

$$\boldsymbol{\epsilon}(\mathbf{k},\lambda)\cdot\boldsymbol{\epsilon}^*(\mathbf{k},\lambda') = \delta_{\lambda\lambda'}\,. \tag{11.352}$$

Because together with $\mathbf{k}/|\mathbf{k}|$ the two vectors $\boldsymbol{\epsilon}(\mathbf{k},\lambda)$ form an orthonormal basis of the three-dimensional space, the following summation rule holds

$$\sum_{\lambda=\pm1} \epsilon^i(\mathbf{k},\lambda)\,\epsilon^{j*}(\mathbf{k},\lambda) = \delta^{ij} - \frac{k^i k^j}{|\mathbf{k}|^2}\,. \tag{11.353}$$

We then postulate the expansion of the operator Π_i^{T} in the form

$$\Pi_i^{\mathrm{T}}(\mathbf{x}) = \frac{1}{i}\int \frac{d^3\mathbf{k}}{(2\pi)^3}\sqrt{\frac{\hbar\omega_{\mathbf{k}}}{2c^2}}\sum_{\lambda=\pm1}\Big[\epsilon^i(\mathbf{k},\lambda)\,a_\lambda(\mathbf{k})\,e^{i\mathbf{k}\cdot\mathbf{x}} - \epsilon^{i*}(\mathbf{k},\lambda)\,a_\lambda^\dagger(\mathbf{k})\,e^{-i\mathbf{k}\cdot\mathbf{x}}\Big]. \tag{11.354}$$

It is then straightforward to check that with the commutation rules

$$\begin{aligned}\Big[a_\lambda(\mathbf{k}),\ a_{\lambda'}^\dagger(\mathbf{k}')\Big] &= (2\pi)^3\delta_{\lambda\lambda'}\delta(\mathbf{k}-\mathbf{k}')\,,\\ [a_\lambda(\mathbf{k}),\ a_{\lambda'}(\mathbf{k}')] = \Big[a_\lambda^\dagger(\mathbf{k}),\ a_{\lambda'}(\mathbf{k}')\Big] &= 0\,,\end{aligned} \tag{11.355}$$

imposed all the fundamental commutation relations between $A^i(\mathbf{x})$ and $\Pi_i^{\mathrm{T}}(\mathbf{x})$ together with the constraints $\Phi_1 = 0$, $\Phi_{2'} = 0$ are satisfied and the free part of the Hamiltonian (11.348) takes (after discarding an infinite constant) the form[92]

$$H_0^{\mathrm{EM}} = \int \frac{d^3\mathbf{k}}{(2\pi)^3}\,\hbar\omega_{\mathbf{k}}\sum_{\lambda=\pm1} a_\lambda^\dagger(\mathbf{k})a_\lambda(\mathbf{k})\,. \tag{11.356}$$

As the full Hamiltonian of the electromagnetic field coupled to a system of N charged particles we therefore get

$$\begin{aligned}H = &\int \frac{d^3\mathbf{k}}{(2\pi)^3}\,\hbar\omega_{\mathbf{k}}\sum_{\lambda=\pm1} a_\lambda^\dagger(\mathbf{k})a_\lambda(\mathbf{k}) + \sum_{i=1}^{N}\frac{1}{2m_i}\Big(\hat{\mathbf{P}}_i - \frac{q_i}{c}\,e\mathbf{A}(\hat{\mathbf{r}}_i)\Big)^2\\ &+ \sum_{i=1}^{N} V(\hat{\mathbf{r}}_i) + \frac{1}{2}\sum_{i\neq j=1}^{N} V(\hat{\mathbf{r}}_i - \hat{\mathbf{r}}_j) + \frac{1}{2}\int d^3\mathbf{x}\int d^3\mathbf{y}\,\frac{e^2}{4\pi c^2}\,\frac{J^0(\mathbf{x})J^0(\mathbf{y})}{|\mathbf{x}-\mathbf{y}|}\,.\end{aligned} \tag{11.357}$$

[92] By the rescaling $\sqrt{2\hbar\omega_{\mathbf{k}}}\,a_\lambda(\mathbf{k}) \to a_\lambda(\mathbf{k})$ the first of the rules (11.355) can be brought into the standard relativistic form

$$[a_\lambda(\mathbf{k}),\ a_{\lambda'}^\dagger(\mathbf{k}')] = \delta_{\lambda\lambda'}\delta_\Gamma(\mathbf{k}-\mathbf{k}')\,.$$

The expansions of the operators A^i and Π_i^{T} then take the simple forms

$$\begin{aligned}A^i(\mathbf{x}) &= \hbar c\int d\Gamma_{\mathbf{k}}\sum_{\lambda=\pm1}\left[\epsilon^i(\mathbf{k},\lambda)\,a_\lambda(\mathbf{k})\,e^{i\mathbf{k}\cdot\mathbf{x}} + \mathrm{H.c.}\right],\\ \Pi_i^{\mathrm{T}}(\mathbf{x}) &= -\frac{i}{c}\int d\Gamma_{\mathbf{k}}\,\hbar\omega_{\mathbf{k}}\left[\epsilon^i(\mathbf{k},\lambda)\,a_\lambda(\mathbf{k})\,e^{i\mathbf{k}\cdot\mathbf{x}} - \mathrm{H.c.}\right],\end{aligned}$$

and $d^3\mathbf{k}/(2\pi)^3$ in (11.356) gets replaced by $d\Gamma_{\mathbf{k}} \equiv d^3\mathbf{k}/(2\pi)^3 2\hbar\omega_{\mathbf{k}}$.

where $\mathbf{A}(\hat{\mathbf{r}}_i)$ is the operator (11.350) depending now on the charged particle position operators $\hat{\mathbf{r}}_i$. This dependence is introduced into $\mathbf{A}$ by the integral in $L_{\rm int}$ (11.320) and the delta functions in the current (11.321). Since in the considered NRQED J^0 is of the form (11.321), the last term in the Hamiltonian (11.357) is just

$$\frac{1}{2}\sum_{i,j=1}^{N}\frac{e^2 q_i q_j}{4\pi|\hat{\mathbf{r}}_i-\hat{\mathbf{r}}_j|}\,,$$

which is just the Coulomb electrostatic energy of a system of charged pointlike particles (in the Heaviside system of units). Note however, that the summation includes infinite terms arising for $i=j$ which have to be subtracted by hand.

There can be no question about the relativistic covariance of the theory just constructed, because matter particles are treated nonrelativistically. Therefore we consider here covariance of the quantum theory of the free (not coupled to matter particles) electromagnetic field. In this case from the Lagrangian density (11.258) one obtains the canonical energy-momentum tensor (we set $\hbar=c=1$)

$$T^{\mu\nu}_{\rm can}=-f^{\mu\lambda}\partial^\nu A_\lambda+\frac{1}{4}\,g^{\mu\nu}f^{\lambda\kappa}f_{\lambda\kappa}\,, \tag{11.358}$$

which is not symmetric and manifestly not gauge invariant (a disqualifying feature according to some). The symmetric Belinfante energy-momentum tensor obtained according to the prescriptions (11.58), (11.60) and (11.61) takes the form

$$T^{\mu\nu}_{\rm symm}=-f^{\mu\lambda}f^{\nu}{}_{\lambda}+\frac{1}{4}\,g^{\mu\nu}f^{\lambda\kappa}f_{\lambda\kappa}\,, \tag{11.359}$$

which is already symmetric and gauge invariant and, in the Coulomb gauge (11.329) gives[93]

$$P^0=\int d^3\mathbf{x}\,T^{00}_{\rm symm}=\frac{1}{2}\int d^3\mathbf{x}\left[\dot{A}^i\dot{A}^i+(\boldsymbol{\nabla}\times\mathbf{A})^2\right], \tag{11.360}$$

which is just (11.348) for zero external current J^μ and

$$P^i=\int d^3\mathbf{x}\,\Pi_k\partial^i A_k\,. \tag{11.361}$$

Of course, the same P^0 and P^i are obtained from the nonsymmetric and not gauge invariant canonical energy-momentum tensor (and in this sense it is not worse than the Belinfante one). The canonical tensor $M^{\mu\nu\lambda}_{\rm can}$ (11.62) of the electromagnetic field also takes the non gauge invariant form

$$M^{\mu\nu\lambda}_{\rm can}=x^\nu\,T^{\mu\lambda}_{\rm can}-x^\lambda\,T^{\mu\nu}_{\rm can}+if^{\mu}{}_{\kappa}\left(\mathcal{J}^{\nu\lambda}_{\rm vec}\right)^{\kappa}{}_{\rho}A^\rho\,. \tag{11.362}$$

[93] Since we consider here the free electromagnetic field, $A^0=0$ and $\Pi^{\rm T}_i=\Pi_i=\dot{A}^i$.

and in the Coulomb gauge yields the generators J^{ij} and J^{0i} of the Lorentz transformations

$$
\begin{aligned}
J^{ij} &= \int d^3\mathbf{x}\,\left\{\dot{A}^k(x^i\partial^j - x^j\partial^i)A^k - \left(\dot{A}^iA^j - \dot{A}^jA^i\right)\right\},\\
J^{0i} &= tP^i - \int d^3\mathbf{x}\,x^i\,\frac{1}{2}\left[\dot{A}^k\dot{A}^k + (\boldsymbol{\nabla}\times\mathbf{A})^2\right].
\end{aligned}
\tag{11.363}
$$

The operators $H \equiv P^0$, $J^i \equiv \frac{1}{2}\epsilon^{kl}J^{kl}$, P^i and $K^i \equiv J^{0i}$ can be shown to satisfy the Poincaré algebra commutation rules (6.21) by virtue of the canonical commutation relations (11.346). However, while

$$
\begin{aligned}
e^{ia_\mu P^\mu}A^k(x)\,e^{-ia_\mu P^\mu} &= A^k(x+a)\,,\\
e^{\frac{i}{2}\omega_{ij}J^{ij}}A^k(x)\,e^{-\frac{i}{2}\omega_{ij}J^{ij}} &= \left(e^{-\frac{i}{2}\omega_{ij}\mathcal{J}^{ij}_{\rm vec}}\right)^k_{\ l}A^l(\Lambda^{-1}(\omega)\cdot x)\,,
\end{aligned}
$$

as usually (here Λ is a pure rotation), performing the transformation of $A^k(x)$ corresponding to a Lorentz boost one finds that

$$
e^{i\omega_{0i}J^{0i}}A^k(x)\,e^{-i\omega_{0i}J^{0i}} = \left(e^{-i\omega_{0i}\mathcal{J}^{0i}_{\rm vec}}\right)^k_{\ l}A^l(\Lambda^{-1}(\omega)\cdot x) + \Delta A^k(x) \equiv A'^k(x)\,,
$$

where the extra term $\Delta A^k(x)$ ensures that the $\partial_k A'^k(x) = 0$ as it must be for consistency: the divergence of the left hand side of the above formula obviously vanishes. This means that transformations of the photon field operator $A^i(x)$ generated by K^i consist of the corresponding Lorentz boost supplemented with a suitable gauge transformation. Thus, transformations generated by the conserved Noether charges (11.358) and (11.362) in the Coulomb gauge automatically preserve this gauge.

Finally we consider the thermodynamical properties of the free electromagnetic field in equilibrium with the walls of a box of volume $V = L^3$. If the field is quantized in the box, the Hamiltonian takes the form

$$
H_{\rm EM} = \sum_{\mathbf{k}}\sum_{\lambda=\pm1}\hbar\omega_{\mathbf{k}}\,a^\dagger_{\mathbf{k}\lambda}a_{\mathbf{k}\lambda}\,,
\tag{11.364}
$$

where now the wave vectors $\mathbf{k}$ are discrete, $\mathbf{k} = (2\pi/L)\mathbf{n}$, $\omega_{\mathbf{k}} = c|\mathbf{k}|$ and

$$
[a_{\mathbf{k}'\lambda'},\ a^\dagger_{\mathbf{k}\lambda}] = \delta_{\mathbf{k}'\mathbf{k}}\delta_{\lambda'\lambda}\,.
\tag{11.365}
$$

The Hilbert space is spanned by the $H_{\rm EM}$ eigenstates $|n_{\mathbf{k}_1\lambda_+}, n_{\mathbf{k}_1\lambda_-}, n_{\mathbf{k}_2\lambda_+}\dots\rangle$, where each occupation number $n_{\mathbf{k}_i\lambda_i}$ can run from 0 to infinity. This allows to easily compute the Gibbs Canonical Ensemble partition function

$$
\begin{aligned}
Z_{\rm stat} &= \prod_{\mathbf{k}}\left\{\sum_{n_{\mathbf{k}\lambda_+}=0}^{\infty}\left(e^{-\hbar\omega_{\mathbf{k}}/k_{\rm B}T}\right)^{n_{\mathbf{k}\lambda_+}}\sum_{n_{\mathbf{k}\lambda_-}=0}^{\infty}\left(e^{-\hbar\omega_{\mathbf{k}}/k_{\rm B}T}\right)^{n_{\mathbf{k}\lambda_-}}\right\}\\
&= \prod_{\mathbf{k}}\left(\frac{1}{1-e^{-\hbar\omega_{\mathbf{k}}/k_{\rm B}T}}\right)^2.
\end{aligned}
\tag{11.366}
$$

The Helmholtz free energy is then[94]

$$F(T,V)=k_{\rm B}T\sum_{\mathbf{k}}2\ln\left(1-e^{-\hbar\omega_{\mathbf{k}}/k_{\rm B}T}\right),$$

or, in the thermodynamical limit,

$$\begin{aligned}F(T,V)&=k_{\rm B}T\,\frac{2V}{(2\pi)^3}\int d^3\mathbf{k}\ln\left(1-e^{-\hbar\omega_{\mathbf{k}}/k_{\rm B}T}\right)\\&=k_{\rm B}T\,\frac{V}{\pi^2c^3}\int_0^\infty d\omega\,\omega^2\ln\left(1-e^{-\hbar\omega/k_{\rm B}T}\right).\end{aligned}\tag{11.367}$$

All well known properties of the electromagnetic radiation in equilibrium, including the celebrated equation of state $p=\frac{1}{3}u(T)$, where $u(T)=U(T)/V$ is the radiation energy density, can be derived from the free energy $F(T,V)$.

11.8 Canonical construction of half-integer spin quantum fields

As demonstrated in Sections 11.2, 11.3, 11.5 and 11.7, the formal procedure of canonical quantization (supplemented in the case of systems subject to constraints by the Dirac methods described in Section 11.6) applied to classical (c-number) fields, which under spatial rotations of the reference frame transform as integer spin representations of the rotation group, allows to build theories of quantum fields elementary excitations ("quanta") of which are bosons. In the interaction picture (see Section 11.9), i.e. when the basic field operators of such quantum theories are represented in Fock spaces of eigenvectors of the free Hamiltonians of such theories, one recovers the same Feynman rules for computing S-matrix elements, which in Chapters 8 and 9 were obtained in the approach based on quantum mechanics of relativistic particles, without any reference to classical fields.

Apart from making (owing to the Noether theorem presented in Section 11.1) extraction of consequences of various possible symmetries more straightforward, the most important virtue of the approach based on field quantization is that being essentially nonperturbative, it leads more straightforwardly to a deeper insight into the structure of the quantum theory (see Chapter 13).

Here we want to lay similar foundations for relativistic interactions of half-integer spin particles. Superficially, it could seem it should suffice to apply the same formalism as previously to classical c-number fields but transforming under changes of the Lorentz frame as half-integer spin representations of the $Spin(1,3)\simeq SL(2,C)$ group (the universal covering of the Lorentz $SO(1,3)$ group). This indeed seems to be so in the case of the best known kind of half-integer spin particles - massive charged fermions accompanied by their

[94]Had we not subtracted the (infinite) energy of the zero point oscillations, we would get in $F(T,V)$ an extra term $\sum_{\mathbf{k}}\hbar\omega_{\mathbf{k}}$.

antifermions: as in many textbooks one can in this case start with a relativistically invariant Lagrangian density of a c-number complex Dirac field ψ transforming as the $(\frac{1}{2},\frac{1}{2})$ representation of the $SL(2,C)$ group arguing that to obtain a sensible quantum theory it is necessary to replace the Poisson brackets with the anticommutators instead of commutators (one usually passes freely over the fact that the ordinary Poisson brackets and aticommutators have different symmetry properties and over the presence of constraints in this case). Yet, relativistic interactions of not all types of half-integer spin particles can be obtained in this way. The simplest counterexample is the massive neutral spin 1/2 particle (called Majorana fermion) the field operator of which was constructed at the end of Section 8.3. This is because already the corresponding classical free Lagrangian density cannot simply be written down: in the expression

$$\mathcal{L} = \bar{\lambda} i\bar{\sigma}^{\mu}\partial_{\mu}\lambda - \frac{1}{2}m(\lambda\lambda + \bar{\lambda}\bar{\lambda}) \tag{11.368}$$

(see Section 8.6 for the Lorentz transformation properties of λ_α and $\bar{\lambda}^{\dot{\alpha}}$) the term proportional to m is simply zero, if λ_α and $\bar{\lambda}^{\dot{\alpha}}$ are treated as c-number spinors. This shows that the consistent field-based approach to constructing quantum theories of fermions must start with Lagrangian densities which are (bosonic) functions of fields taking values in the Grassmann algebras generated by infinite sets of *anticommuting* generators.

To get acquainted with these notions it is convenient to consider first a classical bosonic system having a finite number of degrees of freedom. Dynamics of such a system the state of which is characterized by the values of n generalized (real) variables q^i can be formulated as a mapping from $\mathbb{R}$ (time) into the n-fold direct product $\mathcal{F}\times\ldots\times\mathcal{F}$ of an abstract algebra $\mathcal{F}$ of functions $f(a^1,\ldots,a^{2n})$ of $2n$ variables $a^1,\ldots,a^{2n}$ which are identified with initial values $q_0^1,\ldots,q_0^n,\dot{q}_0^1,\ldots,\dot{q}_0^n$ of the variables $q^i(t)$ and the corresponding generalized velocities $\dot{q}^i(t)$. Indeed, the values $q^i(t)$ at a fixed instant t can be treated as a set of n functions $q^i(t) = f^i(t,a^1,\ldots,a^{2n})$ of the variables $a^1,\ldots,a^{2n}$. With an appropriate topology (to define convergence of sequences) functions like $f^i(t,a^1,\ldots,a^{2n})$ can be viewed as elements of an abstract commutative and associative algebra over $\mathbb{R}$ generated by the set of $2n$ commuting generators $a^1,\ldots,a^{2n}$ ($a^k a^l = a^l a^k$):

$$f(a^1,\ldots,a^{2n}) = f_0 + \sum_k^{2n} f_k\, a^k + \frac{1}{2}\sum_{k_1,k_2}^{2n} f_{k_1k_2}\, a^{k_1}a^{k_2} + \frac{1}{3!}\sum_{k_1,k_2,k_3}^{2n} f_{k_1k_2k_3}\, a^{k_1}a^{k_2}a^{k_3} + \ldots \tag{11.369}$$

Differentiation can in the algebra $\mathcal{F}$ be defined as a linear algebraic operation by the basic rules

$$\frac{\partial}{\partial a^j}\, a^i = \delta^i_j\,, \qquad \text{and} \qquad \frac{\partial}{\partial a^k}\, a^1\ldots a^k\ldots a^l = \frac{\partial}{\partial a^k}\, a^k a^1\ldots a^l = a^1\ldots a^l\,.$$

and their extension to arbitrary elements of $\mathcal{F}$ using linearity. Thus at fixed instant t each $q^i(t)$ is an element of $\mathcal{F}$; a classical trajectory $(q^1(t),\ldots,q^n(t))$ can be, therefore, viewed as a mapping $t\to\mathcal{F}\times\ldots\times\mathcal{F}$.

A Lagrangian which determines the dynamics of such a system is in this picture a mapping of the $2n$-fold direct product of the algebra $\mathcal{F}$ into $\mathcal{F}$ itself: $\mathcal{F}\times\ldots\times\mathcal{F}\to L(f^1,\ldots,f^{2n})\in\mathcal{F}$ and the action functional $I[q(t)]$ is given by

$$\begin{aligned}I[q(t)] &= \int_{t_1}^{t_2} dt\, L(a^1,\ldots,a^{2n})|_{a^1=q^1(t),\ldots,a^{2n}=\dot{q}^{2n}(t)}\\ &\equiv \int_{t_1}^{t_2} dt\; L(a^1,\ldots,a^{2n})\big| \equiv \int_{t_1}^{t_2} dt\, L(q^1(t),\ldots,\dot{q}^{2n}(t))\,.\end{aligned} \tag{11.370}$$

Its variation with respect to the trajectory is to be understood as

$$\delta I[q(t)] = \int_{t_1}^{t_2} dt \sum_{i=1}^{n}\left(\frac{\partial L}{\partial a^i}\bigg|\,\delta q^i(t) + \frac{\partial L}{\partial a^{n+i}}\bigg|\,\frac{d}{dt}\,\delta q^i(t)\right), \tag{11.371}$$

where each $\delta q^i(t)$ is again a mapping from $\mathbb{R}$ into $\mathcal{F}$ which reduces to the zero element of $\mathcal{F}$ for $t=t_1$ and $t=t_2$.

The fermionic counterpart of the algebra $\mathcal{F}$ of functions described above is the Grassmann algebra $\mathcal{G}$ over $\mathbb{C}$ generated by $2m$ elements (generators) $\xi^1,\ldots,\xi^{2m}$ which anticommute:

$$\xi^\alpha\xi^\beta = -\xi^\beta\xi^\alpha\,. \tag{11.372}$$

Each element of $\mathcal{G}$ is of the form

$$g = g_0 + \sum_\alpha g_\alpha\,\xi^\alpha + \sum_{\alpha_1<\alpha_2} g_{\alpha_1\alpha_2}\,\xi^{\alpha_1}\xi^{\alpha_2} + + \sum_{\alpha_1<\ldots<\alpha_{2m}} g_{\alpha_1\ldots\alpha_{2m}}\,\xi^{\alpha_1}\ldots\xi^{\alpha_{2m}}\,. \tag{11.373}$$

In contrast to (11.369), the number of terms in the above sum is finite owing to the property (11.372) which implies that $\xi^\alpha\xi^\alpha=0$ (the zero element of the algebra $\mathcal{G}$). In the natural way the Grassmann algebra splits into the direct sum of $\mathcal{G}_{\rm even}$ and $\mathcal{G}_{\rm odd}$ (the only common element of these two subspaces being the zero element of $\mathcal{G}$) the elements of which have the general forms

$$\begin{aligned}g_{\rm even} &= g_0 + \sum_{\alpha_1<\alpha_2} g_{\alpha_1\alpha_2}\,\xi^{\alpha_1}\xi^{\alpha_2} + \ldots + \sum_{\alpha_1<\ldots<\alpha_{2m}} g_{\alpha_1\ldots\alpha_{2m}}\,\xi^{\alpha_1}\ldots\xi^{\alpha_{2m}}\,,\\ g_{\rm odd} &= \sum_\alpha g_\alpha\,\xi^\alpha + \ldots + \sum_{\alpha_1<\ldots<\alpha_{2m-1}} g_{\alpha_1\ldots\alpha_{2m-1}}\,\xi^{\alpha_1}\ldots\xi^{\alpha_{2m-1}}\,,\end{aligned}$$

respectively. Elements of $\mathcal{G}_{\rm even}$ and $\mathcal{G}_{\rm odd}$ can be distinguished by assigning them the $\mathcal{G}$-algebra parities P_g which assume values 0 and 1 respectively (a general element of $\mathcal{G}$ has no definite parity).

In the algebra $\mathcal{G}$ one defines two linear operations called left- and right-derivatives with respect to the generator ξ^α by the rules

$$\frac{\partial}{\partial_L\xi^\alpha}\,\xi^\beta = \frac{\partial}{\partial_R\xi^\alpha}\,\xi^\beta = \delta^\beta_\alpha\,,$$

$$\frac{\partial}{\partial_L \xi^\alpha}\,\xi^\gamma \dots \xi^\alpha \dots \xi^\sigma = (-1)^L \frac{\partial}{\partial_L \xi^\alpha}\,\xi^\alpha \xi^\gamma \dots \xi^\sigma = (-1)^L \xi^\gamma \dots \xi^\sigma\,,$$
$$\frac{\partial}{\partial_R \xi^\alpha}\,\xi^\gamma \dots \xi^\alpha \dots \xi^\sigma = (-1)^R \frac{\partial}{\partial_R \xi^\alpha}\,\xi^\gamma \dots \xi^\sigma \xi^\alpha = (-1)^R \xi^\gamma \dots \xi^\sigma\,,$$

where L (R) stands for the numbers of interchanges of the generators needed to place ξ^α to the left (right) of the string of the generators.

Analogously to the bosonic case, "classical" dynamics of a fermionic system described by m variables ψ^α is a mapping from $\mathbb{R}$ into an m-fold direct product $\mathcal{G}_{\rm odd} \times \dots \times \mathcal{G}_{\rm odd}$ with the generators $\xi^1, \dots, \xi^m$ and $\xi^{m+1}, \dots, \xi^{2m}$ having the interpretation of the initial "values" ψ_0^α and $\dot{\psi}_0^\alpha$ of $\psi^\alpha(t)$ and $\dot{\psi}^\alpha(t)$, $\alpha = 1, \dots, m$. The Lagrangian is in this case a mapping of the $2m$-fold product $\mathcal{G}_{\rm odd} \times \dots \times \mathcal{G}_{\rm odd}$ into $\mathcal{G}_{\rm even}$:

$$L(\psi^1(t), \dots, \dot{\psi}^1(t), \dots) \equiv L(\xi^1, \dots, \xi^{2m})|_{\xi^1 = \psi^1(t), \dots,\ \xi^{m+1} = \dot{\psi}^1(t), \dots}\,. \tag{11.374}$$

Dynamics (the equations of motion) follows from the condition $\delta I[\psi(t)] = 0$ where the variation is written in terms of the right derivatives[95]

$$\delta I[\psi(t)] = \int_{t_1}^{t_2} dt \sum_{\alpha=1}^{m} \left(\frac{\partial L}{\partial_R \xi^\alpha}\bigg|\, \delta\psi^\alpha(t) + \frac{\partial L}{\partial_R \xi^{m+\alpha}}\bigg|\, \frac{d}{dt}\,\delta\psi^\alpha(t) \right), \tag{11.375}$$

where each $\delta\psi^\alpha(t)$ is an arbitrary mapping from $\mathbb{R}$ into $\mathcal{G}_{\rm odd}$ reducing to the zero element of $\mathcal{G}_{\rm odd}$ for $t = t_1$ and $t = t_2$. The resulting "classical" dynamics is then rather abstract and, as has been discussed in the introduction to this chapter, no picture of fluctuating fields can be associated with the corresponding quantum theory.

Finally, the mathematical structure allowing for a uniform treatment of mixed bosonic and fermionic systems is the Bieriezin algebra $\mathcal{B}$ obtained from the Grassmann algebra by treating the coefficient functions $g_{\alpha_1 \dots}$ in (11.373) as elements of the algebra $\mathcal{F}$ of functions.[96] Thus, the Bieriezin algebra is generated by $2n + 2m$ elements (generators) z^a, $z^i = a^i$, $i = 1, \dots, 2n$ and $z^{2n+\alpha} = \xi^\alpha$, $\alpha = 1, \dots, 2m$, which have the property

$$z^a z^b = (-1)^{P_a P_b} z^b z^a\,, \tag{11.376}$$

with the natural assignments of the $\mathcal{B}$-algebra parities P_a (which as in the Grassmann algebra can assume values 0 and 1). Obviously, the parity of a product $b_1 b_2$ of two elements of $\mathcal{B}$ having well-defined parities P_{b_1} and P_{b_2} is $P_{b_1} + P_{b_2}(\text{mod } 2)$ and $b_1 b_2 = (-1)^{P_{b_1} P_{b_2}} b_2 b_1$. Similarly as the Grassmann algebra, the Bieriezin algebra naturally splits into the direct sum of the even and odd subspaces $\mathcal{B}_{\rm even}$ and $\mathcal{B}_{\rm odd}$ the elements of which have well defined parities (0 and 1, respectively). Left- and right derivatives with respect to the generators

[95] Alternatively it can be written in terms of the left derivatives with the variations placed to the left - see the formula (11.378).

[96] Of course real and imaginary parts of each $g_{\alpha_1 \dots}$ are treated as two independent elements of $\mathcal{F}$.

z^a are linear operations defined by the rules similar to the ones in the $\mathcal{G}$-algebra

$$\frac{\partial}{\partial_L z^a} z^{b_1} \dots z^{b_k} = \sum_{i=1}^{k} (-1)^{L_i} \delta_a^{b_i} z^{b_1} \dots (\text{no}\, z^{b_i}) \dots z^{b_k} ,$$
$$\frac{\partial}{\partial_R z^a} z^{b_1} \dots z^{b_k} = \sum_{i=1}^{k} (-1)^{R_i} \delta_a^{b_i} z^{b_1} \dots (\text{no}\, z^{b_i}) \dots z^{b_k} , \tag{11.377}$$

where L_i (R_i) are the numbers of odd generator interchanges needed to place z^{b_i} on the extreme left (right) of the string of the generators. Left- and right-derivatives with respect to bosonic generators a^i are, of course, identical. A useful property of derivatives with respect to fermionic generators ξ^α are

$$\frac{\partial}{\partial_L \xi^\alpha} (b_{\text{odd}}) = \frac{\partial}{\partial_R \xi^\alpha} (b_{\text{odd}}) ,$$
$$\frac{\partial}{\partial_L \xi^\alpha} (b_{\text{even}}) = -\frac{\partial}{\partial_R \xi^\alpha} (b_{\text{even}}) . \tag{11.378}$$

Dynamics of a mixed bosonic-fermionic system is determined by the action principle

$$\delta I[q, \dot{q}] = \int_{t_1}^{t_2} dt \left(\frac{\partial L}{\partial_R q^a} \delta q^a + \frac{\partial L}{\partial_R \dot{q}^a} \frac{d}{dt} \delta q^a \right) = 0 ,$$

in which $L(z^1, \dots, z^{2n+2m})$ is a function taking values in $\mathcal{B}_{\text{even}}$ and its derivatives with respect to q^a and $\dot{q}^a$ have to be understood as in (11.371) and (11.375). The equations of motion take the standard form

$$\frac{d}{dt} \frac{\partial L}{\partial_R \dot{q}^a} = \frac{\partial L}{\partial_R q^a} \qquad a = 1, \dots, n+m . \tag{11.379}$$

Hamiltonization of the dynamics of a bosonic-fermionic system is analogous to the one of the ordinary bosonic systems except that one has to fix a convention (which is arbitrary) for the definition of the canonical momenta. We chose to define the momenta $p_a(t)$ as *right* derivatives

$$p_a(t) = \frac{\partial L}{\partial_R \dot{q}^a(t)} . \tag{11.380}$$

Consequently, the Hamiltonian must take the form

$$H = \sum_b p_b \, \dot{q}^b - L , \tag{11.381}$$

in order for its differential dH to be independent of the differentials $d\dot{q}^a$ of the velocities. Indeed, with this definition, using

$$dp_b = \frac{\partial}{\partial_R q^a} \left(\frac{\partial L}{\partial_R \dot{q}^b} \right) dq^a + \frac{\partial}{\partial_R \dot{q}^a} \left(\frac{\partial L}{\partial_R \dot{q}^b} \right) d\dot{q}^a,$$

one can write the differential of H

$$dH = \frac{\partial H}{\partial_R q^a}\, dq^a + (-1)^{P_a P_b} \frac{\partial}{\partial_R \dot{q}^a} \left(\frac{\partial L}{\partial_R \dot{q}^b}\right) \dot{q}^b\, d\dot{q}^a,$$

in the form

$$dH = \frac{\partial H}{\partial_R q^a}\, dq^a - \frac{\partial}{\partial_R q^a} \left(\frac{\partial L}{\partial_R \dot{q}^b}\right) dq^a\, \dot{q}^b + dp_b\, \dot{q}^b = dp_b\, \dot{q}^b - \frac{\partial L}{\partial_R q^b}\, dq^b\,.$$

The Poisson bracket of the two functions F and G of definite $\mathcal{B}$-algebra parities must be defined in agreement with the adopted convention. The correct definition is

$$\{F,\ G\}_{\rm PB} = \sum_a \left(\frac{\partial F}{\partial_R q^a} \frac{\partial G}{\partial_L p_a} - (-1)^{P_F P_G} \frac{\partial G}{\partial_R q^a} \frac{\partial F}{\partial_L p_a}\right), \qquad (11.382)$$

with the *left* derivatives with respect to the momenta. It has the easy to check properties:

$$\{F,\ G\}_{\rm PB} = -(-1)^{P_F P_G}\, \{G,\ F\}_{\rm PB}\ ,$$
$$\{F,\ GH\}_{\rm PB} = \{F,\ G\}_{\rm PB}\, H + (-1)^{P_F P_G} G\, \{F,\ H\}_{\rm PB}\ .$$

With the definition (11.382) of the Poisson bracket in the $\mathcal{B}$-algebra the canonical equations of motion $\dot{q}^a = \{q^a,\ H\}_{\rm PB}$, $\dot{p}_a = \{p_a,\ H\}_{\rm PB}$, are, in the case of systems not subject to constraints, equivalent to the Euler-Lagrange equations derived from the action principle $\delta I[q(t)] = 0$.

Canonical quantization of such systems consists of promoting their canonical variables q^i, p_i and ψ^α, π_α to Schrödinger picture operators $\hat{q}^i$, $\hat{p}_i$, $\hat{\psi}^\alpha$ and $\hat{\pi}_\alpha$ satisfying the (anti)commutation relations:

$$[\hat{q}^i,\ \hat{p}_j] = i\hbar\{q^i,\ p_j\}_{\rm PB}\,, \qquad \{\hat{\psi}^\alpha,\ \hat{\pi}_\beta\}_+ = i\hbar\{\psi^\alpha,\ \pi_\beta\}_{\rm PB}\,, \qquad (11.383)$$

in agreements with the symmetry/antisymmetry properties (11.383) of the corresponding Poisson brackets. The resulting equations satisfied by the Heisenberg picture operators $\hat{q}^i_H(t)$ and $\hat{\psi}^\alpha_H(t)$ are then formally identical with the classical Euler-Lagrange equations (11.382).

The Dirac procedure allowing to handle systems subject to second class constraints extends to the case of mixed bosonic-fermionic systems essentially without modifications. The only difference is that the C matrix defined by (11.307) is in this case a supermatrix and has more complicated symmetry/antisymmetry properties (the Poisson bracket (11.382) in the $\mathcal{B}$-algebra is not simply antisymmetric).

Transition from classical mechanics to classical field theory is achieved, as usually, by passing with the number of variables q^a (and therefore also with the number of generators of the Bieriezin algebra) to the continuum, that is by ascribing a certain number of independent variables to each space point $\mathbf{x}$.

As an example of the application of this formalism we consider here quantization of the relativistic Grassmann field ψ transforming as the spinor representation of the Lorentz group, or more precisely, as the representation of the group $Spin(1,3)$ - the universal covering of the Lorentz $SO(1,3)$ group. The extension to $Spin(1,3)$ is possible because (what is clearly reflected in the formalism of the Grassmann and Bieriezin algebras) such fields are not classically measurable; all observables must be bosonic (i.e. belong to $\mathcal{G}_{\rm even}$ or $\mathcal{B}_{\rm even}$) that is bilinear or quadrilinear etc. in the fields ψ and transform as representations (scalar, vector, tensors) of the true Lorentz group.[97] We assume therefore that under the change of the reference frame $x \to x' = \Lambda\cdot x$ the Grassmann algebra valued field ψ transforms as

$$\psi'(x') = e^{-\frac{i}{2}\omega_{\mu\nu}\mathcal{J}^{\mu\nu}_{\rm spin}}\,\psi(x)\,, \tag{11.384}$$

where the matrices $\mathcal{J}^{\mu\nu}_{\rm spin}$ of the spinor representation satisfy the commutation rule

$$\left[\mathcal{J}^{\mu\nu}_{\rm spin},\ \mathcal{J}^{\lambda\rho}_{\rm spin}\right] = i\left(g^{\mu\rho}\mathcal{J}^{\nu\lambda}_{\rm spin} - g^{\mu\lambda}\mathcal{J}^{\nu\rho}_{\rm spin} - g^{\nu\rho}\mathcal{J}^{\mu\lambda}_{\rm spin} + g^{\nu\lambda}\mathcal{J}^{\mu\rho}_{\rm spin}\right). \tag{11.385}$$

As discussed in Section 8.3 the matrices $\mathcal{J}^{\mu\nu}_{\rm spin}$ can be constructed either by means of the Clifford algebras or by exploiting the isomorphism of $Spin(1,3)$ and $SL(2,C)$. Here, as in (8.53) we take

$$\mathcal{J}^{\mu\nu}_{\rm spin} = \frac{i}{4}\,[\gamma^\mu,\ \gamma^\nu] \equiv \frac{1}{2}\,\sigma^{\mu\nu}_{4\times4}\,, \tag{11.386}$$

with the 4×4 matrices γ^μ satisfying the basic Clifford algebra relation (8.52) in the representations (8.64) or (8.66).

The most frequently encountered Lagrangian density (leading to the theory of noninteracting fermions) has the form[98]

$$\mathcal{L} = i\psi^\dagger\gamma^0\gamma^\mu\partial_\mu\psi - m\,\psi^\dagger\gamma^0\psi \equiv \bar\psi(i\gamma^\mu\partial_\mu - m)\psi \tag{11.387}$$

(Lorentz invariance of this Lagrangian density can be shown using the formulae given in Chapter 8). The Grassmann algebra-valued fields ψ_α and $\psi^\dagger_\alpha$ are four-component spinors. They are all treated as independent Grassmann algebra valued variables. The Euler-Lagrange equations derived from (11.387)

$$\begin{aligned}&\gamma^0\,(i\gamma^\mu\partial_\mu - m)\,\psi = 0\,,\\ &\bar\psi\left(-i\gamma^\mu\overleftarrow{\partial}_\mu - m\right) = 0\,,\end{aligned} \tag{11.388}$$

[97] At the "classical" level this is the only argument why to consider fields transforming under a larger symmetry group than $SO(1,3)$ despite the fact that only the Lorentz group of symmetries can be inferred from physical experience. In the quantum case, as was discussed in Chapter 4, going over to the universal covering of the $SO(1,3)$ group is a natural consequence of the possible occurrence of the projective representations of $SO(1,3)$.

[98] Another, less frequently encountered, is the symmetric form:

$$\mathcal{L} = \frac{i}{2}\,\psi^\dagger\gamma^0\gamma^\mu\partial_\mu\psi - \frac{i}{2}\,\partial_\mu\psi^\dagger\gamma^0\gamma^\mu\psi - m\,\psi^\dagger\gamma^0\psi\,.$$

Using the Dirac formalism it is easy to check that it leads to the same quantum theory as (11.387).

are just the Dirac equations.

To set up the canonical formalism for this system we first find the canonical momenta (recall that the derivatives are the right ones)

$$\Pi_\psi = \frac{\partial\mathcal{L}}{\partial_R\dot{\psi}} = i\psi^\dagger\,, \qquad \Pi_{\psi^\dagger} = \frac{\partial\mathcal{L}}{\partial_R\dot{\psi}^\dagger} = 0\,. \tag{11.389}$$

It is clear that it is impossible to express $\dot{\psi}$ and $\dot{\psi}^\dagger$ through Π_ψ and $\Pi_{\psi^\dagger}$. The system is therefore subject to constraints (we suppress spinor indices)

$$\Phi_1(\mathbf{x}) \equiv \Pi_\psi(\mathbf{x}) - i\psi^\dagger(\mathbf{x}) = 0\,, \qquad \Phi_2(\mathbf{x}) \equiv \Pi_{\psi^\dagger}(\mathbf{x}) = 0\,, \tag{11.390}$$

which are of second class. Although in this case the Dirac procedure described in Section 11.6 is not really indispensable, because the canonical equations when restricted to ψ and Π_ψ are fully equivalent to the original Euler-Lagrange equation (11.388), it is instructive to go through this procedure to see how it leads to the well established results.

Constructing the extended Hamiltonian density according to the rules of Section 11.6 one gets

$$\begin{aligned}\mathcal{H}_\mathrm{T} &= \Pi_\psi\dot{\psi} + \Pi_{\psi^\dagger}\dot{\psi}^\dagger - \psi^\dagger i\dot{\psi} - \bar{\psi}i\gamma^i\partial_i\psi + m\,\bar{\psi}\psi \\ &= \psi^\dagger\gamma^0\left(-i\gamma^i\partial_i + m\right)\psi + \Phi_1 u^1 + \Phi_2 u^2\,.\end{aligned} \tag{11.391}$$

The functions u^1 and u^2 are just the velocities $\dot{\psi}$ and $\dot{\psi}^\dagger$. Equating to zero the Poisson brackets of the constraints with the Hamiltonian $H_\mathrm{T} = \int d^3\mathbf{x}\,\mathcal{H}_\mathrm{T}$

$$\begin{aligned}\{\Phi_1,\ H_\mathrm{T}\}_\mathrm{PB} &= -\bar{\psi}(i\gamma^k\overleftarrow{\partial}_k + m) - iu^2\,, \\ \{\Phi_2,\ H_\mathrm{T}\}_\mathrm{PB} &= \gamma^0(-i\gamma^k\partial_k + m)\psi - iu^1\,,\end{aligned}$$

determines these functions. No new (secondary) constraints are needed. Combining u^1 and u^2 obtained in this way with the canonical equations $\dot{\psi} = \{\psi,\ H_\mathrm{T}\}_\mathrm{PB} = u^1$, $\dot{\psi}^\dagger = \{\psi^\dagger,\ H_\mathrm{T}\}_\mathrm{PB} = u^2$ one obtains the equations which are fully equivalent to the Euler-Lagrange ones (11.388).

In order to quantize the theory one computes the Dirac brackets. The nonzero elements of the C matrix are $C_{1\mathbf{x},2\mathbf{y}} = C_{2\mathbf{x},1\mathbf{y}} \equiv \{\Phi_1(\mathbf{x}),\,\Phi_2(\mathbf{y})\}_\mathrm{PB} = -i\delta^{(3)}(\mathbf{x}-\mathbf{y})$ (it is in this case symmetric). Its inverse reads

$$(C^{-1})^{1\mathbf{x},2\mathbf{y}} = (C^{-1})^{2\mathbf{x},1\mathbf{y}} = i\delta^{(3)}(\mathbf{x}-\mathbf{y})\,. \tag{11.392}$$

The Dirac brackets computed as in (11.309) read

$$\begin{aligned}&\{\psi(\mathbf{x}),\ \Pi_\psi(\mathbf{y})\}_\mathrm{D} = \delta^{(3)}(\mathbf{x}-\mathbf{y})\,, \\ &\{\psi(\mathbf{x}),\ \psi^\dagger(\mathbf{y})\}_\mathrm{D} = -i\delta^{(3)}(\mathbf{x}-\mathbf{y})\,, \\ &\{\psi(\mathbf{x}),\ \Pi_{\psi^\dagger}(\mathbf{y})\}_\mathrm{D} = 0\,, \\ &\{\psi^\dagger(\mathbf{x}),\ \Pi_\psi(\mathbf{y})\}_\mathrm{D} = 0\,, \\ &\{\Pi_{\psi^\dagger}(\mathbf{x}),\ \Pi_\psi(\mathbf{y})\}_\mathrm{D} = 0\,, \\ &\{\psi^\dagger(\mathbf{x}),\ \Pi_{\psi^\dagger}(\mathbf{y})\}_\mathrm{D} = 0\,.\end{aligned}$$

The canonical quantization rule is now $\{A,\ B\}_+ = i\hbar\{A,\ B\}_{\rm D}$ which leads to the canonical anticommutators

$$\begin{aligned}\{\psi_\alpha(\mathbf{x}),\ \psi^\dagger_\beta(\mathbf{y})\}_+ &= \hbar\,\delta_{\alpha\beta}\,\delta^{(3)}(\mathbf{x}-\mathbf{y})\,,\\ \{\psi_\alpha(\mathbf{x}),\ \psi_\beta(\mathbf{y})\}_+ &= \{\psi^\dagger_\alpha(\mathbf{x}),\ \psi^\dagger_\beta(\mathbf{y})\}_+ = 0\,.\end{aligned} \tag{11.393}$$

Since in the quantum theory the constraints are realized as operator identities, $\Pi_{\psi^\dagger}$ becomes the zero operator and Π_ψ can everywhere be eliminated in favour of $i\psi^\dagger$. The Hamiltonian density operator therefore reads

$$\mathcal{H} = \psi^\dagger\gamma^0(-i\gamma^k\partial_k + m)\psi\,. \tag{11.394}$$

The next step is to expand the basic operators $\psi_\alpha(\mathbf{x})$ and $\psi^\dagger_\alpha(\mathbf{x})$ into creation and annihilation operators, satisfying simple anticommutation rules. Since ψ_α has four independent components, there must be four annihilation and four creation operators. As the Grassmann variables ψ_α and $\psi^\dagger_\alpha$ are not real c-numbers, there corresponding operators ψ_α and $\psi^\dagger_\alpha$ need not be Hermitian. We write therefore their expansions in the forms

$$\begin{aligned}\psi_\alpha(\mathbf{x}) &= \int d\Gamma_\mathbf{p}\, e^{i\mathbf{p}\cdot\mathbf{x}} \sum_\sigma \left[u_\alpha(\mathbf{p},\sigma)\,a_u(\mathbf{p},\sigma) + v_\alpha(-\mathbf{p},\sigma)\,a_v(-\mathbf{p},\sigma)\right],\\ \psi^\dagger_\alpha(\mathbf{x}) &= \int d\Gamma_\mathbf{p}\, e^{-i\mathbf{p}\cdot\mathbf{x}} \sum_\sigma \left[u^*_\alpha(\mathbf{p},\sigma)\,a^\dagger_u(\mathbf{p},\sigma) + v^*_\alpha(-\mathbf{p},\sigma)\,a^\dagger_v(-\mathbf{p},\sigma)\right],\end{aligned} \tag{11.395}$$

using the functions $u_\alpha(\mathbf{p},\sigma)$ and $v_\alpha(\mathbf{p},\sigma)$ constructed in Section 8.3 and with $d\Gamma_\mathbf{p} = (2\pi)^3 2E_\mathbf{p}$, $E_\mathbf{p} = \sqrt{\mathbf{p}^2+m^2}$. Computing the anticommutators (11.393) and anticipating that $\{a_u,\ a_v\}_+ = \{a^\dagger_u,\ a^\dagger_v\}_+ = 0$ we get

$$\begin{aligned}\{\psi_\alpha(\mathbf{x}),\ \psi^\dagger_\beta(\mathbf{y})\}_+ = \int d\Gamma_\mathbf{p}\int d\Gamma_{\mathbf{p}'}\, e^{i(\mathbf{p}\cdot\mathbf{x}-\mathbf{p}'\cdot\mathbf{y})}\sum_\sigma\sum_{\sigma'}&\\ \big(u_\alpha(\mathbf{p},\sigma)\,u^*_\beta(\mathbf{p}',\sigma')\,\{a_u(\mathbf{p},\sigma),\ a^\dagger_u(\mathbf{p}',\sigma')\}_+&\\ +v_\alpha(-\mathbf{p},\sigma)\,v^*_\beta(-\mathbf{p}',\sigma')\,\{a_v(-\mathbf{p},\sigma),\ a^\dagger_v(-\mathbf{p}',\sigma')\}_+\big)&\,.\end{aligned}$$

Recalling the summation rules (8.105):

$$\begin{aligned}\sum_\sigma u_\alpha(\mathbf{p},\sigma)\,u^*_\beta(\mathbf{p},\sigma) = \sum_s u_\alpha(\mathbf{p},\sigma)\,\bar u_\beta(\mathbf{p},\sigma)\gamma^0 &= \left[(E_\mathbf{p}\gamma^0 - \mathbf{p}\cdot\boldsymbol{\gamma} + m)\gamma^0\right]_{\alpha\beta}\,,\\ \sum_\sigma v_\alpha(-\mathbf{p},\sigma)\,v^*_\beta(-\mathbf{p},\sigma) &= \left[(E_\mathbf{p}\gamma^0 + \mathbf{p}\cdot\boldsymbol{\gamma} - m)\gamma^0\right]_{\alpha\beta}\,,\end{aligned}$$

it is easy to check that the commutation rules (11.393) are satisfied if

$$\begin{aligned}\{a_u(\mathbf{p},\sigma),\ a^\dagger_u(\mathbf{p}',\sigma')\}_+ &= \delta_\Gamma(\mathbf{p}-\mathbf{p}')\,\delta_{\sigma\sigma'}\,,\\ \{a_v(\mathbf{p},\sigma),\ a^\dagger_v(\mathbf{p}',\sigma')\}_+ &= \delta_\Gamma(\mathbf{p}-\mathbf{p}')\,\delta_{\sigma\sigma'}\,,\end{aligned} \tag{11.396}$$

and the other anticommutators are zero as anticipated.

To express the Hamiltonian through the operators a_u etc. we use the equality[99]

$$\gamma^0(-i\gamma^i\partial_i+m)\psi(\mathbf{x}) = \int d\Gamma_{\mathbf{p}}\, E_{\mathbf{p}}\, e^{i\mathbf{p}\cdot\mathbf{x}} \sum_\sigma \left[a_u(\mathbf{p},\sigma)\, u(\mathbf{p},\sigma) - a_v(-\mathbf{p},\sigma)\, v(-\mathbf{p},\sigma)\right],$$

and the normalization conditions

$$\begin{aligned} u^\dagger(\mathbf{p},\sigma)\cdot u(\mathbf{p},\sigma') &= 2E_{\mathbf{p}}\,\delta_{\sigma\sigma'}, \\ v^\dagger(\mathbf{p},\sigma)\cdot v(\mathbf{p},\sigma') &= 2E_{\mathbf{p}}\,\delta_{\sigma\sigma'}, \\ u^\dagger(\mathbf{p},\sigma)\cdot v(-\mathbf{p},\sigma') = v^\dagger(-\mathbf{p},\sigma)\cdot u(\mathbf{p},\sigma') &= 0\,, \end{aligned} \tag{11.397}$$

which can be derived from the rules of constructing the functions $u(\mathbf{p},\sigma)$ and $v(\mathbf{p},\sigma)$ given in Section 8.3. In this way one gets

$$\begin{aligned} H &= \int d^3\mathbf{x}\ \psi^\dagger(\mathbf{x})\gamma^0(-i\gamma^i\partial_i+m)\psi(\mathbf{x}) \\ &= \int d\Gamma_{\mathbf{p}}\, E_{\mathbf{p}} \sum_\sigma \left[a_u^\dagger(\mathbf{p},\sigma)\, a_u(\mathbf{p},\sigma) - a_v^\dagger(\mathbf{p},\sigma)\, a_v(\mathbf{p},\sigma)\right]. \end{aligned} \tag{11.398}$$

If the basic rules (11.393) and consequently (11.396) involved *commutators* (instead of anticommutators) this form of the Hamiltonian would be a disaster: from the *commutation* rules it would unambiguously follow (at least if the system is quantized in a finite box) that there is a state $|\Omega_0\rangle$ annihilated by all a_u and a_v and that all other states are created by acting on $|\Omega_0\rangle$ with $a_u^\dagger$ and $a_v^\dagger$. But the states created by $a_v^\dagger$ would then have negative energy! In the Fock space there would exist states with arbitrarily large negative energy and, therefore, the spectrum of the Hamiltonian would not be bounded from below. Fortunately, with the *anticommutators* such a conclusion does not follow: since the *anticommutation* relations are symmetric with respect to interchanges $a \leftrightarrow a^\dagger$ we can call a_v the *creation* operator. More formally, one can make the substitution

$$\begin{aligned} a_u(\mathbf{p},\sigma) &= b(\mathbf{p},\sigma)\,, \qquad & a_u^\dagger(\mathbf{p},\sigma) &= b^\dagger(\mathbf{p},\sigma)\,, \\ a_v(\mathbf{p},\sigma) &= d^\dagger(\mathbf{p},\sigma)\,, \qquad & a_v^\dagger(\mathbf{p},\sigma) &= d(\mathbf{p},\sigma)\,, \end{aligned} \tag{11.399}$$

which allows to rewrite the Hamiltonian (discarding the infinite constant) as

$$H = \int d\Gamma_{\mathbf{p}}\, E_{\mathbf{p}} \sum_\sigma \left[b^\dagger(\mathbf{p},\sigma)\, b(\mathbf{p},\sigma) + d^\dagger(\mathbf{p},\sigma)\, d(\mathbf{p},\sigma)\right]. \tag{11.400}$$

---

[99]It follows from the equalities

$$\begin{aligned} \gamma^0(-i\gamma^k\partial_k+m)\, u(\mathbf{p},\sigma)\, e^{i\mathbf{p}\cdot\mathbf{x}} &= \gamma^0(\boldsymbol{\gamma}\cdot\mathbf{p}+m)\, u(\mathbf{p},\sigma)\, e^{i\mathbf{p}\cdot\mathbf{x}} = E_{\mathbf{p}}\, u(\mathbf{p},\sigma)\, e^{i\mathbf{p}\cdot\mathbf{x}} \\ \gamma^0(-i\gamma^k\partial_k+m)\, v(-\mathbf{p},\sigma)\, e^{i\mathbf{p}\cdot\mathbf{x}} &= \gamma^0(\boldsymbol{\gamma}\cdot\mathbf{p}+m)\, v(-\mathbf{p},\sigma)\, e^{i\mathbf{p}\cdot\mathbf{x}} = -E_{\mathbf{p}}\, v(-\mathbf{p},\sigma)\, e^{i\mathbf{p}\cdot\mathbf{x}}\,, \end{aligned}$$

which in turn follow from the formulae (8.101) written in the form $(\gamma^0E_{\mathbf{p}} - \boldsymbol{\gamma}\cdot\mathbf{p} - m)u(\mathbf{p}) = 0$ and $(\gamma^0E_{\mathbf{p}} + \boldsymbol{\gamma}\cdot\mathbf{p} + m)v(-\mathbf{p}) = 0$.

The Hamiltonian (11.400) is positive semidefinite and the operators $b^\dagger(\mathbf{p},\sigma)$, $b(\mathbf{p},\sigma)$ and $d^\dagger(\mathbf{p},\sigma)$, $d(\mathbf{p},\sigma)$ satisfying the rules

$$\begin{aligned}\left\{b(\mathbf{p},\sigma),\ b^\dagger(\mathbf{p}',\sigma')\right\}_+ &= \delta_\Gamma(\mathbf{p}-\mathbf{p}')\,\delta_{\sigma\sigma'}\,,\\ \left\{d(\mathbf{p},\sigma),\ d^\dagger(\mathbf{p}',\sigma')\right\}_+ &= \delta_\Gamma(\mathbf{p}-\mathbf{p}')\,\delta_{\sigma\sigma'}\,,\end{aligned}\tag{11.401}$$

(with all other anticommutators zero) are the creation and annihilation operators of particles and antiparticles, respectively. This interpretation follows from the fact that states created by $b^\dagger(\mathbf{p},\sigma)$ and $d^\dagger(\mathbf{p},\sigma)$ have the same energies (and, hence same masses). Moreover, the Lagrangian (11.387) is invariant under the global transformations

$$\psi\to\psi' = e^{-iQ\theta}\psi \qquad \psi^\dagger\to\psi^{\dagger\prime} = e^{iQ\theta}\psi^\dagger\,,\tag{11.402}$$

forming a $U(1)$ symmetry group with Q being the charge (arbitrary in the free field theory) and θ the transformation parameter. The corresponding conserved Noether[100] current is

$$j^\mu = \psi^\dagger\gamma^0\gamma^\mu\psi\,,\tag{11.403}$$

and the states created by $b^\dagger$ and $d^\dagger$ are eigenstates of

$$\hat{Q} = \int d^3\mathbf{x}\, j^0(\mathbf{x}) = Q\int d\Gamma_{\mathbf{p}}\sum_\sigma\left[b^\dagger(\mathbf{p},\sigma)\,b(\mathbf{p},\sigma) - d^\dagger(\mathbf{p},\sigma)\,d(\mathbf{p},\sigma)\right],\tag{11.404}$$

with opposite eigenvalues

$$\begin{aligned}\hat{Q}\ b^\dagger(\mathbf{p},\sigma)|\Omega_0\rangle &= Q\,b^\dagger(\mathbf{p},\sigma)|\Omega_0\rangle\,,\\ \hat{Q}\ d^\dagger(\mathbf{p},\sigma)|\Omega_0\rangle &= -Q\,d^\dagger(\mathbf{p},\sigma)|\Omega_0\rangle\,.\end{aligned}\tag{11.405}$$

Quantization of the theory defined by the Lagrangian (11.368) can be performed along the same lines. Quantization of theories of fermionic and mixed fermionic-bosonic systems of fields the complete Lagrangians of which consist of (11.387) (or (11.368)) plus terms involving higher powers of fields than the second (examples of realistic such theories will be discussed in Section 11.12), also proceed along the same lines.

As in the case of the scalar field φ discussed in Section 11.2, one can now introduce time dependent field operators $\psi(t,\mathbf{x})\equiv\psi(x)$, $\psi^\dagger(t,\mathbf{x})\equiv\psi^\dagger((x)$ by the formulae

$$\psi(t,\mathbf{x}) = e^{iHt}\psi(\mathbf{x})\,e^{-iHt}\,,\qquad \psi^\dagger(t,\mathbf{x}) = e^{iHt}\psi^\dagger(\mathbf{x})\,e^{-iHt}\,.\tag{11.406}$$

They satisfy the canonical equations of motion

$$\frac{d}{dt}\,\psi(t,\mathbf{x}) = i[H,\ \psi(t,\mathbf{x})]\,,\qquad \frac{d}{dt}\,\psi^\dagger(t,\mathbf{x}) = i[H,\ \psi^\dagger(t,\mathbf{x})]\,,\tag{11.407}$$

[100]The Noether theorem is for theories of Grassmann algebra valued fields derived exactly as described in Section 11.1.

which take the form formally identical with the classical equations of motion derived from the underlying Lagrangian (that is with the equations (11.388) in the case of the Lagrangian (11.387)). The time-dependent operators $\psi(t, \mathbf{x})$ and $\psi^\dagger(t, \mathbf{x})$, which in the theory of free fields are the Heisenberg operators and in the theory of interacting fields in which the Hamiltonian (11.394) or (11.400) play the role of the free part H_0 of the full Hamiltonian, are the interaction picture operators, take the forms

$$\psi_\alpha(x) = \int d\Gamma_{\mathbf{p}} \sum_\sigma \left(e^{-ip\cdot x}\, u_\alpha(\mathbf{p}, \sigma)\, b(\mathbf{p}, \sigma) + e^{ip\cdot x}\, v_\alpha(\mathbf{p}, \sigma)\, d^\dagger(\mathbf{p}, \sigma)\right),$$

$$\psi^\dagger_\alpha(x) = \int d\Gamma_{\mathbf{p}} \sum_\sigma \left(e^{-ip\cdot x}\, v^*_\alpha(\mathbf{p}, \sigma)\, d(\mathbf{p}, \sigma) + e^{ip\cdot x}\, u^*_\alpha(\mathbf{p}, \sigma)\, b^\dagger(\mathbf{p}, \sigma)\right).$$

One can also find the canonical energy-momentum tensor

$$T^{\mu\nu}_{\rm can} = \frac{\partial \mathcal{L}}{\partial_R(\partial_\mu \psi)}\, \partial^\nu \psi - g^{\mu\nu} \mathcal{L}\,.$$

As long as the Lagrangian is linear in ψ and $\psi^\dagger$ (as in (11.387)), the time dependent (Heisenberg picture) operators satisfy the equations of motion which make $\mathcal{L}$ vanishing (as the operator) and $T^{\mu\nu}_{\rm can}$ simplifies to the first term only. Thus, in the theory (11.387)

$$T^{\mu\nu}_{\rm can} = i\psi^\dagger \gamma^0 \gamma^\mu \partial^\nu \psi\,. \tag{11.408}$$

When expressed through the field operators $\int d^3\mathbf{x}\, T^{00}_{\rm can}$ is just the Hamiltonian (11.400) and

$$P^i = \int d^3\mathbf{x}\, T^{0i}_{\rm can} = \int d\Gamma_{\mathbf{p}}\, p^i \sum_\sigma \left[b^\dagger(\mathbf{p}, \sigma) b(\mathbf{p}, \sigma) + d^\dagger(\mathbf{p}, \sigma) d(\mathbf{p}, \sigma)\right], \tag{11.409}$$

is the momentum operator commuting with H. Finally, the canonical tensor

$$M^{\mu\nu\kappa} = x^\nu\, T^{\mu\kappa}_{\rm can} - x^\kappa\, T^{\mu\nu}_{\rm can} + \frac{\partial \mathcal{L}}{\partial_R(\partial_\mu \psi)} \left(-i\mathcal{J}^{\nu\kappa}_{\rm spin}\right) \psi\,, \tag{11.410}$$

gives the operators generating rotations and boosts

$$J^{\nu\kappa} = i\int d^3\mathbf{x}\, \psi^\dagger \left(x^\nu \partial^\kappa - x^\kappa \partial^\nu - \frac{i}{2}\, \sigma^{\nu\kappa}_{4\times 4}\right) \psi\,. \tag{11.411}$$

To see that the particles which are "quanta" of the field ψ have spin $s = 1/2$ it is enough to act with $J^{12} = J^z$ on the one-particle states representing particles at rest

$$J^z\, b^\dagger(\mathbf{0}, \pm\tfrac{1}{2})|\Omega_0\rangle = \pm\frac{1}{2}\, b^\dagger(\mathbf{0}, \pm\tfrac{1}{2})|\Omega_0\rangle\,,$$
$$J^z\, d^\dagger(\mathbf{0}, \pm\tfrac{1}{2})|\Omega_0\rangle = \pm\frac{1}{2}\, d^\dagger(\mathbf{0}, \pm\tfrac{1}{2})|\Omega_0\rangle\,. \tag{11.412}$$

More generally, one can check the commutation rules of the operators H, $\mathbf{P}$, $\mathbf{J}$ and $\mathbf{K}$ and verify that the one-particle states $b^\dagger(\mathbf{p}, \pm\frac{1}{2})|\Omega_0\rangle$ and $d^\dagger(\mathbf{p}, \pm\frac{1}{2})|\Omega_0\rangle$ transform under the action of the $U(\Lambda) \approx 1 - \frac{i}{2}\omega_{\mu\nu}J^{\mu\nu} + \dots$ operators in the way appropriate for spin 1/2 particles.

Statistical properties of the quantized fermionic field can also be readily obtained. If it is quantized in the box of volume $V = L^3$ the Hamiltonian (before subtracting the infinite constant) is

$$H = \sum_{\mathbf{p}} \sum_{\sigma} E_{\mathbf{p}} \left(b^\dagger_{\mathbf{p}\sigma} b_{\mathbf{p}\sigma} + d^\dagger_{\mathbf{p}\sigma} d_{\mathbf{p}\sigma} - 1 \right) , \tag{11.413}$$

and the Hilbert space is spanned by the states $|n^b_{\mathbf{p}\sigma}, \dots, n^d_{\mathbf{p}\sigma}, \dots\rangle$ with $n^b_{\mathbf{p}\sigma}, n^d_{\mathbf{p}\sigma} = 0$ or 1. Since there is a conserved charge $\hat{Q}$, one computes the Grand Statistical Ensemble sum

$$\begin{aligned} \Xi(T, V, \mu) &= \mathrm{Tr}\, e^{-\beta(H - \mu\hat{Q})} \\ &= \prod_{\mathbf{p}} \left(\sum_{n^b_{\mathbf{p}}=0}^{1} e^{\frac{1}{2}\beta E_{\mathbf{p}}} e^{-\beta(E_{\mathbf{p}} - \mu) n^b_{\mathbf{p}}} \right)^2 \times \prod_{\mathbf{p}} \left(\sum_{n^d_{\mathbf{p}}=0}^{1} e^{\frac{1}{2}\beta E_{\mathbf{p}}} e^{-\beta(E_{\mathbf{p}} + \mu) n^d_{\mathbf{p}}} \right)^2 \\ &= \prod_{\mathbf{p}} e^{2\beta E_{\mathbf{p}}} \left(1 + e^{-\beta(E_{\mathbf{p}} - \mu)}\right)^2 \left(1 + e^{-\beta(E_{\mathbf{p}} + \mu)}\right)^2 , \end{aligned} \tag{11.414}$$

(the squares come from the two spin states). Thus, for $\Omega(T, V, \mu) = -k_{\mathrm{B}} T \ln \Xi$, after going over to the continuous normalization, one gets the formula

$$\begin{aligned} \Omega(T, V, \mu) = -2V \int \frac{d^3\mathbf{p}}{(2\pi\hbar)^3} \Big\{ & E_{\mathbf{p}} + k_{\mathrm{B}} T \ln\left(1 + e^{-\beta(E_{\mathbf{p}} - \mu)}\right) \\ & + k_{\mathrm{B}} T \ln\left(1 + e^{-\beta(E_{\mathbf{p}} + \mu)}\right) \Big\} , \end{aligned} \tag{11.415}$$

where $E_{\mathbf{p}} = \sqrt{c^2\mathbf{p}^2 + m^2c^4}$ and the factor of 2 accounts for two spin states per each $\mathbf{p}$.

11.9 Transition to the interaction picture

As in the approach to quantum field theory based on quantum mechanics of relativistic interacting particles, also in the approach relying on quantizing fields one can assume that the full Hamiltonians of interacting fields possesses (in the infinite space) *in* and *out* particle-like generalized eigenvectors $|\alpha_\pm\rangle$ which, owing to ther Lorentz transformation properties, can be interpreted as representing particles entering or emerging after interactions. One is therefore primarily interested in obtaining S-matrix elements

$$S_{\beta\alpha} = \langle \beta_- | \alpha_+ \rangle ,$$

allowing to compute various transition rates using the recipes formulated in Chapter 10. It is therefore necessary to discuss how to recover in the approach based on quantization of interacting classical fields the perturbative expansion (and the related Feynman rules) which in Chapter 9 was formulated within the approach based on constructing interactions of relativistic particles. As we will see, the procedure by which this is achieved, called *transition to the interaction picture*, when applied to classical relativistic fields automatically produces all the additional interactions which in the previous approach had to be included by hands in order to obtain a Lorentz covariant S-matrix.

We begin with the simplest example of the theory of a single scalar field $\varphi(x)$ defined by the Lagrangian density

$$\mathcal{L} = \frac{1}{2}\partial_\mu\varphi\partial^\mu\varphi - \frac{1}{2}M^2\varphi^2 - \mathcal{H}_{\rm int}(\varphi)\,. \tag{11.416}$$

in which $\mathcal{H}_{\rm int}(\varphi)$ is some interaction Hamiltonian density. After performing quantization as in Section 11.2, the Hamiltonian of this system, expressed through the Schrödinger picture (time independent) operators $\varphi(\mathbf{x})$, $\Pi(\mathbf{x})$ satisfying the rules (11.78) takes the form of the sum of the free Hamiltonian

$$H_0 = \frac{1}{2}\int d^3\mathbf{x}\left[\Pi^2(\mathbf{x}) + (\boldsymbol{\nabla}\varphi(\mathbf{x}))^2 + M^2\varphi^2(\mathbf{x})\right], \tag{11.417}$$

which is quadratic in the field operators and the of interaction[101]

$$V_{\rm int} = \int d^3\mathbf{x}\,\mathcal{H}_{\rm int}(\varphi(\mathbf{x}),\Pi(\mathbf{x}))\,. \tag{11.418}$$

The formula

$$S_{\beta\alpha} = \langle\beta_0|T\exp\left(-i\int_{-\infty}^{+\infty}dt\,V^I_{\rm int}(t)\right)|\alpha_0\rangle\,, \tag{11.419}$$

for S-matrix elements established in Section 7.3 using general rules of quantum mechanics should apply also to the field theory case. In (11.419), as in (7.64),

$$\begin{aligned}V^I_{\rm int}(t) &= e^{iH_0t}\int d^3\mathbf{x}\,\mathcal{H}_{\rm int}(\varphi(\mathbf{x}),\Pi(\mathbf{x}))\,e^{-iH_0t}\\ &= \int d^3\mathbf{x}\,\mathcal{H}_{\rm int}(\varphi_I(t,\mathbf{x}),\Pi_I(t,\mathbf{x})) \equiv \int d^3\mathbf{x}\,\mathcal{H}^I_{\rm int}(t,\mathbf{x})\,.\end{aligned} \tag{11.420}$$

and for the states $|\alpha_0\rangle$ one should take the generalized eigenvectors of H_0 (11.417) constructed in Section 11.2. The time-dependent operators $\varphi_I(x) \equiv \varphi_I(t,\mathbf{x})$ and $\Pi_I(x) \equiv \Pi_I(t,\mathbf{x})$ given by

$$\begin{aligned}\varphi_I(t,\mathbf{x}) &= e^{iH_0t}\,\varphi(\mathbf{x})\,e^{-iH_0t}\,,\\ \Pi_I(t,\mathbf{x}) &= e^{iH_0t}\,\Pi(\mathbf{x})\,e^{-iH_0t}\,,\end{aligned} \tag{11.421}$$

[101] To be more general we write the formulae as if $\mathcal{H}_{\rm int}$ depended also on $\Pi(\mathbf{x})$.

are called in this context the interaction picture operators and have the structure of the free-field operators constructed in Chapter 8:

$$\varphi_I(x) = \int d\Gamma_{\mathbf{k}} \left(a(\mathbf{k})\, e^{-ik\cdot x} + a^\dagger(\mathbf{k})\, e^{ik\cdot x}\right),$$
$$\Pi_I(x) = \frac{1}{i}\int d\Gamma_{\mathbf{k}}\, E(\mathbf{k}) \left(a(\mathbf{k})\, e^{-ik\cdot x} - a^\dagger(\mathbf{k})\, e^{ik\cdot x}\right). \tag{11.422}$$

with $k\cdot x \equiv E_{\mathbf{k}}t - \mathbf{k}\cdot\mathbf{x}$. They satisfy (by construction) the equal-time commutation rules

$$[\varphi_I(t,\mathbf{x}),\ \Pi_I(t,\mathbf{y})] = i\delta^{(3)}(\mathbf{x}-\mathbf{y}), \tag{11.423}$$

etc. and the "Klein - Gordon" equation

$$(\partial_\mu\partial^\mu + M^2)\varphi_I(t,\mathbf{x}) = 0. \tag{11.424}$$

It should also be clear that the free Hamiltonian H_0 (11.417) can be also written in terms of the time dependent interaction picture operators

$$H_0 = e^{iH_0t}H_0\, e^{-iH_0t} = \frac{1}{2}\int d^3\mathbf{x}\left[\Pi_I^2(t,\mathbf{x}) + (\boldsymbol{\nabla}\varphi_I(t,\mathbf{x}))^2 + M^2\varphi_I^2(t,\mathbf{x})\right].$$

In the case of the theory (11.416) it is obvious that $\dot{\varphi}_I(t,\mathbf{x}) = \Pi_I(t,\mathbf{x})$. In more complicated cases like the one of the electromagnetic field interacting with other fields (to be discussed below) it is useful to remember that $\varphi_I(t,\mathbf{x})$ and $\Pi_I(t,\mathbf{y})$ satisfy the equations

$$\dot{\varphi}_I(t,\mathbf{x}) = i[H_0,\ \varphi_I(t,\mathbf{x})],$$
$$\dot{\Pi}_I(t,\mathbf{x}) = i[H_0,\ \Pi_I(t,\mathbf{x})], \tag{11.425}$$

with the initial conditions $\varphi_I(0,\mathbf{x}) = \varphi_H(0,\mathbf{x}) = \varphi(\mathbf{x})$, $\Pi_I(0,\mathbf{x}) = \Pi_H(0,\mathbf{x}) = \Pi(\mathbf{x})$. The first of the equations (11.425) unambiguously fixes the relation of $\dot{\varphi}_I(t,\mathbf{x})$ to $\Pi_I(t,\mathbf{x})$.

In this way the perturbation expansion for the S-matrix (which leads to the Feynman rules discussed in Sections 9.3-9.6) derived from principles of quantum mechanics in the framework of Chapters 8 and 9 is recovered in the approach based on quantization of relativistic fields. It should be obvious, that this perturbative expansion is based on the same assumptions as those formulated in Section 7.3, namely that the particle-like *in* and *out* eigenstates of the complete Hamiltonian are in the strict one-to-one correspondence, specified by (7.39), with the eigenstates $|\alpha_0\rangle$ of the free Hamiltonian (11.417).

Transition to the interacting picture is more subtle in the case of vector (massive and massless) fields. Let us first consider the Lagrangian density

$$\mathcal{L} = \frac{1}{2}\partial_\mu\varphi\partial^\mu\varphi - \frac{1}{2}M^2\varphi^2 - \mathcal{H}_{\rm int}(\varphi) - J^\mu\partial_\mu\varphi + \dots, \tag{11.426}$$

in which J^μ is some four-vector constructed out of fields other than φ itself (the ellipses stands for terms depending on these other fields). The canonical momentum is given by

$$\Pi = \frac{\partial \mathcal{L}}{\partial \dot{\varphi}} = \partial_0 \varphi - J_0 \,, \tag{11.427}$$

and the Hamiltonian constructed according to the rules takes the form

$$\begin{aligned} H &= \int d^3\mathbf{x}\, \Pi(\mathbf{x})\dot{\varphi}(\mathbf{x}) - \int d^3\mathbf{x}\, \mathcal{L}(\varphi(\mathbf{x}), \dot{\varphi}(\mathbf{x})) \\ &= \int d^3\mathbf{x} \left\{ \Pi\,(\Pi + J_0) - \frac{1}{2}(\Pi + J_0)^2 + \frac{1}{2}\,(\boldsymbol{\nabla}\varphi)^2 \right. \\ &\qquad\qquad \left. + \frac{1}{2}M^2\varphi^2 + \mathcal{H}_{\rm int} + \mathbf{J}\cdot\boldsymbol{\nabla}\varphi + J^0(\Pi + J_0) \right\}. \end{aligned} \tag{11.428}$$

Upon canonical quantization $\varphi(x)$ and $\Pi(x)$ become Schrödinger picture operators $\varphi(\mathbf{x})$ and $\Pi(\mathbf{x})$ satisfying the standard canonical commutation rules. Splitting the full Hamiltonian into the free part H_0 and the interaction $V_{\rm int}$ one gets $H_0(\Pi(\mathbf{x}), \varphi(\mathbf{x}))$ of the same form as in (11.417) and

$$V_{\rm int} = \int d^3\mathbf{x}\, [\Pi(\mathbf{x})J^0(\mathbf{x}) + \mathbf{J}(\mathbf{x})\cdot\boldsymbol{\nabla}\varphi(\mathbf{x}) + \frac{1}{2}\big(J^0(\mathbf{x})\big)^2 + \mathcal{H}_{\rm int}(\varphi(\mathbf{x}))]\,. \tag{11.429}$$

As in the preceding case, H_0 gets diagonalized upon introducing the creation and annihilation operators and the Fock basis of H_0 eigenvectors is constructed to play the role of the $|\alpha_0\rangle$ states in the formula (11.419). The interaction picture operators $\varphi_I(t, \mathbf{x})$ and $\Pi_I(t, \mathbf{x})$ are obtained as in (11.421). The important point is that the first of the equations (11.425) tells that the operator $\Pi_I(t, \mathbf{x})$ is related to $\partial_0\varphi_I(t, \mathbf{x})$ by

$$\Pi_I(t, \mathbf{x}) = \partial_0 \varphi_I(t, \mathbf{x})\,, \tag{11.430}$$

and *not* by $\partial_0\varphi_I(t, \mathbf{x}) - J_I^0(t, \mathbf{x})$ (as in the case of the corresponding Heisenberg picture operators). For this reason the interaction $V^I_{\rm int}(t)$ used in the perturbation expansion for the S-matrix takes the form

$$V^I_{\rm int}(t) = \int d^3\mathbf{x} \left[J^\mu_I(t, \mathbf{x})\, \partial_\mu \varphi_I(t, \mathbf{x}) + \frac{1}{2}\left(J^0_I(t, \mathbf{x})\right)^2 + \mathcal{H}_{\rm int}(\varphi_I(t, \mathbf{x})) \right], \tag{11.431}$$

i.e. it acquires a noncovariant term $(J_I^0)^2$, so that the expression under the integral over $d^3\mathbf{x}$ in (11.431) is *not* a Lorentz invariant interaction density (as was assumed in (7.101)). However, as we have seen in Section 9.5, precisely such a noncovariant term had to be added (by hands in the approaches of Chapters 7, 8) to compensate for the noncovariant term in the propagator of the vector field $\partial_\mu\varphi_I(x)$ (arising due to the singular nature of products of field operators taken at the same space-time point) and to restore the Poincaré covariance of the S-matrix. Similarly, in the case of the Proca theory (quantized in Section 11.5) the Hamiltonian density of which takes (after exploiting the constraint

I Electromagnetic units

In comparing electromagnetic quantities in the SI and Gauss systems one should keep in mind that this is not only the question of using different units (in mechanics the difference between the SI and the cgs systems is only the one of units): electromagnetic quantities in the two systems have different physical dimensions. Hence we use below the sub(super)scripts distinguishing quantities in different systems.

The Maxwell equations in the SI system read

$$\boldsymbol{\nabla}\times\mathbf{E}_{\rm SI}+\frac{\partial\mathbf{B}_{\rm SI}}{\partial t}=\mathbf{0}\,,\qquad\qquad \boldsymbol{\nabla}\cdot\mathbf{B}_{\rm SI}=0\,, \tag{I.1}$$

$$\boldsymbol{\nabla}\times\mathbf{B}_{\rm SI}-\mu_0\varepsilon_0\,\frac{\partial\mathbf{E}_{\rm SI}}{\partial t}=\mu_0\,\mathbf{j}_{\rm SI}\,,\qquad\qquad \varepsilon_0\boldsymbol{\nabla}\cdot\mathbf{E}_{\rm SI}=\rho_{\rm SI}\,, \tag{I.2}$$

with $\varepsilon_0\mu_0=1/c^2$. The action giving rise to (I.2) is

$$I[A_{\rm SI},J]=\int dt\int d^3\mathbf{x}\left(-\frac{1}{4}\,\varepsilon_0c^2f^{\rm SI}_{\mu\nu}f^{\mu\nu}_{\rm SI}-A^{\rm SI}_{\nu}\,e_{\rm SI}J^{\nu}\right). \tag{I.3}$$

Here $f^{\rm SI}_{\mu\nu}=\partial_\mu A^{\rm SI}_\nu-\partial_\nu A^{\rm SI}_\mu$ (recall that $x^0\equiv ct$) and

$$f^{\rm SI}_{\mu\nu}f^{\mu\nu}_{\rm SI}=-\frac{2}{c^2}\,\mathbf{E}^2_{\rm SI}+2\mathbf{B}^2_{\rm SI}\,. \tag{I.4}$$

In this system of units the following identifications[1] hold:

$$A^{\mu}_{\rm SI}=\left(\frac{1}{c}\,\varphi_{\rm SI},\ \mathbf{A}_{\rm SI}\right),\qquad e_{\rm SI}J^{\mu}=(c\,\rho_{\rm SI},\ \mathbf{j}_{\rm SI})\,, \tag{I.5}$$

and

$$\frac{1}{c}\mathbf{E}_{\rm SI}=-\boldsymbol{\nabla}\frac{\varphi_{\rm SI}}{c}-\frac{1}{c}\frac{\partial\mathbf{A}_{\rm SI}}{\partial t}\,,\qquad \mathbf{B}_{\rm SI}=\boldsymbol{\nabla}\times\mathbf{A}_{\rm SI}\,. \tag{I.6}$$

The dimensionless fine structure constant $\alpha_{\rm EM}=1/137.03599$ is given by

$$\alpha_{\rm EM}=\frac{e^2_{\rm SI}}{4\pi\varepsilon_0\hbar c}\,. \tag{I.7}$$

In the Gauss' system of electromagnetic units

$$\boldsymbol{\nabla}\times\mathbf{E}_{\rm Gauss}+\frac{1}{c}\,\frac{\partial\mathbf{B}_{\rm Gauss}}{\partial t}=0\,,\qquad\qquad \boldsymbol{\nabla}\cdot\mathbf{B}_{\rm Gauss}=0\,, \tag{I.8}$$

$$\boldsymbol{\nabla}\times\mathbf{B}_{\rm Gauss}-\frac{1}{c}\,\frac{\partial\mathbf{E}_{\rm Gauss}}{\partial t}=\frac{4\pi}{c}\,\mathbf{j}_{\rm Gauss}\,,\qquad\qquad \boldsymbol{\nabla}\cdot\mathbf{E}_{\rm Gauss}=4\pi\rho_{\rm Gauss}\,. \tag{I.9}$$

[1]All components of a four-vector have the same physical dimension.

The action giving rise to (I.9) is

$$I[A_{\rm Gauss}, J] = \int dt \int d^3\mathbf{x} \left(-\frac{1}{16\pi} f^{\rm Gauss}_{\mu\nu} f^{\mu\nu}_{\rm Gauss} - \frac{1}{c} A^{\rm Gauss}_\nu \, e_{\rm Gauss} J^\nu \right) . \tag{I.10}$$

$$f^{\rm Gauss}_{\mu\nu} f^{\mu\nu}_{\rm Gauss} = -2\,\mathbf{E}^2_{\rm Gauss} + 2\mathbf{B}^2_{\rm Gauss} . \tag{I.11}$$

Here the identifications are

$$A^\mu_{\rm Gauss} = (\varphi_{\rm Gauss},\ \mathbf{A}_{\rm Gauss}), \qquad e_{\rm Gauss} J^\mu = (c\,\rho_{\rm Gauss},\ \mathbf{j}_{\rm Gauss}) . \tag{I.12}$$

and

$$\mathbf{E}_{\rm Gauss} = -\boldsymbol{\nabla}\varphi_{\rm Gauss} - \frac{1}{c}\frac{\partial \mathbf{A}_{\rm Gauss}}{\partial t} , \qquad \mathbf{B}_{\rm Gauss} = \boldsymbol{\nabla} \times \mathbf{A}_{\rm Gauss} . \tag{I.13}$$

In this system

$$\alpha_{\rm EM} = \frac{e^2_{\rm Gauss}}{\hbar c} . \tag{I.14}$$

Outside electromagnetically active media (dielectrics or magnetics) in the SI units one has $\mathbf{B}_{\rm SI} = \mu_0 \mathbf{H}_{\rm SI}$, whereas in the Gauss' system $\mathbf{B}_{\rm Gauss} = \mathbf{H}_{\rm Gauss}$. For this reason in most textbooks using the latter system one uses $\mathbf{H}$ in place of $\mathbf{B}$.

The following pairs of quantities in the two systems have the same physical dimension

$$\begin{aligned}
[e_{\rm Gauss}] \ &= \ \frac{[e_{\rm SI}]}{[\varepsilon_0]^{1/2}} = \frac{[M]^{1/2}[L]^{3/2}}{[T]} , \\
[\varphi_{\rm Gauss}] = \frac{[e_{\rm Gauss}]}{[L]} \ &= \ [\varepsilon_0]^{1/2}[\varphi_{\rm SI}] = \frac{[e_{\rm SI}]}{[\varepsilon_0][L]} , \\
[\mathbf{E}_{\rm Gauss}] = \frac{[e_{\rm Gauss}]}{[L]^2} \ &= \ \big[\varepsilon_0\big]^{1/2}[\mathbf{E}_{\rm SI}\big] = \frac{[e_{\rm SI}]}{[\varepsilon_0]^{1/2}[L]^2} , \\
[\mathbf{B}_{\rm Gauss}] = [\mathbf{E}_{\rm Gauss}] &= [c][\varepsilon_0]^{1/2}\,[\mathbf{B}_{\rm SI}] = \big[\varepsilon_0\big]^{1/2}[\mathbf{E}_{\rm SI}\big] .
\end{aligned} \tag{I.15}$$

In both systems the physical dimension $[I]$ of the action I is obviously $[I] = [\hbar] \equiv [M][L]^2[T]^{-1}$.

The SI system quantities are related to their Gauss's system conterparts by

$$\begin{aligned}
e_{\rm SI}/\sqrt{4\pi\varepsilon_0} &= e_{\rm Gauss} , \\
\sqrt{4\pi\varepsilon_0}\,\varphi_{\rm SI} &= \varphi_{\rm Gauss} , \\
\sqrt{4\pi\varepsilon_0}\,c\,A^0_{\rm SI} &= A^0_{\rm Gauss} , \\
\sqrt{4\pi\varepsilon_0}\,c\,\mathbf{A}_{\rm SI} &= \mathbf{A}_{\rm Gauss} , \\
\sqrt{4\pi\varepsilon_0}\,\mathbf{E}_{\rm SI} &= \mathbf{E}_{\rm Gauss} , \\
\sqrt{4\pi\varepsilon_0}\,c\,\mathbf{B}_{\rm SI} &= \mathbf{B}_{\rm Gauss} .
\end{aligned} \tag{I.16}$$

It is easy to check that inserting these relations converts Maxwell's equations (I.8), (I.9) into (I.1), (I.2).

The Heaviside-Lorentz system of electromagnetic units, the one which is implicitly used in quantum field theory formulae, differs from the Gauss' system only by rescaling the physical quantities back by the $\sqrt{4\pi}$ factor (see section 11.4). Thus, in the Heaviside-Lorentz system electromagnetic quantities have the same physical dimension as in the Gauss' system.

In $d = D+1$ space-time dimensions the electromagnetic quantities have (in the system of units corresponding to the Heaviside-Lorentz one for $d = 4$) the units

$$\begin{aligned} [A^\mu] &= [M]^{1/2}[T]^{-1}[L]^{\frac{4-D}{2}}\,, \\ [\mathbf{E}] = [\mathbf{B}] &= [M]^{1/2}[T]^{-1}[L]^{\frac{2-D}{2}}\,, \\ [e] &= [M]^{1/2}[T]^{-1}[L]^{\frac{D}{2}}\,. \end{aligned} \tag{I.17}$$

The dimension of A^μ follows from $[I] = [\hbar]$, the dimension of $\mathbf{E}$ and $\mathbf{B}$ from the fact that they are elements of $f_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$ and that of e from the Gauss' law in (I.8), given that $[\rho] = [e][L]^{-D}$.

The non-Abelian gauge fields and coupling constants have the same physical dimensions (I.17) as the electromagnetic fields (i.e. $[g] = [e]$, $[A^a_\mu] = [A_\mu]$). The full field strength tensor then is

$$F^{\mu\nu}_a = \partial^\mu A^\nu_a - \partial^\nu A^\mu_a - (g/\hbar c) f_a{}^{bc} A^\mu_b A^\nu_c\,,$$

and the covariant derivatives with the dimensionfull constants restored read

$$D^\mu = \partial/\partial x_\mu + i(g/\hbar c) T^a A^\mu_a\,.$$

For the basis of the mechanical units one takes c and $\hbar$ (the two fundamental units). The natural third unit would be the Planck mass

$$M_{\rm Pl} = \left(\frac{\hbar c}{G_{\rm N}}\right)^{1/2}\,,$$

set by the third fundamental constant of Nature - the Newton constant G_N - but since it is not very practical from the point of view of elementary particle physics, one normally uses for the third unit a conveniently chosen mass unit like e.g. eV/c^2 or GeV/c^2.

Unlike the electromagnetic fields which do have a classical limit the physical dimension of fermionic fields can be chosen arbitrarily. In ordinary nonrelativistic quantum mechanics the probability $\int d^D\mathbf{x}\,|\psi|^2$ is dimensionless and the dimension of ψ is $[L]^{-D/2}$. However, since the fermionic fields ψ are not directly related to the nonrelativistic wave functions, it is more convenient to chose the dimension of the fermionic fields so that the kinetic part of the action

$$I = \int dt \int d^D\mathbf{x}\, \bar{\psi}\, i\gamma^0\, \frac{\partial}{\partial t}\, \psi\,,$$

has the right dimension without any compensating factors of c or $\hbar$. Similarly, it is convenient to take

$$I = \int dt \int d^D\mathbf{x} \left[\left(\frac{\partial \varphi}{\partial t} \right)^2 - c^2 \left(\nabla \varphi \right)^2 \right],$$

for the kinetic part of the action of a scalar field φ. The physical dimensions of the fields (canonical dimensions of the corresponding field operators) then are

$$[\psi] = [M]^{1/2}[T]^{-1/2}[L]^{\frac{2-D}{2}}\,, \qquad [\varphi] = [M]^{1/2}[L]^{\frac{2-D}{2}}\,.$$

Dimensions of the other terms in the field theory Lagrangians can be then easily established. For example, the action describing a multiplet of complex scalar fields ϕ of mass M interacting with the gauge fields A^a_μ with all dimensionfull constants written down explicitly has form

$$I = \int dt \int d^D\mathbf{x} \left\{ c^2 \left(\phi^* \overleftarrow{\partial}^\mu - i(g/\hbar c)\phi^* T^a A^{a\mu} \right) \left(\partial_\mu \phi + i(g/\hbar c) T^a A^a_\mu \phi \right) \right.$$
$$\left. - \frac{M^2 c^4}{\hbar^2}\, \phi^* \phi - \frac{c^3}{\hbar}\, \lambda \left(\phi^* \phi \right)^2 \right\}.$$

Notice that the coupling λ has dimension $[L]^{D-3}$, i.e. it is dimensionless only in $D = 3$ ($d = 4$); in $D \neq 3$ this dimension cannot be removed by adjusting the powers of c and $\hbar$ - for this a mass unit is necessary.

J Primary constraints

Here we give the derivation of the starting point for quantization of systems subject to constraints. To this end we consider a system described by some variables q^i and their generalized velocities $\dot{q}^i$, $i = 1, \ldots, n$. The dynamics is assumed to be determined by a Lagrangian $L = L(q,\, \dot{q})$. It is convenient to treat it as a function of q^i and v^i: $L^v(q,\, v) \equiv L(q, v)$. Standard transition to the Hamilton's formulation is obstructed if the matrix

$$M^v_{ij} \equiv \frac{\partial^2 L^v}{\partial v^i \partial v^j}\,, \tag{J.1}$$

is singular, because then some of the generalized velocities v^i cannot be expressed through the canonical momenta p_i. Hamiltonization of the system requires then a special approach.

The standard Euler-Lagrange equations corresponding to the Lagrangian L^v can be written in the equivalent form as the set of first order equations

$$\begin{aligned} M^v_{ij}\dot{v}^j &= \frac{\partial L^v}{\partial q^i} - \frac{\partial^2 L^v}{\partial v^i \partial q^j}\, v^j\,, \\ \dot{q}^i &= v^i\,, \\ p_i &= \frac{\partial L^v}{\partial v^i}\,. \end{aligned} \tag{J.2}$$

The last group of the equations (J.2) is at the moment redundant - it only serves to define the momenta p_i. The equations (J.2) can be rewritten in another, equivalent form

$$\begin{aligned} \dot{p}_i &= \frac{\partial L^v}{\partial q^i}\,, \\ \dot{q}^i &= v^i\,, \\ p_i &= \frac{\partial L^v}{\partial v^i}\,. \end{aligned} \tag{J.3}$$

The equivalence follows by differentiating the last set of equations with respect to time and substituting the result in the first set of equations.

One now introduces the quasi-Hamiltonian

$$H^v(q, p, v) \equiv p_i v^i - L^v(q, v)\,. \tag{J.4}$$

which is treated as a function of three sets of independent variables: q, p and v. This allows to cast the equations (J.3) in the form:

$$\begin{aligned} \dot{q}^i &= \{q^i,\ H^v\}_{\mathrm{PB}} \equiv \frac{\partial H^v}{\partial p_i}\,, \\ \dot{p}_i &= \{p_i,\ H^v\}_{\mathrm{PB}} \equiv -\frac{\partial H^v}{\partial q^i}\,, \\ \frac{\partial H^v}{\partial v^i} &= 0\,, \end{aligned} \tag{J.5}$$

in which the Poisson brackets are taken with respect to the variables q^i and p_i treating the velocities v^i as independent variables. The quasi-Hamiltonian system (J.5) is fully equivalent to the equations (J.3) and, hence, to the original Euler-Lagrange equations. Moreover, if the last group of equations (J.5) is satisfied, one can treat v^i in the first two groups of equations as (unknown) functions of q^i and p_i even in computing the Poisson brackets.

Suppose now the matrix M^v_{ij} (J.1) is of rank $r < n$. It is then possible to express r velocities v^i in terms of the variables q^i, the momenta p_i and the remaining $n-r$ velocities. Labeling the variables so that it is the first r velocities which can be expressed in this way, it is convenient to introduce the following notation:

$$\begin{aligned} &\Pi_i \equiv p_i\,, &&V^i \equiv v^i\,, &&\text{for}\quad i = 1, \dots, r,\\ &\pi_M \equiv p_{r+M}\,, &&u^M \equiv v^{r+M}\,, &&\text{for}\quad M = 1, \dots, n-r \end{aligned}$$

Of the last group of equations (J.5), which in this notation read

$$\frac{\partial H^v}{\partial V^i} \equiv \Pi_i - \frac{\partial L^v}{\partial V^i} = 0\,, \qquad \frac{\partial H^v}{\partial u^M} \equiv \pi_M - \frac{\partial L^v}{\partial u^M} = 0\,,$$

the first r ones can be solved yielding $V^i = \overline{V}^i(q, \Pi, u) \equiv \overline{V}^i(q, \Pi, u)$, while the remaining $n - r$ equations become the primary constraints.

Next we introduce the total Hamiltonian $H_{\mathrm{T}} \equiv H^v|_{V=\overline{V}}$:

$$H_{\mathrm{T}}(q,\, \Pi,\, \pi,\, u) = \Pi_i \overline{V}^i(q, \Pi, u) + \pi_M u^M - L^v(q,\, \overline{V}(q, \Pi, u),\, u)\,. \tag{J.6}$$

After substituting $\overline{V}^i(q, \Pi, u)$ for V^i the primary constraints take the form

$$\Phi_M \equiv \left.\frac{\partial H^v}{\partial u^M}\right|_{V=\overline{V}} \equiv \pi_M - f_M(q, \Pi)\,, \tag{J.7}$$

with $f_M(q, \Pi) \equiv (\partial L^v(q, V, u)/\partial u^M)_{V=\overline{V}}$. The constraints Φ_M can be also obtained as $\Phi_M = (\partial H_{\mathrm{T}}/\partial u^M)$ because the additional contributions vanish for $V^i = \overline{V}^i$. It is important to realize that the constraints (J.7), $\Phi_M = 0$, do not depend on u^N - if they did, one could solve for more velocities, contary to what has been assumed.

The system of equations (J.5) is now fully equivalent to the following one

$$\begin{aligned} &\dot{q}^i\; = \{q^i,\ H_{\mathrm{T}}\}_{\mathrm{PB}}\,, &&q^i = (X^i,\, x^M)\,,\\ &\dot{\Pi}_i = \{\Pi_i,\ H_{\mathrm{T}}\}_{\mathrm{PB}}\,, &&\\ &\dot{\pi}_M = \{\pi_M,\ H_{\mathrm{T}}\}_{\mathrm{PB}}\,, &&\\ &\Phi_M = 0\,, && \end{aligned} \tag{J.8}$$

because, as has been noted, it is admissible to substitute in (J.5) $\overline{V}^i(q, \Pi, u)$ in place of V^i in computing the Poisson brackets.

The final step is to make clear the structure of H_{T}. To this end one can use the identity[1]

$$H^v = E^v + \frac{\partial H^v}{\partial v^i}\, v^i \equiv \left(\frac{\partial L^v}{\partial v^i}\, v^i - L^v\right) + \frac{\partial H^v}{\partial v^i}\, v^i \,.$$

Using it one can write

$$\begin{aligned} H_{\mathrm{T}} &= H^v|_{V=\overline{V}} = \left(E^v + \frac{\partial H^v}{\partial V^i}\, V^i + \frac{\partial H^v}{\partial u^M}\, u^M\right)_{V=\overline{V}} \\ &= E^v|_{V=\overline{V}} + \Phi_M u^M \equiv H + \Phi_M u^M \,, \end{aligned} \qquad \text{(J.9)}$$

because $(\partial H^v/\partial V^i)_{V=\overline{V}} = 0$. The Hamiltonian H defined in this way does not depend on the primarily unsolvable velocities u^M. Indeed, on one hand,

$$\frac{\partial H_{\mathrm{T}}}{\partial u^M} = \frac{\partial H}{\partial u^M} + \Phi_M \,.$$

On the other hand, as has been noted below the formula (J.7),

$$\frac{\partial H_{\mathrm{T}}}{\partial u^M} \equiv \frac{\partial}{\partial u^M}\left(H^v|_{V=\overline{V}}\right) = \Phi_M \,.$$

Hence, $(\partial H/\partial u^M)_{V=\overline{V}}$ must vanish.

[1] To justify it just insert in the right hand side H^v as given by (J.4).