The extra non-invariant term in Eq. (7.5.8) is just what we saw in Section 6.2 is needed to cancel a non-invariant term in the propagator of $\partial\phi$.

Vector Field, Spin One

Similar results are obtained in the canonical quantization of the vector field V_μ for a particle of spin one. Let's here keep an open mind, and write the Lagrangian density in a fairly general form

$$\mathscr{L} = -\tfrac{1}{2}\alpha\,\partial_\mu V_\nu\,\partial^\mu V^\nu - \tfrac{1}{2}\beta\,\partial_\mu V_\nu\,\partial^\nu V^\mu - \tfrac{1}{2}m^2\,V_\mu V^\mu - J_\mu V^\nu\,, \qquad (7.5.9)$$

where α, β, and m^2 are so far arbitrary constants, and J_μ is either a c-number external current, or an operator depending on fields other than V^μ, in which case additional terms involving these fields must be added to $\mathscr{L}$. The Euler–Lagrange field equations for V_μ read

$$\alpha\Box V^\nu + \beta\,\partial_\nu(\partial_\mu V^\mu) + m^2 V_\mu = -J_\mu\,.$$

Taking the divergence gives

$$(\alpha+\beta)\Box\partial_\lambda V^\alpha + m^2\partial_\lambda V^\lambda = -\partial_\lambda J^\lambda\,. \qquad (7.5.10)$$

This is the equation for an ordinary scalar field with mass $m^2/(\alpha+\beta)$ and source $\partial_\lambda J^\lambda/(\alpha+\beta)$. We want to describe a theory containing only particles of spin one, not spin zero, so to avoid the appearance of $\partial_\lambda V^\lambda$ as an independently propagating scalar field, we take $\alpha = -\beta$, in which case $\partial_\lambda V^\lambda$ can be expressed in terms of an external current or other fields, as $-\partial_\lambda J^\lambda/m^2$. The constant α can be absorbed in the definition of V_μ, so we can take $\alpha = -\beta = 1$, and therefore

$$\mathscr{L} = -\tfrac{1}{4}\,F_{\mu\nu}F^{\mu\nu} - \tfrac{1}{2}m^2\,V_\mu V^\mu - J_\mu V^\mu\,, \qquad (7.5.11)$$

where

$$F_{\mu\nu} \equiv \partial_\mu V_\nu - \partial_\nu V_\mu\,. \qquad (7.5.12)$$

The derivative of the Lagrangian with respect to the time-derivative of the vector field is

$$\frac{\partial\mathscr{L}}{\partial\dot{V}^\mu} = -F^{0\mu}\,. \qquad (7.5.13)$$

This is non-vanishing for μ a spatial index i, so the V^i are canonical fields, with conjugates

$$\Pi^i = F^{i0} = \dot{V}^i + \partial_i V^0\,. \qquad (7.5.14)$$

On the other hand $F^{00} = 0$, so $\dot{V}^0$ does not appear in the Lagrangian, and V^0 is therefore an auxiliary field. This causes no serious difficulty: the fact that $\partial\mathscr{L}/\partial\dot{V}^0$ vanishes means that the field equation for V^0 involves no second time-derivatives, and can therefore be used as a constraint that eliminates a field variable. Specifically, the Euler–Lagrange equation for $\nu = 0$ is

$$\partial_i F^{i0} = m^2 V^0 + J^0 \tag{7.5.15}$$

or using Eq. (7.5.14)

$$V^0 = \frac{1}{m^2}(\nabla \cdot \boldsymbol{\Pi} - J^0) . \tag{7.5.16}$$

Now let us calculate the Hamiltonian $H = \int d^3x \, (\boldsymbol{\Pi} \cdot \dot{\mathbf{V}} - \mathscr{L})$ for this theory. Eq. (7.5.14) allows us to write $\dot{\mathbf{V}}$ in terms of $\boldsymbol{\Pi}$ and J^0:

$$\dot{\mathbf{V}} = -\nabla V^0 + \boldsymbol{\Pi} \;=\; \boldsymbol{\Pi} - \frac{1}{m^2}\nabla(\nabla \cdot \boldsymbol{\Pi} - J^0),$$

so

$$\begin{aligned} H = \int d^3x \Big[& \boldsymbol{\Pi}^2 + m^{-2}(\nabla \cdot \boldsymbol{\Pi})(\nabla \cdot \boldsymbol{\Pi} - J^0) \\ & - \tfrac{1}{2}\boldsymbol{\Pi}^2 + \tfrac{1}{2}(\nabla \cdot \mathbf{V})^2 + \tfrac{1}{2}m^2 \mathbf{V}^2 \\ & - \tfrac{1}{2}m^{-2}(\nabla \cdot \boldsymbol{\Pi} - J^0)^2 + \mathbf{J} \cdot \mathbf{V} - m^{-2}J^0(\nabla \cdot \boldsymbol{\Pi} - J^0) \Big] . \end{aligned}$$

Again, we split this up into a free-particle term H_0 and interaction V:

$$H = H_0 + V , \tag{7.5.17}$$

and pass to the interaction picture by replacing the Heisenberg-picture quantities $\mathbf{V}$ and $\boldsymbol{\Pi}$ with their interaction-picture counterparts $\mathbf{v}$ and $\boldsymbol{\pi}$ (and, though not shown explicitly, likewise for whatever fields and conjugates are present in J^μ):

$$H_0 = \int d^3x \left[\frac{1}{2}\boldsymbol{\pi}^2 + \frac{1}{2m^2}(\nabla \cdot \boldsymbol{\pi})^2 + \frac{1}{2}(\nabla \times \mathbf{v})^2 + \frac{m^2}{2}\mathbf{v}^2 \right] , \tag{7.5.18}$$

$$V = \int d^3x \left[\mathbf{J} \cdot \mathbf{v} - m^{-2} J^0 \nabla \cdot \boldsymbol{\pi} + \frac{1}{2m^2}(J^0)^2 \right] . \tag{7.5.19}$$

The relation between $\boldsymbol{\pi}$ and $\mathbf{v}$ is then

$$\dot{\mathbf{v}} = \frac{\delta H_0(\mathbf{v}, \boldsymbol{\pi})}{\delta \boldsymbol{\pi}} = \boldsymbol{\pi} - m^{-2}\nabla(\nabla \cdot \boldsymbol{\pi}) \tag{7.5.20}$$

and the 'field equation' is

$$\dot{\boldsymbol{\pi}} = -\frac{\delta H_0(\mathbf{v}, \boldsymbol{\pi})}{\delta \mathbf{v}} = +\nabla^2 \mathbf{v} - \nabla(\nabla \cdot \mathbf{v}) - m^2 \mathbf{v} . \tag{7.5.21}$$

Since V^0 is not an independent field variable, it is not related by a similarity transformation to any interaction-picture object v^0. Instead, we can *invent* a quantity

$$v^0 \equiv m^{-2}\nabla \cdot \boldsymbol{\pi} . \tag{7.5.22}$$

Eq. (7.5.20) then allows us to write $\boldsymbol{\pi}$ as

$$\boldsymbol{\pi} = \dot{\mathbf{v}} + \nabla v^0 \,. \tag{7.5.23}$$

Inserting this in Eqs. (7.5.22) and (7.5.21) gives our field equations in the form

$$\nabla^2 v^0 + \nabla \cdot \dot{\mathbf{v}} - m^2 v^0 = 0 \,,$$
$$\nabla^2 \mathbf{v} - \nabla(\nabla \cdot \mathbf{v}) - \ddot{\mathbf{v}} - \nabla \dot{v}^0_i - m^2 \mathbf{v} = 0 \,.$$

These can be combined in the covariant form

$$\Box v^\mu - \partial^\mu \partial_\nu v^\nu - m^2 v^\mu = 0 \,. \tag{7.5.24}$$

Taking the divergence gives

$$\partial_\mu v^\mu = 0 \tag{7.5.25}$$

and hence

$$(\Box - m^2) v^\mu = 0 \,. \tag{7.5.26}$$

A real vector field satisfying Eqs. (7.5.25) and (7.5.26) can be expressed as a Fourier transform

$$v^\mu(x) = (2\pi)^{-3/2} \sum_\sigma \int d^3p \, (2p^0)^{-1/2} \Big\{ e^\mu(\mathbf{p}, \sigma) a(\mathbf{p}, \sigma) e^{ip\cdot x} + e^{\mu *}(\mathbf{p}, \sigma) a^\dagger(\mathbf{p}, \sigma) \, e^{-ip\cdot x} \Big\} \,, \tag{7.5.27}$$

where $p^0 = \sqrt{\mathbf{p}^2 + m^2}$; the $e^\mu(\mathbf{p}, \sigma)$ for $\sigma = +1, 0, -1$ are three independent vectors satisfying

$$p_\mu e^\mu(\mathbf{p}, \sigma) = 0 \tag{7.5.28}$$

and normalized so that

$$\sum_\sigma e^\mu(\mathbf{p}, \sigma) e^{\nu *}(\mathbf{p}, \sigma) = \eta^{\mu\nu} + p^\mu p^\nu / m^2 \quad ; \tag{7.5.29}$$

and the $a(\mathbf{p}, \sigma)$ are operator coefficients. It is straightforward using Eqs. (7.5.23), (7.5.27), and (7.5.29) to calculate that $\mathbf{v}$ and $\boldsymbol{\pi}$ satisfy the correct commutation relations

$$\Big[v^i(\mathbf{x}, t), \, \pi^j(\mathbf{y}, t) \Big] = i \, \delta_{ij} \, \delta^3(\mathbf{x} - \mathbf{y}) \,,$$
$$\Big[v^i(\mathbf{x}, t), \, v^j(\mathbf{x}, t) \Big] = \Big[\pi^i(\mathbf{x}, t), \, \pi^j(\mathbf{x}, t) \Big] = 0 \,, \tag{7.5.30}$$

provided that $a(\mathbf{p}, \sigma)$ and $a^\dagger(\mathbf{p}, \sigma)$ satisfy the commutation relations

$$\Big[a(\mathbf{p}, \sigma), \, a^\dagger(\mathbf{p}', \sigma') \Big] = \delta^3(\mathbf{p}' - \mathbf{p}) \delta_{\sigma'\sigma} \,, \tag{7.5.31}$$

$$\Big[a(\mathbf{p}, \sigma), \, a(\mathbf{p}', \sigma') \Big] = 0 \,. \tag{7.5.32}$$

We already know that the vector field for a spin one particle must take the form (7.5.27), so our derivation of these results serves to verify that Eq. (7.5.18) gives the correct free-particle Hamiltonian for a massive particle of spin one. It is easy to check also that Eq. (7.5.18) may be written (up to a constant term) in the standard form of a free-particle energy, as $\sum_\sigma \int d^3p\, p^0\, a^\dagger(\mathbf{p},\sigma)a(\mathbf{p},\sigma)$. Finally, using Eq. (7.5.22) in Eq. (7.5.19) yields the interaction in the interaction picture

$$V(t) = \int d^3x \left[J_\mu v^\mu \; + \; \frac{1}{2m^2}(J^0)^2\right] . \qquad (7.5.33)$$

The extra non-invariant term in Eq. (7.5.33) is just what we found in Chapter 6 is needed to cancel a non-invariant term in the propagator of the vector field.

Dirac field, Spin One Half

For the Dirac field of a particle of spin 1/2, we tentatively take the Lagrangian as

$$\mathscr{L} = -\bar{\Psi}(\gamma^\mu\partial_\mu + m)\Psi - \mathscr{H}(\bar{\Psi},\Psi) \qquad (7.5.34)$$

with $\mathscr{H}$ a real function of $\bar{\Psi}$ and Ψ. This is not real, but the action is, because

$$\bar{\Psi}\gamma^\mu\partial_\mu\Psi - (\bar{\Psi}\gamma^\mu\partial_\mu\Psi)^\dagger = \bar{\Psi}\gamma^\mu\partial_\mu\Psi + (\partial_\mu\bar{\Psi})\gamma^\mu\Psi = \partial_\mu(\bar{\Psi}\gamma^\mu\Psi) .$$

Hence the field equations obtained by requiring the action to be stationary with respect to $\bar{\Psi}$ are the adjoints of those obtained by requiring the action to be stationary with respect to Ψ, as necessary if we are to avoid having too many field equations. The canonical conjugate to Ψ is

$$\Pi = \frac{\partial\mathscr{L}}{\partial\dot{\Psi}} = -\bar{\Psi}\gamma^0 , \qquad (7.5.35)$$

so we should not regard $\bar{\Psi}$ as a field like Ψ, but rather as proportional to the canonical conjugate of Ψ. The Hamiltonian is

$$H = \int d^3x[\Pi\dot{\Psi} - \mathscr{L}] = \int d^3x\Big[\Pi\gamma^0[\boldsymbol{\gamma}\cdot\nabla + m]\Psi + \mathscr{H}\Big] .$$

We write this as

$$H = H_0 + V , \qquad (7.5.36)$$

where

$$H_0 = \int d^3x\, \Pi\gamma_0[\boldsymbol{\gamma}\cdot\nabla + m]\Psi , \qquad (7.5.37)$$

$$V = \int d^3x\, \mathscr{H}(\bar{\Psi},\Psi) . \qquad (7.5.38)$$

We now pass to the interaction picture. Since Eq. (7.5.35) does not involve the time, the similarity transformation (7.1.28), (7.1.29) yields immediately

$$\pi = -\bar{\psi}\,\gamma^0 \,. \tag{7.5.39}$$

Likewise, H_0 and $V(t)$ can be calculated by replacing Ψ and Π with ψ and π in Eqs. (7.5.37) and (7.5.38). This gives the equation of motion

$$\dot{\psi} = \frac{\delta H_0}{\delta \pi} = \gamma_0(\boldsymbol{\gamma}\cdot\nabla + m)\psi \tag{7.5.40}$$

or more neatly

$$(\gamma^\mu \partial_\mu + m)\psi = 0\,. \tag{7.5.41}$$

(The other equation of motion, $\dot{\pi} = -\delta H_0 \big/ \delta\psi$, yields just the adjoint of this one.) Any field satisfying Eq. (7.5.41) can be written as a Fourier transform

$$\psi(x) = (2\pi)^{-3/2}\int d^3p \sum_\sigma \left\{u(\mathbf{p},\sigma)\,e^{ip\cdot x}a(\mathbf{p},\sigma) + \; v(\mathbf{p},\sigma)e^{-ip\cdot x}b^\dagger(\mathbf{p},\sigma)\right\}, \tag{7.5.42}$$

where $p^0 \equiv \sqrt{\mathbf{p}^2 + m^2}$; $a(\mathbf{p},\sigma)$ and $b^\dagger(\mathbf{p},\sigma)$ are operator coefficients; and $u(\mathbf{p}, \pm\frac{1}{2})$ are the two independent solutions of

$$(i\gamma^\mu p_\mu + m)\,u(\mathbf{p},\sigma) = 0 \tag{7.5.43}$$

and likewise

$$(-i\gamma^\mu p_\mu + m)\,v(\mathbf{p},\sigma) = 0 \tag{7.5.44}$$

normalized so that*

$$\sum_\sigma u(\mathbf{p},\sigma)\bar{u}(\mathbf{p},\sigma) = \frac{(-i\gamma^\mu p_\mu + m)}{2p^0}\,, \tag{7.5.45}$$

$$\sum_\sigma v(\mathbf{p},\sigma)\bar{v}(\mathbf{p},\sigma) = -\frac{(i\gamma^\mu p_\mu + m)}{2p^0}\,. \tag{7.5.46}$$

In order to obtain the desired anticommutators

$$\begin{aligned}\Big[\psi_\alpha(\mathbf{x},t),\,\bar{\psi}_\beta(\mathbf{y},t)\Big]_+ &= \Big[\psi_\alpha(\mathbf{x},t),\,\pi_\gamma(\mathbf{y},t)\Big]_+(\gamma^0)_{\gamma\beta}\\ &= i\,(\gamma^0)_{\alpha\beta}\delta^3(\mathbf{x}-\mathbf{y})\,,\end{aligned} \tag{7.5.47}$$

$$\Big[\psi_\alpha(\mathbf{x},t),\,\psi_\beta(\mathbf{y},t)\Big]_+ = 0\,, \tag{7.5.48}$$

* The matrix $i\gamma^\mu p_\mu$ has eigenvalues $\pm m$, so $\Sigma u\bar{u}$ and $\Sigma v\bar{v}$ must be proportional to the projection matrices $(-i\gamma^\mu p_\mu + m)/2m$ and $(i\gamma^\mu p_\mu + m)/2m$, respectively. The proportionality factor may be adjusted up to a sign by absorbing it in the definition of u and v. The overall sign is determined by positivity: Tr $\Sigma u\bar{u}\beta = \Sigma u^\dagger u$ and Tr $\Sigma v\bar{v}\beta = \Sigma v^\dagger v$ must be positive.

we must adopt the anticommutation relations

$$\Big[a(\mathbf{p},\sigma),\, a^\dagger(\mathbf{p}',\sigma')\Big]_+ = \Big[b(\mathbf{p},\sigma),\, b^\dagger(\mathbf{p}',\sigma')\Big]_+ = \delta^3(\mathbf{p}'-p)\delta_{\sigma'\sigma}\,, \qquad (7.5.49)$$

$$\Big[a(\mathbf{p},\sigma),\, a(\mathbf{p}',\sigma')\Big]_+ = \Big[b(\mathbf{p},\sigma),\, b(\mathbf{p}',\sigma')\Big]_+ =$$

$$\Big[a(\mathbf{p},\sigma),\, b(\mathbf{p}',\sigma')\Big]_+ = \Big[a(\mathbf{p},\sigma),\, b^\dagger(\mathbf{p}',\sigma')\Big]_+ = 0\,, \qquad (7.5.50)$$

and their adjoints. These agree with the results obtained in Chapter 5, thus verifying that (7.5.37) is the correct free-particle Hamiltonian for spin $\frac{1}{2}$. In terms of the as and bs, this Hamiltonian is

$$H_0 = \sum_\sigma \int d^3p\; p^0\Big(a^\dagger(\mathbf{p},\sigma)a(\mathbf{p},\sigma) - b(\mathbf{p},\sigma)b^\dagger(\mathbf{p},\sigma)\Big)\,. \qquad (7.5.51)$$

We can rewrite this as a more conventional free-particle Hamiltonian, plus another infinite c-number**

$$H_0 = \sum_\sigma \int d^3\mathbf{p}\, p^0\Big[a^\dagger(\mathbf{p},\sigma)a(\mathbf{p},\sigma) + b^\dagger(\mathbf{p},\sigma)b(\mathbf{p},\sigma) - \delta^3(\mathbf{p}-\mathbf{p})\Big]\,. \quad (7.5.52)$$

The c-number term in Eq. (7.5.52) is only important if we worry about gravitational phenomena; otherwise here, as for the scalar field, we can throw it away, since it only affects the zero of energy with respect to which all energies are measured. With this understanding, H_0 is a positive operator, just as for bosons.

7.6 Constraints and Dirac Brackets

The chief obstacle to deriving the Hamiltonian from the Lagrangian is the occurrence of constraints. The standard analysis of this problem is that of Dirac,[5] whose terminology we will follow here. Dirac's analysis is not really needed for the simple theories discussed in this chapter, where it is easy to identify the unconstrained canonical variables. We shall use the theory of a real massive vector field for illustration here, returning to Dirac's approach in the next chapter, where it will be actually useful.

Primary constraints are either imposed on the system (as when in the next chapter we choose a gauge for the electromagnetic field) or arise from the structure of the Lagrangian itself. For an example of the latter type, consider the Lagrangian (7.5.11) of a massive vector field V^μ interacting

** Note the negative sign of the c-number term. The conjectured symmetry known as *supersymmetry*[4] connects the numbers of boson and fermion fields, in such a way that the c-numbers in H_0 all cancel.

with a current J_μ:

$$\mathscr{L} = -\tfrac{1}{4} F_{\mu\nu} F^{\mu\nu} - \tfrac{1}{2} m^2 V_\mu V^\mu - J_\mu V^\mu \tag{7.6.1}$$

where

$$F_{\mu\nu} \equiv \partial_\mu V_\nu - \partial_\nu V_\mu \,. \tag{7.6.2}$$

Suppose we try to treat all four components of V^μ on the same basis. We should then define the conjugates

$$\Pi_\mu \equiv \frac{\partial \mathscr{L}}{\partial(\partial_0 A^\mu)} = -F^{0\mu} \,. \tag{7.6.3}$$

We immediately find the primary constraint:

$$\Pi_0 = 0 \,. \tag{7.6.4}$$

More generally, we encounter primary constraints whenever the equations $\Pi_\ell = \delta L/\delta\partial_0\Psi^\ell$ cannot be solved to give all the $\partial_0\Psi^\ell$ (at least locally) in terms of Π_ℓ and Ψ^ℓ. This will be the case if and only if the matrix $\delta^2 L/\delta(\partial_0\Psi^\ell)\,\delta(\partial_0\Psi^m)$ has vanishing determinant. Such Lagrangians are called *irregular*.

Then there are *secondary constraints*, which arise from the requirement that the primary constraints be consistent with the equations of motion. For the massive vector field, this is just the Euler–Lagrange equation (7.5.16) for V^0:

$$\partial_i \Pi_i = m^2 V^0 + J^0 \,. \tag{7.6.5}$$

Here we are finished, but in other theories we might encounter further constraints by requiring consistency of the secondary constraints with the field equations, and so on. The distinction between primary, secondary, etc. constraints is not important; we will treat them all together here.

There is another distinction between certain types of constraint that is more important. The constraints we have found for the massive vector field are of a type known as *second class*, for which there is a universal prescription for the commutation relations. To explain the distinction between first and second class constraints, and the prescription used to deal with second class constraints, it is useful first to recall the definition of the Poisson brackets of classical mechanics.

Consider any Lagrangian $L(\Psi, \dot\Psi)$ that depends on a set of variables $\Psi^a(t)$ and their time-derivatives $\dot\Psi^a(t)$. (The Lagrangians of quantum field theory are a special case, with the index a running over all pairs of ℓ and $\mathbf{x}$.) We can define canonical conjugates for *all* of these variables by

$$\Pi_a \equiv \frac{\partial L}{\partial \dot\Psi^a} \,. \tag{7.6.6}$$

The Πs and Ψs will in general not be independent variables, but may instead be related by various constraint equations, both primary and secondary. The Poisson bracket is then defined by

$$[A, B]_{\rm P} \equiv \frac{\partial A}{\partial \Psi^a}\frac{\partial B}{\partial \Pi_a} - \frac{\partial B}{\partial \Psi^a}\frac{\partial A}{\partial \Pi_a} \tag{7.6.7}$$

with the constraints ignored in calculating the derivatives with respect to Ψ^a and Π_a. In particular, we always have $[\Psi^a, \Pi_b]_{\rm P} = \delta^a_b$. (Here and below all fields are taken at the same time, and time arguments are everywhere dropped.) These brackets have the same algebraic properties as commutators:

$$[A, B]_{\rm P} = -[B, A]_{\rm P}, \tag{7.6.8}$$

$$[A, BC]_{\rm P} = [A, B]_{\rm P}C + B[A, C]_{\rm P}, \tag{7.6.9}$$

including the Jacobi identity

$$[A, [B, C]_{\rm P}]_{\rm P} + [B, [C, A]_{\rm P}]_{\rm P} + [C, [A, B]_{\rm P}]_{\rm P} = 0\,. \tag{7.6.10}$$

If we could adopt the usual commutation relations $[\Psi^a, \Pi_b] = i\delta^a_b$, $[\Psi^a, \Psi^b] = [\Pi_a, \Pi_b] = 0$, then the commutator of any two functions of the Ψs and Πs would be just $[A, B] = i[A, B]_{\rm P}$. But the constraints do not always allow this.

The constraints may in general be expressed in the form $\chi_N = 0$, where the χ_N are a set of functions of the Ψs and Πs. Because we are including secondary constraints along with the primary constraints, the set of all the constraints is necessarily consistent with the equations of motion $\dot{A} = [A, H]_{\rm P}$, and therefore

$$[\chi_N, H]_{\rm P} = 0 \tag{7.6.11}$$

when the constraint equations $\chi_N = 0$ are satisfied.

We call a constraint *first class* if its Poisson bracket with all the other constraints vanishes when (*after* calculating the Poisson brackets) we impose the constraints. We shall see a simple example of such a constraint in the quantization of the electromagnetic field in the next chapter, where the first class constraint arises from a symmetry of the action, electromagnetic gauge invariance. In fact, the set of first class constraints $\chi_N = 0$ is always associated with a group of symmetries, under which an arbitrary quantity A undergoes the infinitesimal transformation

$$\delta_N A \equiv \sum_N \epsilon_N\, [\chi_N, A]_{\rm P}\,. \tag{7.6.12}$$

(In field theory these are local transformations, because the index N contains a spacetime coordinate.) Eq. (7.6.11) shows that this transformation leaves the Hamiltonian invariant, and for first class constraints it also respects all other constraints. Such first class constraints can be eliminated by a choice of gauge, or treated by gauge-invariant methods described in Volume II.

After all of the first class constraints have been eliminated, the remaining constraint equations $\chi_N = 0$ are such that no linear combination $\sum_N u_N [\chi_N, \chi_M]_{\rm P}$ of the Poisson brackets of these constraints with each other vanishes. It follows that the matrix of the Poisson brackets of the remaining constraints is non-singular:

$$\text{Det}\, C \neq 0\,, \tag{7.6.13}$$

where

$$C_{NM} \equiv [\chi_N, \chi_M]_{\rm P}\,. \tag{7.6.14}$$

Constraints of this sort are called *second class*. Note that there must always be an even number of second class constraints, because an antisymmetric matrix of odd dimensionality necessarily has vanishing determinant.

As we have seen, in the case of the massive real vector field the constraints are

$$\chi_{1\mathbf{x}} = \chi_{2\mathbf{x}} = 0\,, \tag{7.6.15}$$

where

$$\chi_{1\mathbf{x}} = \Pi_0(\mathbf{x})\,, \qquad \chi_{2\mathbf{x}} = \partial_i \Pi_i(\mathbf{x}) - m^2 V^0(\mathbf{x}) - J^0(\mathbf{x})\,. \tag{7.6.16}$$

The Poisson bracket of these constraints is

$$C_{1\mathbf{x},2\mathbf{y}} = -C_{2\mathbf{y},1\mathbf{x}} = [\chi_{1\mathbf{x}}, \chi_{2\mathbf{y}}]_{\rm P} = m^2 \delta^3(\mathbf{x} - \mathbf{y}) \tag{7.6.17}$$

and, of course,

$$C_{1\mathbf{x},1\mathbf{y}} = C_{2\mathbf{x},2\mathbf{y}} = 0\,. \tag{7.6.18}$$

This 'matrix' is obviously non-singular, so the constraints (7.6.15) are second class.

Dirac suggested that when all constraints are second class, the commutation relations will be given by

$$[A, B] = i[A, B]_{\rm D}\,, \tag{7.6.19}$$

where $[A, B]_{\rm D}$ is a generalization of the Poisson bracket known as the *Dirac bracket*:

$$[A, B]_{\rm D} \equiv [A, B]_{\rm P} - [A, \chi_N]_{\rm P} \,(C^{-1})^{NM}\, [\chi_M, B]_{\rm P}\,. \tag{7.6.20}$$

(Here N and M are compound indices including the position in space, taking values like $1,\mathbf{x}$ and $2,\mathbf{x}$ in the vector field example.) He noted that the Dirac bracket like the Poisson bracket satisfies the same algebraic relations as the commutators

$$[A, B]_{\rm D} = -[B, A]_{\rm D} , \tag{7.6.21}$$

$$[A, BC]_{\rm D} = [A, B]_{\rm D} C + B[A, C]_{\rm D} , \tag{7.6.22}$$

$$[A, [B, C]_{\rm D}]_{\rm D} + [B, [C, A]_{\rm D}]_{\rm D} + [C, [A, B]_{\rm D}]_{\rm D} = 0 , \tag{7.6.23}$$

and also the relations

$$[\chi_N, B]_{\rm D} = 0 \tag{7.6.24}$$

which make the commutation relations (7.6.19) consistent with the constraints $\chi_N = 0$. Also, the Dirac brackets are unchanged if we replace the χ_N with any functions χ'_N for which the equations $\chi'_N = 0$ and $\chi_N = 0$ define the same submanifold of phase space. But all these agreeable properties do not prove that the commutators are actually given by Eq. (7.6.19) in terms of the Dirac brackets.

This issue is illuminated if not settled by a powerful theorem proved by Maskawa and Nakajima.[6] They showed that for any set of canonical variables Ψ^a, Π_a governed by second class constraints, it is always possible by a canonical transformation* to construct two sets of variables Q^n, $\mathscr{Q}^r$ and their respective conjugates P_n, $\mathscr{P}_r$, such that the constraints read $\mathscr{Q}^r = \mathscr{P}_r = 0$. Using these coordinates to calculate Poisson brackets, and redefining the constraint functions as $\chi_{1r} = \mathscr{Q}^r$, $\chi_{2r} = \mathscr{P}_r$, we have

$$C_{1r,2s} = [\mathscr{Q}^r, \mathscr{P}_s]_P = \delta^r_s ,$$

$$C_{1r,1s} = [\mathscr{Q}^r, \mathscr{Q}^s]_P = 0 , \qquad C_{2r,2s} = [\mathscr{P}_r, \mathscr{P}_s]_P = 0 ,$$

and for any functions A, B

$$[A, \chi_{1r}]_P = -\frac{\partial A}{\partial \mathscr{P}^r} , \qquad [A, \chi_{2r}]_P = \frac{\partial A}{\partial \mathscr{Q}_r} ,$$

This C-matrix has inverse $C^{-1} = -C$, so the Dirac brackets (7.6.20) are

* Recall that by a canonical transformation, we mean a transformation from a set of phase space coordinates Ψ^a, Π_a to some other phase space coordinates $\tilde{\Psi}^a$, $\tilde{\Pi}_a$, such that $[\tilde{\Psi}^a, \tilde{\Pi}_b]_P = \delta^a_b$ and $[\tilde{\Psi}^a, \tilde{\Psi}^b]_P = [\tilde{\Pi}_a, \tilde{\Pi}_b]_P = 0$, the Poisson brackets being calculated in terms of the Ψ^a and Π_a. It follows that the Poisson brackets for any functions A, B are the same whether calculated in terms of Ψ^a and Π_a or in terms of $\tilde{\Psi}^a$ and $\tilde{\Pi}_a$. It also follows that if Ψ^a and Π_a satisfy the Hamiltonian equations of motion, then so do $\tilde{\Psi}^a$ and $\tilde{\Pi}_a$, with the same Hamiltonian. The Lagrangian is changed by a canonical transformation, but only by a time-derivative, which does not affect the action.

here

$$[A,B]_D = [A,B]_P + [A,\chi_{1r}]_P[\chi_{2r},B]_P - [A,\chi_{2r}]_P[\chi_{1r},B]_P$$
$$= [A,B]_P - \frac{\partial A}{\partial \mathcal{Q}^r}\frac{\partial B}{\partial \mathcal{P}_r} + \frac{\partial B}{\partial \mathcal{Q}^r}\frac{\partial A}{\partial \mathcal{P}_r}$$
$$= \frac{\partial A}{\partial Q^n}\frac{\partial B}{\partial P_n} - \frac{\partial B}{\partial Q^n}\frac{\partial A}{\partial P_n} . \tag{7.6.25}$$

In other words, *the Dirac bracket is equal to the Poisson bracket calculated in terms of the reduced set of unconstrained canonical variables Q^n, P_n*. If we assume that these unconstrained variables satisfy the canonical commutation relations, then the commutators of general operators A, B are given by Eq. (7.6.19) in terms of the Dirac brackets.**

We now return to the massive vector field, to see how it can be quantized using Dirac brackets. This is a case where it is easy to express the constrained variables V^0 and Π_0 in terms of the unconstrained ones† V_i and Π_i; we have simply $\Pi_0 = 0$, and V^0 is given by Eq. (7.6.5). From Eqs. (7.6.17) and (7.6.18), we see that C_{NM} here has the inverse

$$(C^{-1})^{1\mathbf{x},2\mathbf{y}} = -(C^{-1})^{2\mathbf{y},1\mathbf{x}} = -m^{-2}\delta^3(\mathbf{x}-\mathbf{y}), \tag{7.6.26}$$

$$(C^{-1})^{1\mathbf{x},1\mathbf{y}} = (C^{-1})^{2\mathbf{x},2\mathbf{y}} = 0 . \tag{7.6.27}$$

Therefore the Dirac prescription (7.6.19), (7.6.20) yields the equal-time commutators

$$[A,B] = i[A,B]_{\rm P} + im^{-2}\int d^3z\,\Big([A,\Pi_0(\mathbf{z})]_{\rm P}\,[\partial_i\Pi_i(\mathbf{z}) - m^2V^0(\mathbf{z}) - J^0(\mathbf{z}),B]_{\rm P} - A\leftrightarrow B\Big). \tag{7.6.28}$$

By definition, we have

$$[V^\mu(\mathbf{x}),\Pi_\nu(\mathbf{y})]_{\rm P} = \delta^3(\mathbf{x}-\mathbf{y})\delta^\mu_\nu , \quad [V^\mu(\mathbf{x}),V^\nu(\mathbf{y})]_{\rm P} = [\Pi_\mu(\mathbf{x}),\Pi_\nu(\mathbf{y})]_{\rm P} = 0 . \tag{7.6.29}$$

Hence

$$[V^i(\mathbf{x}),V^j(\mathbf{y})] = [V^0(\mathbf{x}),V^0(\mathbf{y})] = 0 ,$$

** It is still an open question whether we should adopt canonical commutation relations for the unconstrained variables Q^n, P_n constructed by the Maskawa-Nakajima canonical transformation. Ultimately, the test of such canonical commutation relations is their consistency with the free-field commutation relations derived in Chapter 5, but to apply this test we need to know what the Q^n and P_n are. In the Appendix to this chapter we display two large classes of theories in which we can identify a set of unconstrained Qs and Ps, such that the Dirac commutation relations (7.6.19) follow from the ordinary canonical commutation relations of the Qs and Ps. We shall also show that in these cases, the Hamiltonian defined in terms of the unconstrained Ψs and Πs may be written just as well in terms of the constrained variables.

† This is a special case of the theories discussed in Part A of the Appendix.

$$\begin{aligned}
[V^i(\mathbf{x}), V^0(\mathbf{y})] &= -im^{-2}\partial_i\delta^3(\mathbf{x}-\mathbf{y}) , \\
[V^i(\mathbf{x}), \Pi_j(\mathbf{y})] &= i\delta^i_j\delta^3(\mathbf{x}-\mathbf{y}) , \\
[V^0(\mathbf{x}), \Pi_j(\mathbf{y})] &= [V^\mu(\mathbf{x}), \Pi_0(\mathbf{y})] = 0 , \\
[\Pi^\mu(\mathbf{x}), \Pi^\nu(\mathbf{y}] &= 0 .
\end{aligned} \tag{7.6.30}$$

These are indeed just the commutation relations that we would find by assuming that the unconstrained variables satisfy the usual canonical commutation relations $[V^i(\mathbf{x}), \Pi_j(\mathbf{y})] = i\delta^i_j\delta^3(\mathbf{x}-\mathbf{y})$, $[V^i(\mathbf{x}), V^j(\mathbf{y})] = [\Pi_i(\mathbf{x}), \Pi_j(\mathbf{y})] = 0$, and using the constraints to evaluate the commutators involving Π_0 and V^0.

7.7 Field Redefinitions and Redundant Couplings*

Observables like masses and S-matrix elements are independent of some of the coupling parameters in any action, known as the *redundant* parameters. This is because changes in these parameters can be undone by simply redefining the field variables. A continuous redefinition of the fields, such as an infinitesimal local transformation $\Psi^\ell(x) \to \Psi^\ell(x) + \epsilon F^\ell(\Psi(x), \partial_\mu\Psi(x), \cdots)$, clearly cannot affect any *observable* of the theory,** though, of course, it would change the values of matrix elements of the fields themselves.

How can we tell whether some variation in the parameters of a theory can be cancelled by a field redefinition? A continuous local field redefinition will produce a change in the action of the form

$$\delta I[\Psi] = \epsilon \sum_\ell \int d^4x \, \frac{\delta I[\Psi]}{\delta\Psi^\ell(x)} \, F^\ell(\Psi(x), \partial\Psi(x), \cdots) \quad . \tag{7.7.1}$$

So any change δg_i in the coupling parameters g_i, for which the change in the action is of the form

$$\sum_i \frac{\partial I}{\partial g_i} \, \delta g_i = -\epsilon \sum_\ell \int d^4x \, \frac{\delta I[\Psi]}{\delta\Psi^\ell(x)} \, F^\ell(\Psi(x), \partial\Psi(x), \cdots) , \tag{7.7.2}$$

may be compensated by a field redefinition

$$\Psi^\ell(x) \to \Psi^\ell(x) + \epsilon F^\ell(\Psi(x), \partial_\mu\Psi(x), \cdots) ,$$

* This section lies somewhat out of the book's main line of development, and may be omitted in a first reading.

** For instance, the theorem of Section 10.2 shows that as long as we multiply by the correct field renormalization constants, S-matrix elements can be obtained from the vacuum expectation value of a time-ordered product of *any* operators that have non-vanishing matrix elements between the vacuum and the one-particle states of the particles participating in the reaction.