International Journal of Theoretical Physics, Vol. 6, No. 4 (1972), pp. 307–320

Gauge-Independent Quantization of the Schrödinger and Radiation Fields†

EGON MARX

Drexel University, Philadelphia, Pennsylvania 19104, U.S.A.‡

Received: 11 *October* 1971

Abstract

The quantization of several Schrödinger fields interacting with the electromagnetic field is carried out without reference to a particular gauge. The canonical formalism requires a modification introduced by Dirac and Bergmann for constraints. The Coulomb interaction is separated from the radiation and it gives rise to bound states of atoms and molecules. Particle operators are represented in the usual manner in Fock space, while the radiation field can be described by state functionals. Constraints can be included in the canonical formalism by Lagrange multipliers, leading to results equivalent to those of Dirac and Bergmann.

1. *Introduction*

There are a number of possible approaches to the problem of emission and absorption of radiation by atoms and molecules (Kramers, 1957). The most realistic treatment involves a quantized radiation field, and theories involving only a classical electromagnetic field are usually accepted only as approximations. It is then quite logical to extend to the non-relativistic Schrödinger field the well-developed formalism of quantum electrodynamics. Nevertheless, it is difficult to take into account bound states in a theory primarily developed for scattering problems, and these states are obviously important for the interaction of radiation with atoms and molecules. It then follows that the Coulomb interaction should be separated from the radiation field, and this is usually carried out in a Coulomb or radiation gauge.

It is well known that the electromagnetic field has only two degrees of freedom, and that a representation by four potentials or six field components is redundant, leading to constraint equations. We also have to require that all physical results be independent of the gauge chosen to perform the calculations. In order to help the understanding of the theory and to

† This work was supported in part by Drexel University.

‡ Current address: Harry Diamond Laboratories, Electromagnetic Effects Laboratory, Dawson Beach Rd., Woodbridge, Va 22191, U.S.A.

facilitate comparisons with the relativistic analogues, we derive the equations which often serve as a starting point in this problem in a systematic way from the simplest assumptions of quantum field theory. And we do so without reference to any particular gauge.

We start from a Lagrangian density for several free fields that obey the Schrödinger equation, we introduce what is often called the minimal interaction with the electromagnetic field and we add a gauge invariant Lagrangian density for the potentials. The canonical quantization procedure is based on the Hamiltonian formulation of the dynamics, and here we find some complications in the form of first-class constraints for the electromagnetic field and second-class constraints for the matter fields. It is not necessary, though, to choose a particular gauge in order to carry out the quantization. Instead, we use the general procedures developed by Dirac (1950b, 1951, 1964, 1965) and Bergmann & Goldberg (1955). We also change variables and use the gauge-independent or physical matter fields, that create and annihilate particles together with their static electromagnetic fields.

We carry out the derivations for the classical fields in Section 2, where we find a Hamiltonian in which the Coulomb interaction is separated from the radiation field. In Section 3 we calculate the Dirac brackets that lead to the appropriate commutation or anticommutation relations for the field operators. We then go over to the Fock space representation for the particles and that of state functionals for the fields, and obtain the corresponding equations of motion in the Schrödinger picture. We conclude with some remarks in Section 4, and add an Appendix where we show how Lagrange multipliers can be used in the transformation theory of dynamical systems with constraints, together with the relation to the formalisms of Dirac and Bergmann.

We use natural units such that $\hbar$, c, ϵ_0 and μ_0 are all equal to 1. We follow the summation convention for repeated indices, modified in the case of Greek subindices that range from 0 to 3 to reflect the time-favoring indefinite metric in spacetime.

2. *The Classical Interacting Fields*

A real Lagrangian density for n free fields ϕ_k that represent noninteracting particles of mass m_k when spin and relativistic effects are ignored is

$$\mathscr{L}_0 = \sum_{k=1}^{n} \left[\tfrac{1}{2} i(\phi_k^* \dot{\phi}_k - \dot{\phi}_k^* \phi_k) - \frac{1}{2m_k} (\nabla \phi_k^*) . \nabla \phi_k \right] \tag{2.1}$$

The equation of motion of each field ϕ_k is the corresponding Schrödinger equation. We now assume that each particle has a charge q_k, and introduce the interaction with the electromagnetic field through the usual gauge invariant substitution

$$\partial_\mu \to D_{k\mu} = \partial_\mu + iq_k A_\mu \tag{2.2}$$

We add a free-field Lagrangian density for A_μ and obtain

$$\mathscr{L}=\sum_{k=1}^{n}\left[\tfrac{1}{2}i(\phi_k^* D_{k0}\phi_k-\phi_k D_{k0}^*\phi_k^*)-\frac{1}{2m_k}(\mathbf{D}_k^*\phi_k^*).\mathbf{D}_k\phi_k\right]-\tfrac{1}{4}F_{\mu\nu}F_{\mu\nu} \quad (2.3)$$

where

$$F_{\mu\nu}=A_{\mu,\nu}-A_{\nu,\mu} \quad (2.4)$$

This Lagrangian density is invariant under arbitrary gauge transformations of the second kind. We now change the matter field variables to the gauge invariant (Dirac, 1950a; Goldberg & Marx, 1968; Marx, 1970) fields

$$\psi_k=\phi_k\exp(-iq_k\xi) \quad (2.5)$$

where

$$\xi(\mathbf{x},t)=\int d^3x' G(\mathbf{x}-\mathbf{x}')\nabla'.\mathbf{A}(\mathbf{x}',t) \quad (2.6)$$

$$G(\mathbf{x})=(4\pi|\mathbf{x}|)^{-1} \quad (2.7)$$

It is obvious that ξ vanishes when the potentials are transverse, that is, when we use a Coulomb or radiation gauge; this is the reason why many of the special procedures we use are trivial in such a gauge, to the extent that they are usually ignored. We now find that, for each value of k,

$$D_{k\mu}[\psi_k\exp(iq_k\xi)]=(D'_{k\mu}\psi_k)\exp(iq_k\xi) \quad (2.8)$$

where

$$D'_{k0}=\partial_0+iq_k\chi \quad (2.9)$$

$$\mathbf{D}'_k=-\nabla+iq_k\mathbf{A}_S \quad (2.10)$$

$$\chi=A_0-\dot{\xi} \quad (2.11)$$

$$\mathbf{A}_S=\mathbf{A}+\nabla\xi \quad (2.12)$$

We have pointed out (Marx, 1970) that χ and $\mathbf{A}_S$ are the gauge invariant parts of the potentials, which can be expressed in terms of the fields by

$$\chi(\mathbf{x},t)=\int d^3x'\,\nabla G(\mathbf{x}-\mathbf{x}').\mathbf{E}(\mathbf{x}',t) \quad (2.13)$$

$$\mathbf{A}_S(\mathbf{x},t)=\int d^3x'\,\nabla G(\mathbf{x}-\mathbf{x}')\wedge\mathbf{B}(\mathbf{x}',t) \quad (2.14)$$

In this manner we express $\mathscr{L}$ in terms of gauge-independent quantities only,

$$\mathscr{L}=\sum_{k=1}^{n}\left[\tfrac{1}{2}i(\psi_k^*\dot{\psi}_k-\dot{\psi}_k^*\psi_k)-\rho\chi-\frac{1}{2m_k}(\mathbf{D}_k'^*\psi_k^*).(\mathbf{D}'_k\psi_k)\right]-\tfrac{1}{4}F_{\mu\nu}F_{\mu\nu} \quad (2.15)$$

where

$$\rho=\sum q_k|\psi_k|^2 \quad (2.16)$$

We now find that the momenta conjugate to ψ_k, ψ_k^*, A_0 and $\mathbf{A}$ (not χ and $\mathbf{A}_S$) are

$$\Pi_k = \tfrac{1}{2} i \psi_k^*, \qquad \Pi_k^* = -\tfrac{1}{2} i \psi_k \tag{2.17}$$

$$\Pi_0 = 0 \tag{2.18}$$

and

$$\mathbf{\Pi}(\mathbf{x}, t) = \dot{\mathbf{A}}(\mathbf{x}, t) + \nabla A_0(\mathbf{x}, t) + \int d^3x' \, \nabla G(\mathbf{x} - \mathbf{x}') \rho(\mathbf{x}', t) \tag{2.19}$$

respectively. The Hamiltonian is given by

$$H = \int d^3x (\Pi_k \dot{\psi}_k + \Pi_k^* \dot{\psi}_k^* + \dot{\mathbf{\Pi}} . \mathbf{A} - \mathscr{L}) \tag{2.20}$$

and the constraint (2.18) leads to a secondary constraint. The condition

$$\dot{\Pi}_0 = -\delta H / \delta A_0 = 0 \tag{2.21}$$

gives

$$\nabla . \mathbf{\Pi} = 0 \tag{2.22}$$

which, as can be seen from equation (2.19), is one of Maxwell's equations,

$$\nabla . \mathbf{E} = \rho \tag{2.23}$$

It is easy to see that the set of first-class constraints (2.18) and (2.22) and second-class constraints (2.17) is complete. Elimination of the generalized velocities and integration by parts then gives the Hamiltonian in the form

$$\begin{aligned} H = \sum_{k=1}^{n} \int d^3x \, \frac{1}{2m_k} [(\nabla + i q_k \mathbf{A}_S) \psi_k^*] . (\nabla - i q_k \mathbf{A}_S) \psi_k \\ + \tfrac{1}{2} \int d^3x \, d^3x' \, \rho(\mathbf{x}, t) \rho(\mathbf{x}', t) G(\mathbf{x} - \mathbf{x}') \\ + \tfrac{1}{2} \int d^3x [\mathbf{\Pi}^2 + (\nabla \wedge \mathbf{A}_S)^2] \end{aligned} \tag{2.24}$$

The first term contains the free-field Hamiltonian for the Schrödinger fields and their interaction with the radiation field, the second term represents the Coulomb interaction and the last term is the Hamiltonian for the free electromagnetic field.

We can compare this derivation to the corresponding one for the Dirac field in the radiation gauge (Lurié, 1968) and to the gauge-independent formulation for that field (Goldberg & Marx, 1968), where we should choose the vector n that represents the state of motion of the observer along the time axis.

3. *Canonical Quantization*

In order to obtain the commutation relations from their classical analogues, we find that the constraints in the theory require the computation of Dirac brackets, some of which differ from the corresponding Poisson brackets. We follow the procedure explained in the Appendix.

The functionals that vanish due to the constraints are

$$Y_{1k} = \Pi_k - \tfrac{1}{2} i\psi_k^* \tag{3.1}$$

$$Y_{2k} = \Pi_k^* + \tfrac{1}{2} i\psi_k \tag{3.2}$$

$$Y_3 = \Pi_0 \tag{3.3}$$

$$Y_4 = \nabla . \mathbf{\Pi} \tag{3.4}$$

and the matrix of the Poisson brackets of these functionals is

$$\theta = \begin{pmatrix} 0 & -i1 & 0 & 0 \\ i1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \tag{3.5}$$

where 1 stands for an $n \times n$ unit matrix multiplied by the unit matrix in Hilbert space, that is, a Dirac δ-function. There are two null vectors, which can be chosen with one nonvanishing element, a δ-function, in either of the last two positions, and the observables for the radiation field are $\mathbf{\Pi}$ and $\mathbf{A}_S$. The quasi-inverse of θ is equal to θ, and we can calculate the Dirac brackets of the observables.

It is convenient to express the solenoidal fields $\mathbf{\Pi}$ and $\mathbf{A}_S$ in terms of the momentum space expansions

$$\mathbf{A}_S(\mathbf{x}, t) = (2\pi)^{-3/2} \int d^3k\, \boldsymbol{\epsilon}^s(\mathbf{k})\, a_s(\mathbf{k}, t) \exp(i\mathbf{k}.\mathbf{x}) \tag{3.6}$$

$$\mathbf{\Pi}(\mathbf{x}, t) = (2\pi)^{-3/2} \int d^3k\, \boldsymbol{\epsilon}^s(\mathbf{k})\, \pi_s(\mathbf{k}, t) \exp(-i\mathbf{k}.\mathbf{x}) \tag{3.7}$$

where the $\boldsymbol{\epsilon}^s$ are two polarization vectors perpendicular to $\mathbf{k}$. We choose the circular polarization vectors that satisfy

$$\boldsymbol{\epsilon}^s(-\mathbf{k})^* = \boldsymbol{\epsilon}^s(\mathbf{k}) \tag{3.8}$$

so that the reality conditions for the fields $\mathbf{A}_S$ and $\mathbf{\Pi}$ are expressed by

$$a_s(-\mathbf{k}, t)^* = a_s(\mathbf{k}, t), \qquad \pi_s(-\mathbf{k}, t)^* = \pi_s(\mathbf{k}, t) \tag{3.9}$$

The Dirac bracket we have to determine is that of a_s with $\pi_{s'}$, and it is clear that it is equal to the Poisson bracket. We use

$$a_s(\mathbf{k}, t) = \boldsymbol{\epsilon}^s(\mathbf{k})^* . (2\pi)^{-3/2} \int d^3x\, \mathbf{A}(\mathbf{x}, t) \exp(-i\mathbf{k}.\mathbf{x}) \tag{3.10}$$

$$\pi_{s'}(\mathbf{k}', t) = \boldsymbol{\epsilon}^{s'}(\mathbf{k}') . (2\pi)^{-3/2} \int d^3x\, \mathbf{\Pi}(\mathbf{x}, t) \exp(i\mathbf{k}'.\mathbf{x}) \tag{3.11}$$

and we find the Poisson bracket

$$\{a_s(\mathbf{k}, t), \pi_{s'}(\mathbf{k}', t)\} = \delta_{ss'}\, \delta(\mathbf{k} - \mathbf{k}') \tag{3.12}$$

The commutators then are†

$$[a_s(\mathbf{k}), a_{s'}(\mathbf{k}')]_- = 0 \tag{3.13}$$

$$[\pi_s(\mathbf{k}), \pi_{s'}(\mathbf{k}')]_- = 0 \tag{3.14}$$

$$[a_s(\mathbf{k}), \pi_{s'}(\mathbf{k}')]_- = i\delta_{ss'}\,\delta(\mathbf{k}-\mathbf{k}') \tag{3.15}$$

The Dirac brackets that we need for the matter fields are

$$[\psi_k(\mathbf{x},t), \psi_{k'}^*(\mathbf{x}',t)]_D = -i\delta_{kk'}\,\delta(\mathbf{x}-\mathbf{x}') \tag{3.16}$$

differing from the Poisson brackets which vanish. The corresponding commutators or anticommutators‡ then are

$$[\psi_k(\mathbf{x}), \psi_{k'}^\dagger(\mathbf{x}')]_\pm = \delta_{kk'}\,\delta(\mathbf{x}-\mathbf{x}') \tag{3.17}$$

We can now write the Hamiltonian operator in the form

$$H = H_m + H_{em} + H_I \tag{3.18}$$

where

$$H_m = -\sum_{k=1}^{n} \int d^3x \frac{1}{2m_k} \psi^\dagger(\mathbf{x}) \nabla^2 \psi_k(\mathbf{x}) + \tfrac{1}{2} \sum_{k=1}^{n} \sum_{k'=1}^{n} \int d^3x\, d^3x'\, \psi_{k'}^\dagger(\mathbf{x}') \psi_k^\dagger(\mathbf{x}) \frac{q_k q_{k'}}{4\pi|\mathbf{x}-\mathbf{x}'|} \psi_k(\mathbf{x}) \psi_{k'}(\mathbf{x}') \tag{3.19}$$

$$H_{em} = \tfrac{1}{2} \int d^3k [\pi_s(\mathbf{k})\, \pi_s(-\mathbf{k}) + \mathbf{k}^2 a_s(\mathbf{k})\, a_s(-\mathbf{k})] \tag{3.20}$$

$$H_I = \int d^3x\, \mathbf{A}_S(\mathbf{x}) . \sum_{k=1}^{n} \frac{q_k}{m_k} \psi_k^\dagger(\mathbf{x})\, i\nabla \psi_k(\mathbf{x}) + \int d^3x\, \mathbf{A}_S^2(\mathbf{x}) \sum_{k=1}^{n} \frac{q_k^2}{2m_k} \psi_k^\dagger(\mathbf{x})\, \psi_k(\mathbf{x}) \tag{3.21}$$

† The reason the sign in equation (3.15) is opposite to that in equation (92) in Goldberg & Marx (1968) is our choice there of $\Pi^i = -\Pi_i$ as the momentum conjugate to A_i. We also note that this formalism does not lead to the assumption that the commutator of A_{Si} and Π_j is proportional to δ_{ij}, a choice that is first proposed (Bjorken & Drell, 1965; Lurie, 1968) and then discarded because it leads to a contradiction with the solenoidal nature of the operators.

‡ The relationship between Dirac brackets and anticommutators is somewhat ambiguous, since the former change sign under an interchange of the fields and the latter do not. In quantum mechanics, the observables for fermion fields have an even number of field operators, and their commutators do not depend on the choice of anticommutation instead of commutation relations. These particular (anti)-commutation relations are often 'derived' by using an equivalent complex Lagrangian density, such that $\Pi_k = i\psi_k^*$ and ignoring the fact that $\Pi_k^* = 0$.

Although many calculations can be performed by algebraic manipulations, it is often convenient to use explicit representations for the state vectors and operators, especially when doing numerical calculations. A convenient representation for the particle operators is a nonrelativistic Fock space (Schweber, 1961); in this theory only one wave function with fixed numbers of particles needs to be considered at a time. The analogous representation for the radiation field in the Schrödinger picture is based on state functionals (Marx, 1969), in which the momenta are represented by functional derivatives,

$$\pi_s(\mathbf{k}) = -i\delta/\delta a_s(\mathbf{k}) \tag{3.22}$$

The practical convenience of this choice is limited by our lack of familiarity with functional differential equations and functional integrals, but we do know the ground state functional and the raising and lowering operators, and we can always go back to the usual algebraic calculations.

The components of the Fock space vector are then functions of the appropriate number of particle variables and functionals of the two independent fields. When we consider a dynamical problem, these probability amplitudes also depend on time. Thus, the amplitude for a state with N particles, of which ν_k are of the type k, has the form

$$\Psi_N = \Psi^{(\nu_1 \nu_2 \ldots \nu_n)}(\mathbf{x}_{11}, \mathbf{x}_{12}, \ldots, \mathbf{x}_{1\nu_1}, \mathbf{x}_{21}, \ldots, \mathbf{x}_{n\nu_n}; a_1, a_2] \tag{3.23}$$

The equation of motion

$$i\dot{\Psi}(t) = H\Psi(t) \tag{3.24}$$

leads to the Schrödinger equation for the amplitude

$$\begin{aligned} i\dot{\Psi}_N(t) = \Bigg\{ &-\sum_{k=1}^{n}\sum_{r=1}^{\nu_k} \frac{1}{2m_k}\nabla_{kr}^2 + \sum_{k=1}^{n}\sum_{k'=1}^{k-1}\sum_{r=1}^{\nu_k}\sum_{r'=1}^{\nu_{k'}} \frac{q_k q_{k'}}{4\pi|\mathbf{x}_{kr} - \mathbf{x}_{k'r'}|} \\ &+ \sum_{k=1}^{n}\sum_{r=1}^{\nu_k}\sum_{r'=1}^{r-1} \frac{q_k^2}{4\pi|\mathbf{x}_{kr} - \mathbf{x}_{kr'}|} + \tfrac{1}{2}\int d^3k \left[-\frac{\delta^2}{\delta a_s(\mathbf{k})\,\delta a_s(-\mathbf{k})} \right. \\ &\left. + \mathbf{k}^2 a_s(\mathbf{k})\,a_s(-\mathbf{k}) \right] + \sum_{k=1}^{n}\sum_{r=1}^{\nu_k} \frac{q_k}{m_k} \mathbf{A}_S(\mathbf{x}_{kr}) . i\nabla_{kr} \\ &+ \sum_{k=1}^{n}\sum_{r=1}^{\nu_k} \frac{q_k^2}{2m_k} \mathbf{A}_S^2(\mathbf{x}_{kr}) \Bigg\} \Psi_N(t) \end{aligned} \tag{3.25}$$

So far we have not made any approximations in our presentation and we obtain equation (3.25), which usually serves as a point of departure for practical calculations. By choosing the right masses and charges for the fields ψ_k to represent nuclei and electrons and the amplitude with the right numbers of particles, we can deal with one or more atoms or molecules. The eigenstates of the first three terms of the Hamiltonian in equation (3.25) correspond to the bound states for such systems when the Coulomb interaction alone is considered; they can be determined either exactly or by different approximation methods. The translational degrees of freedom lead to nonnormalizable stationary states, unless we put the system in a box. The stationary states for the free radiation system are usually interpreted in terms of photons, and are not normalizable either when we extend the field over all of space. The terms coming from H_I are then usually considered as a perturbation that induces transitions between the energy levels of the atoms and at the same time creates or annihilates photons. We could consider, for instance, an atom in an excited state and the radiation field in its ground state initially, and look for transitions to other states in a time interval.

Although the separation of the terms in the Hamiltonian was made in a form suited for a discussion of bound states, it should be clear that the same equations are valid for scattering of particles and radiation.

4. *Concluding Remarks*

We have given a derivation of the equation of motion for a quantized radiation field interacting with an arbitrary number of charged particles, possibly arranged as atoms or molecules. We started from the free Schrödinger and electromagnetic fields and added the so-called minimal interaction. We have kept explicit gauge invariance throughout the whole derivation, and the procedure is independent of any choice of gauge.

We followed a canonical quantization procedure properly modified to take into account the constraints that are present in the theory. The Coulomb interaction separated in a natural way from the truly dynamical electromagnetic field, represented by the transverse part of the vector potential. In this manner we do not quantize any redundant, gauge-dependent fields. This explains why our procedure resembles closely the quantization in a Coulomb or radiation gauge, although the additional constraint $\nabla . \mathbf{A} = 0$ should receive special attention in such a gauge.

We have restricted ourselves to the Schrödinger picture, which is more readily interpreted in terms of probability amplitudes, both for the particles and the radiation field. These occur naturally when we use the Fock space representation for states and operators, combined with the state functionals for the radiation. The Dirac picture is often used to simplify the equations in a perturbation theory, but the same expressions can be obtained working in the Schrödinger picture.

We did not include an external electromagnetic field, which is a given classical field, in our presentation, but such a field can be easily added in the gauge invariant substitution (2.2).

What was done with the Schrödinger equation can be repeated with the

Pauli equation for nonrelativistic particles with spin, or with the relativistic equations of Dirac and Klein–Gordon. We have seen how a covariant treatment in the relativistic case still involves a choice of an inertial observer (Goldberg & Marx, 1968; Marx, 1970), which also plays a role in the separation of the Coulomb and radiation fields. Although this detracts from the formal elegance of the formulation, we believe that it is nevertheless desirable for a more meaningful physical interpretation, and that such a distinction is probably unavoidable for a consideration of bound states in a relativistic theory.

The equations derived here can also serve as a guide for a similar relativistic theory with a fixed number of 'particles'. In such a theory, pair creation and annihilation is taken into account by changing the direction of propagation in time. Propagation forward in time corresponds to a particle state, and propagation backward in time, to an antiparticle state. This formalism avoids some of the usual difficulties with infinities, such as those related to closed fermion loops in Feynman diagrams, but is not sufficiently developed yet.

We have also left for an Appendix the general theory of Hamiltonian dynamics and infinitesimal transformations for classical fields with constraints. The use of Lagrange multipliers simplifies the derivations, and the results so obtained are equivalent to those of Dirac and Bergmann.

Appendix

The canonical quantization procedure is only one way to obtain the equations of motion and the commutators of operators in quantum mechanics, but it provides a certain amount of guidance and reliability that make its application desirable whenever it is possible.

The relationship between the classical and quantum theories is best understood in terms of the infinitesimal canonical transformations of the former and the unitary ones of the latter. The commutator $\mathscr{T}_2^{-1}\mathscr{T}_1^{-1}\mathscr{T}_2\mathscr{T}_1$ of two transformations $\mathscr{T}_1$ and $\mathscr{T}_2$ is generated by the Poisson bracket of the generators in the classical theory, and by the commutator in the quantum theory.

The presence of constraints in a theory can lead to inconsistencies unless the canonical transformations are restricted to those that leave the constraints invariant. As in many similar problems, the use of Lagrange multipliers provides a convenient way of implementing the constraints.

We consider a system described by n generalized coordinates, the fields ϕ_k. The dynamics is determined by a Lagrangian, which is a functional of the generalized coordinates and velocities,

$$L = L[\phi, \dot{\phi}] \tag{A.1}$$

The momenta conjugate to these coordinates are defined by the functional derivatives

$$\pi_k = \delta L / \delta \dot{\phi}_k \tag{A.2}$$

and the Hamiltonian is obtained from

$$H = \pi_k . \dot{\phi}_k - L \tag{A.3}$$

by elimination of the generalized velocities. The dot indicates a scalar product in the Hilbert space, that is, an integration over the continuous indices. The canonical equations of motion have the form

$$\dot{\phi}_k = \delta H/\delta\pi_k, \qquad \dot{\pi}_k = -\delta H/\delta\phi_k \tag{A.4}$$

The Poisson bracket of two functionals is defined by

$$\{F, G\} = \frac{\delta F}{\delta\phi_k} . \frac{\delta G}{\delta\pi_k} - \frac{\delta F}{\delta\pi_k} . \frac{\delta G}{\delta\phi_k} \tag{A.5}$$

It is convenient to designate the coordinates and momenta collectively by ξ_r, so that

$$\xi_r = \phi_r, \qquad r = 1, \ldots, n \tag{A.6}$$

$$\xi_r = \pi_{r-n}, \qquad r = n+1, \ldots, 2n \tag{A.7}$$

and define the matrix κ in terms of $n \times n$ submatrices by

$$\kappa = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \tag{A.8}$$

We can then rewrite equations (A.4) and (A.5) in the form

$$\dot{\xi}_r = \{\xi_r, H\} \tag{A.9}$$

$$\{F, G\} = \kappa_{rt} \frac{\delta F}{\delta\xi_r} . \frac{\delta G}{\delta\xi_t} \tag{A.10}$$

Canonical transformations of coordinates and momenta are then defined by the requirement of invariance of the form of equations (A.4). We can easily extend the formalism used in mechanics (Lanczos, 1962) to the theory of classical fields. These transformations can be specified by a generating functional, and the choice of independent variables is indicated by later applications to infinitesimal transformations. The form of the canonical equations is preserved when the differential forms constructed out of the old and new variables satisfy

$$\pi_k . \delta\phi_k + \bar{\phi}_k . \delta\bar{\pi}_k = \delta U[\phi, \bar{\pi}] \tag{A.11}$$

We now assume that we have p constraints of the form†

$$Y_l[\xi] = 0 \tag{A.12}$$

† We assume that we have 'vector' constraints, such as those in equations (3.1) to (3.4), that is, both the constraints and the arguments are vectors in the Hilbert space. The consideration of 'scalar' constraints requires only a few changes in notation.

and we restrict the transformation to those that leave the constraints form invariant. In this case, the variations $\delta\phi_k$ and $\delta\bar{\pi}_k$ are no longer independent, since equation (A.12) has to be satisfied both for the ξ_r and $\bar{\xi}_r$. We now replace U in equation (A.11) by

$$V[\phi, \bar{\pi}] = U[\phi, \bar{\pi}] + \Lambda_l[\phi, \bar{\pi}] \,.\, Y_l[\phi, \pi[\phi, \bar{\pi}]] + \bar{\Lambda}_l[\phi, \bar{\pi}] \,.\, Y_l[\bar{\phi}[\phi, \bar{\pi}], \bar{\pi}] \tag{A.13}$$

where the Lagrange multipliers Λ_l and $\bar{\Lambda}_l$ can be used to write down the equations

$$\pi_k = \delta V/\delta\phi_k, \qquad \bar{\phi}_k = \delta V/\delta\bar{\pi}_k \tag{A.14}$$

as if the variations were now independent. These are $2n$ equations, which together with the two sets of constraint equations determine the $2n + 2p$ vector functionals π_k, $\bar{\phi}_k$, Λ_l and $\bar{\Lambda}_l$ in terms of ϕ_k and $\bar{\pi}_k$, and we can then solve for the new variables in terms of the old or vice versa. Since the Y_l vanish, we have no functional derivatives of the Λ_l and $\bar{\Lambda}_l$ in the equations.

We now restrict ourselves to infinitesimal transformations. They are generated by functionals F related to U by

$$U = \phi_k \,.\, \bar{\pi}_k + \epsilon F \tag{A.15}$$

We obtain a set of equations for

$$\delta\xi_r = \bar{\xi}_r - \xi_r \tag{A.16}$$

from equations (A.12) and (A.14). To lowest order in ϵ, they are

$$-\delta\pi_k = \epsilon\left(\frac{\delta F}{\delta\phi_k} + \lambda_l \,.\, \frac{\delta Y_l}{\delta\phi_k}\right) \tag{A.17}$$

$$\delta\phi_k = \epsilon\left(\frac{\delta F}{\delta\pi_k} + \lambda_l \,.\, \frac{\delta Y_l}{\delta\pi_k}\right) \tag{A.18}$$

$$\frac{\delta Y_l}{\delta\phi_k} \,.\, \delta\phi_k + \frac{\delta Y_l}{\delta\pi_k} \,.\, \delta\pi_k = 0 \tag{A.19}$$

where, in this order, we can replace $\bar{\pi}_k$ by π_k in F and we have set

$$\epsilon\lambda_l = \Lambda_l + \bar{\Lambda}_l \tag{A.20}$$

Substituting equations (A.17) and (A.18) into (A.19), we obtain

$$\{Y_l, F\} + \{Y_l, Y_{l'}\} \,.\, \lambda_{l'} = 0 \tag{A.21}$$

If the 'matrix' θ of the Poisson brackets

$$\theta_{ll'} \equiv \{Y_l, Y_{l'}\} \tag{A.22}$$

has an inverse, we can solve for the λ_l and determine the $\delta\xi_r$. But when this matrix is singular,† we have a solution only if certain consistency

† This is obviously the case when one or more constraints are first class. The Poisson brackets of a first-class constraint with *all* other constraints vanish; other constraints are called second class. But the matrix can still be singular when all constraints are second class.

conditions are satisfied, and this solution then has arbitrary parameters. If $\{u^{(n)}\}$ is a complete set of independent null vectors† of the antisymmetric matrix θ, we multiply both sides of equation (A.21) by each $u^{(n)}$ to obtain the consistency conditions

$$u_l^{(n)}.\{Y_l, F\} = 0 \tag{A.23}$$

A functional F that satisfies all these conditions is called an allowed generator, and the corresponding solutions are

$$\lambda_l = -\gamma_{ll'}.\{Y_{l'}, F\} + \alpha_n u_l^{(n)} \tag{A.24}$$

where the α_n are arbitrary functions and γ is a quasi-inverse of θ, and satisfies

$$\gamma_{ll'}.\theta_{l'l''}.\gamma_{l''l'''} = \gamma_{ll'''}, \qquad \theta_{ll'}.\gamma_{l'l''}.\theta_{l''l'''} = \theta_{ll'''} \tag{A.25}$$

This procedure is particularly simple when the constraints are expressed in such a way that a maximum number is first class, then the nonzero submatrix has an inverse. This is the procedure followed by Dirac, but the use of the quasi-inverse makes this rearrangement of constraints unnecessary. We now have that

$$\delta\pi_k = -\epsilon\left(\frac{\delta F}{\delta\phi_k} - \frac{\delta Y_l}{\delta\phi_k}.\gamma_{ll'}.\{Y_{l'}, F\} + \alpha_n \frac{\delta Y_l}{\delta\phi_k}.u_l^{(n)}\right) \tag{A.26}$$

$$\delta\phi_k = \epsilon\left(\frac{\delta F}{\delta\pi_k} - \frac{\delta Y_l}{\delta\pi_k}.\gamma_{ll'}.\{Y_{l'}, F\} + \alpha_n \frac{\delta Y_l}{\delta\pi_k}.u_l^{(n)}\right) \tag{A.27}$$

and the change in a functional G is

$$\delta G = \epsilon(\{G, F\} - \{G, Y_l\}.\gamma_{ll'}.\{Y_{l'}, F\} + \alpha_n\{G, Y_l\}.u_l^{(n)}) \tag{A.28}$$

A functional G is an observable if its changes under the restricted infinitesimal canonical transformations are independent of any arbitrary parameters α_n. In the case of the electromagnetic field, these (classical) observables are the gauge-independent quantities. The conditions that G has to satisfy to be an observable are clearly the same as the consistency conditions (A.23), and observables and allowed generators are the same functionals. We now define the Dirac bracket of two observables G and F as

$$[G, F] = \{G, F\} - \{G, Y_l\}.\gamma_{ll'}.\{Y_{l'}, F\} \tag{A.29}$$

and equation (A.28) reduces to

$$\delta G = \epsilon[G, F] \tag{A.30}$$

Since the procedure to maximize the number of first-class constraints, as demanded by Dirac, is simply a new form to express the same constraints,

† In this case they really are families of null vectors with a continuous index that can be left implicit in n to avoid complicating the notation. Thus, the α_n are functions and when n is repeated we sum over the discrete values and integrate over the continuous index.

there is no difference in the brackets that are obtained. Bergmann restricts the canonical transformations in the same way we have done here, but the mathematical calculations are quite different. We give the main steps below in a direct proof of the equivalence of the two procedures.

When we follow Bergmann's approach, we define new coordinates in phase space by completing the set $\{Y_l\}$ by any $2n-p$ other coordinates $\{y_m\}$. The matrix η is given by

$$\eta_{mm'} = -\kappa_{rt}\frac{\delta\xi_r}{\delta y_m}\cdot\frac{\delta\xi_t}{\delta y_{m'}} \tag{A.31}$$

and the consistency conditions for a functional F' related to F by

$$F' = F + \lambda_l \,.\, Y_l \tag{A.32}$$

are

$$U_m^{(n)} \,.\, \delta F'/\delta y_m = 0 \tag{A.33}$$

where the $\{U^{(n)}\}$ are a complete set of null vectors of η. The brackets are then defined by Bergmann as

$$[G', F'] = \frac{\delta G'}{\delta y_m}\cdot\beta_{mm'}\cdot\frac{\delta F'}{\delta y_{m'}} \tag{A.34}$$

where β is a quasi-inverse of η. The two sets of null vectors are related by

$$U_m^{(n)} = \{y_m, Y_l\} \,.\, u_l^{(n)} \tag{A.35}$$

$$u_l^{(n)} = \zeta_{lm} \,.\, U_m^{(n)} \tag{A.36}$$

where

$$\zeta_{lm} = -\kappa_{rt}\frac{\delta\xi_r}{\delta Y_l}\cdot\frac{\delta\xi_t}{\delta y_m} \tag{A.37}$$

It is now straightforward to show the equivalence of the consistency conditions that define the observables. The equality of the brackets of two observables in both formalisms can be concluded from the fact that equation (A.30) is satisfied in both. A direct proof seems to require a few lengthy calculations. We first note that the product of a matrix and a quasi-inverse differs from the unit matrix by linear combinations of products of null vectors, and we define $c_{nn'}$ and $c'_{nn'}$ in

$$\theta_{ll'} \,.\, \gamma_{l'l''} = \delta_{ll''}\,\delta + c_{nn'}\, v_l^{(n)}\, u_{l''}^{(n')} \tag{A.38}$$

$$\gamma_{ll'} \,.\, \theta_{l'l''} = \delta_{ll''}\,\delta + c'_{nn'}\, u_l^{(n)}\, v_{l''}^{(n')} \tag{A.39}$$

where δ stands for the Dirac δ-function and the $\{v^{(n)}\}$ are a complete set of null vectors of γ.

We then find that

$$[F, G] = \frac{\delta F'}{\delta y_m}\cdot\beta'_{mm'}\cdot\frac{\delta G'}{\delta y_{m'}} \tag{A.40}$$

where

$$\beta'_{mm'} = \{y_m, y_{m'}\} - \{y_m, Y_l\} \cdot \gamma_{ll'} \cdot \{Y_{l'}, y_{m'}\} \tag{A.41}$$

and that a quasi-inverse of η is given by

$$\beta_{mm'} = \beta'_{mm'} + a_{nn'} U_m^{(n)} U_{m'}^{(n')} \tag{A.42}$$

where

$$a_{nn'} = -c'_{nn''} c_{n'''n'} \kappa_{rt} v_l^{(n'')} \cdot \frac{\delta \xi_r}{\delta Y_l} \cdot \frac{\delta \xi_t}{\partial Y_{l'}} \cdot v_{l'}^{(n''')} \tag{A.43}$$

The consistency conditions (A.33) then show that the second term in equation (A.42) does not contribute when we substitute it into equation (A.40), which completes the proof.

References

Bergmann, P. G. and Goldberg, I. (1955). *Physical Review*, **98**, 531.

Bjorken, J. D. and Drell, S. D. (1965). *Relativistic Quantum Fields*, p. 71. McGraw-Hill, Inc., New York.

Dirac, P. A. M. (1950a). *Proceedings of the Royal Society* (*London*), **235**, 138.

Dirac, P. A. M. (1950b). *Canadian Journal of Mathematics*, **2**, 129.

Dirac, P. A. M. (1951). *Canadian Journal of Mathematics*, **3**, 1.

Dirac, P. A. M. (1964). *Lectures on Quantum Mechanics*, Belfer Graduate School of Science, Yeshiva University, New York.

Dirac, P. A. M. (1965). *Lectures on Quantum Field Theory*, Belfer Graduate School of Science, Yeshiva University, New York.

Goldberg, I. and Marx, E. (1968). *Nuovo Cimento*, **57B**, 485.

Kramers, H. A. (1957). *Quantum Mechanics*, Chapter 8. North Holland Publishing Company, Amsterdam.

Lanczos, C. (1962). *The Variational Principles of Mechanics*, 2nd Ed., Chapter 7. University of Toronto Press, Toronto.

Lurié, D. (1968). *Particles and Fields*, pp. 151 and 174. Interscience Publishers, New York.

Marx, E. (1969). *Nuovo Cimento*, **60A**, 683.

Marx, E. (1970). *International Journal of Theoretical Physics*, Vol. 3, p. 467.

Schweber, S. S. (1961). *An Introduction to Relativistic Quantum Field Theory*, p. 142. Row, Peterson and Company, Evanston, Illinois.