Classical and Quantum Field Theories in the Lagrangian Formalism*

PETER G. BERGMANN AND RALPH SCHILLER†
Department of Physics, Syracuse University, Syracuse, New York
(Received August 12, 1952)

In this paper, we have first examined the relationship between the transformation properties of a (nonquantum) covariant field theory and its constraints, generating functionals, conservation laws, and "superpotentials" purely within the Lagrangian formalism and indicated the relevance of these quantities for the problem of motion of particles (singularities) in the field. This discussion includes a presentation of actual methods of computation of these important quantities suitable for a very wide class of theories. In the second part of the paper, we have discussed the probable structure of a quantum covariant field theory, both in the Hamiltonian and in the Lagrangian formalism. In the Hamiltonian formalism, it is suggested that those field variables canonically conjugate to constraints are not observables in the physical sense nor operators in Hilbert space, and that the states of a system which alone can be regarded as Hilbert vectors are those consistent with all the constraints inherent in the theory and its transformation properties. This approach permits the characterization of legitimate observables even if the isolation of the "constraint variables" is not feasible, as in the general theory of relativity; observables must commute with all the constraints. They are, thus, invariants under the group of invariant transformations. It is asserted that this selection of observables, which is mathematically self-consistent, does not lead to the discard of quantities of physical interest. On the other hand, all the so-called paradoxes between constraints (subsidiary conditions) and commutation relations are thereby avoided. In the development of the Lagrangian quantization, we are proposing a new set of field equations which are different from the usual ones but which can be shown to permit the transition to the canonical formulation if desired. Our proposal is to assert the stationary character of the Feynman-Schwinger action operator not with respect to all, but only with respect to those variations that correspond to invariant transformations. As a result, the number of operator equations at each point of space-time is finite, though it is different from the number of equations in the nonquantum theory. These equations, though considerably weaker than what would be obtained if the action integral were to be made stationary with respect to all conceivable variations, are sufficient to yield all the usual conservation laws and also to permit the transition to the Hamiltonian form of the theory if desired. Commutation relations can be obtained for the field variables and their time derivations on the same hypersurface, simply by requiring that the field variables and their derivatives be algebraically independent of each other. The procedure employed breaks down if applied to a variable that is canonically conjugate to a constraint, an indication that in the Lagrangian formulation, too, the set of observables must be selected if paradoxes are to be avoided. Altogether it appears that the Lagrangian and the Hamiltonian quantizations, if set up properly, are largely equivalent; but this does not preclude the possibility that one may be more useful heuristically than the other.

1. INTRODUCTION

IN a number of preceding papers[1–4] we have ascertained the properties of field theories that are covariant with respect to groups of transformations depending on arbitrary functions. The examination of such theories is suggested by the physical importance of such transformation groups in the general theory of relativity (i.e., the theory of gravitation) and in the theory of the electromagnetic field. We found that very generally invariance of a Hamiltonian principle from which the field equations can be derived results in the existence of certain differential identities between the field equations and of "strong conservation laws."[1]

If theories of this type are brought into the canonical scheme, we found that the canonical field variables satisfy a certain number of constraints (also known as subsidiary conditions in quantum electrodynamics) and that the Hamiltonian functional of such a theory is determined only up to a linear combination of the so-called primary constraints, the coefficients of this linear combination being arbitrary functions.[2] This degree of arbitrariness is equivalent to Dirac's proposal to leave a certain number of velocities remain in the Hamiltonian density.[5] Recently we found that all the constraints play a role in the so-called functional of the transformation group, which generates the infinitesimal canonical transformations with respect to which the theory is covariant.[4] In fact, the circumstance that this functional is a constant of the motion and that the commutator of two such canonical transformations must again be a member of the group leads to a powerful and convenient method of determining the structure of the whole function group encompassing the constraints and the Hamiltonian.

It had been recognized for some time that the constraints of the Hamiltonian formalism were related to the conservation laws of the Lagrangian formalism. The first purpose of the present paper is to trace out this relationship in specific detail. The existence of the strong conservation laws—the vanishing of a certain divergence—implies the existence of "superpotentials," i.e., a set of quantities of which the components of the energy-momentum-stress "tensor" form the curl. We have succeeded in obtaining these superpotentials directly from the Lagrangian and its transformation properties (Sec. 2). On the other hand, the transfor-

* This work was supported by the ONR. This article contains, *inter alia*, the results of a doctoral dissertation submitted by one of us (R.S.) to the Graduate School of Syracuse University.
† Now at Universidade de Sao Paulo, Brasil.
[1] P. G. Bergmann, Phys. Rev. **75**, 680 (1949).
[2] P. G. Bergmann and J. H. M. Brunings, Revs. Modern Phys. **21**, 480 (1949).
[3] Bergmann, Penfield, Schiller, and Zatzkis, Phys. Rev. **80**, 81 (1950).
[4] J. L. Anderson and P. G. Bergmann, Phys. Rev. **83**, 1018 (1951).
[5] P. A. M. Dirac, Can. J. Math. **2**, 129 (1950).

mation law of the Lagrangian may also be formulated in terms of the generating functional, and Sec. 3 deals with the relationship between the generating potential in the Hamiltonian theory and the transformation law of the Lagrangian in the Lagrangian theory. The generating functional, and with it the constraints, will be constructed from within the Lagrangian formalism; this derivation leads to a considerable simplification of the expressions given earlier.[4]

In Sec. 4, we have endeavored to construct a quantum field theory that is covariant with respect to the same transformation group as the corresponding classical field theory. In the Hamiltonian theory, it appears fairly obvious what the general scheme must be, because the pertinent quantum theoretical formations, in particular commutators, which characterize the structure of the invariant transformation group, all have their classical analogs, e.g., Poisson brackets. However, we have attempted to show that from among all the classical canonical field variables only those may be considered as Hilbert operators in the quantum field theory (and therefore as physically meaningful observables) which commute with all the constraints and which, therefore, are invariant with respect to the invariant transformation group.

We have also dealt with the Lagrangian theory, largely on the basis of the results obtained in the analysis of the classical Lagrangian theory and the possibility of introducing canonical transformations and generators into the Lagrangian formalism. While we follow to some extent the Feynman-Schwinger developments,[6,7] to relate the unitary mapping operator $U(t_2, t_1)$ to the action integral, we have found it impossible to adopt field equations exactly analogous to the classical Euler-Lagrange equations, as Schwinger does. Instead, we are putting forward, as a conjecture for the time being, a new set of "field equations" whose number does not depend on the number of field variables (as in the classical theory), but on the structure of the invariant transformation group. Our conjectured "field equations" lead to the usual conservation laws. While they may not go over into the usual Euler-Lagrange equations in the classical limit for all conceivable covariant field theories, they will do so for the general theory of relativity with electromagnetic field. Our field equations are obtained by the requirement that the action integral be stationary with respect to variations that correspond to members of the invariant transformation group, but not necessarily to all other conceivable variations. This requirement, we have found, leads to relatively simple covariant differential equations for the observables without special assumptions concerning the commutation properties of the infinitesimal variations. Our theory is sufficiently definite to lead to the Schrödinger equation (or its equivalent in the Heisenberg representation). Thus it appears to lead to a "Lagrangian" theory that is equivalent to a Hamiltonian quantum field theory.

[6] R. P. Feynman, Revs. Modern Phys. **20**, 367 (1948).
[7] J. Schwinger, Phys. Rev. **82**, 914 (1951).

2. THE SUPERPOTENTIALS

Instead of introducing the components of the metric tensor as our fundamental field variables, as is done in the general theory of relativity, we shall characterize our field variables by the symbol y_A $(A=1, \cdots, N)$, where N is the number of algebraically independent components. These are equal to ten in the general theory of relativity and four in electromagnetic theory. We shall further assume that the field equations are derivable from a variational principle of the form

$$\delta S=\delta\int L(y_A, y_{A,\sigma})d^4x, \tag{2.1}$$

where L is a function of the field variables y_A and their first partial derivatives with respect to the space-time coordinates.

The transformation properties of L are of some importance, and we shall specify them more fully than has been done in the past.[1]

If we desire to have covariant field equations (the field equations to transform as densities of weight one and contragrediently to the field variables), then we need only choose an invariant as our action integral. In fact, if we were to construct a new covariant theory, it is hard to envisage how one could find the correct field equations without using such an invariant action integral, for the covariance of our field equations and their compatibility is always assured if they are derivable from such a variational principle. However, the only known invariant density in the prototype of any new theory that we might propose, the general theory of relativity, contains second derivatives of the field variables. These higher-order derivatives appear in the Lagrangian in terms that have the form of a pure divergence. If we desire to go over from the Lagrangian to the usual Hamiltonian theory, it is desirable to subtract this divergence from the invariant density. The field equations will not be altered by this change in the Lagrangian, as the addition of a pure divergence to the Lagrangian density does not contribute to the variation of the action integral as long as the variations of the field variables are confined to the interior of the four-dimensional domain of integration. Thus, our Lagrangian density will differ from an invariant density by a pure divergence, and we shall always assume that the precise form of this divergence is known. In the general theory of relativity, this divergence has important transformation properties which we shall describe below [Eq. (2.17)].

We shall assume that our field variables transform in the following manner:

$$\bar{\delta}y_A=c_{Ai}{}^{\sigma}\xi^i{}_{,\sigma}+c_{Ai}\xi^i-y_{A,\rho}\delta x^\rho. \tag{2.2}$$

The $c_{Ai}{}^{\sigma}$ and the c_{Ai} are functions of the y_A, and their exact form depends upon the specific transformation properties of the field variables. The index σ, as well as all other Greek indices, refers to the space-time coordinates. The index i identifies the arbitrary functions ξ^i (descriptors) that appear in the transformation law.[8] For coordinate-covariant theories, i becomes the Greek index ι. The infinitesimal variations of the space-time coordinates depend on some or all of the arbitrary ξ^i,

$$\delta x^{\nu} = a_i{}^{\nu} \xi^i, \tag{2.3}$$

where the $a_i{}^{\nu}$ are constants, 0 and 1.

The complicated character of the infinitesimal transformation law (2.2), (2.3) (more general than the coordinate transformation law for tensors) arises only because we want to encompass within our formalism a wider class of transformations than mere coordinate transformations. It is necessary to have the variations of the space-time coordinates depend on the set of constants $a_i{}^{\nu}$ since in some cases we are going to vary the dependent variables y_A and the independent variables x^{ρ} independently of one another. We have introduced the variation $\bar{\delta} y_A$ because this transformation law obeys the group property. This requirement is met if the commutator of two successive infinitesimal $\bar{\delta}$-transformations is a transformation of the same type. If this commutator is calculated for a given field variable, the arbitrary constants that appear in $c_{Ai}{}^{\sigma}$ and c_{Ai} will satisfy certain identities.[9] $\bar{\delta} y_A$ is the infinitesimal difference $y_A'(\mathbf{x}) - y_A(\mathbf{x})$, i.e., y_A' is compared with y_A at a point that has the same coordinate values, rather than with y_A at the same point in space-time.

We have assumed that our Lagrangian density differs from an invariant density by a known divergence. The transformation law of the invariant density is

$$\bar{\delta} K + (K\xi^{\rho})_{,\rho} = 0, \quad K = L + S^{\rho}{}_{,\rho}, \tag{2.4}$$

where K, the invariant density, has been separated into the Lagrangian density, L, and a divergence, $S^{\rho}{}_{,\rho}$. The transformation law of the Lagrangian density is then

$$\bar{\delta} L = Q^{\rho}{}_{,\rho}, \tag{2.5}$$

where

$$Q^{\rho} = -(L\xi^{\rho} + S^{\sigma}{}_{,\sigma}\xi^{\rho} + \bar{\delta} S^{\rho}). \tag{2.6}$$

If we now actually carry out the variation indicated on the left-hand side of Eq. (2.5) and utilize Eq. (2.6), then it is possible to show that the right-hand side of Eq. (2.5) will hold only if the following identities are satisfied:

$$(L^A c_{Ai}{}^{\rho})_{,\rho} + a_i{}^{\nu} L^A y_{A,\nu} + [c_{Ai}(\partial^{A\rho} L + \partial^A S^{\rho}) + c_{Ai,\sigma} \partial^{A\sigma} S^{\rho}]_{,\rho} \equiv 0. \tag{2.7}$$

In the general theory of relativity, a coordinate-covariant theory, they are known as the contracted Bianchi identities. We shall call them Bianchi identities even in the more general theory proposed in this paper, since our generalization is clearly based on the example of the gravitational field.

As a consequence of the existence of the Bianchi identities, it is possible to show that the divergence of the stress "tensor," $T_i{}^{\mu}$, vanishes identically,

$$\begin{aligned} &T_i{}^{\mu}{}_{,\mu} \equiv 0, \\ &T_i{}^{\mu} = a_i{}^{\nu}(\partial^{A\mu} L y_{A,\nu} - \delta_{\nu}{}^{\mu} L) \\ &\qquad - c_{Ai}{}^{\mu} L^A - [c_{Ai}(\partial^{A\mu} L + \partial^A S^{\mu}) + c_{Ai,\sigma} \partial^{A\sigma} S^{\mu}], \end{aligned} \tag{2.8}$$

even when the field equations are not satisfied. Equations (2.8) are known as the *strong conservation laws* of our theory. The stress "tensor" is not a geometric object; for the invariant class of transformations in which we are interested many of its components do not transform in any simple manner. One can only say that the identities hold in every coordinate, gauge, or "bein" (n-uple) system, and for all other frames of reference for which the field variables have a corresponding invariant significance. Closer inspection of these strong laws shows that the first few terms are similar in form to the conservation laws of energy and momentum that appear in Lorentz covariant field theories, where the conservation laws hold only if the field equations are satisfied ("weak" conservation laws). For our general theory these laws are

$$\begin{aligned} &t_i{}^{\mu}{}_{,\mu} = 0, \\ &t_i{}^{\mu} = a_i{}^{\nu}(y_{A,\nu} \partial^{A\mu} L - \delta_{\nu}{}^{\mu} L) \\ &\qquad - [c_{Ai}(\partial^{A\mu} L + \partial^A S^{\mu}) + c_{Ai,\sigma} \partial^{A\sigma} S^{\mu}]. \end{aligned} \tag{2.9}$$

The essential mathematical distinction between weak and strong laws is that the former arise because of covariance with respect to a finite set of arbitrary *constants*, while the strong laws are related to covariance with respect to a finite set of arbitrary *functions*. One can look at the question of strong *versus* weak laws from another viewpoint. It is possible to formulate strong laws in weak theories, but they will be integral laws. On the other hand, generally covariant theories will always yield differential strong laws, i.e., they are satisfied at each space-time point.

Because of Eq. (2.8), it must be possible to write the strong stress "tensor" in the form

$$T_i{}^{\mu} \equiv U_i{}^{[\mu\sigma]}{}_{,\sigma}, \tag{2.10}$$

where the $U_i{}^{[\mu\sigma]}$ are antisymmetric in the indices μ and σ. Henceforth, we shall call the $U_i{}^{[\mu\sigma]}$ the superpotentials of our theory. For any applications of the theory they must be explicitly calculated. In the general theory of relativity this was done by Freud,[10] but by a method that is only applicable in this case. We shall

[8] We shall use the Latin indices $i, k, \cdots$ to number the descriptors of the invariant transformation group. Coordinate indices running from 1 to 3 where needed will be identified by letters $r, s, \cdots$.

[9] Reference 1, Eq. (2.4).

[10] Ph. von Freud, Ann. Math. **40**, 417 (1939).

now show that it is possible to find an explicit solution for the super-potentials.

If we carry out the operations indicated by Eqs. (2.5), (2.6), the terms will group themselves as the coefficients of the various differential orders of the descriptors, ξ^i,

$$M_i{}^{\mu\rho\sigma}\xi^i{}_{,\mu\rho\sigma}+L_i{}^{\mu\rho}\xi^i{}_{,\mu\rho}-(T_i{}^\mu+V_i{}^{\mu\sigma}{}_{,\sigma})\xi^i{}_{,\mu}-T_i{}^\mu{}_{,\mu}\xi^i\equiv 0, \quad (2.11)$$

where

$$\begin{aligned} M_i{}^{\mu\rho\sigma}&=c_{Ai}{}^\mu\partial^{A\rho}S^\sigma,\\ L_i{}^{\rho\sigma}&=c_{Ai}{}^\rho\partial^{A\sigma}L+a_i{}^\iota(\delta_\iota{}^\sigma S^\rho-\delta_\iota{}^\rho S^\sigma)\\ &\quad+c_{Ai}{}^\sigma{}_{,\iota}\partial^{A\iota}S^\rho+c_{Ai}{}^\sigma\partial^A S^\rho+c_{Ai}\partial^{A\sigma}S^\rho \quad (2.12)\\ &\quad-a_i{}^\iota y_{A,\iota}\partial^{A\sigma}S^\rho+(c_{Ai}{}^\rho\partial^{A\sigma}S^\iota)_{,\iota},\\ V_i{}^{\rho\sigma}&=L_i{}^{\rho\sigma}-(c_{Ai}{}^\rho\partial^{A\sigma}S^\iota)_{,\iota}. \end{aligned}$$

Since the descriptors and their derivatives at any one space-time point are arbitrary, their coefficients must vanish separately.[11] For the third-order terms in the descriptors the coefficients must vanish when they are completely symmetrized in the indices μ, ρ, and σ,

$$[c_{Ai}{}^\mu\partial^{A\rho}S^\sigma]_{(\mu\rho\sigma)}=0. \quad (2.13)$$

The parentheses indicate that the expression inside the brackets is to be summed over six terms that are completely symmetric with respect to the indices inside the parentheses, $(\mu\rho\sigma)$.

For the second-order term, we have likewise

$$L_i{}^{\rho\sigma}+L_i{}^{\sigma\rho}=0, \quad (2.14)$$

while the first and zeroth-order terms tell us that although the divergence of $T_i{}^\mu$ vanished, $T_i{}^\mu$ is apparently not equal to the divergence of an antisymmetric form: comparison of $V_i{}^{\rho\sigma}$ with $L_i{}^{\rho\sigma}$ shows that $V_i{}^{\mu\sigma}$ is antisymmetric except for the divergence of the coefficient of the third-order term. However, by making use of the symmetry properties of this term, we can antisymmetrize $V_i{}^{\rho\sigma}$ completely by adding to it the divergence of another skew-symmetric density,

$$W_i{}^{\mu[\rho\sigma]}=\tfrac{1}{3}(c_{Ai}{}^\sigma\partial^{A\mu}S^\rho-c_{Ai}{}^\rho\partial^{A\mu}S^\sigma). \quad (2.15)$$

The final form of the super-potentials is then

$$U_i{}^{[\mu\sigma]}=V_i{}^{\mu\sigma}+W_i{}^{\mu[\sigma\rho]}{}_{,\rho}. \quad (2.16)$$

The expressions for the super-potentials, (2.16), can be simplified even further if we restrict ourselves to the general theory of relativity. In that case the c_{Ai} vanish, and explicit calculation shows that

$$a_i{}^\rho S^\sigma-a_i{}^\sigma S^\rho=c_{Ai}{}^\sigma\partial^A S^\rho+c_{Ai}{}^\sigma{}_{,\iota}\partial^{A\iota}S^\rho-a_i{}^\iota y_{A,\iota}\partial^{A\sigma}S^\rho. \quad (2.17)$$

It is not clear whether the conditions (2.17) are to be required for all generally covariant theories. The existence of these relations certainly simplifies the general theory of relativity as compared with the very general theory dealt with in this paper, especially if we make the transition to the Hamiltonian formalism. However, the transition can be made even if (2.17) is not satisfied, only the general formalism then becomes quite involved.

3. LAGRANGIAN FORMALISM AND GENERATING FUNCTIONAL

In what follows, we shall consider infinitesimal transformations of the (Lagrangian) field variables in the sense that at any one point in space-time the field variables y_A may be replaced by new field variables which are algebraic functions of the y_A and their first space-time derivations $y_{A,\rho}$. While eventually we shall focus our attention on transformations in which the Lagrangian does not change its form, this assumption will not be made at first. Given some function F of the field variables and their derivatives, we shall make a distinction between two infinitesimal expressions. We shall designate by the symbol $\bar{\delta}F$ the change in the value of F at the space-time point with the same coordinates x^ρ, on the assumption that we are describing in terms of our new variables the same original physical situation. We shall designate by the symbol $\delta' F$ the change in F as a function of its arguments y_A, $y_{A,\rho}$. The relationship between these two infinitesimal transformations is:

$$\begin{aligned} \delta' F&=\bar{\delta}F-\partial^A F\bar{\delta}y_A-\partial^{A\rho}F\bar{\delta}y_{A,\rho},\\ \bar{\delta}y_{A,\rho}&=(\bar{\delta}y_A)_{,\rho}=\bar{\delta}(y_{A,\rho}). \end{aligned} \quad (3.1)$$

This notation is the same as that used in earlier papers.[1,4]

Let us now consider what happens to the Lagrangian density L as a result of the infinitesimal transformation

$$\bar{\delta}y_A=f_A(y_B,y_{B,\rho}). \quad (3.2)$$

If the Lagrangian density were not to change its value at all, we should have $\bar{\delta}L=0$. It is, however, possible to add to the Lagrangian density a complete divergence, $Q^\rho{}_{,\rho}$ without affecting the Euler-Lagrange equations. We set, therefore,

$$\begin{aligned} \delta' L&=Q^\rho{}_{,\rho}-f_A\partial^A L-f_{A,\rho}\partial^{A\rho}L,\\ f_{A,\rho}&=y_{B,\rho}\partial^B f_A+y_{B,\rho\sigma}\partial^{B\sigma}f_A, \quad (3.3)\\ Q^\rho&=Q^\rho(y_A,y_{A,\rho}), \end{aligned}$$

and hence,

$$\delta' L=y_{A,\rho}\partial_A Q^\rho-f_A\partial^A L-y_{B,\rho}\partial^B f_A\partial^{A\rho}L+(\partial^{B\sigma}Q^\rho-\partial^{B\sigma}f_A\partial^{A\rho}L)y_{B,\rho\sigma}. \quad (3.4)$$

This change in the functional dependence of the Lagrangian density on its arguments implies that even though our Lagrangian density originally depends only on the field variables and their first derivatives, the transformed L will depend on higher derivatives as well unless we can choose Q^ρ and f_A so that the last term in Eq. (3.4) cancels, i.e.,

$$(\partial^{A\rho}Q^\sigma+\partial^{A\sigma}Q^\rho)-(\partial^{A\rho}f_B\partial^{B\sigma}L+\partial^{A\sigma}f_B\partial^{B\rho}L)=0. \quad (3.5)$$

We shall now consider more particularly that class of

[11] J. Heller, Phys. Rev. 81, 946 (1951).

transformations which will not produce second time derivatives. In that case, Eq. (3.5) must be satisfied merely with both ρ and σ equal to 4, and we have as a restriction on both the f_A and Q^4 the following:

$$\partial^A \cdot Q^4 - \partial^A \cdot f_B \cdot \pi^B = 0, \qquad \pi^B \equiv \partial^B \cdot L, \quad \partial^A \cdot \equiv \partial/\partial y_{A,4} \equiv \partial/\partial \dot{y}_A, \tag{3.6}$$

which may also be put in the form

$$f_B \partial^A \cdot \partial^B \cdot L = \partial^A \cdot (\pi^B f_B - Q^4). \tag{3.7}$$

This restriction may be put into a convenient form if we introduce instead of the $\dot{y}_A$ the π^A as new variables. If the matrix,

$$\Lambda^{AB} \equiv \partial^A \cdot \partial^B \cdot L, \tag{3.8}$$

is nonsingular, all that need to be done is to multiply Eq. (3.7) by the factor $(\partial \dot{y}_A / \partial \pi^C)$ to get

$$f_C = \partial_C(\pi^B f_B - Q^4), \quad \partial_C \equiv \partial/\partial \pi^C; \tag{3.9}$$

and we come to the conclusion that the class of transformations we have introduced are the *canonical* transformations, with the expression $\int (\pi^B f_B - Q^4) d^3x$ being the generating functional. In covariant field theories, however, the matrix (3.8) is always singular, and hence, Eq. (3.9) cannot be obtained that easily.

A theory is *covariant* (not Lorentz-covariant) if there exists a group of transformations, depending on one or several arbitrary *functions* of the four coordinates, which do not change the form of the Euler-Lagrange equations. The form of the Lagrangian may also be considered to remain unchanged under any transformation belonging to the group. In such a covariant theory, we can always produce a formally new solution out of an existing one by carrying out one of the *invariant* transformations. While this is also true of Lorentz-covariant theories, the peculiar feature of covariant theories is that the transformation may leave the solution unchanged for all values of $t = x^4 \leqslant t_0$, but change the solution *formally* for $t > t_0$. That means that the equations cannot determine the solution uniquely from a set of properly chosen initial conditions; it must be impossible to solve the equations with respect to the highest (i.e., second) time derivatives of all the field variables.

The coefficients of the highest time derivatives are just the Λ^{AB} of Eq. (3.8). We conclude that this matrix must have (at least) as many null-vectors as the transformation group of the theory possesses arbitrary functions. We conclude that there exists n algebraically independent quantities u_{iA}, $i = 1, \cdots, n$,

$$u_{iA} \Lambda^{AB} = 0. \tag{3.10}$$

Moreover, if we consider the π^A as functions of the velocities $\dot{y}_A$, then, because

$$\delta \pi^A = \Lambda^{AB} \delta \dot{y}_B, \tag{3.11}$$

the momentum densities will not change their values if the velocities are changed by amounts that are arbitrary linear combinations of the u_{iA},

$$\delta \pi^A = 0, \quad \delta \dot{y}_A = s^i u_{iA}. \tag{3.12}$$

We have, therefore, n algebraically independent "primary constraints," i.e., relationships between field variables, their spatial derivatives, and momentum densities that do not involve velocities,

$$g_i(y_A, y_{A,s}, \pi^A) = 0. \tag{3.13}$$

Because of Eq. (3.12), the u_{iA} are nothing but the partial derivatives of the expressions (3.13) with respect to the momentum densities:

$$u_{iA} = \partial_A g_i. \tag{3.14}$$

It is clear that the velocities cannot be uniquely determined functions of the canonical field variables, but even if (at one space-time point) values are adopted for the canonical variables in accordance with the primary constraints, then in addition to a particular solution for the velocities, say k_A, we have solutions depending on a set of arbitrary functions w^i,

$$\dot{y}_A = k_A(y_B, y_{B,s}, \pi^B) + w^i u_{iA}. \tag{3.15}$$

Values of the momentum densities are ruled out when not in accord with Eqs. (3.13). It makes no difference if we add to the functions k_A arbitrary linear combinations of the u_{iA} and also of the constraint expressions (3.13), because the latter additions vanish for all pertinent values of the canonical field variables. We shall utilize this freedom of choice later to make a desirable normalization.

In Eq. (3.15), we may consider the $\dot{y}_A$ as functions of the canonical field variables plus the functions w^i. In this sense, we may introduce partial derivatives of the velocities with respect to the momentum densities and with respect to these new arguments w^i. We may and shall require that

$$\Lambda^{AC} \partial_C \dot{y}_B + v^{iA} u_{iB} = \delta_B{}^A, \quad v^{iA} = \partial^A \cdot w^i. \tag{3.16}$$

In this set of conditions, the changes in the momentum densities by which the newly defined partial derivatives are multiplied are automatically consistent with the constraints (3.13). Hence, Eqs. (3.16) hold irrespective of the normalization of the functions k_A mentioned above. From them it follows that

$$u_{iA} v^{kA} = \delta_i{}^k. \tag{3.17}$$

Through normalization it is possible to accomplish further that

$$v^{iA} \partial_B \dot{y}_A = 0, \quad v^{iA} \partial_A \dot{y}_B = 0, \quad \partial_A \dot{y}_B - \partial_B \dot{y}_A = 0. \tag{3.18}$$

If these conditions are adopted, the chain rule of differentiation will hold for any functions that depend on the velocities, and also for any variables that depend on the momentum densities and the w^i in such a manner that

$$v^{iA} \partial_A F = 0. \tag{3.19}$$

Equation (3.19) represents no real restriction on the functional dependence of F, because it may always be satisfied by the addition of an appropriate combination of the constraints. That is, if some function F^* does not satisfy Eq. (3.19) initially, we may construct F,

$$F=F^*-g_i v^{iA}\partial_A F^*, \tag{3.20}$$

which will have the same value as F^* and will satisfy condition (3.19).

We shall now return to Eq. (3.6). Even though the matrix Λ^{AB} is singular, we shall obtain the relationship (3.9). We assume that we have defined the derivatives of the velocities with respect to the momentum densities in accordance with Eqs. (3.16) and (3.18) and also that the functions Q^4 and f_A all have been made to satisfy the condition (3.19) with the help of the procedure indicated in Eq. (3.20). Now we may multiply Eq. (3.6) by $\partial_C \dot{y}_A$, with the result,

$$\partial_C Q^4-\pi^B\partial_C f_B=0. \tag{3.21}$$

Integrating by parts yields

$$\begin{gathered} f_A=\partial_A C, \\ C=\pi^B f_B-Q^4, \quad v^{iA}\partial_A Q^4=0, \quad v^{iA}\partial_A f_B=0. \end{gathered} \tag{3.22}$$

If we multiply Eq. (3.6) by u_{iA}, we obtain the further relationship,

$$\partial C/\partial w^i=0. \tag{3.23}$$

It is interesting to note that the generating functional is not simply proportional to the integral of Q^4, as one might have suspected. It is possible to have canonical transformations that do not change the value of the action integral, i.e., transformations in which Q^4 is zero. This type of transformation is, of course, what is commonly called a point transformation (in configuration space), a transformation in which the transformed field variables are independent of the time derivatives of the original field variables. In that case, the generating functional is linear in the canonical momentum densities, a well-known fact; this special case is also consistent with Eq. (3.22).

The change in the action integral may be calculated either as the change if the physical situation remains unchanged and the values of the field variables are changed in accordance with Eq. (3.2); the four-dimensional action integral then changes by the amount,

$$\begin{gathered} \bar{\delta}S=\oint Q^\rho d\Sigma_\rho, \\ d\Sigma_\rho=\delta_{\rho\alpha_1\alpha_2\alpha_3}\prod_{s=1}^{3}\frac{\partial x^{\alpha_s}}{\partial u^s}du^s. \end{gathered} \tag{3.24}$$

Or, we may determine the change in S if under the transformation (3.2) we consider a new physical situation in which the transformed field variables have the same values as the untransformed field variables in the original situation. This calculation leads, of course, to the determination of $\delta' S$, the change in S as a function of its arguments. According to Eq. (3.3), this change is

$$\begin{gathered} \delta' S=\oint (Q^\rho-f_A\partial^{A\rho}L)d\Sigma_\rho-\int f_A L^A d^4x, \\ L^A\equiv\partial^A L-(\partial^{A\rho}L)_{,\rho}, \end{gathered} \tag{3.25}$$

where L^A is short for the left-hand side of the field equations. If we favor the x^4-directions by choosing a domain of integration which spatially extends to infinity and which extends between two $x^4=$constant surfaces, then Eq. (3.25) becomes

$$\delta' S=\left[\int_{x^s=-\infty}^{\infty}(Q^4-\pi^A f_A)d^3x\right]_{x^4=t_1}^{t_2} -\int_{x^4=t_1}^{t_2}\int_{x^s} f_A L^A d^4x. \tag{3.26}$$

If the field equations are satisfied, the second term vanishes, of course, and the change in action equals the difference between the generating functionals at the two time surfaces.

It remains to determine the change in the Hamiltonian functional as a function of *its* arguments. Just as we denote by δ' the change in a quantity when the values of y_A and $\dot{y}_A$ are held fixed, we shall introduce another symbol, δ'', to denote changes of quantities in which y_A and π^A are held fixed. We have, then

$$\delta'' L=\delta' L+\pi^A\delta''\dot{y}_A, \quad \delta''\pi^A\equiv 0, \tag{3.27}$$

and as a result

$$\begin{aligned} \delta'' H&=-\delta'' L+\pi^A\delta''\dot{y}_A=-\delta' L \\ &=-\dot{C}+f_A L^A. \end{aligned} \tag{3.28}$$

To summarize these results, Eq. (3.25) indicates the role of a four-"vector" as the generator of a canonical transformation in the Lagrangian formalism if all four coordinates are treated on the same footing; while Eqs. (3.27) and (3.28) show the relationship to the single generating functional of the canonical formalism.

Among all the canonical transformations, there is a set of transformations that leave the form of the Lagrangian invariant. These transformations are those in which $\delta' L$ vanishes (or equals a divergence, this latter possibility being of no consequence). According to Eq. (3.3) we have

$$0\equiv\delta' L=(Q^\rho-f_A\partial^{A\rho}L)_{,\rho}-f_A L^A. \tag{3.29}$$

For a transformation that does not change the form of the field equations we have, therefore

$$f_A L^A=(Q^\rho-f_A\partial^{A\rho}L)_{,\rho}=-C^\rho{}_{,\rho}. \tag{3.30}$$

Among all the transformations satisfying Eq. (3.30), there are some that correspond to a general invariance property of the theory and which by themselves form a significant subgroup. The term "invariant transformations" ordinarily applies to the members of that subgroup. In the case of invarient transformations, we call generating functionals more particularly the (three-dimensional) integrals over C^4, defined by Eq. (3.30). In the future we shall write this functional as

$$\bar{C}\equiv\int C^4 d^3x. \tag{3.31}$$

It remains to indicate the relationship between variations of the "path" (i.e., the field), which involve changes in the actual field, and transformations, which involve changes in representation. In a variation, the frame of reference remains fixed, and the values of the field variables are changed as functions of the four coordinates. But actually, no frame of reference can be identified without its contents, and thus the two operations are physically and mathematically equivalent. There is, however, a distinction between the trivial statement that the action integral S under a transformation changes only by a surface integral (that is how we arrange its transformation properties), and the requirement that the integral be stationary under an arbitrary variation. The difference is that in the first instance we change L by $\bar{\delta}L$, which leads to the result (3.24) and does not involve restrictions either on the original field or on the functions f_A. In the second instance, we ask for the change in S while we consider L a fixed function of its arguments y_A, $y_{A,\rho}$. This change is obtained if we subtract from the expression (3.24) the contribution that is due to the change of L as a function of its arguments, $\delta' L$, i.e., the expression (3.25). The requirement that the difference should also be a divergence, regardless of the choice of the variations f_A,

$$\begin{aligned}\bar{\delta}S-\delta' S&\equiv\oint f_A\partial^{A\rho}L d\textstyle\sum_\rho+\int f_A L^A d^4x\\&=\oint N^\rho d\textstyle\sum_\rho,\end{aligned} \tag{3.32}$$

leads to the field equations $L^A=0$.

We now proceed to examine the group of invariant variations described by (2.2) and (2.5). Under these transformations the C^ρ of (3.30) take the form

$$\begin{aligned}C^\rho=\partial^{A\rho}L(c_{Ai}{}^\sigma\xi^i{}_{,\sigma}+c_{Ai}\xi^i-y_{A,\sigma}a_i{}^\sigma\xi^i)&\\+a_i{}^\rho(L+S^\mu{}_{,\mu})\xi^i+\bar{\delta}S^\rho,&\end{aligned} \tag{3.33}$$

or, more simply,

$$C^\rho=-U_i{}^{[\rho\sigma]}\xi^i{}_{,\mu}-t_i{}^\rho\xi^i, \tag{3.34}$$

where $t_i{}^\rho$ and $U_i{}^{[\rho\sigma]}$ are defined by (2.9) and (2.16).

It is easily seen from (3.34), or more directly from (3.30) that when the field equations are satisfied, the divergence of C^ρ vanishes,

$$C^\rho{}_{,\rho}=0. \tag{3.35}$$

The time derivative of $\bar{C}$ (the generator of the invariant infinitesimal transformation) then vanishes if we assume that there are no variations of the field variables on two-dimensional spatial surfaces at infinity; for if we integrate the divergence of C^ρ over all of space and discard two-dimensional surface integrals, we get

$$\int C^\rho{}_{,\rho}d^3x=\int \dot{C}^4 d^3x\equiv d\bar{C}/dt. \tag{3.36}$$

However, the generating functional contains arbitrary functions (the descriptors) and their time derivatives. Thus, the coefficients of the various orders of the time derivatives of the descriptors, ξ^i, in $\bar{C}$ and $d\bar{C}/dt$ vanish. In the Lagrangian form of the theory these coefficients will vanish identically or as a result of the field equations. In the Hamiltonian theory those coefficients that were formerly field equations do not appear as any of the Hamiltonian equations of motion, but as secondary constraints. The appearance of these missing equations in the generating functional, where they mush vanish, insures the complete equivalence of both formalisms.

The terms that vanished identically now become the primary constraints of the Hamiltonian theory. There are also those expressions of the field variables and the momentum densities that vanish and continue to do so in the course of the motion only if the missing field equations are satisfied. In the Hamiltonian theory these field equations are called the secondary constraints. Furthermore, if particles are present as singularities of the field, the primary constraints will continue to be satisfied in the course of time only if restrictions are placed on the motion of these bodies. The number of such restrictions per particle (they appear in the form of two-dimensional surface integrals that vanish when the field equations are satisfied on the surface of integration) is equal to the number of arbitrary functions that appear in the transformation law for the field variables. In electromagnetic theory, a gauge invariant theory, there is but one restriction, the conservation of charge. In a coordinate covariant theory there are four restrictions, the conservation laws of energy and linear momentum for particles. We shall now examine the generating functional in detail to see how these restrictions arise.

We write $\bar{C}$ in the form

$$\bar{C}=\int[-U_i{}^{[44]}\dot{\xi}^i+(U_i{}^{[4s]}{}_{,s}-t_i{}^4)\xi^i]d^3x, \tag{3.37}$$

and its time derivative

$$\begin{aligned}\frac{d\bar{C}}{dt}=\int[-U_i{}^{[44]}\ddot{\xi}^i+(U_i{}^{[4s]}{}_{,s}-t_i{}^4-\dot{U}_i{}^{[44]})\dot{\xi}^i&\\+\frac{\partial}{\partial t}(U_i{}^{[4s]}{}_{,s}-t_i{}^4)\xi^i]d^3x,&\end{aligned} \tag{3.38}$$

where $\dot{\xi}^i$ stands for the time derivative of ξ^i. We know from our previous considerations that the various terms in (3.37) and (3.38) vanish separately. The first equation tells us that

$$U_i{}^{[44]}=0 \tag{3.39}$$

and

$$U_i{}^{[4s]}{}_{,s}-t_i{}^4=0. \tag{3.40}$$

(3.39) are the primary constraints of our theory. (3.40) are the missing field equations, the secondary constraints.

From (3.38) we deduce that

$$\dot{U}_i{}^{[44]}=U_i{}^{[4s]}{}_{,s}-t_i{}^4 \tag{3.41}$$

and

$$\frac{\partial}{\partial t}(U_i{}^{[4s]}{}_{,s}-t_i{}^4)=0. \tag{3.42}$$

The first of these equations states that the primary constraints will continue to be satisfied in the course of time only if the secondary constraints vanish. Equation (3.42) will vanish identically if the field equations are satisfied everywhere. However, if we integrate (3.42) over space and convert the expression into a three-dimensional divergence by means of the weak conservation laws, (2.9), we get from Gauss' law:

$$\frac{d}{dt}\oint U_i{}^{[4s]}n_s dS+\oint t_i{}^s n_s dS=0. \tag{3.43}$$

If the field equations are satisfied on the surface of integration, but not everywhere inside (at the position of particles), then Eq. (3.43) (which will still hold) represents restrictions on the motion of these particles. We define the conserved particle qualities by the surface integrals,

$$T_i=\oint U_i{}^{[4s]}n_s dS. \tag{3.44}$$

The restrictions (3.43) are in the form of conservation laws which state that the rate of change of the particle characteristic T_i depends upon the amount of field flux $t_i{}^s$ that passes through the surface. Thus, in the most general case, the particle and field are in constant interaction and such qualities of the particle as energy and linear momentum are strongly dependent on the motion of the field.

Since the motion of particles should be completely determined by a knowledge of the field equations on any hypersurface surrounding the particle, the values of the sums of the two surface integrals in each of Eqs. (3.43) should not depend on the surfaces of integration chosen. The proof of this independence is most easily accomplished by showing that the three-dimensional divergence of the integrand of (3.43) vanishes when the field equations are satisfied. Then it immediately follows from Gauss' law that the integrals are independent of the surfaces of integration.

As a consequence of the strong conservation laws it is possible to show that when the field equations are satisfied

$$U_i{}^{[rs]}{}_{,s}=t_i{}^r+\dot{U}_i{}^{[4r]}. \tag{3.45}$$

The divergence of this expression certainly vanishes, and since the right-hand side is the integrand of (3.43), our proof is complete.

4. QUANTUM FIELD THEORY IN THE CANONICAL FORMALISM

In attempting to quantize covariant field theories, we have been guided by the assumption that the basic structure of the transformation group should remain unaffected by quantization. This assumption has been justified by the history of quantum theory to date. In every successful quantum theory all invariance properties, such as symmetry properties, Lorentz covariance, and gauge covariance have been the same as in the corresponding classical theory. This is true even for theories involving spinors. While the spin has, strictly speaking, no classical analog, the covariance group of the spin field is the same as that encountered in similar classical theories. Merely the representation is different. But that means that commutators and other combinations of individual transformations lead to the same transformations as they do in nonquantum theories. That is why we are convinced that it is at least reasonable to modify the classical results obtained in previous papers very slightly if at all, as far as the transformation theory is concerned.

Prior to the formal simplification of Feynman's Lagrangian quantization[6] by Schwinger,[7] we concentrated primarily on quantization through the canonical formalism. Whether direct quantization of the Lagrangian theory leads to an appreciable simplification remains to be seen.

In the classical theory, a covariant theory in the canonical formalism comprises a Hamiltonian and a number of constraints, classified as primary and secondary constraints. The group of invariant transformations is generated by functionals that are three-dimensional integrals over generating densities, the latter being linear combinations of the constraints. The constraints are multiplied by arbitrary functions, the *descriptors*, which characterize each individual transformation,[4] in such a manner that the highest-order constraints are multiplied by the undifferentiated descriptors, and the primary constraints by the highest time derivatives that occur. Under any of these invariant transformations, the Hamiltonian density remains unchanged in value, but it will add a linear combination of the primary constraints whose coefficients depend on the descriptors. The totality of the constraints by themselves form a system of involutions; together with the Hamiltonian they form a function group. For the transition to the quantum theory, it appears essential that Poisson brackets are not merely formal expressions but that they have a natural group-theoretical significance; they represent the group-theoretical commutator of two infinitesimal transformations.

It may then be said that not only the Hamiltonian but each one of the constraints generates an infinitesimal canonical transformation which leaves the form of each constraint and of the Hamiltonian unchanged (modulo some constraints). Among these, the members of the original invariant transformation group form a subgroup which adds to the Hamiltonian only primary constraints. Furthermore, the primary constraints by themselves form another subgroup which, however, adds to the Hamiltonian higher-order constraints.

It is to be assumed, then, that the transformation theory of the quantized theory will run exactly along the same lines. Instead of with canonical transformations we shall have to deal with unitary transformations. Poisson brackets will have to be replaced by commutators of the corresponding Hermitian operators. In that respect, then, we must require that the commutation rules between the Hamiltonian density and the various constraints be exactly the same as between their nonquantum analogs.

In the full formulation of the quantized theory, there arises a well-known difficulty, which is somewhat more serious in an essentially nonlinear theory than in a theory like quantum electrodynamics. That difficulty is that there appears a discrepancy between the commutation relations on the one hand and the constraints on the other. Briefly, there are, classically speaking, variables that are canonically conjugate to constraints. Now it is a well-established principle that if two operators have a c-number commutator, then neither of them possesses discrete eigenvalues nor normalizable eigenfunctions. On the other hand, physical situations must be characterized by wave functionals (Hilbert vectors) that are eigenfunctions of every constraint, belonging to the eigenvalue 0.[12] A possible way out of this paradox is to build up the Hilbert space from only those states that satisfy all constraints and to make it, thus, deliberately, a small subspace of the functional space of all conceivable wave functionals without regard to constraints.

Clearly, this subspace contains all states that can possibly characterize a physical situation. It is also certain that the continuous unitary transformation generated by the Hamiltonian does not lead outside this subspace, though this operator induces some motion throughout the "big" functional space as well. All unitary transformations generated by any of the constraints map the Hilbert space on itself, and this includes those particular unitary transformations forming the group of transformations with respect to which the (quantized) theory is to be invariant.

But how does transition to the smaller space solve our difficulties? It does so by severely restricting the number of linear operations that may properly be called Hilbert operators. A Hilbert operator must map at least a dense set of states belonging to the Hilbert space into that same Hilbert space. But that will be the case only for such linear operations in the "big" space that commute with all constraints (modulo the constraints themselves). If we adopt this condition as a necessary property of all Hilbert operators (this condition can be sensibly formulated as a restriction only in the "big" space; operators not satisfying it will simply not be defined on the subspace), our difficulties will be removed for the simple reason that the troublesome operations canonically conjugate to constraints are not Hilbert operators in our sense! The constraints themselves reduce to zero operators. They do not represent significant physical quantities.

Our new convention saves us from formal embarrassment. Will the elimination of a large number of operations from consideration not embarrass us as physicists, by eliminating the mathematical description of physically meaningful quantities?

In answer, we find that the observables not ruled out are invariants, quantities that remain unchanged under all the infinitesimal transformations with respect to which the theory is assumed to be invariant. For electrodynamic quantities, for instance, commutability with the subsidiary conditions of quantum electrodynamics implies gauge invariance. And truly, any quantity that can be given a well-defined numerical value must be an invariant!

Frequently, it appears that we operate with non-invariant quantities as if they were directly observable, such as a particular component of a four-vector or tensor. On closer examination, we must admit that such a quantity has a meaning only with respect to a specified frame of reference. But how do we determine a particular frame of reference? Generally, we define a frame of reference in terms of physical realities; we speak of a "laboratory frame of reference," etc. A (local) frame of reference is, for instance, frequently defined in terms of a prevailing local velocity. This local velocity is, however, the time-like eigenvector of the matter tensor (if observed in a particular fashion). If we observe the scalar product of some four-vector (e.g., electric charge-current density) with this local velocity vector, we actually measure an invariant that depends in a complicated fashion, but uniquely, on both the matter tensor and the charge-current vector. We have gone through a number of examples to convince ourselves that any physically observable quantity

[12] The remark by G. Wentzel in his *Quantentheorie der Wellenfelder* (Franz Deuticke, Wien, 1943), p. 111, that while it is not permissible to set a constraint operator equal to zero outright, one may restrict oneself to Schrödinger functions which are made to vanish by the application of a constraint operator, refers to the general function space of all conceivable wave functionals, whether they satisfy the constraints or not, and holds, therefore, only without regard to normalizability. On the other hand, unless there exists a norm and with it the concepts of Hermiticity and unitariness, Dirac brackets are generally meaningless, and one cannot ascribe a precise meaning to the expectation value of a dynamical variable.

is an invariant, but this point is so obviously of major importance that it should be more fully investigated. Suffice it here to say that the point of view we have adopted is a generalization of and consistent with accepted practices in quantum electrodynamics and elsewhere.

So far, we have indicated the general form of a quantized covariant field theory. While we have not assumed the introduction of parameters, their introduction[2,3] does not lead to serious modifications. Most important, the Hamiltonian becomes itself a constraint. There is, therefore, no "motion." Any state that obeys all constraints is a solution of the Schrödinger equation, provided we do not permit it to change in the course of "time." This apparent freezing is, however, a purely formal result of the introduction of parameters. The Hilbert operators in the parameter formalism are perforce all constants of the motion, and the distinction between Schrödinger representation and Heisenberg representation disappears. In the parameter formalism the same physical quantity, observed at two different times, simply appears as two distinct operators.

A conjecture we made some time ago,[2] to the effect that the ordinary coordinates might turn into quantum-theoretical observables in a parameter formalism cannot be maintained. The coordinates do not commute with the "parameter constraints"[2] and are, therefore, not Hilbert operators. In fact, parametrization does not add to or subtract from the observables of the parameter-free theory.[13] It is a matter of purely formal convenience whether or not they should be used in any investigation.

Suppose a classical canonical covariant field theory is available. Can its quantum-theoretical analog be obtained by a definite algorithm? At present the answer is no. Primary constraints are generally of very simple structure, and it is possible to construct them by simple symmetrization with respect to the position of the momentum densities, which occur only linearly. They will satisfy among themselves the same commutation relations as their nonquantum analogs. The Hamiltonian is generally quadratic in the momentum densities, and so are the secondary constraints. It will be necessary to discover such an arrangement of factors in the Hamiltonian that the higher-order constraints, which are essentially defined as commutators between the primary constraints and the Hamiltonian, will in turn satisfy the correct commutation relations with the primary constraints and with the Hamiltonian. We have not yet carried out a specific example to show how difficult a problem it is to discover the right arrangement or whether this arrangement is in fact uniquely determined.[14]

[13] R. Penfield, Phys. Rev. 84, 737 (1951).

[14] In informal discussion, J. L. Anderson has suggested that the sequence of factors be settled by the introduction of "locally geodesic" variables and retransformation to ordinary physical variables. This suggestion is closely related to a recent paper by B. S. DeWitt, Phys. Rev. 85, 653 (1952).

5. LAGRANGIAN QUANTIZATION

One of the purposes of characterizing canonical transformations within the Lagrangian formalism in Sec. 3 was to prepare a formulation that would be sufficient to quantize in the Lagrangian formalism. We are here reformulating Schwinger's ideas[7] with this end in mind. In all that follows it is understood that equations involving operators are to be read as equations for the matrix elements leading from one fixed surface to another.

We shall assume that the elements of the unitary matrix $U(t_2, t_1)$ lead from one hypersurface in space-time belonging to the coordinate value t_1 to another hypersurface that belongs to the coordinate value $x^4 = t_2$. In a parameter formalism, x^4 must be replaced by the parameter t. For any transformation, invariant or otherwise, we shall set

$$\delta' U = \frac{i}{\hbar}\delta' S = \frac{i}{\hbar}\int \delta' L d^4x. \tag{5.1}$$

We adopt $\delta' S$ rather than $\bar{\delta} S$ on the right-hand side, because all operators and matrix elements are defined in terms of "a complete set of commuting operators," by assumption the field variables $y_A(x^1, x^2, x^3)$ for any fixed x^4. For an invariant transformation, in which the Lagrangian as a function of its arguments, i.e., the form of the theory, does not change, the elements of the matrix U do not change, either.

Following closely the arguments of Sec. 3, we now set $\delta' L$ equal to

$$\begin{aligned} \delta' L &= Q^\rho{}_{,\rho} - \{\partial^A L \cdot f_A\} - \{\partial^{A\rho} L \cdot f_{A,\rho}\} \\ &= [Q^\rho - \{\partial^{A\rho} L \cdot f_A\}]_{,\rho} - \{L^A \cdot f_A\}, \\ L^A &\equiv \partial^A L - (\partial^{A\rho} L)_{,\rho}. \end{aligned} \tag{5.2}$$

In this equation, "symbolic" products of operators are indicated by braces and dots and are to be understood so that the factor f_A is to be inserted in each product at the point where the partial derivative of L with respect to one of its arguments produces a vacancy. It follows that

$$\begin{aligned} \delta' U &= -\frac{i}{\hbar}\oint C^\rho d\textstyle\sum_\rho - \frac{i}{\hbar}\displaystyle\int \{L^A \cdot f_A\} d^4x, \\ C^\rho &\equiv \{\partial^{A\rho} L \cdot f_A\} - Q^\rho, \end{aligned} \tag{5.3}$$

if we make no assumptions concerning the field equations.

It is tempting to assume that for all conceivable transformations the volume integral on the extreme right-hand side must vanish and that this be the natural form of the quantized field equations, in view of Eqs. (3.31). But unless the choice of f_A is restricted, such a requirement is much too stringent. It would be equivalent to a "double-operator" equation at each space-time point. In other words, if we call an ordinary operator a "Hilbert tensor of rank 2," then the requirement that the expression $(L^A\cdot f_A)$ should vanish for arbitrary f_A would represent a Hilbert tensor equation of rank 4. Such a theory would obviously have little in common with quantum field theory as we know it. (Exactly the same argument may be carried out, with the same results, in quantum mechanics.)

Instead, we propose to require that $(L^A\cdot f_A)$ vanish only for such choices of f_A that correspond to *invariant* transformations. This proposal can be justified as follows. First of all, this choice is itself an invariant requirement, since every invariant transformation maps the set of invariant transformations on itself. Second, with this mild requirement we obtain all the conservation laws that are associated with the invariant transformations of a given theory; thus, the major results of Schwinger's paper[7] will hold. Third, among the invariant transformations is, of course, the motion in the course of time, i.e., a coordinate transformation in which only x^4 (or t) changes. For this particular transformation, the generating density becomes directly the Hamiltonian density, and thus the Schrödinger equation is obtained.

Thus, for a theory that is generally covariant we should have

$$\{L^A\cdot\bar{\delta}y_A\}\equiv\{L^A\cdot(F_{A\mu}{}^{B\nu}y_B\xi^\mu{}_{,\nu}-y_{A,\mu}\xi^\mu)\}=0. \tag{5.4}$$

In this expression, the choice of the "descriptors" (which we shall assume are c-numbers) is arbitrary.[1] We separate into a complete divergence and, in addition, a set of terms that are multiplied by undifferentiated descriptors,

$$\{L^A\cdot\bar{\delta}y_A\}\equiv\{L^A\cdot F_{A\mu}{}^{B\nu}\xi^\mu y_B\}_{,\nu} -\xi^\mu(\{L^A\cdot F_{A\mu}{}^{B\nu}y_B\}_{,\nu}+\{L^A\cdot y_{A,\mu}\}). \tag{5.5}$$

In view of the assumption that the theory is covariant, it follows from the structure of the Lagrangian alone that the expression computed must be a divergence, though it is not a consequence of the theory that it must vanish. However, we may infer that the second term vanishes, i.e., that the Bianchi identities are satisfied, purely as a result of the transformation properties of the Lagrangian:

$$\{L^A\cdot F_{A\mu}{}^{B\nu}y_B\}_{,\nu}+\{L^A\cdot y_{A,\mu}\}\equiv 0. \tag{5.6}$$

We may not infer, but we shall assume that Eq. (5.4) holds, and that means that the first term on the right-hand side of Eq. (5.5) must vanish. If we integrate this expression over a four-dimensional domain, we may convert it into a surface integral and find

$$\oint \xi^\mu\{L^A\cdot F_{A\mu}{}^{B\nu}y_B\}d\Sigma_\nu=0. \tag{5.7}$$

Finally, we may choose $\xi^\mu(x^\rho)$ so that only a small (cap) surface element can contribute. We finally arrive at the result

$$\{L^A\cdot F_{A\mu}{}^{B\nu}y_B\}=0. \tag{5.8}$$

More generally, we may conclude from Eq. (5.3) that for invariant transformations, where $\delta' L$ vanishes, our conjectured field equations,

$$\{L^A\cdot f_A\}=0, \tag{5.9}$$

are equivalent to the requirement that the divergence $C^\rho{}_{,\rho}$ vanishes. In a covariant theory, in which C^ρ is given by an expression like (3.34), Eq. (5.9) leads to the requirement,

$$T_i{}^\rho-t_i{}^\rho=0=\{L^A\cdot c_{Ai}{}^\rho\}. \tag{5.10}$$

Furthermore, if the difference between the strong and the weak "stress tensor" vanishes, it follows that the weak stress tensor by itself satisfies a conservation law,

$$t_i{}^\rho{}_{,\rho}=0. \tag{5.11}$$

In a simple theory like the theory of gravitation, where the C_{Ai} vanish, the weak stress tensor belonging to the coordinate descriptors is given by the expression

$$t_\mu{}^\rho=\{\partial^{A\rho}L\cdot y_{A,\mu}\}-L\delta_\mu{}^\rho. \tag{5.12}$$

We can now show that $t_4{}^4$ is the Hamiltonian density. If we carry out a coordinate transformation in which $\xi^4=1$ on one space-like three-dimensional surface (t_2), but vanishes on the other (t_1), we have, because of Eq. (5.3),

$$\delta' U=-\frac{i}{\hbar}\int_{x^4=t_2} C^4 d^3x,\qquad C^4=\{\partial^A\cdot L\cdot\dot{y}_A\}-L\equiv H. \tag{5.13}$$

This last equation is the Schrödinger equation. It shows at the same time how the Lagrangian and the Hamiltonian are related to each other, including the sequence of factors. The usual rule, that the time derivative of an observable is determined by its commutator with the Hamiltonian, also follows directly from Eq. (5.13).

With the assumptions made, it is possible to work out commutation relations. These commutation relations will be between the field variables and their time derivatives and do not involve by themselves momentum densities (which we have not introduced). These commutation relations are obtained as follows. We determine the generating functions for each of the invariant transformations of the theory and use them to calculate the changes in the various field variables. These changes will appear on both sides of the equation and will generally depend on the descriptors or on derivatives of the descriptors. The requirement that these equations should not lead to restrictions on the descriptors and their derivatives, nor permit a determination of the time derivatives of the field variables in terms of the field variables themselves, or some similar restriction, leads to the commutation relations.

Generally, there will be some relations that cannot be satisfied. In such cases, it will turn out invariably that the variables involved are "constraint variables" and therefore not true observables. This circumstance shows that in the Lagrangian formalism the Hilbert space will be very similar to (and as involved as in) the Hamiltonian formalism described in the beginning of this section.

Two examples will suffice. First, consider a problem in ordinary quantum mechanics. Assume the problem possesses a mixed quadratic Lagrangian of the form,

$$L=\tfrac{1}{2}\dot{q}^k g_{kl}\dot{q}^l - V, \tag{5.14}$$

where g_{kl} and V are independent of the velocities. A brief calculation shows that the Hamiltonian will be:

$$H=\tfrac{1}{2}\dot{q}^k g_{kl}\dot{q}^l + V. \tag{5.15}$$

We now attempt to determine the velocities by the standard procedure. The resulting equation is

$$\begin{aligned}\dot{q}^m &= \frac{i}{2\hbar}([\dot{q}^k, q^m]g_{kl}\dot{q}^l + \dot{q}^l g_{kl}[\dot{q}^k, q^m], \\ [\dot{q}^k, q^m] &\equiv [\dot{q}^m, q^k].\end{aligned} \tag{5.16}$$

If the form g_{kl} is nonsingular, the last equation will lead to no algebraic restriction on the velocities only if we assume

$$\frac{i}{\hbar}[\dot{q}^k, q^m] = g^{km}, \quad g^{km}g_{ml} = \delta_l{}^k. \tag{5.17}$$

This relationship is equivalent to the usual commutation relation between q_k and p_l. On the other hand, if g_{kl} is singular, then each of its null vectors corresponds to a "constraint variable," and the commutators involving them cannot be determined in a satisfactory manner.

As a second example take the commutation relations of the electromagnetic field in the vacuum. The generating function of a gauge transformation is

$$\bar{C} = \frac{1}{4\pi}\int \phi^{4s}\xi_{,s}d^3x. \tag{5.18}$$

We can derive the equation

$$\begin{aligned}\delta\phi_\rho(\mathbf{x}) &= \xi_{,\rho} \\ &= \frac{i}{4\pi\hbar}\int [\phi^{4s}(\mathbf{x}'), \phi_\rho(\mathbf{x})]\xi_{,s}(\mathbf{x}')d^3x',\end{aligned} \tag{5.19}$$

which leads to

$$[\phi^{4s}(\mathbf{x}), \phi_r(\mathbf{x}')] = \frac{4\pi\hbar}{i}\delta_r{}^s\delta(\mathbf{x}, \mathbf{x}'), \tag{5.20}$$

but does not lead to any definite commutation relation involving ϕ_4.

It appears very definitely that in the Lagrangian formalism, too, the quantum state of the system, or the matrix U, for that matter, must not depend on the "constraint variables," i.e., the variables canonically conjugate to the constraints. There is another argument which makes this result appear reasonable. It has been pointed out previously that the variables canonically conjugate to constraints will change in the face of one of the invariant transformations, the remainder not. That means that the value of one of the "canonical" variables, or, rather, its probability distribution at some time t_2, depends only on the state of the system at another time t_1, but the distribution of a "constraint" variable both on the state previously and on the frame of reference chosen. Since the choice of frame of reference in this context involves no physical activity on the part of the observer (such as modifications of his measuring instruments), but rather his "state of mind," the matrix U can hardly give information on "constraint variables." Thus, we come to the conclusion that our original assumption, that the field variables y_A represent a complete set of commuting operators, must be modified in that all those combinations of y_A that are canonically conjugate to the constraints must be excluded.

6. CONCLUSION

We have conjectured that the field equations of a covariant quantized theory are not simply the analogs of the Euler-Lagrange equations of the nonquantum theory, but of the form (5.9), where f_A stands for the most general invariant $\bar{\delta}y_A$ of the theory. If this conjecture should turn out to be correct, then the resulting Lagrangian quantum field theory will be equivalent to a canonical field theory (because there will exist a single "Hamiltonian" functional generating the motion), but possibly more manageable. The total number of algebraically independent "field equations" (5.9) is determined not by the number of field variables but by the structure of the transformation group of the theory. In the case of coordinate-covariant theories, we are led to the set of 16 equations (5.8); in theories of the electromagnetic field the gauge group generates 4 equations. In the case of the general theory of relativity and in the case of electrodynamics, it can be shown that in the classical limit these relations are exactly equivalent to the usual Euler-Lagrange field equations, but it is conceivable that in a theory in which the number of field variables greatly exceeds the number of field equations of the type (5.10) such a correspondence breaks down. We have not yet examined this question in any detail.

The significance of the conjectured equations (5.8) or (5.10) is closely related to the so-called secondary constraints. While in a Lagrangian theory the primary constraints are empty, the secondary constraints are not, but represent certain of the conjectured field equations, e.g., $\{L^A\cdot F_{A\mu}{}^{B4}y_B\}$, which are free of second time derivatives. We may, conversely, obtain all of the Eqs. (5.8) from the secondary constraints simply by giving the three-dimensional surface with respect to which they are formed every possible orientation.

In our Lagrangian theory, the commutation relations between the y_A and the $\dot{y}_A$ are not determined completely. As a result, the Hamiltonian, though formally it generates the motion, actually does not determine the time derivatives of all field variables. Preliminary examination shows, however, that the variables whose time derivatives remain indeterminate are precisely the ones whose time derivatives are also indeterminate in the classical theory, those which in the Hamiltonian theory are canonically conjugate to the primary constraints. In a nonlinear covariant theory, the separation of those variables whose time derivatives are indeterminate from those which are completely determined is a mathematical task of almost insurmountable difficulty. In all probability, the construction of the Hilbert space of permissible states and legitimate observables will be just as difficult in the Lagrangian formalism as it is in the canonical formalism (see Sec. 4).

In future work we plan to pursue the application of both the Hamiltonian and the Lagrangian theory to actual physical theories and to ascertain the usefulness of either.

PHYSICAL REVIEW VOLUME 89, NUMBER 1 JANUARY 1, 1953

Meson Theory of β-Decay and the ΔL-Forbidden Transitions

SEITARO NAKAMURA
Department of Physics, University of Tokyo, Tokyo, Japan
(Received April 21, 1952)

Existence of the β-meson of vector type which is neither a π- nor μ-meson and which transmits the β-decay of nuclei has been tantatively assumed, the results thus derived being compared with experiments on β-decay. The differences between the β-meson theory and the phenomenological Fermi theory lie in the natural deduction of the vector interaction as well as the introduction of new nuclear matrix elements which are characteristic of the meson dipole. The selection rules for the ΔL-forbidden transitions are involved in these matrix elements, the orders of which are expected to be that of the unfavored parity transitions in each degree of forbidden transitions. The observability of a free β-meson is also discussed.

1. THE MESON THEORY OF β-DECAY

THE theory of β-decay, originally formulated by Yukawa,[1] predicted the virtual emission and reabsorption of charged mesons in the β-decay of nucleons. After the two mesons, π and μ, were discovered, one was not successful in identifying either of them as an intermediary agent of nuclear β-decay, in spite of extensive analyses.[2–5] Recently, Friedman and Rainwater[6] showed that a free π-meson will not decay into an electron and neutrino with a probability more than 1/1419 times the probability of $\pi-\mu$ decay. Sasaki, Hayakawa, and the present author[7] once proposed the existence of a vector meson, which is neither π nor μ, and could be the agent of Yukawa's original idea, that of transmitting the nuclear force and β-decay. We have named it the β-meson. Tanikawa[8] and Caianiello[9] suggested the meson theory of β-decay through a τ-meson, the former assuming it to be pseudoscalar while the latter taking it to be vector. However, it is almost certain that the fate of these theories will primarily be dependent on the β-ray analysis. To clarify this point, we shall derive the β-meson theory in some detail.

For the purpose of this discussion, the β-meson is assumed to be of vector type, as in the Fermi theory the tensor interaction is indispensable in the explanation of the results of the recent experiments on the β-ray spectra and $\beta-\gamma$ angular correlations.[10] Its mass and coupling constants with nucleons and leptons will be tentatively taken arbitrary, not referring to those of any observed meson. Taking Konopinski's[11,12] Hamiltonian for the vector meson theory of β-decay, we have, in the case of an allowed transition, (as regards notation, see refer-

[1] H. Yukawa, Proc. Phys. Math. Soc. Japan **17**, 48 (1935), third series; S. Sakata, Proc. Phys. Math. Soc. Japan **23**, 291 (1941); **24**, 843 (1942); **25**, 86 (1943); H. A. Bethe and L. W. Nordheim, Phys. Rev. **57**, 998 (1940).
[2] Taketani, Nakamura, Ono, and Sasaki, Phys. Rev. **76**, 60 (1949); Nakamura, Fukuda, Ono, Sasaki, and Taketani, Prog. Theor. Phys. **5**, 740 (1950).
[3] H. Yukawa, Revs. Modern Phys. **21**, 474 (1949).
[4] J. Steinberger, Phys. Rev. **76**, 1180 (1949).
[5] M. Ruderman and R. Finkelstein, Phys. Rev. **76**, 1458 (1949).
[6] H. L. Friedman and J. Rainwater, Phys. Rev. **84**, 684 (1951).
[7] Sasaki, Nakamura, and Hayakawa, Prog. Theor. Phys. **3**, 454 (1948).
[8] T. Tanikawa, Prog. Theor. Phys. **3**, 315 (1948).
[9] E. R. Caianiello, Phys. Rev. **81**, 625 (1951).
[10] C. S. Wu, Revs. Modern Phys. **22**, 386 (1950); I. Shaknov, Phys. Rev. **82**, 333 (1951).
[11] E. J. Konopinski, Revs. Modern Phys. **15**, 209 (1945).
[12] E. J. Konopinski and E. Uhlenbeck, Phys. Rev. **60**, 308 (1941).