A First-Order Variational Principle for Classical Electrodynamics

F. J. BELINFANTE
*The University of British Columbia, Vancouver, British Columbia, Canada**
(Received April 21, 1948)

The fact that the Schrödinger and the Dirac equations for the wave function of an electron are differential equations of the first order with respect to time, while in classical theories differential equations of the second order are common, has necessitated a slightly different setup of the canonical theory in both fields. Under influence of the classical methods, second-order equations are also often used in the quantum theory of particles of integer spin, thus causing a difference between the treatment of Fermi-Dirac and of Einstein-Bose particles.

It is shown here that conformity between classical and quantum-mechanical methods can be achieved easily by use of first-order equations throughout, thus avoiding a superfluous distinction between integral and half-odd-integral spin fields. The classical theory of a point charge in an electromagnetic field of force is set up here from this point of view.

IN classical mechanics, it is usual to derive the equations of motion from a variational principle $\delta\int\mathcal{L}(q_k, \dot{q}_k)dt=0$ of such a type that the differential equations obtained are of the second order. When the Hamiltonian, $\mathcal{H}=\sum p_k\dot{q}_k-\mathcal{L}$, is expressed in terms of the variables q_k and their canonical conjugates $p_k=\partial\mathcal{L}/\partial\dot{q}_k$ afterwards, the $\dot{q}_k$ are eliminated then by expressing them in terms of the momenta p_k. A similar procedure is also used for Maxwell's theory, when $(1/8\pi)\int\int\int\int(\mathbf{E}^2-\mathbf{H}^2)dxdydzdt$ or some other convenient expression is used as a Lagrangian, with the equations $\mathbf{H}=\text{curl}\mathbf{A}$ and $\mathbf{E}=-\nabla\Phi-\dot{\mathbf{A}}/c$ added as definitions of $\mathbf{H}$ and of $\mathbf{E}$.

In wave mechanics, the procedure is slightly different. The variational principle used is here $\delta\int dt\mathcal{L}(Q_k(x), \dot{Q}_k(x))=0$, where the functional $\mathcal{L}$ is given by $\mathcal{L}=\mathcal{S}\psi^\dagger(i\hbar\partial/\partial t-\text{H})\psi$. (The symbol $\mathcal{S}$ means integration over x, y, and z and summation over components.) This variational principle gives equations of motion $i\hbar\partial\psi/\partial t=\text{H}\psi$ and $-i\hbar\partial\psi^\dagger/\partial t=(\text{H}\psi)^\dagger$, which are of the *first* order, at least in t. (In Dirac's relativistic theory H is also linear in the gradient operator ∇.) Further, the "momentum" (P) canonically conjugate to the "variable" ($Q\equiv)\psi$ is here $(P\equiv)i\hbar\psi^\dagger$, which is one of the variables, on which the Lagrangian $\mathcal{L}$ itself depends. The Hamiltonian is now $\mathcal{H}=\mathcal{S}P\dot{Q}-\mathcal{L}=\mathcal{S}\psi^\dagger\text{H}\psi$, so that the $\dot{Q}$ drop out automatically. (Remark also that P is no longer some simple function of $\dot{Q}$.)

It seems surprising that there is so much difference between the wave mechanical and the classical procedure. This has led many authors, in particular when a field of particles of integral spin had to be described, to the use of a "second-order Lagrangian" also in wave mechanics, in close analogy with the procedure commonly used for Maxwell's theory. In the case of particles of spin $\frac{1}{2}$, however, this method remains unsatisfactory, so that a complete unification of methods would seem impossible.

This note serves to point out that the unification can be sought and found in the opposite direction. The application of first-order Lagrangians to the description of wave fields of particles with integral as well as with half-odd-integral spin has been recommended and demonstrated by the author several times.[1] Here I want to point out that this method is not confined to wave mechanics, but can be used just as well in classical point mechanics. The interesting point is the close resemblance one finds between the Hamiltonian in this classical formalism and the Hamiltonian of Dirac's theory of electrons.

We shall show here this method for a classical electron (for the sake of simplicity considered here as a point charge e at the position $\mathbf{x}_e$ at the time t) in an arbitrary Maxwell field. We start from the relativistic variational principle

$$\delta\{\int\Lambda d\tau+\mathcal{S}\int L_f dxdydzdt\}=0, \tag{1}$$

* Now at Purdue University, Lafayette, Indiana, U. S. A.

[1] F. J. Belinfante, *Theory of Heavy Quanta* (thesis) (M. Nijhoff, The Hague, 1939). See also: F. J. Belinfante, Physica **6**, 887 (1939); Physica **7**, 449, 765 (1940); Physica **12**, 1 (1946). Also: H. A. Kramers, F. J. Belinfante, and J. K. Lubański, Physica **8**, 597 (1941).

with

$$-c^2d\tau^2=dx^\mu dx_\mu,\quad (x^0=-x_0=ct,\ x^1=x_1=x_e,\ x^2=x_2=y_e,\ x^3=x_3=z_e),\tag{2}$$

$$\left.\begin{aligned}\Lambda&=-mc(-u^\lambda u_\lambda)^{\frac{1}{2}}+p_\lambda dx^\lambda/d\tau\\&\qquad-(p_\lambda-[e/c]A_\lambda{}^e)u^\lambda,\\L_f&=(1/4\pi)(\tfrac{1}{4}F^{\mu\lambda}F_{\mu\lambda}-F^{\mu\lambda}\nabla_\mu A_\lambda),\\\text{with}&\qquad F_{\mu\lambda}=-F_{\lambda\mu}.\end{aligned}\right\}\tag{3}$$

Here u_λ and p_λ are four-vectors, the meaning of which will be found later. $A_\lambda{}^e$ is the value of the potential four-vector of the (external) electromagnetic field at the point x^μ of the electron.

We introduce three-dimensional notation by $F_{10}=F^{01}=\mathbf{E}_x$, $F_{12}=\mathbf{H}_z$, $A^0=-A_0=\Phi$, $p^0=-p_0=E/c$, etc., while $\mathbf{x}_e$, $\mathbf{p}$, and $\mathbf{A}$ denote the spatial parts of the vectors x^λ, p^λ, and A^λ, respectively. Then, by $d\tau=(1-\dot{\mathbf{x}}_e{}^2/c^2)^{\frac{1}{2}}dt$, and also putting

$$\mathbf{u}(1-\dot{\mathbf{x}}_e{}^2/c^2)^{\frac{1}{2}}=\mathbf{v},\quad u^0(1-\dot{\mathbf{x}}_e{}^2/c^2)^{\frac{1}{2}}=cv_t,\tag{4}$$

one can write (1) in the form $\delta\int\mathfrak{L}dt=0$ with

$$\left.\begin{aligned}\mathfrak{L}&=\Lambda\cdot(d\tau/dt)+\textstyle\int L_f=-mc^2(v_t{}^2-v^2/c^2)^{\frac{1}{2}}\\&\quad+\mathbf{p}\cdot\dot{\mathbf{x}}_e-E-\mathbf{p}\cdot\mathbf{v}+Ev_t+(e/c)\mathbf{v}\cdot\mathbf{A}^e\\&\quad-ev_t\Phi^e+(1/8\pi)\textstyle\int\{(\mathbf{H}^2-\mathbf{E}^2)\\&\quad-2(\mathbf{H}\cdot\mathrm{curl}\mathbf{A}+\mathbf{E}\cdot\nabla\Phi+\mathbf{E}\cdot\dot{\mathbf{A}}/c)\}\\&=-mc^2(v_t{}^2-v^2/c^2)^{\frac{1}{2}}+\mathbf{p}\cdot(\dot{\mathbf{x}}_e-\mathbf{v})\\&\quad+E(v_t-1)+(1/4\pi)\textstyle\int\{\tfrac{1}{2}\mathbf{H}^2\\&\quad-\mathbf{A}\cdot(\mathrm{curl}\mathbf{H}-4\pi\mathbf{i}/c)-\mathbf{E}\cdot\dot{\mathbf{A}}/c-\tfrac{1}{2}\mathbf{E}^2\\&\qquad+\Phi(\mathrm{div}\mathbf{E}-4\pi\rho)\}.\end{aligned}\right\}\tag{5}$$

Here, we put $\rho(\mathbf{x})=ev_t\delta^{(3)}(\mathbf{x}-\mathbf{x}_e)$ and $\mathbf{i}(\mathbf{x})=e\mathbf{v}\delta^{(3)}(\mathbf{x}-\mathbf{x}_e)$.

In this variational principle we shall now consider the variables $\mathbf{p}$, E, $\mathbf{x}_e$, $\mathbf{v}$, v_t, $\mathbf{A}(\mathbf{x})$, $\Phi(\mathbf{x})$, $\mathbf{H}(\mathbf{x})$, and $\mathbf{E}(\mathbf{x})$ as functions of t **to be varied independently** in $\delta\int\mathfrak{L}dt=0$. Thus, we get the following first-order equations:

$$\dot{\mathbf{x}}_e-\mathbf{v}=0;\quad v_t-1=0;\tag{6}$$

$$-dp_x/dt+(e/c)\mathbf{v}\cdot\partial\mathbf{A}^e/\partial x_e-ev_t\partial\Phi^e/\partial x_e=0;\tag{7}$$

$$m\mathbf{v}(v_t{}^2-v^2/c^2)^{-\frac{1}{2}}-\mathbf{p}+(e/c)\mathbf{A}^e=0;\tag{8a}$$

$$-mc^2v_t(v_t{}^2-v^2/c^2)^{-\frac{1}{2}}+E-e\Phi^e=0;\tag{8b}$$

and $(1/4\pi)$ times the equations

$$-\mathrm{curl}\mathbf{H}+\dot{\mathbf{E}}/c+4\pi\mathbf{i}/c=0;\quad \mathrm{div}\mathbf{E}-4\pi\rho=0;\tag{9}$$

$$\mathbf{H}-\mathrm{curl}\mathbf{A}=0\quad\text{and}\quad-(\mathbf{E}+\nabla\Phi+\dot{\mathbf{A}}/c)=0.\tag{10}$$

The first two equations give the meaning of the vector $\mathbf{v}$ and v_t, (thus indirectly the meaning of the four-vector u^λ which obviously is now $dx^\lambda/d\tau$). The third Eq. (7) is the equation of motion for the electron. The next two Eqs. (8a–b) determine $\mathbf{p}$ and E in the usual way as functions of the velocity $\mathbf{v}$ and the potentials Φ and $\mathbf{A}$. The last four equations, (9) and (10), give Maxwell's equations and the expressions of the field strengths in terms of the potentials. The equation of motion of the electron (7) can be written in the more usual form

$$\frac{d}{dt}\left\{\frac{m\mathbf{v}}{(1-v^2/c^2)^{\frac{1}{2}}}\right\}=e\mathbf{E}+\frac{e}{c}[\mathbf{v}\times\mathbf{H}],\tag{11}$$

if use is made of $\mathbf{v}\cdot\partial\mathbf{A}^e/\partial x_e=[\mathbf{v}\times\mathrm{curl}\mathbf{A}^e]_x+\mathbf{v}\cdot\nabla A_x{}^e$, of $dA_x{}^e/dt=\nabla A_x{}^e\cdot d\mathbf{x}_e/dt+\partial A_x{}^e/\partial t$ and of Eqs. (6), (8a), and (10). Thus we see that Maxwell's equations and the usual classical equations of motion for a point charge can be obtained from one "first-order" variational principle.

From the Lagrangian (5) we obtain the Hamiltonian in the usual way. The canonical conjugates of $\mathbf{x}_e$ and of $\mathbf{A}(\mathbf{x})$ are given by

$$\frac{\partial\mathfrak{L}}{\partial\dot{\mathbf{x}}_e}=\mathbf{p}\quad\text{and}\quad\frac{\delta\mathfrak{L}}{\delta\dot{\mathbf{A}}(\mathbf{x})}=-\frac{\mathbf{E}(\mathbf{x})}{4\pi c},\tag{12}$$

respectively. Hence

$$\mathcal{H}=\mathbf{p}\cdot\dot{\mathbf{x}}_e-\textstyle\int(\mathbf{E}\cdot\dot{\mathbf{A}}/4\pi c)-\mathfrak{L}.\tag{13}$$

The derivatives of variables with respect to time drop out automatically, and one obtains

$$\begin{aligned}\mathcal{H}&=mc^2(v_t{}^2-v^2/c^2)^{\frac{1}{2}}+\mathbf{p}\cdot\mathbf{v}+E(1-v_t)+\mathcal{W}\\&\quad+(1/4\pi)\textstyle\int\{\mathbf{A}\cdot\mathrm{curl}\mathbf{H}-\tfrac{1}{2}\mathbf{H}^2\\&\qquad+\tfrac{1}{2}\mathbf{E}^2-\Phi\mathrm{div}\mathbf{E}\},\end{aligned}\tag{14a}$$

with

$$\begin{aligned}\mathcal{W}&=ev_t\Phi(\mathbf{x}_e)-[e/c]\mathbf{A}(\mathbf{x}_e)\\&=e\textstyle\int\{\Phi v_t-\mathbf{A}\cdot\mathbf{v}/c\}\delta^{(3)}(\mathbf{x}-\mathbf{x}_e).\end{aligned}\tag{14b}$$

The canonical equations

$$\frac{dP}{dt}=-\frac{\delta\mathcal{H}}{\delta Q}\quad\text{(15a)}\qquad\text{and}\qquad\frac{dQ}{dt}=\frac{\delta\mathcal{H}}{\delta P}\quad\text{(15b)}$$

are now easily verified as far as their right-hand members have a meaning. That is, for $Q=\mathbf{x}_e$ or $=\mathbf{A}(\mathbf{x})$, and $P=\mathbf{p}$ or $=\{-\mathbf{E}(\mathbf{x})/4\pi c\}$, respectively, both (15a) and (15b) are valid. But, if we take for Q one of the quantities $\mathbf{v}$, v_t, E, $\Phi(\mathbf{x})$, or $\mathbf{H}(\mathbf{x})$, the Eqs. (15b) are obviously meaningless, as the canonical conjugates P of these variables are zero. The Eqs. (15a) remain valid; they give the so-called "identities"[2]

$$\mathbf{p}=m\mathbf{v}(v_t^2-v^2/c^2)^{-\frac{1}{2}}+[e/c]\mathbf{A}^e,$$
$$E=mc^2v_t(v_t^2-v^2/c^2)^{-\frac{1}{2}}+e\Phi^e,$$
$$v_t=1,\quad \mathrm{div}\mathbf{E}=4\pi\rho,\quad \text{and}\quad \mathbf{H}=\mathrm{curl}\mathbf{A}. \tag{16}$$

They do not contain any differentiation with respect to time and, when the equations of motion (15a–b) are to be integrated say from $t=0$, the initial conditions for the variables have to be such that Eqs. (16) are satisfied at $t=0$.

The Hamiltonian (14) can be simplified by elimination of some of the redundant ("derived") variables that have no canonical conjugate. For instance, by substitution of $v_t\equiv 1$ and $\mathbf{H}\equiv\mathrm{curl}\mathbf{A}$ one gets

$$\begin{aligned}\mathcal{H}=mc^2(1-v^2/c^2)^{\frac{1}{2}}+\mathbf{p}\cdot\mathbf{v}&\\ +(1/8\pi)\int\{\mathbf{E}^2+(\mathrm{curl}\mathbf{A})^2\}&\\ -(e/c)\int(\mathbf{A}\cdot\mathbf{v})\delta^{(3)}(\mathbf{x}-\mathbf{x}_e)&\\ +\int\Phi\{e\delta^{(3)}(\mathbf{x}-\mathbf{x}_e)-\mathrm{div}\mathbf{E}/4\pi\},&\end{aligned} \tag{17}$$

which gives again canonical Eqs. (15a) for $Q=$ one of the variables $\mathbf{x}_e$, $\mathbf{A}(\mathbf{x})$, $\mathbf{v}$ or $\Phi(\mathbf{x})$, while (15b) is valid again for $\mathbf{x}_e$ and $\mathbf{A}(\mathbf{x})$ only. Expression (17) is remarkable for its resemblance to the Hamiltonian of Dirac's relativistic wave-mechanical theory of the electron. There, the Dirac matrix β simply replaces $(1-v^2/c^2)^{\frac{1}{2}}$ in the first term of (17), while $c\alpha$ replaces $\mathbf{v}$ in the other terms. This is in complete accordance with the expectation values, which these Dirac matrices have in a situation in which the electron possesses a definite momentum.

A corresponding simplification can be made in the Lagrangian:

$$\begin{aligned}\mathcal{L}=-mc^2(1-v^2/c^2)^{\frac{1}{2}}+\mathbf{p}\cdot(\dot{\mathbf{x}}_e-\mathbf{v})&\\ +e\int\{(\mathbf{A}\cdot\mathbf{v}/c)-\Phi\}\delta^{(3)}(\mathbf{x}-\mathbf{x}_e)&\\ -(1/4\pi)\int\{\tfrac{1}{2}(\mathrm{curl}\mathbf{A})^2+\mathbf{E}\cdot\dot{\mathbf{A}}/c&\\ +\tfrac{1}{2}\mathbf{E}^2+\mathbf{E}\cdot\nabla\Phi\},&\end{aligned} \tag{18}$$

[2] The variables $\mathbf{v}$, v_t, E, and $\mathbf{H}(\mathbf{x})$ can be regarded as so-called "derived variables." Compare F. J. Belinfante, Physica **7**, 765 (1940).

where E, v_t and $\mathbf{H}(\mathbf{x})$ have been eliminated, so that only $\mathbf{x}_e$, $\mathbf{p}$, $\mathbf{v}$, $\mathbf{A}(\mathbf{x})$, $\mathbf{E}(\mathbf{x})$, and $\Phi(\mathbf{x})$ are to be varied independently. Equations (6)–(10) following from (18) can obviously be made relativistically covariant by adding to them the missing Eqs. (6b), (8b), and (10a) as definitions for v_t, E, and $\mathbf{H}(\mathbf{x})$. The expression (18) may have the advantage of being simpler than (5), but, on the other hand, this Lagrangian (18) cannot be written in a covariant form as (5) was in Eqs. (1)–(3), without re-introducing the quantities eliminated.

Further elimination of redundant variables would call for substitution of

$$\mathbf{v}\equiv\frac{c\mathbf{p}-e\mathbf{A}^e}{(m^2c^2+(\mathbf{p}-[e/c]\mathbf{A}^e)^2)^{\frac{1}{2}}}.$$

This would make the Hamiltonian a rather complicated expression, and does not lead to any simplification.

Finally, the variable Φ cannot be eliminated in a similar way, as among the identities (16) there is no equation that gives $\Phi(\mathbf{x})$ in terms of the other variables. In quantum electrodynamics, this fact necessitates the introduction of a new variable S, canonically conjugate to Φ. This quantity S has no physical meaning, of course. It is, therefore, usually stated that the "situation functions" χ, that describe situations which can occur in nature, must satisfy the additional condition $S\chi=0$. This, however, makes it impossible to normalize χ in the usual way.[3] It is, therefore, then necessary to revise the definition of the method of normalization of a situation function and of calculation of matrix elements of observables in general. It has been shown that, when this is done, then the condition $S\chi=0$ even becomes superfluous.[3]

[3] F. J. Belinfante, Physica **12**, 17 (1946).