First-order Lagrangians and the limit of normal Lagrangians

John Collins*
Physics Department, Penn State University, 104 Davey Laboratory, University Park PA 16802, U.S.A.
(Dated: 21 January 2021)

The rules for quantizing Lagrangians first order in time derivatives differ from the standard textbook rules for Lagrangians quadratic in time derivatives. In particular, the coordinates do not commute with each other, contrary to the standard case. This leads to a paradox if a first-order Lagrangian is obtained from a second-order Lagrangian in the limit of vanishing coefficients for the terms quadratic in time derivatives. I show by means of a simple pedagogical example that the paradox is resolved by removing modes whose frequency goes to infinity in the limit. This can be implemented by a suitable averaging over time for the coordinate variables.

PACS numbers:

I. INTRODUCTION

There are well-known rules for quantizing Lagrangians, such as those for non-relativistic particle mechanics, which are quadratic in velocities. These rules were discovered in the earliest days of real quantum mechanics [1–4]. However, Lagrangians that are first order in time derivatives of coordinates and fields also occur in physics, notably for the Schrödinger and Dirac fields (e.g., [6, 7]), and the quantization rules are different. A clear and simple formulation was given by Floreanini, Faddeev, and Jackiw [8, 9].

Now a first-order Lagrangian can be obtained from a suitable second-order Lagrangian in limit of vanishing coefficients for the terms quadratic in time derivatives. However, the rules for equal-time commutation relations (ETCR) for the first-order system differ from the limit those for the second-order system. Notably, the generalized coordinates in the second-order case have exactly vanishing ETCR with each other, whereas these same commutators are non-vanishing in the first-order case.

This paper shows how to resolve the paradox.

II. QUANTIZATION RULES

A. Why the rules are what they are: Specification of a quantum mechanical system

Why should we apply particular rules for quantization to a system specified by some given Lagrangian? Why not use some other rules? How are modifications to the rules to be determined in cases like a first-order Lagrangian? When one studies quantum field theory, it can often *appear* that the rules are being made up as one goes along. It is therefore important to understand to what extent there are overall principles at work that do not change and are applicable to all cases, to understand to what extent new kinds of theory are being constructed, and to find the justifications.

In fact, the following principles invariantly cover all the cases studied. We assume throughout that we use the Heisenberg picture. Then

1. The system is specified by a set of basic variables ("generalized coordinates") and a formula for the Lagrangian or Lagrangian density, or equivalently for the action.

2. The equations of motion are the Euler-Lagrange equations.

3. The equations of motion are also given by the Heisenberg equation

$$i\hbar \frac{dA(t)}{dt} = [A, H], \tag{1}$$

for the basic variables. Here the Hamiltonian H is the Noether charge for time-translation invariance.

These principles cover all the elementary systems considered in the founding of quantum mechanics, and they apply equally to the fundamental microscopic laws of all systems, e.g., quantum field theories. But it should be noted that in condensed matter systems, particularly, it is also common to specify a system directly by a formula for the Hamiltonian.

Originally the procedures embodied in the above rules arose [1–4] as a method that would give a quantum mechanical system having a given classical limit. The key insight was Heisenberg's proposal [1] that formulas should remain unchanged between the classical and quantum systems; what should change is the mathematical nature of the quantities represented by symbols for positions of particles, for fields, etc. An analysis of the Fourier components of radiation from systems led Heisenberg to propose that multiplication of these quantities is matrix multiplication and that they should obey Eq. (1). Not all of these ideas were fully explicit in [1], but a complete formalization was given almost immediately in the work of Born, Jordan and Heisenberg [2, 3], and of Dirac [4].

In this approach, the ETCR do not arise as completely independent postulates. Rather they are highly

* collins@phys.psu.edu

constrained to whatever gives consistency between the two forms of equation of motion: Euler-Lagrange and Heisenberg. Indeed for the simplest case of a single degree of freedom, with $L = m\dot{q}^2/2 - V(q)$, the ETCR $[q(t), p(t)] = i\hbar$ arises as a *theorem* [2], where $p = \partial L/\partial\dot{q}$. In this case, the standard ETCR gives the only way of gaining consistency between the two forms of equation of motion.

For more complicated systems, the standard ETCR, (4) below, are unique, as far as I know.

Of course, given the above considerations, further calculations and the quantum-theoretical solution of a theory can all be done purely and completely within the Hamiltonian formulation, taking the following as the starting axioms: (a) The formula for the Hamiltonian. (b) The ETCR. (c) The Heisenberg equation of motion. These are fully self consistent.

The above considerations were appropriate to the initial development of quantum theory, which effectively amounted to reverse-engineering a quantum system from its classical limit. For a free-standing quantum theory, it is obviously inappropriate to base the formulation on a correspondence with a classical limit. If nothing else, one needs the freedom to modify the specification of a quantum system to better agree with reality. We note a couple of well-known examples of such modifications: the introduction of electron spin, and the replacement of commutators by anticommutators for theories of fermionic fields.

In fact the above three postulates, which give primacy to the Lagrangian formulation, do have great significance within quantum theory, in the functional integral method (e.g., [7]). The integrand in this method includes a factor of e^{iS}, where S is the theory's action functional. The action therefore is the fundamental quantitative specification of the theory. This formulation implies the Euler-Lagrange equations for the theory's (operator-valued) variables (generalized coordinates in the terminology of Lagrangian mechanics). The functional integral formulation is a compact formulation particularly appropriate to quantum fields, which at the microscopic level form the most fundamental formulation of realistic quantum theories.

It is useful to ask what the corresponding operator formulation is. The Heisenberg equation of motion arises as giving the relation between the implementation of a symmetry on states and on operators. The ETCR can also be derived from the functional integral, as well as the Euler-Lagrange equations.

In this view, the quantization rules give a method of formulating a quantum mechanical theory purely in terms of properties of the theory's operators, given that the quantum theory has been specified by its action functional, i.e., by its Lagrangian or Lagrangian density.

B. Second-order case

In this section we review the standard quantization rules. We suppose that the theory under consideration is specified by a set of "generalized coordinates" $\underline{q} = (q_1, q_2, \ldots)$ and a Lagrangian $L(\underline{\dot{q}}, \underline{q})$. Here we use an underline to represent the whole array of coordinates. Although certain generalizations are possible, we restrict in this section to the case that L is of the form of a quadratic kinetic-energy-like term minus a potential-energy-like term

$$L = K(\underline{\dot{q}}) - V(\underline{q}). \tag{2}$$

The Hamiltonian is the Noether charge for time-translation symmetry, and is given by

$$H = \sum_j \dot{q}_j \frac{\partial L}{\partial \dot{q}_j} - L, \tag{3}$$

which we usually express in terms of the q_js and the canonical momenta $p^j \equiv \partial L/\partial\dot{q}_j$.

The solution is specified by the canonical ETCR [1–4]

$$[q_j(t), q_k(t)] = [p^j(t), p^k(t)] = 0, \quad [q_j(t), p^k(t)] = i\hbar\delta^k_j, \tag{4}$$

together with the Heisenberg equations of motion (1) for the time-dependence of the operators. Self-consistency conditions follow as theorems: Given that the ETCR hold at one time $t = t_0$, it is implied by the equations of motion that they hold at all other times as well. Given that (1) applies to the basic variables q_j and p^j, it also applies when the operator A is given as a function of the basic variables that is in the form of an algebraic formula with time-independent coefficients.

It is readily verified that the Euler-Lagrange equations follow from the Heisenberg equations, once the ETCR are of the given form.

C. Floreanini, Faddeev, and Jackiw rules for first-order case

We now consider a system specified by the following first-order Lagrangian

$$L = \frac{1}{2}\sum_{j,k} q^j f_{jk} \dot{q}^k - F(\underline{q}). \tag{5}$$

Here f is an invertible *anti*symmetric[12] numerical matrix, $f_{jk} = -f_{kj}$, and $F(\underline{q})$ is some function of the coordinates. Thus F appears to be like the potential in the usual second-order Lagrangians of classical mechanics etc.

a. Euler-Lagrange equations and Hamiltonian There is *exactly* no change from the usual case either in the form of the Euler-Lagrange equations,

$$0 = \frac{d}{dt}\frac{\partial L}{\partial \dot{q}^j} - \frac{\partial L}{\partial q^j} = -\sum_k f_{jk}\dot{q}_k + \frac{\partial F}{\partial q^j}, \tag{6}$$

or in the definition of the Hamiltonian,

$$H(\underline{q}) \equiv \sum_j \frac{\partial L}{\partial \dot{q}^j}\dot{q}^j - L = F(\underline{q}). \tag{7}$$

Thus H is the Noether charge for time-translation invariance.

b. Equal-time commutation relations The canonical momenta, with the usual definition, are not independent of the coordinates, which necessitates some change in the commutation relations. One possibility is to use constrained quantization [10, 11]. But this is quite an elaborate formulation.

Here, instead, we use the observation that the equal-time commutation relations (ETCR) between the q^js are to be such that the ordinary Heisenberg equations of motion in the quantum theory,

$$i\hbar\frac{dq^j(t)}{dt} = [q^j(t), H], \tag{8}$$

agree with the Euler-Lagrange equations Eq. (6). We do this by imposing [8, 9]

$$[q^j(t), q^k(t)] = i\hbar(f^{-1})^{jk}. \tag{9}$$

Here f^{-1} is defined to obey $\sum_k (f^{-1})^{jk} f_{kl} = \delta^j_l$. These commutation relations comprise the only changes compared with the usual case. It is readily verified that the Euler-Lagrange equations of motion and the Heisenberg equations of motion do in fact agree, up to operator ordering issues.

III. MODEL AND ITS LIMIT

A. First-order version

A minimal first-order system has 2 coordinates (Q_1, Q_2) and

$$L = \frac{1}{2}\left(Q_2\dot{Q}_1 - Q_1\dot{Q}_2\right) - \frac{\omega}{2}\left(Q_1^2 + Q_2^2\right). \tag{10}$$

We use upper-case letters for the coordinates to distinguish them from those in a second-order Lagrangian that gives Eq. (10) as a limit. The Hamiltonian is

$$H = \frac{\omega}{2}\left(Q_1^2 + Q_2^2\right), \tag{11}$$

and the Euler-Lagrange equations are

$$0 = \frac{d}{dt}\frac{\partial L}{\partial \dot{Q}_1} - \frac{\partial L}{\partial Q_1} = \dot{Q}_2 + \omega Q_1, \tag{12}$$

$$0 = \frac{d}{dt}\frac{\partial L}{\partial \dot{Q}_2} - \frac{\partial L}{\partial Q_2} = -\dot{Q}_1 + \omega Q_2. \tag{13}$$

Their most general solution is that of a simple harmonic oscillator of angular frequency ω:

$$Q_1 = \sqrt{\frac{\hbar}{2}}\left(ae^{-i\omega t} + a^\dagger e^{i\omega t}\right), \tag{14}$$

$$Q_2 = \sqrt{\frac{\hbar}{2}}\left(-iae^{-i\omega t} + ia^\dagger e^{i\omega t}\right), \tag{15}$$

where the normalization factor $\sqrt{\hbar/2}$ was chosen for later convenience. The Floreanini-Faddeev-Jackiw rules give

$$[Q_1(t), Q_2(t)] = i\hbar, \tag{16}$$

and this is the unique ETCR that reproduces the Heisenberg equations for this system. It follows that the a and $a^\dagger$ operators obey

$$[a, a^\dagger] = 1, \tag{17}$$

and that the Hamiltonian is $H = \hbar\omega(a^\dagger a + 1/2)$, i.e., that the system is the conventional quantum-mechanical simple harmonic oscillator.

B. Complex coordinates

To make the connection with the Schrödinger field theory, we can rewrite the system with a complex variable $\psi = (Q_1 + iQ_2)/\sqrt{2\hbar}$, to give

$$L = \frac{i\hbar}{2}\left(\psi^\dagger\dot{\psi} - \dot{\psi}^\dagger\psi\right) - \hbar\omega\psi^\dagger\psi, \tag{18}$$

up to operator-ordering. This is just the Lagrangian for a Schrödinger quantum *field* [6, 7] with an external potential $\hbar\omega$ but restricted to a zero-dimensional space, i.e., a situation where space consists of one point.

C. Second-order system

Next we consider a system differing from (10) by a small second-order term,

$$L_2 = \frac{\epsilon(1+\epsilon)}{2\omega}\left(\dot{q}_1^2 + \dot{q}_2^2\right) + \frac{1}{2}\left(\dot{q}_1 q_2 - \dot{q}_2 q_1\right) - \frac{\omega}{2}\left(q_1^2 + q_2^2\right), \tag{19}$$

where we will take the parameter ϵ to zero to obtain (10). The presence of $1/\omega$ in the coefficient of the second-order term is dictated by dimensional analysis, and we have chosen to write the numerator of the coefficient as $\epsilon(1+\epsilon)$ rather than ϵ so as to simplify the parameterization of the results.

Standard canonical quantization applies to this new Lagrangian, which is actually that of coupled harmonic oscillators in unusual coordinates. The canonical momenta are

$$p^1 = \frac{\epsilon(1+\epsilon)}{\omega}\dot{q}_1 + \frac{1}{2}q_2, \quad p^2 = \frac{\epsilon(1+\epsilon)}{\omega}\dot{q}_2 - \frac{1}{2}q_1, \tag{20}$$

obeying the standard ETCR (4). In this case $q_1(t)$ and $q_2(t)$ unambiguously commute with each other. If we take the limit $\epsilon \to 0$ and use Q_j to denote the limiting values of q_j, then these standard ETCR are evidently incompatible with the nonzero commutator (16) for $Q_1(t)$ and $Q_2(t)$.

D. Resolution of paradox

To understand what is happening, we solve the equations of motion for the second order system by decomposing in normal modes:

$$q_1(t) = \sqrt{\frac{\hbar}{2(1+2\epsilon)}}\Big[a_+ e^{-i\omega t/\epsilon} + a_- e^{-i\omega t/(1+\epsilon)} + a_+^\dagger e^{i\omega t/\epsilon} + a_-^\dagger e^{i\omega t/(1+\epsilon)}\Big] \quad (21)$$

$$q_2(t) = \sqrt{\frac{\hbar}{2(1+2\epsilon)}}\Big[ia_+ e^{-i\omega t/\epsilon} - ia_- e^{-i\omega t/(1+\epsilon)} - ia_+^\dagger e^{i\omega t/\epsilon} + ia_-^\dagger e^{i\omega t/(1+\epsilon)}\Big] \quad (22)$$

It can be verified explicitly that the ETCR for the q_j and p^j are equivalent to

$$[a_\alpha, a_\beta^\dagger] = \delta_{\alpha\beta}, \qquad [a_\alpha, a_\beta] = 0, \quad (23)$$

where the indices take on the values + and −. Thus the operators $a_\pm$ and $a_\pm^\dagger$ are raising and lowering for two independent oscillators, of angular frequencies ω/ϵ and $\omega/(1+\epsilon)$.

In the limit $\epsilon \to 0$, the a_- part of the solution for q_j agrees exactly with the solution (14) and (15) for Q_j from the first-order Lagrangian. But there also remains the a_+ mode whose frequency goes to infinity in the limit. The zero commutator of $q_1(t)$ with $q_2(t)$ is obtained by equal and opposite nonzero terms for each mode.

Evidently, the correct way of taking the limit, to get the first-order system, is to remove the infinite frequency modes. This can be done, for example, by a suitable average over time:

$$Q_j(t) = \lim_{\epsilon\to 0} \int q_j(t')\, f(t'-t;\Delta t)\, dt'. \quad (24)$$

Here f is a function that performs an average over a range of time of size Δt, e.g.,

$$f(t-t';\Delta t) = \frac{e^{-(t-t')^2/\Delta t^2}}{\Delta t\sqrt{\pi}}. \quad (25)$$

The width of the averaging should go to zero with ϵ, but less rapidly than ϵ, for example $\Delta t = \sqrt{\epsilon}/\omega$, so that in the limit infinitely many oscillations of the high-frequency mode are averaged over.

By removing the contribution of the high-frequency mode, the averaging also removes its contribution to the commutator of $Q_1(t)$ with $Q_2(t)$. This results in exactly the nonzero commutator (16) for the first-order model. The raising and lowering operators a and $a^\dagger$ of the first-order model are evidently to be identified with those for the low frequency mode of the second-order model.

ACKNOWLEDGMENTS

This work was supported in part by the U.S. Department of Energy under grant number DE-FG02-90ER-40577.

[1] W. Heisenberg, "Quantum-theoretical Re-interpretation of Kinematic and Mechanical Relations", Zeit. f. Phys. **33** 879–893 (1925). English translation in [5].

[2] M. Born and P. Jordan, "On Quantum Mechanics", Zeit. f. Phys. **34**, 858–888 (1925). English translation in [5], but with the part on quantum fields omitted.

[3] M. Born, W. Heisenberg and P. Jordan, "On Quantum Mechanics II", Zeit. f. Phys. **35**, 557–615 (1926). English translation in [5].

[4] P.A.M. Dirac, "The Fundamental Equations of Quantum Mechanics", Proc. Roy. Soc. A **109**, 642–653 (1926).

[5] B.L. van der Waerden, "Sources of Quantum Mechanics" (Dover, 1968).

[6] L.S. Brown "Quantum Field Theory" (Cambridge University Press, 1992).

[7] M. Srednicki, "Quantum Field Theory" (Cambridge University Press, 2007).

[8] R. Floreanini and R. Jackiw, "Selfdual Fields As Charge Density Solitons", Phys. Rev. Lett. **59**, 1873 (1987).

[9] L.D. Faddeev and R. Jackiw, "Hamiltonian Reduction of Unconstrained and Constrained Systems", Phys. Rev. Lett. **60**, 1692 (1988).

[10] P.A.M. Dirac "Lectures on Quantum Mechanics" (Yeshiva University, New York, 1964).

[11] M.E.V. Costa and H.O. Girotti, "Comment on 'Selfdual Fields as Charge Density Solitons'", Phys. Rev. Lett. **60**, 1771 (1988).

[12] Note that by adding a time derivative, other forms may be obtained, e.g., $q^1\dot{q}^2$ instead of $\frac{1}{2}(q^1\dot{q}^2 - q^2\dot{q}^1)$. Such changes do not affect the equations of motion or the Hamiltonian. But the form (5) is the most symmetric, so we will take it as the canonical one.