Modified Hamiltonian formalism for higher-derivative theories

K. Andrzejewski, J. Gonera, P. Machalski, and P. Maślanka*

Department of Theoretical Physics II, University of Łódź, Pomorska 149/153, 90–236 Łódź, Poland

(Received 17 June 2010; published 6 August 2010)

An alternative version of Hamiltonian formalism for higher-derivative theories is proposed. As compared with the standard Ostrogradski approach, it has the following advantages: (i) The Lagrangian, when expressed in terms of new variables, yields proper equations of motion; no additional Lagrange multipliers are necessary. (ii) The Legendre transformation can be performed in a straightforward way, provided the Lagrangian is nonsingular in the Ostrogradski sense. The generalizations to singular Lagrangians as well as field theory are presented.

DOI: 10.1103/PhysRevD.82.045008 PACS numbers: 11.10.Lm, 04.60.Kz

I. INTRODUCTION

It is a long-standing problem whether and why it is sufficient to use in physics the Lagrangians containing only first-order time derivatives. It is more intriguing that adding higher derivatives may improve our models in some respects, like ultraviolet behavior [1,2] (in particular, making modified gravity renormalizable [3] or even asymptotically free [4]); also, higher-derivative Lagrangians appear to be a useful tool to describe some interesting models, like relativistic particles with rigidity, curvature, and torsion [5]. Moreover, almost any effective theory obtained by integrating out some degrees of freedom (usually, but not always, those related to high-energy excitations) of the underlying "microscopical" theory contains higher derivatives. One can argue that the effective theory, being an approximation to perfectly consistent quantum theory, need not be considered and quantized separately. However, we are never sure if our theory is the basic or the effective one; therefore, it is important to know whether it is at all possible to quantize the effective theory in a way which would correctly reproduce some aspects of the microscopic one.

The first step toward the quantum theory is to put its classical counterpart in Hamiltonian form. The standard framework for dealing with higher-derivative theories on the Hamiltonian level is provided by the Ostrogradski formalism [6–10]. The main disadvantage of this approach is that the Hamiltonian, being a linear function of some momenta, is necessarily unbounded from below. In general, this cannot be cured by trying to devise an alternative canonical formalism. In fact, any Hamiltonian is an integral of motion, while it is by far not obvious that a generic system described by higher-derivative Lagrangians possesses globally defined integrals of motion, except the one related to time translation invariance. Moreover, the instability of the Ostrogradski Hamiltonian is not related to finite domains in phase space, which implies that it will survive the standard quantization procedure (i.e., it cannot be cured by the uncertainty principle).

The Ostrogradski approach also has some other disadvantages. There is no straightforward transition from the Lagrangian to the Hamiltonian formalism. In fact, the Ostrogradski approach is based on the idea that the consecutive time derivatives of initial coordinate(s) form new coordinates $q_i \sim q^{(i-1)}$. It appears then that the Lagrangian cannot be viewed as a function on the tangent bundle to coordinate manifold because it leads to incorrect equations of motion. Also, the Legendre transformation to the cotangent bundle (phase space) cannot be performed. One deals with this problem by adding Lagrange multipliers, enforcing the proper relation between new coordinates and time derivatives of the original ones. This results in further enlarging the coordinate manifold; moreover, the theory becomes constrained [in spite of the fact that the initial theory may be nonsingular in the Ostrogradski sense, cf. Eq. (2.2) below], and the Hamiltonian formalism is obtained by applying Dirac constraint theory, i.e., by reduction of the cotangent bundle to a submanifold endowed with symplectic structure defined by Dirac brackets.

In the present paper an alternative approach is proposed. It leads directly to the Lagrangian which, being a function on the tangent manifold, gives correct equations of motion; no new coordinate variables need to be added. Furthermore, for Lagrangians that are nonsingular in the Ostrogradski sense, the Legendre transformation takes the standard form. Our approach is also applicable to the most interesting case of singular Lagrangians [for example, those defining $f(R)$ gravity [11]].

The paper is organized as follows. In Sec. II we consider nonsingular Lagrangians containing second- and third-order time derivatives. Constrained theories are discussed in Sec. III. The general formalism is applied to the minisuperspace formulation of $f(R)$ gravity [12] in Sec. IV. In Sec. V, the modifications necessary to cover the field-theoretic case are given. In the Appendix we describe (for one degree of freedom) the generalization of our formalism to Lagrangians containing arbitrarily high derivatives.

*pmaslan@uni.lodz.pl

II. NONSINGULAR LAGRANGIANS OF SECOND AND THIRD ORDER

In this section we consider the Lagrangians containing second and third time derivatives which are nonsingular in the Ostrogradski sense. The Ostrogradski approach is based on the idea that the consecutive time derivatives of the initial coordinate form new coordinates, $q_i \sim q^{(i-1)}$. However, it has been suggested [13–17] that one can use every second derivative as a new variable, $q_i \sim q^{(2i-2)}$. We generalize this idea by introducing new coordinates as some functions of the initial ones and their time derivatives. Our paper is inspired by the results obtained in Ref. [14].

A. The case of second derivatives

Let us start with Lagrangians containing time derivatives up to the second order,

$$L = L(q, \dot{q}, \ddot{q}); \tag{2.1}$$

here $q = (q^\mu)$, $\mu = 1, \ldots, N$, denotes the set of generalized coordinates. The nonsingularity condition of the Ostrogradski approach reads

$$\det\left(\frac{\partial^2 L}{\partial \ddot{q}^\mu \partial \ddot{q}^\nu}\right) \neq 0. \tag{2.2}$$

In order to put our theory in the first-order form, we define new coordinates q_1^μ, q_2^ν:

$$q^\mu = q_1^\mu, \qquad \dot{q}^\mu = \dot{q}_1^\mu, \qquad \ddot{q}^\mu = \chi^\mu(q_1, \dot{q}_1, q_2), \tag{2.3}$$

where χ^μ are the functions specified below.

We select an arbitrary function

$$F = F(q_1, \dot{q}_1, q_2), \tag{2.4}$$

subjected to the single condition

$$\det\left(\frac{\partial^2 F}{\partial \dot{q}_1^\mu \partial q_2^\nu}\right) \neq 0. \tag{2.5}$$

Now, χ^μ are defined as a unique [at least locally due to (2.2)] solution to the following set of equations:

$$\frac{\partial L(q_1, \dot{q}_1, \chi)}{\partial \chi^\mu} + \frac{\partial F(q_1, \dot{q}_1, q_2)}{\partial \dot{q}_1^\mu} = 0. \tag{2.6}$$

The new Lagrangian, which is now a standard Lagrangian of first order, is given by

$$\begin{aligned} \mathcal{L}(q_1, \dot{q}_1, q_2, \dot{q}_2) = {} & L(q_1, \dot{q}_1, \chi(q_1, \dot{q}_1, q_2)) \\ & + \frac{\partial F(q_1, \dot{q}_1, q_2)}{\partial q_1^\mu}\dot{q}_1^\mu \\ & + \frac{\partial F(q_1, \dot{q}_1, q_2)}{\partial q_2^\mu}\dot{q}_2^\mu \\ & + \frac{\partial F(q_1, \dot{q}_1, q_2)}{\partial \dot{q}_1^\mu}\chi^\mu(q_1, \dot{q}_1, q_2). \end{aligned} \tag{2.7}$$

It differs from the initial one by an expression which becomes a total time derivative "on shell."

The equation of motion for q_2^μ yields

$$\frac{\partial^2 F}{\partial \dot{q}_1^\nu \partial q_2^\mu}(\chi^\nu - \ddot{q}_1^\nu) = 0, \tag{2.8}$$

which, by virtue of (2.5), implies

$$\ddot{q}^\mu = \chi^\mu(q_1, \dot{q}_1, q_2). \tag{2.9}$$

For the remaining variables q_1^μ one obtains

$$\frac{\partial L}{\partial q_1^\mu} - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_1^\mu}\right) + \frac{d^2}{dt^2}\left(\frac{\partial L}{\partial \chi^\mu}\right) = 0, \tag{2.10}$$

and taking into account (2.9), one gets the initial Euler-Lagrange equations.

It is worth noting that, contrary to the original Ostrogradski approach, the formalism presented above leads directly to the standard picture of the Lagrangian as a function defined on the tangent bundle to coordinate space (with no need for enlarging the latter by adding the appropriate Lagrange multipliers).

Our Lagrangian (2.7) is nonsingular in the usual sense, so one can directly pass to the Hamiltonian picture by performing Legendre transformation, leading to canonical dynamics on the cotangent bundle.

To this end, we define the canonical momenta

$$\begin{aligned} p_{1\mu} \equiv {} & \frac{\partial \mathcal{L}}{\partial \dot{q}_1^\mu} \\ = {} & \frac{\partial L}{\partial \dot{q}_1^\mu} + \frac{\partial^2 F}{\partial q_1^\nu \partial \dot{q}_1^\mu}\dot{q}_1^\nu + \frac{\partial^2 F}{\partial \dot{q}_1^\mu \partial \dot{q}_1^\nu}\chi^\nu + \frac{\partial^2 F}{\partial \dot{q}_1^\mu \partial q_2^\nu}\dot{q}_2^\nu \\ & + \frac{\partial F}{\partial q_1^\mu}, \end{aligned} \tag{2.11}$$

$$p_{2\mu} \equiv \frac{\partial \mathcal{L}}{\partial \dot{q}_2^\mu} = \frac{\partial F(q_1, \dot{q}_1, q_2)}{\partial q_2^\mu}. \tag{2.12}$$

By virtue of (2.5) the second set of equations can be uniquely solved (at least locally) for $\dot{q}_1^\mu$,

$$\dot{q}_1^{\mu} = \dot{q}_1^\mu(q_1, q_2, p_2). \tag{2.13}$$

As for the first set (2.11), we note that $\dot{q}_2^\mu$ appears (linearly) only in the fourth term on the right-hand side (RHS).

Again, the same condition (2.5) allows us to solve (2.11) for $\dot{q}_2^\mu$,

$$\dot{q}_2^{\mu} = q_2^\mu(q_1, q_2, p_1, p_2). \tag{2.14}$$

The Hamiltonian H is computed in the standard way, and the final result reads

$$H = p_{1\mu}\dot{q}_1^\mu - L - \frac{\partial F}{\partial q_1^\mu}\dot{q}_1^\mu - \frac{\partial F}{\partial \dot{q}_1^\mu}\chi^\mu, \tag{2.15}$$

where everything is expressed in terms of q_1, q_2, p_1, and p_2. We have checked, by direct calculation, that the canonical equations following from H are equivalent to the initial Lagrangian ones.

There exists a canonical transformation which relates our Hamiltonian to the Ostrogradski one. It reads

$$\begin{aligned} \tilde{q}_1^\mu &= q_1^\mu, \qquad \tilde{q}_2^\mu = f^\mu(q_1, q_2, p_2),\\ \tilde{p}_{1\mu} &= p_{1\mu} - \frac{\partial F}{\partial q_1^\mu}(q_1, f, q_2),\\ \tilde{p}_{2\mu} &= -\frac{\partial F}{\partial f^\mu}(q_1, f, q_2), \end{aligned} \tag{2.16}$$

where tildes refer to Ostrogradski variables and $f^\mu(q_1, q_2, p_2)$ solve Eq. (2.12), i.e., $f^\mu = \dot{q}_1^\mu(q_1, q_2, p_2)$. The corresponding generating function Φ has the form

$$\Phi(q_1, \tilde{p}_1, q_2, \tilde{q}_2) = q_1^\mu \tilde{p}_{1\mu} + F(q_1, \tilde{q}_2, q_2). \tag{2.17}$$

However, it should be stressed that the Ostrogradski Hamiltonian is singular in the sense that the inverse Legendre transformation cannot be performed (contrary to our case). This means that the structure of the symplectic manifold (phase space) as a cotangent bundle to the coordinate manifold is not transparent if Ostrogradski variables are used.

Let us conclude this part with a very simple example. The Lagrangian

$$L = \lambda\epsilon_{\mu\nu}\dot{q}^\mu\ddot{q}^\nu + \frac{\beta}{2}(\ddot{q}^\nu)^2, \qquad \beta \neq 0, \qquad \mu, \nu = 1, 2 \tag{2.18}$$

is nonsingular in the Ostrogradski sense, provided $\beta \neq 0$. We take

$$F = \alpha\dot{q}_1^\mu q_2^\mu, \qquad \alpha \neq 0. \tag{2.19}$$

Then

$$\chi^\mu = \frac{\lambda}{\beta}\epsilon_{\mu\nu}\dot{q}_1^\nu - \frac{\alpha}{\beta}q_2^\mu, \tag{2.20}$$

and

$$\mathcal{L} = -\frac{\alpha^2}{2\beta}(q_2^\mu)^2 - \frac{\lambda^2}{2\beta}(\dot{q}_1^\mu)^2 - \frac{\alpha\lambda}{\beta}\epsilon_{\mu\nu}\dot{q}_1^\mu q_2^\nu + \alpha\dot{q}_1^\mu\dot{q}_2^\mu. \tag{2.21}$$

Finally, the Hamiltonian reads

$$H = \frac{1}{\alpha}p_{1\mu}p_{2\mu} + \frac{\lambda^2}{2\alpha^2\beta}(p_{2\mu})^2 + \frac{\lambda}{\beta}\epsilon_{\mu\nu}p_{2\mu}q_2^\nu + \frac{\alpha^2}{2\beta}(q_2^\mu)^2. \tag{2.22}$$

It depends on an arbitrary parameter α. One can ask whether any relevant physical quantity may depend on α. The answer is no: All physical quantities are α independent. Formally this can be shown using Eqs. (2.16) and (2.17). Indeed, the function generating the canonical transformation to Ostrogradski variables reads

$$\Phi(q_1^\mu, \tilde{p}_{1\mu}, q_2^\mu, \tilde{q}_2^\mu) = q_1^\mu\tilde{p}_{1\mu} + \alpha\tilde{q}_2^\mu q_2^\mu. \tag{2.23}$$

The corresponding canonical transformation takes the form

$$\begin{aligned} p_{1\mu} &= \tilde{p}_{1\mu}, \qquad q_1^\mu = \tilde{q}_1^\mu,\\ q_2^\mu &= -\frac{1}{\alpha}\tilde{p}_{2\mu}, \qquad p_{2\mu} = \alpha\tilde{q}_2^\mu; \end{aligned} \tag{2.24}$$

when inserted into the Hamiltonian (2.22), it yields the standard Ostrogradski Hamiltonian

$$H = \tilde{p}_{1\mu}\tilde{q}_2^\mu + \frac{1}{2\beta}(\tilde{p}_{2\mu})^2 - \frac{\lambda}{\beta}\epsilon_{\mu\nu}\tilde{q}_2^\mu\tilde{p}_{2\nu} + \frac{\lambda^2}{2\beta}(\tilde{q}_2^\mu)^2. \tag{2.25}$$

It does not depend on α. Therefore, the energy (energy spectrum in quantum theory) does not depend on α. The role of our α-dependent modification is to provide the formalism which yields standard Lagrangian dynamics and a regular Legendre transformation.

The above explanation is slightly formal. We shall now look at the problem of α dependence from a slightly different point of view. Let us note that the classical state of our system is uniquely determined once the values of $q(t)$, $\dot{q}(t)$, $\ddot{q}(t)$, $\dddot{q}(t)$ at some moment t are given. Moreover, most physically relevant quantities are constructed via the Noether procedure (they are either conserved or partially conserved; i.e., their time derivatives are defined by transformation properties of symmetry breaking terms in the action). As such, they are expressible in terms of q, $\dot{q}$, $\ddot{q}$, and $\dddot{q}$. Therefore the latter are the basic variables. We can find their quantum counterparts, provided we compute the relevant Poisson brackets.

To this end we write out the canonical equations of motion following from Eq. (2.22):

$$\begin{aligned} \dot{q}_1^\mu &= \frac{1}{\alpha}p_{2\mu}, \qquad \dot{q}_2^\mu = \frac{1}{\alpha}p_{1\mu} + \frac{\lambda^2}{\alpha^2\beta}p_{2\mu} + \frac{\lambda}{\beta}\epsilon_{\mu\nu}q_2^\nu,\\ \dot{p}_{1\mu} &= 0, \qquad \dot{p}_{2\mu} = \frac{\lambda}{\beta}\epsilon_{\mu\nu}p_{2\nu} - \frac{\alpha^2}{\beta}q_2^\mu. \end{aligned} \tag{2.26}$$

They lead to the following relations

$$
\begin{aligned}
q^\mu &= q_1^\mu, \qquad \dot{q}^\mu = \frac{1}{\alpha} p_{2\mu},\\
\ddot{q}^\mu &= \frac{\lambda}{\alpha\beta}\epsilon_{\mu\nu}p_{2\nu} - \frac{\alpha}{\beta}q_2^\mu,\\
\dddot{q}^\mu &= -\frac{2\lambda^2}{\alpha\beta^2}p_{2\mu} - \frac{2\lambda\alpha}{\beta^2}\epsilon_{\mu\nu}q_2^\nu - \frac{1}{\beta}p_{1\mu}.
\end{aligned}
\tag{2.27}
$$

One can now find the Poisson brackets among q, $\dot{q}$, $\ddot{q}$, and $\dddot{q}$. The nonvanishing ones read

$$
\begin{aligned}
\{q^\mu, \dddot{q}^\nu\} &= -\frac{1}{\beta}\delta_{\mu\nu}, \qquad \{\dot{q}^\mu, \ddot{q}^\nu\} = \delta_{\mu\nu},\\
\{\dot{q}^\mu, \dddot{q}^\nu\} &= -\frac{2\lambda}{\beta^2}\epsilon_{\mu\nu}, \qquad \{\ddot{q}^\mu, \ddot{q}^\nu\} = \frac{2\lambda}{\beta}\epsilon_{\mu\nu},\\
\{\dot{q}^\mu, \ddot{q}^\nu\} &= \frac{4\lambda^2}{\beta^3}\delta_{\mu\nu}, \qquad \{\dot{q}^\mu, \ddot{q}^\nu\} = \frac{8\lambda^3}{\beta^4}\epsilon_{\mu\nu}.
\end{aligned}
\tag{2.28}
$$

Note that they are α independent. Upon quantizing, we get four observables obeying α-independent algebra. Any other observable including energy can be constructed out of them so its spectrum and other properties do not depend on α.

B. The case of third derivatives

Let us consider a nonsingular Lagrangian of the form

$$L = L(q, \dot{q}, \ddot{q}, \dddot{q}). \tag{2.29}$$

It is slightly surprising that this case (and, in general, the case when the highest time derivatives are of odd order—see the Appendix) is simpler. We define the new variables

$$q^\mu = q_1^\mu, \quad \dot{q}^\mu = \dot{q}_1^\mu, \quad \ddot{q}^\mu = q_2^\mu, \quad \dddot{q}^\mu = \dot{q}_2^\mu. \tag{2.30}$$

Next, the function $F(q_1, \dot{q}_1, q_2, q_3)$ is selected, which obeys

$$\det\left(\frac{\partial^2 F}{\partial \dot{q}_1^\mu \partial q_3^\nu}\right) \neq 0; \tag{2.31}$$

here q_3^μ are additional variables. The modified Lagrangian reads

$$
\begin{aligned}
\mathcal{L}(q_1, \dot{q}_1, q_2, \dot{q}_2, q_3, \dot{q}_3) = {}& L(q_1, \dot{q}_1, q_2, \dot{q}_2) + \frac{\partial F(q_1, \dot{q}_1, q_2, q_3)}{\partial q_1^\mu}\dot{q}_1^\mu + \frac{\partial F(q_1, \dot{q}_1, q_2, q_3)}{\partial q_2^\mu}\dot{q}_2^\mu + \frac{\partial F(q_1, \dot{q}_1, q_2, q_3)}{\partial q_3^\mu}\dot{q}_3^\mu\\
&+ \frac{\partial F(q_1, \dot{q}_1, q_2, q_3)}{\partial \dot{q}_1^\mu}q_2^\mu.
\end{aligned}
\tag{2.32}
$$

It can be easily shown that the Euler-Lagrange equations for $\mathcal{L}$ yield the initial equations for the original variable $q^\mu \equiv q_1^\mu$. Again, as in the second-order case, the Legendre transformation can be directly performed due to the condition (2.31). The momenta read

$$
\begin{aligned}
p_{1\mu} = {}& \frac{\partial L}{\partial \dot{q}_1^\mu} + \frac{\partial^2 F}{\partial q_1^\nu \partial \dot{q}_1^\mu}\dot{q}_1^\nu + \frac{\partial^2 F}{\partial \dot{q}_1^\mu \partial \dot{q}_1^\nu}q_2^\nu + \frac{\partial^2 F}{\partial \dot{q}_1^\mu \partial q_2^\nu}\dot{q}_2^\nu\\
&+ \frac{\partial^2 F}{\partial \dot{q}_1^\mu \partial q_3^\nu}\dot{q}_3^\nu + \frac{\partial F}{\partial q_1^\mu},
\end{aligned}
\tag{2.33}
$$

$$p_{2\mu} = \frac{\partial L}{\partial \dot{q}_2^\mu} + \frac{\partial F}{\partial q_2^\mu}, \tag{2.34}$$

$$p_{3\mu} = \frac{\partial F(q_1, \dot{q}_1, q_2, q_3)}{\partial q_3^\mu}. \tag{2.35}$$

By virtue of (2.31) one can solve (2.35) for $\dot{q}_1^\mu$,

$$\dot{q}_1^\mu = \dot{q}_1^\mu(q_1, q_2, q_3, p_3). \tag{2.36}$$

Inserting this solution into Eq. (2.34), one computes

$$\dot{q}_2^\mu = \dot{q}_2^\mu(q_1, q_2, q_3, p_2, p_3); \tag{2.37}$$

the solution is (at least locally) unique because L is, by assumption, nonsingular in the Ostrogradski sense. Similarly, (2.33) can be solved in terms of $\dot{q}_3^\mu$:

$$\dot{q}_3^\mu = \dot{q}_3^\mu(q_1, q_2, q_3, p_1, p_2, p_3). \tag{2.38}$$

Finally, the Hamiltonian is of the form

$$
\begin{aligned}
H = {}& p_{1\mu}\dot{q}_1^\mu + p_{2\mu}\dot{q}_2^\mu - L - \frac{\partial F}{\partial q_1^\mu}\dot{q}_1^\mu - \frac{\partial F}{\partial \dot{q}_1^\mu}q_2^\mu\\
&- \frac{\partial F}{\partial q_2^\mu}\dot{q}_2^\mu,
\end{aligned}
\tag{2.39}
$$

where everything is expressed in terms of q's and p's (the terms containing $\dot{q}_3^\mu$ cancel). As above, we have checked that the canonical equations of motion yield the initial equation. The canonical transformation which relates our formalism to the Ostrogradski one reads

$$
\begin{aligned}
\tilde{q}_1^\mu &= q_1^\mu, \qquad \tilde{q}_2^\mu = f^\mu(q_1, q_2, q_3, p_3), \qquad \tilde{q}_3^\mu = q_2^\mu,\\
\tilde{p}_{1\mu} &= p_{1\mu} - \frac{\partial F}{\partial q_1^\mu}(q_1, f(q_1, q_2, q_3, p_3), q_2, q_3),\\
\tilde{p}_{2\mu} &= -\frac{\partial F}{\partial f^\mu}(q_1, f(q_1, q_2, q_3, p_3), q_2, q_3),\\
\tilde{p}_{3\mu} &= p_{2\mu} - \frac{\partial F}{\partial q_2^\mu}(q_1, f(q_1, q_2, q_3, p_3), q_2, q_3),
\end{aligned}
\tag{2.40}
$$

where f^μ is the solution of Eq. (2.35), i.e., $f^\mu = \dot{q}_1^\mu$. The relevant generating function reads

$$\Phi(q_1, \tilde{p}_1, q_2, \tilde{q}_2, q_3, \tilde{p}_3) = q_1^\mu \tilde{p}_{1\mu} + q_2^\mu \tilde{p}_{3\mu} + F(q_1, \tilde{q}_2, q_2, q_3). \quad (2.41)$$

Again, the advantage of our Hamiltonian over the Ostrogradski one is that the former is nonsingular in the sense that the inverse Legendre transformation can be performed directly.

C. The second-order Lagrangian once more

By comparing Secs. II A and II B we see that the modified Hamiltonian formalism is somewhat simpler in the case of the third-order Lagrangian (actually, as shown in the Appendix, this is the case for all Lagrangians of odd order). Namely, in the latter case no counterpart of the condition (2.6) is necessary. This will appear to play the crucial role in the case of singular (in the Ostrogradski sense) Lagrangians (see Sec. III below). Therefore, as a preliminary step, we consider here the second-order Lagrangians as a special, singular case of third-order ones. The resulting Hamiltonian formalism is then constrained. However, with an additional assumption that the function F does not depend on q_2^μ, one can perform a complete reduction of phase space, obtaining the structure described in Sec. II A.

Let

$$L = L(q, \dot{q}, \ddot{q}), \quad (2.42)$$

and $F = F(q_1, \dot{q}_1, q_3)$ obeys (2.31). We define

$$\mathcal{L}(q_1, \dot{q}_1, q_2, q_3, \dot{q}_3) = L(q_1, \dot{q}_1, q_2) + \frac{\partial F(q_1, \dot{q}_1, q_3)}{\partial q_1^\mu}\dot{q}_1^\mu + \frac{\partial F(q_1, \dot{q}_1, q_3)}{\partial q_3^\mu}\dot{q}_3^\mu + \frac{\partial F(q_1, \dot{q}_1, q_3)}{\partial \dot{q}_1^\mu} q_2^\mu. \quad (2.43)$$

The relevant momenta read

$$p_{1\mu} = \frac{\partial L}{\partial \dot{q}_1^\mu} + \frac{\partial^2 F}{\partial q_1^\nu \partial \dot{q}_1^\mu}\dot{q}_1^\nu + \frac{\partial^2 F}{\partial \dot{q}_1^\mu \partial \dot{q}_1^\nu} q_2^\nu + \frac{\partial^2 F}{\partial \dot{q}_1^\mu \partial q_3^\nu}\dot{q}_3^\nu + \frac{\partial F}{\partial q_1^\mu}, \quad (2.44)$$

$$p_{2\mu} = 0, \quad (2.45)$$

$$p_{3\mu} = \frac{\partial F}{\partial q_3^\mu}. \quad (2.46)$$

There is one set of primary constraints (2.45). On the other hand, due to the condition (2.31) $\dot{q}_1^\mu$ and $\dot{q}_3^\mu$ can be expressed in terms of q_1, q_2, q_3, p_1, p_3. The Dirac Hamiltonian takes the form

$$H = p_{1\mu}\dot{q}_1^\mu - L - \frac{\partial F}{\partial q_1^\mu}\dot{q}_1^\mu - \frac{\partial F}{\partial \dot{q}_1^\mu} q_2^\mu + c^\mu p_{2\mu}, \quad (2.47)$$

where c^μ are Lagrange multipliers enforcing the constraints $\Phi_{1\mu} \equiv p_{2\mu} \approx 0$.

The stability of primary constraints implies

$$0 \approx \dot{\Phi}_{1\mu} \equiv \Phi_{2\mu} = \frac{\partial L(q_1, \dot{q}_1(q_1, q_3, p_3), q_2)}{\partial q_2^\mu} + \frac{\partial F(q_1, \dot{q}_1(q_1, q_3, p_3), q_3)}{\partial \dot{q}_1^\mu}. \quad (2.48)$$

In order to check the stability of secondary constraints $\Phi_{2\mu}$, we note that, as can be verified by direct computation,

$$\{\dot{q}_1^\mu, \dot{q}_1^\nu\} = 0. \quad (2.49)$$

Using (2.49) together with

$$0 \approx \dot{\Phi}_{2\mu} = \{\Phi_{2\mu}, H\}, \quad (2.50)$$

we arrive at the following condition:

$$\frac{\partial^2 L}{\partial q_2^\mu \partial q_2^\nu} c^\nu + \frac{\partial^2 L}{\partial q_1^\mu \partial q_1^\nu}\dot{q}_1^\nu + \frac{\partial^2 L}{\partial q_2^\mu \partial \dot{q}_1^\nu} q_2^\nu + p_{1\mu} - \frac{\partial L}{\partial \dot{q}_1^\mu} - \frac{\partial F}{\partial q_1^\mu} = 0. \quad (2.51)$$

The initial Lagrangian is nonsingular, and Eq. (2.51) can be used to determine the Lagrange multipliers c^ν uniquely. Therefore, there are no further constraints.

In order to convert our constraints into strong equations, we define Dirac brackets. To this end, we compute

$$\{\phi_{1\mu}, \phi_{1\nu}\} = 0, \quad (2.52)$$

$$\{\phi_{1\mu}, \phi_{2\nu}\} = -\frac{\partial^2 L}{\partial q_2^\mu \partial q_2^\nu} \equiv -W_{\mu\nu}. \quad (2.53)$$

Moreover,

$$\left\{\frac{\partial L}{\partial q_2^\mu}, \frac{\partial L}{\partial q_2^\nu}\right\} = 0, \qquad \left\{\frac{\partial F}{\partial \dot{q}_1^\mu}, \frac{\partial F}{\partial \dot{q}_1^\nu}\right\} = 0, \qquad \left\{\frac{\partial F}{\partial \dot{q}_1^\mu}, \frac{\partial L}{\partial q_2^\nu}\right\} = \frac{\partial^2 L}{\partial q_2^\nu \partial \dot{q}_1^\mu}, \quad (2.54)$$

which implies

$$\{\phi_{2\mu}, \phi_{2\nu}\} = \frac{\partial^2 L}{\partial \dot{q}_1^\mu \partial q_2^\nu} - \frac{\partial^2 L}{\partial \dot{q}_1^\nu \partial q_2^\mu} \equiv V_{\mu\nu}. \quad (2.55)$$

By assumption, W is a nonsingular matrix. Consequently,

$$C = \begin{pmatrix} \{\phi_{1\mu}, \phi_{1\nu}\} & \{\phi_{1\mu}, \phi_{2\nu}\} \\ \{\phi_{2\mu}, \phi_{1\nu}\} & \{\phi_{2\mu}, \phi_{2\nu}\} \end{pmatrix} = \begin{pmatrix} 0 & -W \\ W & V \end{pmatrix} \quad (2.56)$$

is also nonsingular, and

$$C^{-1} = \begin{pmatrix} W^{-1}VW^{-1} & W^{-1} \\ -W^{-1} & 0 \end{pmatrix}. \quad (2.57)$$

The Dirac brackets take the following form:

$$\begin{aligned}\{\cdot,\cdot\}_D = \{\cdot,\cdot\} - \{\cdot,\phi_{1\mu}\}(W^{-1}VW^{-1})_{\mu\nu}\{\phi_{1\nu},\cdot\}\\ - \{\cdot,\phi_{1\mu}\}(W^{-1})_{\mu\nu}\{\phi_{2\nu},\cdot\}\\ + \{\cdot,\phi_{2\mu}\}(W^{-1})_{\mu\nu}\{\phi_{1\nu},\cdot\}.\end{aligned} \tag{2.58}$$

The constraints $\Phi_{1\mu}$ depend on $p_{2\mu}$ only. We conclude from (2.58) that the Dirac brackets for q_1^μ, q_3^μ, $p_{1\mu}$, $p_{3\mu}$ take the canonical form. Moreover, $p_{2\mu} = 0$, while q_2^μ can be determined from (2.48). Note that the solution for q_2^μ, by virtue of Eq. (2.6), reads

$$q_2^\mu = \chi^\mu(q_1, \dot{q}_1(q_1, q_3, p_3), q_3). \tag{2.59}$$

So, up to renumbering $q_2 \leftrightarrow q_3$, we arrived at the same scheme as in Sec. II A.

In order to illustrate the above approach, we use the same example as before:

$$L = \lambda\epsilon_{\mu\nu}\dot{q}^\mu\ddot{q}^\nu + \frac{\beta}{2}(\ddot{q}^\mu)^2, \qquad \beta \neq 0, \tag{2.60}$$

and

$$F = \alpha\dot{q}_1^\mu q_3^\mu, \qquad \alpha \neq 0. \tag{2.61}$$

Then H takes the form

$$\begin{aligned}H = \frac{1}{\alpha}p_{1\mu}p_{3\mu} - \frac{\lambda}{\alpha}\epsilon_{\mu\nu}p_{3\mu}q_2^\nu - \frac{\beta}{2}(q_2^\mu)^2 - \alpha q_2^\mu q_3^\nu\\ + \frac{2\lambda}{\beta}\epsilon_{\mu\nu}p_{2\mu}q_2^\nu - \frac{1}{\beta}p_{1\mu}p_{2\mu},\end{aligned} \tag{2.62}$$

while the constraints are

$$\phi_{1\mu} = p_{2\mu}, \qquad \phi_{2\mu} = -\frac{\lambda}{\alpha}\epsilon_{\mu\nu}p_{3\nu} + \beta q_2^\mu + \alpha q_3^\mu, \tag{2.63}$$

and serve to eliminate $p_{2\mu}$ and q_2^μ,

$$p_{2\mu} = 0, \qquad q_2^\mu = \frac{\lambda}{\alpha\beta}\epsilon_{\mu\nu}p_{3\nu} - \frac{\alpha}{\beta}q_3^\mu. \tag{2.64}$$

Inserting this back into the Hamiltonian, we arrive at the following expression:

$$H = \frac{1}{\alpha}p_{1\mu}p_{3\mu} + \frac{\lambda^2}{2\alpha^2\beta}(p_{3\mu})^2 + \frac{\lambda}{\beta}\epsilon_{\mu\nu}p_{3\mu}q_3^\nu + \frac{\alpha^2}{2\beta}(q_3^\mu)^2, \tag{2.65}$$

which coincides with the one given by Eq. (2.22), provided the replacement $q_2 \leftrightarrow q_3$, $p_2 \leftrightarrow p_3$ has been made.

III. SINGULAR LAGRANGIANS OF THE SECOND ORDER

In this section we consider the second-order Lagrangians

$$L = L(q, \dot{q}, \ddot{q}), \tag{3.1}$$

which are singular in the Ostrogradski sense, i.e.,

$$\det(W_{\mu\nu}) \equiv \det\left(\frac{\partial^2 L}{\partial\ddot{q}^\mu\partial\ddot{q}^\nu}\right) = 0. \tag{3.2}$$

For the standard Ostrogradski approach to such singular Lagrangians see, for example, Refs. [18,19].

The formalism of Sec. II A is not directly applicable because, due to Eq. (3.2), Eq. (2.6) cannot be solved to determine the functions χ^μ. Moreover, in this case Eq. (2.6) puts further restrictions on the form of F.

In order to get rid of these problems we will follow the method of Sec. II B and consider L as a third-order singular Lagrangian. From this point of view its singularity comes both from Eq. (3.2) and from the fact that the third-order time derivatives are absent. Given a singular Lagrangian L, we select a function $F = F(q_1, \dot{q}_1, q_3)$ obeying (2.31) and define

$$\begin{aligned}\mathcal{L}(q_1, \dot{q}_1, q_2, q_3, \dot{q}_3)\\ = L(q_1, \dot{q}_1, q_2) + \frac{\partial F(q_1, \dot{q}_1, q_3)}{\partial q_1^\mu}\dot{q}_1^\mu + \frac{\partial F(q_1, \dot{q}_1, q_3)}{\partial q_3^\mu}\dot{q}_3^\mu\\ + \frac{\partial F(q_1, \dot{q}_1, q_3)}{\partial \dot{q}_1^\mu}q_2^\mu.\end{aligned} \tag{3.3}$$

As before, the canonical momenta given by (2.45) provide the primary constraints, while (2.44) and (2.46) allow us to compute $\dot{q}_1^\mu$ and $\dot{q}_3^\mu$. The Hamiltonian is given by Eq. (2.47). The secondary constraints read again

$$\begin{aligned}0 \approx \Phi_{2\mu}\\ = \frac{\partial L(q_1, \dot{q}_1(q_1, q_3, p_3), q_2)}{\partial q_2^\mu}\\ + \frac{\partial F(q_1, \dot{q}_1(q_1, q_3, p_3), q_3)}{\partial \dot{q}_1^\mu}.\end{aligned} \tag{3.4}$$

Now we have to investigate the stability of $\Phi_{2\mu}$. To this end, we assume that W has rank K, $K < N$; this implies the existence of $J = N - K$ linearly independent null eigenvectors $\gamma_a^\mu(q_1, \dot{q}_1, q_2)$, $a = 1, 2, \ldots, J$,

$$W_{\mu\nu}\gamma_a^\nu = 0. \tag{3.5}$$

Equation (2.51) does not determine uniquely the Lagrange multipliers c^μ; on the contrary, we get new constraints of the form

$$\begin{aligned}0 \equiv \Phi_{3a}\\ = \gamma_a^\mu\left(\frac{\partial^2 L}{\partial q_1^\mu\partial q_1^\nu}\dot{q}_1^\nu + \frac{\partial^2 L}{\partial q_2^\mu\partial \dot{q}_1^\nu}q_2^\nu + p_{1\mu} - \frac{\partial L}{\partial \dot{q}_1^\mu} - \frac{\partial F}{\partial q_1^\mu}\right);\end{aligned} \tag{3.6}$$

here, as previously, $\dot{q}_1^\mu = \dot{q}_1^\mu(q_1, q_3, p_3)$, so the above constraints contain q_1, q_2, q_3, p_1, and p_3.

We have started with the third-order formalism; therefore, our phase space is $6N$ dimensional. As in the nonsingular case (Sec. II) we would like to eliminate the q_2's and p_2's. The latter are equal to zero by the primary constraints $\Phi_{1\mu}$. As far as the q_2's are considered, the situation is more involved.

First, by virtue of the assumption (3.2) about W, we can determine from Eq. (3.4) K variables q_2^μ in terms of q_1, p_1, q_3, p_3, and the remaining q_2's. By substituting the resulting expression back into Eq. (3.4), we arrive at J constraints on q_1, p_1, q_3, and p_3. We denote these new constraints by $\psi_a(q_1, q_3, p_1, p_3)$. Let us now concentrate on the constraints (3.6). In general, they contain the q_2^μ variables and imply the constraints on q_1, q_3, p_1, p_3 only provided the q_2's enter in the combinations which can be determined from Eq. (3.4). In order to decide if this happens, consider the variations δq_2^μ which do not change the RHS of (3.4). From the definition of $W_{\mu\nu}$ we conclude that such δq_2^μ are linear combinations of γ_a^μ [see (3.5)]. If the RHS of (3.6) is stationary under such variations δq_2^μ, Eqs. (3.4) and (3.6) can be combined to yield the constraints which do not depend on the q_2's. The relevant condition reads

$$\frac{\partial \Phi_{3a}}{\partial q_2^\mu}\gamma_b^\mu = 0, \qquad b = 1, \ldots, J, \tag{3.7}$$

where a takes M values, which, without loss of generality, can be chosen as $a = 1, \ldots, M$. In this way, we obtain M new constraints on q_1, p_1, q_3, p_3.

One can check that

$$\frac{\partial \Phi_{3a}}{\partial q_2^\mu}\gamma_b^\mu = \gamma_b^\mu \gamma_a^\nu\left(\frac{\partial^2 L}{\partial \dot{q}_1^\mu \partial q_2^\nu} - \frac{\partial^2 L}{\partial \dot{q}_1^\nu \partial q_2^\mu}\right). \tag{3.8}$$

By virtue of (2.55) we find

$$\{\psi_a, \psi_b\} \approx \gamma_a^\mu \gamma_b^\nu \{\phi_{2\mu}, \phi_{2\nu}\}, \qquad a = 1, \ldots, M, b = 1, \ldots, J. \tag{3.9}$$

Let us summarize. For the nonsingular second-order Lagrangian viewed as a singular third-order one, (q_1, p_1, q_3, p_3) forms the reduced phase space; no further constraints exist. On the contrary, in the singular case q_1, p_1, q_3, p_3 are still constrained. First, there exist J constraints $\psi_a(q_1, p_1, q_3, p_3)$; moreover, if some (say, M) ψ's are in involution (on the constraint surface) with all ψ's, there exist additional M constraints following from Eqs. (2.6) and (3.4). This agrees with the conclusions of Ref. [18].

In general, for a singular Lagrangian it is not possible to determine uniquely all Lagrange multipliers c^μ. However, we are in fact interested only in dynamical equations for q_1, q_3, p_1, and p_3. Therefore, we can use the following Hamiltonian:

$$H = p_{1\mu}\dot{q}_1^\mu - L - \frac{\partial F}{\partial q_1^\mu}\dot{q}_1^\mu - \frac{\partial F}{\partial \dot{q}_1^\mu}q_2^\mu. \tag{3.10}$$

On the constraint surface it does not depend on the q_2's,

$$\frac{\partial H}{\partial q_2^\mu} = -\frac{\partial L}{\partial q_2^\mu} - \frac{\partial F}{\partial \dot{q}_1^\mu} \approx 0. \tag{3.11}$$

The existence of further secondary constraints depends on the particular form of the Lagrangian.

Finally, let us note that the canonical transformation (2.40) leads to the form of dynamics presented in Ref. [18]. However, within our procedure the Legendre transformation from the tangent bundle of the configuration manifold to the phase manifold is again straightforward (if one takes into account standard modifications due to the existence of constraints).

Singular higher-derivative Lagrangians were also considered in [20]. The authors considered the physically important case of reparametrization invariant theories (higher-derivative reparametrization invariant Lagrangians appear, for example, in the description of the radiation reaction [21]). In their geometrical approach the image of the Legendre transformation forms a submanifold of some cotangent bundle. This suggests that in the case of higher-derivative singular theories, it is advantageous to start with enlarged phase space; this agrees with our conclusions.

To conclude this section with a simple example, consider the following Lagrangian:

$$L = \lambda\epsilon_{\mu\nu}\dot{q}^\mu\ddot{q}^\nu + \frac{\beta}{2}(\ddot{q}^1)^2, \qquad \mu, \nu = 1, 2. \tag{3.12}$$

It is singular and the matrix W [Eq. (2.53)] is of rank 1 for $\beta \neq 0$ and 0 for $\beta = 0$. We take F as

$$F = \alpha\dot{q}_1^\mu q_3^\mu. \tag{3.13}$$

Assume first $\beta \neq 0$. Then

$$\mathcal{L} = \lambda\epsilon_{\mu\nu}\dot{q}_1^\mu q_2^\nu + \frac{\beta}{2}(q_2^1)^2 + \alpha q_3^\mu q_2^\mu + \alpha\dot{q}_1^\mu\dot{q}_3^\mu, \tag{3.14}$$

and

$$p_{1\mu} = \lambda\epsilon_{\mu\nu}q_2^\nu + \alpha\dot{q}_3^\mu, \qquad p_{2\mu} = 0, \qquad p_{3\mu} = \alpha\dot{q}_1^\mu. \tag{3.15}$$

The primary constraints are

$$\Phi_{1\mu} = p_{2\mu} \approx 0, \tag{3.16}$$

while the Hamiltonian reads

$$H = \frac{1}{\alpha}p_{1\mu}p_{3\mu} - \frac{\lambda}{\alpha}\epsilon_{\mu\nu}p_{3\mu}q_2^\nu - \frac{\beta}{2}(q_2^1)^2 - \alpha q_2^\mu q_3^\mu + c^\mu p_{2\mu}. \tag{3.17}$$

One easily derives the secondary constraints

$$0 \approx \Phi_{21} = \frac{\lambda}{\alpha}p_{32} - \beta q_2^1 - \alpha q_3^1, \qquad 0 \approx \Phi_{22} = \frac{\lambda}{\alpha}p_{31} + \alpha q_3^2. \tag{3.18}$$

The stability for $\Phi_{2\mu}$ yields

$$0 \approx \{\Phi_{21}, H\} = 2\lambda q_2^2 - \beta c^1 - p_{11}, \tag{3.19}$$

$$0 \approx \{\Phi_{22}, H\} = p_{12} + 2\lambda q_2^1 = \Phi_3. \tag{3.20}$$

Equation (3.19) allows us to compute c^1,

$$c^1 = \frac{1}{\beta}(2\lambda q_2^2 - p_{11}), \tag{3.21}$$

while (3.20) provides a new constraint. Its stability enforces $c^1 = 0$, which together with (3.21) yields the further constraint

$$0 \approx \Phi_4 = \frac{1}{\beta}(2\lambda q_2^2 - p_{11}). \tag{3.22}$$

Finally, differentiating the above equation with respect to time, we get $c^2 = 0$. The resulting Hamiltonian is

$$H = \frac{1}{\alpha} p_{1\mu} p_{3\mu} - \frac{\lambda}{\alpha}\epsilon_{\mu\nu} p_{3\mu} q_2^\nu - \frac{\beta}{2}(q_2^1)^2 - \alpha q_2^\mu q_3^\mu. \tag{3.23}$$

Still we have to take into account the constraints $\Phi_{2\mu}$, Φ_3, and Φ_4. The latter two can be rewritten as

$$\Phi_{3\mu} = p_{1\mu} - 2\lambda\epsilon_{\mu\nu} q_2^\nu. \tag{3.24}$$

$\Phi_{2\mu}$ and $\Phi_{3\mu}$ are now used in order to eliminate all variables except q_1^μ, $p_{1\mu}$ and q_3^2, p_{32}. The only nonstandard Dirac bracket reads

$$\{q_3^2, p_{32}\}_D = \tfrac{1}{2}. \tag{3.25}$$

The Hamiltonian, when expressed in terms of unconstrained variables, takes the form

$$H = \frac{1}{\alpha} p_{12} p_{32} - \frac{\alpha}{\lambda} p_{11} q_3^2 + \frac{\beta}{8\lambda^2}(p_{12})^2. \tag{3.26}$$

Let us note that the limit $\beta \longrightarrow 0$ is smooth. Of course, we could put $\beta = 0$ from the very beginning and arrive at the same conclusion.

IV. AN EXAMPLE: MINISUPERSPACE FORMULATION OF $f(R)$ GRAVITY

As a more elaborate but still toy example, we consider the minisuperspace Hamiltonian formulation of $f(R)$ gravity [12]. We consider the following (Lemaitre-Friedmann-Robertson-Walker–type) metrics:

$$ds^2 = -N^2 dt^2 + a^2 d\vec{x}^2. \tag{4.1}$$

Under such reduction the Lagrangian of $f(R)$ gravity takes the form

$$L(a, N) = \tfrac{1}{2} N a^3 f(R), \tag{4.2}$$

where the curvature is given by

$$R = 6\left(\frac{\dot{a}}{NA}\right)^{\cdot} + 12\left(\frac{\dot{a}}{Na}\right)^2. \tag{4.3}$$

We see that L depends on second time derivatives. We proceed along the lines described in Sec. II. The basic dynamical variables are chosen as follows:

$$a_1 = a, \quad \dot{a}_1 = \dot{a}, \quad N_1 = N, \quad \dot{N}_1 = \dot{N}, \quad a_2 = R, \tag{4.4}$$

while

$$\ddot{a} = \chi(a_1, \dot{a}_1, N_1, \dot{N}_1, a_2) \tag{4.5}$$

is determined by Eq. (2.6) once an appropriate F is selected. We take

$$F = -3a_1^2 f'(a_2)\dot{a}_1; \tag{4.6}$$

under the assumption $f'' \neq 0$, Eqs. (2.6) and (4.6) yield

$$a_2 = R. \tag{4.7}$$

Solving (4.7) with respect to $\ddot{a}$ we find

$$\ddot{a} = \frac{a_1 N_1}{6}\left(R - \frac{6}{N_1^2 a_1^2}((2 - N_1)\dot{a}_1^2 - a_1\dot{a}_1\dot{N}_1)\right). \tag{4.8}$$

The modified Lagrangian reads

$$\mathcal{L} = \frac{1}{2}a_1^3 N_1 f(a_2) + f'(a_2)\left(-9a_1\dot{a}_1^2 + \frac{6a_1\dot{a}_1^2}{N_1} - \frac{1}{2}a_1^3 N_1 a_2 - \frac{3a_1^2\dot{a}_1\dot{N}_1}{N_1}\right) - 3f''(a_2)a_1^2\dot{a}_1\dot{a}_2. \tag{4.9}$$

It is straightforward to check that $\mathcal{L}$ leads to the correct equations of motion. In order to simplify our considerations we introduce a new variable,

$$n_1 = N_1 f'(a_2). \tag{4.10}$$

In terms of the new variable $\mathcal{L}$ reads

$$\mathcal{L} = \frac{1}{2}a_1^3 n_1 \frac{f(a_2)}{f'(a_2)} - 9a_1\dot{a}_1^2 f'(a_2) + \frac{6a_1\dot{a}_1^2 f'^2(a_2)}{n_1} - \frac{1}{2}a_1^3 n_1 a_2 - \frac{3a_1^2\dot{a}_1\dot{n}_1 f'(a_2)}{n_1}. \tag{4.11}$$

Now, we compute the canonical momenta:

$$p_1 \equiv \frac{\partial\mathcal{L}}{\partial\dot{n}_1} = -\frac{3a_1^2}{n_1} f'(a_2)\dot{a}_1, \tag{4.12}$$

$$\pi_1 \equiv \frac{\partial\mathcal{L}}{\partial\dot{a}_1} = -18a_1\dot{a}_1 f'(a_2) + \frac{12a_1\dot{a}_1 f'^2(a_2)}{n_1} - 3\frac{a_1^2}{n_1} f'(a_2)\dot{n}_1, \tag{4.13}$$

$$\pi_2 \equiv \frac{\partial\mathcal{L}}{\partial\dot{a}_2} = 0. \tag{4.14}$$

One can solve (4.12) and (4.13) in terms of $\dot{a}_1$ and $\dot{n}_1$. We form the Hamiltonian

$$H = -\frac{n_1 p_1 \pi_1}{3a_1^2 f'(a_2)} - \frac{n_1 a_1^3 f(a_2)}{2f'(a_2)} + \frac{n_1^2 p_1^2}{a_1^3 f'(a_2)} - \frac{2n_1 p_1^2}{3a_1^3} + \frac{1}{2} a_1^3 a_2 n_1 + \mu\pi_2 \equiv \tilde{H} + \mu\pi_2. \tag{4.15}$$

Now, we investigate the stability of the $\Phi_1 \equiv \pi_2$ constraint,

$$0 \approx \dot{\Phi}_1 = \{\Phi_1, H\} = \frac{f''(a_2)}{f'(a_2)}\left(\tilde{H} + \frac{2n_1 p_1^2}{3a_1^3} - \frac{a_1^3 a_2 n_1}{2}\right) \equiv \Phi_2. \tag{4.16}$$

The stability condition for Φ_2 determines μ; an explicit expression for μ is irrelevant for what follows. In fact, (Φ_1, Φ_2) are second class constraints,

$$\{\Phi_1, \Phi_2\} \approx \frac{f''(a_2) a_1^3 N_1}{2f'(a_2)}. \tag{4.17}$$

Thus, the constraints can be solved, provided we use Dirac brackets. In particular, the Hamiltonian takes a simple form,

$$H = \tilde{H} = \frac{1}{2} a_1^3 a_2 n_1 - \frac{2}{3}\frac{n_1 p_1^2}{a_1^3}, \tag{4.18}$$

where

$$a_2 = f^{-1}\left(-\frac{2p_1\pi_1}{3a_1^5} + \frac{2n_1 p_1^2}{a_1^6}\right). \tag{4.19}$$

Moreover, Dirac brackets for the variables a_1, n_1, π_1, p_1 remain canonical. Therefore, Eqs. (4.18) and (4.19) give the complete Hamiltonian description. We have checked explicitly that this leads to the correct equations of motion. In the case under consideration our formalism, when compared with the Ostrogradski version, seems to be more complicated. However, it has an advantage that the curvature R is one of basic variables.

V. FIELD THEORY

Our formalism has a straightforward generalization to the field theory case. For simplicity, we consider only the Lagrangian densities depending on first and second derivatives. Such a density can be written in the form

$$\mathcal{L} = \mathcal{L}(\Phi, \partial_k\Phi, \partial_k\partial_l\Phi, \dot{\Phi}, \partial_k\dot{\Phi}, \ddot{\Phi}). \tag{5.1}$$

Again, we put $\Phi = \Phi_1$ and select a function $F = F(\Phi_1, \dot{\Phi}_1, \Phi_2)$ obeying

$$\frac{\partial^2 F}{\partial\dot{\Phi}_1 \partial\Phi_2} \neq 0; \tag{5.2}$$

in the case of a multicomponent field the relevant matrix should be nonsingular. We define, as previously, the function

$$\chi = \chi(\Phi_1, \partial_k\Phi_1, \partial_k\partial_l\Phi_1, \dot{\Phi}_1, \partial_k\dot{\Phi}_1, \Phi_2) \tag{5.3}$$

as the [locally unique by virtue of (5.2)] solution to the equation

$$\frac{\partial\mathcal{L}(\Phi_1, \partial_k\Phi_1, \partial_k\partial_l\Phi_1, \dot{\Phi}_1, \partial_k\dot{\Phi}_1, \chi)}{\partial\chi} + \frac{\partial F(\Phi_1, \dot{\Phi}_1, \Phi_2)}{\partial\dot{\Phi}_1} = 0. \tag{5.4}$$

Finally, the new Lagrangian density reads

$$\tilde{\mathcal{L}} = \mathcal{L}(\Phi_1, \partial_k\Phi_1, \partial_k\partial_l\Phi_1, \dot{\Phi}_1, \partial_k\dot{\Phi}_1, \chi(\ldots)) + \frac{\partial F(\Phi_1, \dot{\Phi}_1, \Phi_2)}{\partial\Phi_1}\dot{\Phi}_1 + \frac{\partial F(\Phi_1, \dot{\Phi}_1, \Phi_2)}{\partial\Phi_2}\dot{\Phi}_2 + \frac{\partial F(\Phi_1, \dot{\Phi}_1, \Phi_2)}{\partial\dot{\Phi}_1}\chi(\ldots). \tag{5.5}$$

It is now straightforward to check that the Lagrange equations

$$\frac{\partial\tilde{\mathcal{L}}}{\partial\Phi_i} - \partial_k\frac{\partial\tilde{\mathcal{L}}}{\partial(\partial_k\Phi_i)} + \partial_k\partial_l\frac{\partial\tilde{\mathcal{L}}}{\partial(\partial_k\partial_l\Phi_i)} - \frac{d}{dt}\left(\frac{\partial\tilde{\mathcal{L}}}{\partial\dot{\Phi}_i} - \partial_k\frac{\partial\tilde{\mathcal{L}}}{\partial(\partial_k\dot{\Phi}_i)}\right) = 0 \tag{5.6}$$

yield the initial equation for the original variable $\Phi \equiv \Phi_1$; as in Sec. II A $\ddot{\Phi} = \chi(\ldots)$. One can now perform the Legendre transformation. The canonical momenta read

$$\pi_i(x) = \frac{\delta\tilde{L}}{\delta\dot{\Phi}_i(x)}, \qquad \tilde{L} \equiv \int d^3x\tilde{\mathcal{L}}. \tag{5.7}$$

Equations (5.7) can be solve [due to (5.2)] with respect to $\dot{\Phi}_i$:

$$\dot{\Phi}_1 = \dot{\Phi}_1(\Phi_1, \Phi_2, \Pi_2), \tag{5.8}$$

$$\dot{\Phi}_2 = \dot{\Phi}_2(\Phi_1, \partial_k\Phi_1, \partial_k\partial_l\Phi_1, \partial_k\partial_l\partial_m\Phi_1, \Phi_2, \partial_k\Phi_2, \partial_k\partial_l\Phi_2, \Pi_1, \Pi_2, \partial_k\Pi_2, \partial_k\partial_l\Pi_2). \tag{5.9}$$

H is defined in a standard way,

$$H = \int d^3x(\Pi_1(x)\dot{\Phi}_1(x) + \Pi_2(x)\dot{\Phi}_2(x)) - \tilde{L}, \tag{5.10}$$

and leads to the correct canonical equations of motions.

ACKNOWLEDGMENTS

We are thankful to Professor P. Kosiński for interesting discussions. We are grateful to the anonymous referee for very useful remarks.

APPENDIX: EXTENSION TO THE CASE OF ARBITRARILY HIGH DERIVATIVES

Here we generalize the approach proposed in Sec. II to the case of Lagrangians containing time derivatives of arbitrary order [17]. We restrict ourselves to the case of one degree of freedom. We start with the Lagrangian depending on time derivatives up to some even order,

$$L = L(q, \dot{q}, \ddot{q}, \ldots, q^{(2n)}), \tag{A1}$$

which is assumed to be nonsingular in the Ostrogradski sense, $\frac{\partial^2 L}{\partial q^{(2n)2}} \neq 0$. Define the new variables

$$\begin{aligned} q_i &\equiv q^{(2i-2)}, \qquad i = 1, \ldots, n+1, \\ \dot{q}_i &\equiv q^{(2i-1)}, \qquad i = 1, \ldots, n, \end{aligned} \tag{A2}$$

so that

$$L = L(q_1, \dot{q}_1, q_2, \dot{q}_2, \ldots, q_n, \dot{q}_n, q_{n+1}). \tag{A3}$$

Further, let F be any function of the following variables:

$$F = F(q_1, \dot{q}_1, \ldots, q_n, \dot{q}_n, q_{n+1}, q_{n+2}, \ldots, q_{2n}), \tag{A4}$$

obeying

$$\frac{\partial L}{\partial q_{n+1}} + \frac{\partial F}{\partial \dot{q}_n} = 0 \tag{A5}$$

and

$$\det\left[\frac{\partial^2 F}{\partial q_i \partial \dot{q}_j}\right]_{\substack{n+2 \le i \le 2n \\ 1 \le j \le n-1}} \neq 0, \qquad n \geq 2 \tag{A6}$$

[for $n = 1$ only (A5) remains].

Finally, we define a new Lagrangian,

$$\mathcal{L} \equiv L + \sum_{k=1}^{n}\left(\frac{\partial F}{\partial q_k}\dot{q}_k + \frac{\partial F}{\partial \dot{q}_k} q_{k+1}\right) + \sum_{j=n+1}^{2n} \frac{\partial F}{\partial q_j}\dot{q}_j. \tag{A7}$$

Let us have a look at the Lagrange equations

$$\frac{\partial \mathcal{L}}{\partial q_i} - \frac{d}{dt}\left(\frac{\partial \mathcal{L}}{\partial \dot{q}_i}\right) = 0, \qquad i = 1, \ldots, 2n. \tag{A8}$$

Using (A3), (A4), (A7), and (A8) one finds

$$\sum_{k=1}^{n} \frac{\partial^2 F}{\partial q_i \partial \dot{q}_k}(q_{k+1} - \ddot{q}_k) = 0, \qquad i = n+1, \ldots, 2n. \tag{A9}$$

Consider the matrix $[\frac{\partial^2 F}{\partial q_i \partial \dot{q}_j}]_{\substack{n+1 \le i \le 2n \\ 1 \le j \le n}}$ entering the left-hand side of Eq. (A9). By virtue of (A5), $\frac{\partial^2 F}{\partial q_i \partial \dot{q}_n} = 0$ for $i = n + 2, \ldots, 2n$, while $\frac{\partial^2 F}{\partial q_{n+1} \partial \dot{q}_n} = -\frac{\partial^2 L}{\partial q_{n+1}^2} \neq 0$ due to the Ostrogradski nonsingularity condition. Therefore, the first column of our matrix has only one nonvanishing element. This, together with the condition (A6), implies that it is invertible. Therefore, Eq. (A9) gives

$$q_{k+1} = \ddot{q}_k, \qquad k = 1, \ldots, n. \tag{A10}$$

Let us now consider (A8) for $1 \leq i \leq n$. We find

$$\frac{\partial L}{\partial q_i} - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_i}\right) + \frac{\partial F}{\partial \dot{q}_{i-1}} - \frac{d^2}{dt^2}\left(\frac{\partial F}{\partial \dot{q}_i}\right) = 0, \qquad i = 1, \ldots, n, \tag{A11}$$

where, by definition, $\frac{\partial F}{\partial \dot{q}_0} = 0$. By combining these equations and using (A5) and (A10), we finally arrive at the initial Lagrange equation. We conclude that, contrary to the case of the Ostrogradski Lagrangian, our modified Lagrangian leads to the proper equation of motion. Let us now consider the Hamiltonian formalism. Again, the Legendre transformation can be immediately performed; neither additional Lagrange multipliers nor constraints analysis are necessary. In fact, let us define the canonical momenta in a standard way,

$$p_i = \frac{\partial \mathcal{L}}{\partial \dot{q}_i}, \tag{A12}$$

so that

$$p_i = \frac{\partial F}{\partial q_i}, \qquad i = n+1, \ldots, 2n, \tag{A13}$$

$$p_i = \frac{\partial L}{\partial \dot{q}_i} + \sum_{k=1}^{n}\left(\frac{\partial^2 F}{\partial \dot{q}_i \partial q_k}\dot{q}_k + \frac{\partial^2 F}{\partial \dot{q}_i \partial \dot{q}_k} q_{k+1}\right) + \sum_{j=n+1}^{2n} \frac{\partial^2 F}{\partial \dot{q}_i \partial q_j}\dot{q}_j + \frac{\partial F}{\partial q_i}, \qquad i = 1, \ldots, n. \tag{A14}$$

Because of the nonsingularity of $[\frac{\partial^2 F}{\partial q_i \partial \dot{q}_j}]_{\substack{n+1 \le i \le 2n \\ 1 \le j \le n}}$ Eq. (A13) can be solved for $\dot{q}_1, \dot{q}_2, \ldots, \dot{q}_n$,

$$\dot{q}_i = f_i(q_1, \ldots, q_{2n}, p_{n+1}, \ldots, p_{2n}), \qquad i = 1, \ldots, n. \tag{A15}$$

Now, Eq. (A14) is linear with respect to $\dot{q}_i$, $i = n+1, \ldots, 2n$, and can be easily solved. Finally, the Hamiltonian is calculated according to the standard prescription.

In order to compare the present formalism with the Ostrogradski approach, let us note that they must be related by a canonical transformation. To see this we define new (Ostrogradski) variables $\tilde{q}_k$, $\tilde{p}_k$, $1 \leq k \leq 2n$:

$$\tilde{q}_{2i-1} = q_i, \qquad i = 1, .., n, \tag{A16}$$

$$\tilde{q}_{2i} = f_i(q_1, \ldots, q_{2n}, p_{n+1}, \ldots, p_{2n}), \qquad i = 1, \ldots, n, \tag{A17}$$

$$\tilde{p}_{2i-1} = p_i - \frac{\partial F}{\partial q_i}(q_1, f_1(\ldots), \ldots, q_n, f_n(\ldots), q_{n+1}, \ldots, q_{2n}), \qquad i = 1, \ldots, n, \tag{A18}$$

$$\tilde{p}_{2i} = -\frac{\partial F}{\partial f_i}(q_1, f_1(\ldots), \ldots, q_n, f_n(\ldots), q_{n+1}, \ldots, q_{2n}), \qquad i = 1, \ldots, n. \tag{A19}$$

It is easily seen that the above transformation is a canonical one; i.e., the Poisson brackets are invariant. It is not hard to find the relevant generating function,

$$\Phi(q_1, \ldots, q_{2n}, \tilde{p}_1, \tilde{q}_2, \tilde{p}_3, \tilde{q}_4, \ldots, \tilde{p}_{2n-1}, \tilde{q}_{2n}) = \sum_{k=1}^{n} q_k \tilde{p}_{2k-1} + F(q_1, \tilde{q}_2, q_2, \tilde{q}_4, \ldots, q_n, \tilde{q}_{2n}, q_{n+1}, \ldots, q_{2n}). \tag{A20}$$

Let us now consider the case of the Lagrangian depending on time derivatives up to some odd order,

$$L = L(q, \dot{q}, \ddot{q}, \ldots, q^{(2n+1)}). \tag{A21}$$

Again, we define

$$q_i \equiv q^{(2i-2)}, \qquad i = 1, \ldots, n+1, \tag{A22}$$

$$\dot{q}_i \equiv q^{(2i-1)}, \qquad i = 1, \ldots, n+1, \tag{A23}$$

so that

$$L = L(q_1, \dot{q}_1, q_2, \dot{q}_2, \ldots, q_{n+1}, \dot{q}_{n+1}). \tag{A24}$$

Now, we select a function F,

$$F = F(q_1, \dot{q}_1, q_2, \dot{q}_2, \ldots, q_n, \dot{q}_n, q_{n+1}, \ldots, q_{2n+1}), \tag{A25}$$

subject to the single condition

$$\det\left[\frac{\partial^2 F}{\partial q_i \partial \dot{q}_k}\right]_{\substack{n+2\le i\le 2n+1\\ 1\le k\le n}} \neq 0, \tag{A26}$$

and define the Lagrangian

$$\mathcal{L} = L + \sum_{k=1}^{n}\left(\frac{\partial F}{\partial q_k}\dot{q}_k + \frac{\partial F}{\partial \dot{q}_k} q_{k+1}\right) + \sum_{j=n+1}^{2n+1} \frac{\partial F}{\partial q_j}\dot{q}_j. \tag{A27}$$

Consider the Lagrange equation (A8). First, we have

$$\sum_{k=1}^{n} \frac{\partial^2 F}{\partial q_i \partial \dot{q}_k}(q_{k+1} - \ddot{q}_k) = 0, \qquad i = n+2, \ldots, 2n+1, \tag{A28}$$

and, by virtue of (A26),

$$q_{k+1} = \ddot{q}_k, \qquad k = 1, \ldots, n. \tag{A29}$$

The remaining equations read

$$\frac{\partial L}{\partial q_i} - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_i}\right) + \frac{\partial F}{\partial \dot{q}_{i-1}} - \frac{d^2}{dt^2}\left(\frac{\partial F}{\partial \dot{q}_i}\right) = 0, \quad i = 1, \ldots, n+1, \tag{A30}$$

with $\frac{\partial F}{\partial \dot{q}_0} = 0$, $\frac{\partial F}{\partial \dot{q}_{n+1}} = 0$. Combining (A29) and (A30) one gets

$$\sum_{k=0}^{2n+1} (-1)^k \frac{d^k}{dt^k}\left(\frac{\partial L}{\partial q^{(k)}}\right) = 0. \tag{A31}$$

Let us note that no condition of the form (A5) is necessary here.

Also, in the odd case the present formalism is related to that of Ostrogradski by a canonical transformation. Indeed, the canonical momenta read

$$p_i = \frac{\partial F}{\partial q_i}, \qquad i = n+2, \ldots, 2n+1, \tag{A32}$$

$$p_i = \frac{\partial L}{\partial \dot{q}_i} + \sum_{k=1}^{n}\left(\frac{\partial^2 F}{\partial \dot{q}_i \partial q_k}\dot{q}_k + \frac{\partial^2 F}{\partial \dot{q}_i \partial \dot{q}_k} q_{k+1}\right) + \sum_{j=n+1}^{2n+1} \frac{\partial^2 F}{\partial \dot{q}_i \partial q_j}\dot{q}_j, \qquad i = 1, \ldots, n+1; \tag{A33}$$

by virtue of (A26), one can solve Eq. (A32) for $\dot{q}_1, \ldots, \dot{q}_n$. The remaining $n+1$ equations (A33) are used to compute the velocities $\dot{q}_{n+1}, \ldots, \dot{q}_{2n+1}$. In fact, using Eqs. (A24)–(A26) as well as the Ostrogradski nonsingularity condition $\frac{\partial^2 L}{\partial \dot{q}_{n+1}^2} \neq 0$, one easily finds

$$\det\left[\frac{\partial p_i}{\partial \dot{q}_i}\right]_{\substack{1\le i\le n+1\\ n+1\le j\le 2n+1}} \neq 0. \tag{A34}$$

In particular,

$$\dot{q}_i = f_i(q_1, \ldots, q_{2n+1}, p_{n+2}, \ldots, p_{2n+1}), \qquad i = 1, \ldots, n. \tag{A35}$$

Now, one can define the canonical transformation to Ostrogradski variables,

$$\tilde{q}_{2i-1} = q_i, \qquad i = 1, \ldots, n+1, \tag{A36}$$

$$\tilde{q}_{2i} = f_i(q_1, \ldots, q_{2n+1}, p_{n+2}, \ldots, p_{2n+1}), \qquad i = 1, \ldots, n, \tag{A37}$$

$$\tilde{p}_{2i-1} = p_i - \frac{\partial F}{\partial q_i}(q_1, f_1(\ldots), \ldots, q_n, f_n(\ldots), q_{n+1}, \ldots, q_{2n+1}), \qquad i = 1, \ldots, n+1, \tag{A38}$$

$$\tilde{p}_{2i} = -\frac{\partial F}{\partial f_i}(q_1, f_1(\ldots), \ldots, q_n, f_n(\ldots), q_{n+1}, \ldots, q_{2n+1}), \qquad i = 1, \ldots, n. \tag{A39}$$

The relevant generating function reads

$$\Phi(q_1, q_2, \ldots, q_{2n+1}, \tilde{p}_1, \tilde{q}_2, \tilde{p}_3, \tilde{q}_4, \ldots, \tilde{q}_{2n}, \tilde{p}_{2n+1}) = \sum_{k=1}^{n+1} \tilde{p}_{2k-1} q_k + F(q_1, \tilde{q}_2, \ldots, q_n, \tilde{q}_{2n}, q_{n+1}, \ldots, q_{2n+1}). \tag{A40}$$

Summarizing, we have found modified Lagrangian and Hamiltonian formulations of higher-derivative theories. They are equivalent to the Ostrogradski formalism in the sense that, on the Hamiltonian level, they are related to the latter by a canonical transformation. However, the advantage of this approach is that the Legendre transformation can be performed in a straightforward way.

[1] W. Thiring, Phys. Rev. **77**, 570 (1950).
[2] A. Pais and G. E. Uhlenbeck, Phys. Rev. **79**, 145 (1950).
[3] K. S. Stelle, Phys. Rev. D **16**, 953 (1977).
[4] E. S. Fradkin and A. A. Tseytlin, Nucl. Phys. **B201**, 469 (1982).
[5] M. S. Plyushchay, Mod. Phys. Lett. A **4**, 837 (1989); Mod. Phys. Lett. A **3**, 1299 (1988); Phys. Lett. B **243**, 383 (1990); Yu. A. Kuznetsov and M. S. Plyushchay, Nucl. Phys. **B389**, 181 (1993).
[6] M. Ostrogradski, Mem. Acad. St. Petersburg VI **4**, 385 (1850).
[7] J. M. Pons, Lett. Math. Phys. **17**, 181 (1989).
[8] T. Govaerts and M. S. Rashid, arXiv:hep-th/9403009.
[9] T. Nakamura and S. Hamamoto, Prog. Theor. Phys. **95**, 469 (1996).
[10] M. Henneaux and C. Teitelboim, *Quantization of Gauge Systems* (Princeton University Press, Princeton, NJ, 1992).
[11] A. De Felice and S. Tsujikawa, Living Rev. Relativity **13**, 3 (2010).
[12] N. Derulle, Y. Sendouda, and A. Youssef, Phys. Rev. D **80**, 084032 (2009).
[13] H. J. Schmidt, Phys. Rev. D **49**, 6354 (1994).
[14] H. J. Schmidt, arXiv:gr-qc/9501019.
[15] T.-C. Cheng, P.-M. Ho, and M.-C Yeh, Phys. Rev. D **66**, 085015 (2002).
[16] S. Hawking and T. Hertog, Phys. Rev. D **65**, 103515 (2002).
[17] K. Andrzejewski, J. Gonera, and P. Maślanka, arXiv:0710.2976.
[18] C. Battle, J. Gomis, J. M. Pons, and N. Roman-Roy, J. Phys. A **21**, 2693 (1988).
[19] V. V. Nesterenko, J. Phys. A **22**, 1673 (1989).
[20] P. Dunin-Barkowski and A. Sleptsov, arXiv:0801.4293.
[21] A. Mironov and A. Morozov, Theor. Math. Phys. **156**, 1209 (2008).A. Mironov and A. Morozov, Int. J. Mod. Phys. A **23**, 4677 (2008).D. Galakhov, JETP Lett. **87**, 452 (2008).