Stability and Hamiltonian formulation of higher derivative theories

Hans-Jürgen Schmidt
Universität Potsdam, Mathematisch-naturwiss. Fakultät Projektgruppe Kosmologie D-14415 Potsdam, PF 601553, Am Neuen Palais 10, Germany
(Received 7 September 1993)

We analyze the presuppositions leading to instabilities in theories of order higher than second. The type of fourth-order gravity which leads to an inflationary (quasi–de Sitter) period of cosmic evolution by inclusion of one curvature-squared term (i.e., the Starobinsky model) is used as an example. The corresponding Hamiltonian formulation (which is necessary for deducing the Wheeler-DeWitt equation) is found both in the Ostrogradski approach and in another form. As an example, a closed form solution of the Wheeler-DeWitt equation for a spatially flat Friedmann model and $L=R^2$ is found. The method proposed by Simon to bring fourth order gravity to second order can be (if suitably generalized) applied to bring sixth-order gravity to second order.

PACS number(s): 98.80.Cq, 04.50.+h

I. INTRODUCTION

It is a quite general belief that curvature-squared terms, if added to the Einstein-Hilbert action, describe semiclassically quantum corrections to general relativity. Further, there is no doubt that the existence of an inflationary period (exponential expansion of the cosmic scale factor) solves a lot of problems connected with the standard big bang model of the Universe. So it is no wonder that the Starobinsky model (curvature-squared terms lead automatically to the desired inflationary period) enjoyed so much interest in recent years. Now Simon and others have formulated some reasons against the Starobinsky model; the main reason is the fact that the field equation underlying the Starobinsky model is of fourth order.

It is the aim of the present paper to analyze those arguments which are connected with higher (higher than second) derivative theories.

Reference [1] discusses in its Sec. 2 the "fundamental problems of nonlocality through the higher derivative limiting procedure." The principal result of its Sec. 2.1 is that at least $N-1$ of the solutions of a nondegenerate theory of order $2N$ carry negative energy. Eliezer and Woodard write "The energy is therefore unbounded below for all nondegenerate, higher derivative theories." This leads to the instability observed in almost all fourth- and higher-order theories. The Starobinsky cosmological model follows from fourth-order gravity, and so it seems to be a candidate for such an unstable theory (see, e.g., Ref. [2]).

We analyze that part of the arguments which is connected with the higher order. To this end we specialize in the Ostrogradski approach [3] (which is a method to bring a higher-order Lagrangian into Hamiltonian form—more recent work on this topic can be found in Ref. [4]) to fourth-order theories in Sec. II and give some intuitive examples. In Sec. III we discuss the question of whether fourth-order theories lead to a minimum or only to a saddle point of the action. In Sec. IV a method different from Ostrogradski's is proposed to bring fourth-order equations in a Hamiltonian form.

Then we are prepared to consider the Starobinsky model [5] in Sec. V. The main problem comes from the R^2 term, and so we simplify in Sec. V A by discussing the high-curvature limit and derive the corresponding Wheeler-DeWitt equation by the method described in Sec. IV.

Section V B discusses the question of the superfluous degrees of freedom of fourth-order gravity, and Sec. V C gives the Starobinsky model in form of a power series not yet found in the literature.

Section VI is on sixth- and higher-order gravity. It is included to show which kinds of problems additionally appear, if $L=R+c_0R^2$ gravity is intended to be the $k=0$ truncation of a power series:

$$L=R+\sum_{i=0}^{k} c_i R\Box^i R \ . \tag{1.1}$$

We look for the Newtonian limit of that theory and generalize Simon's approach [2] to this Lagrangian (1.1) truncated at $k=1$.

Section VII discusses the results.

II. OSTROGRADSKI'S METHOD FOR A FOURTH-ORDER SYSTEM

We follow Ostrogradski [3], but use the notation published in Ref. [1], which is more familiar to the present reader, and we specialize always in fourth-order theories which follow from a nondegenerate Lagrangian of second order. So we consider a one-dimensional point particle with position $q(t)$ at time t. A dot denotes d/dt, and the Lagrangian is of the type

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

where $q\in I$, $I\neq\emptyset$ being a connected open subset of the space R of all reals, and $\dot{q},\ddot{q}$ are allowed to cover all the reals. The momentum P_2 is defined by

$$P_2 = \frac{\partial L}{\partial \ddot{q}} \ . \tag{2.2}$$

In [1] the Lagrangian L is defined to be nondegenerate if this Eq. (2.2) can be solved for $\ddot{q}$, which takes place, loosely speaking, iff $\partial P_2/\partial\ddot{q} \neq 0$. To avoid discussions for the case that (2.2) can be solved, but not uniquely, we additionally require that

$$\frac{\partial P_2}{\partial \ddot{q}} = F(q,\dot{q}) \ .$$

Under this circumstance the Lagrangian is nondegenerate if and only if F does not have any zeros; i.e., it is a map $F: I \times R \to R \setminus \{0\}$. This we shall assume in the following. Then the Lagrangian (2.1) can be written as

$$L = \tfrac{1}{2}(\ddot{q})^2 F(q,\dot{q}) + \ddot{q}G(q,\dot{q}) + K(q,\dot{q}) \ . \tag{2.3}$$

To avoid discussions of differentiability, we simply require the three functions $F \neq 0, G, K$ to be real analytic ones. Then Eq. (2.2) becomes

$$P_2 = \ddot{q}F(q,\dot{q}) + G(q,\dot{q}) \tag{2.4}$$

and it can be uniquely inverted to

$$\ddot{q} = \frac{P_2 - G(q,\dot{q})}{F(q,\dot{q})} \ . \tag{2.5}$$

The Euler-Lagrange equation following from Eq. (2.1) reads

$$\frac{\delta L}{\delta q} \equiv \frac{\partial L}{\partial q} - \frac{d}{dt}\frac{\partial L}{\partial \dot{q}} + \frac{d^2}{dt^2}\frac{\partial L}{\partial \ddot{q}} = 0 \ . \tag{2.6}$$

Subsequently, we write $q^{(0)} = q$ and $q^{(n+1)} = \dot{q}^{(n)}$. Inserting Eq. (2.3) into Eq. (2.6), we get an equation of the structure,

$$0 = q^{(4)}F(q^{(0)},q^{(1)}) + J(q^{(0)},q^{(1)},q^{(2)},q^{(3)}) \ , \tag{2.7}$$

where J is a real analytic function composed of F, G, and K. From Eq. (2.7) the notion of nondegeneracy becomes apparent: The second-order Lagrangian (2.1) is nondegenerate iff the Euler-Lagrange equation is a regular fourth-order equation. The fact that we restricted the domain of q to the subset I of R is no real restriction because by a real analytic redefinition $\tilde{q}(q)$ we could get $\tilde{I} = R$.

To get a Hamiltonian for this system, one needs a first-order formulation with two coordinates. We define them as

$$Q^1 = q, \quad Q^2 = \dot{q} \ . \tag{2.8}$$

The two conjugate momenta are

$$P_1 = \frac{\partial L}{\partial \dot{q}} - \frac{d}{dt}\frac{\partial L}{\partial \ddot{q}} \ , \tag{2.9}$$

and P_2 is defined by Eqs. (2.2) and (2.4). The Hamiltonian $H = H(Q^1,Q^2,P_1,P_2)$ is obtained via

$$H = -L + \sum_{n=1}^{2} P_n \dot{Q}^n \ . \tag{2.10}$$

With these definitions the canonical equations

$$\dot{Q}^n = \frac{\partial H}{\partial P_n}, \quad \dot{P}_n = -\frac{\partial H}{\partial Q^n} \tag{2.11}$$

take place, and their validity implies the validity of the Euler-Lagrange equation (2.6). Now the arguments of F,G,K are (Q^1,Q^2) and we get

$$\dot{Q}^1 = Q^2, \quad \dot{Q}^2 = (P_2 - G)/F \tag{2.12}$$

and, after some calculus,

$$H = P_1 Q^2 + \frac{1}{2F}(P_2 - G)^2 - K \ , \tag{2.13}$$

where $dH/dt = 0$ follows from Eq. (2.6). So H can be called the energy of the system. The essential point is that the energy is unbounded both below and above. This is directly seen from Eq. (2.13) because H is a linear function in P_1.

Remark: In Ref. [1] it was argued that the problem lies in the fact that energy is unbounded below. More exactly, one should say unbounded below and above; suppose the energy is unbounded below and bounded above; then, we simply change the signs of both L and H and get the energy bounded below. This is possible because no sign of H is preferred *a priori*, in contrast with classical mechanics where the sign of H is defined by the condition that kinetic energy be non-negative.

$H =$ const represents a first integral of Eq. (2.6). It is, as must be the case, a third-order equation for $q(t)$, and it has the structure

$$H = -q^{(1)}q^{(3)}F + \text{lower order terms} \ . \tag{2.14}$$

There is a singular point at $\dot{q} = 0$. The definition Eq. (2.13) of H is essentially (i.e., up to invertible linear transformations of L and H which do not change the dynamics) unique, because time-independent canonical transformations do not change it. That H is unbounded both below and above can be seen from Eq. (2.14): Fixing the initial values $q, \dot{q} \neq 0, \ddot{q}$, one can freely choose $q^{(3)}$ [cf. Eq. (2.7)] and get H as unbounded.

The instability following from H being unbounded below and above can be described as follows. (A) In the particle picture one gets particles with positive energy and particles with negative energy. Then the unlimited production of pairs of such particles is not prevented by energy conservation. (B) In the four-parameter set of solutions of the Euler-Lagrange equation, one gets a subset of dimension ≥ 1 of solutions with negative energy. Let us make this last point more explicit.

To this end we first consider what happens if we add such a total derivative to the Lagrangian that the functional dependence does not change. This is done by

$$\tilde{L} = L + \frac{d}{dt}[M - Et] \ , \tag{2.15}$$

where E is a constant and M depends on q and $\dot{q}$. One gets

$$\tilde{P}_1 = P_1 + \frac{\partial M}{\partial q}, \quad \tilde{P}_2 = P_2 + \frac{\partial M}{\partial \dot{q}} \ , \tag{2.16}$$

so that the condition of invertibility of P_2 does not change, and

$$\tilde{F}=F, \quad \tilde{H}=H+E \; . \tag{2.17}$$

So this transformation, too, does not change the properties discussed.

Let us continue with the discussion of negative energy solutions. We assume that $q=0$ is a solution, and we fix E such that $q=0$ is a zero-energy solution.

Remark: In the first-order Lagrangian $L=\frac{1}{2}\dot{q}^2-(A/2)q^2$, one has $p=\dot{q}$ and $H=\frac{1}{2}p^2+(A/2)q^2$. For $A=1$, one has the solutions $q=\sin t$ and $q=\cos t$, which both have energy $H=\frac{1}{2}$. For $A=-1$, however, one has the solutions $q=\sinh t$ and $q=\cosh t$, which have energy $H=\frac{1}{2}$ and $-\frac{1}{2}$, respectively. They sum up to the zero-energy solution $q=e^t$. This is the instability meant.

For the second-order Lagrangian, we consider only the terms up to second degree in the arguments. The terms $\dot{q}$, $\ddot{q}$, $\dot{q}\ddot{q}$, $q\dot{q}$, and $q\ddot{q}+\dot{q}^2$ represent total derivatives, and we use them to bring the general form to

$$L=\frac{1}{2}\ddot{q}^2+\frac{A}{2}\dot{q}^2-\frac{B}{2}q^2 \; . \tag{2.18}$$

The corresponding Hamiltonian becomes

$$H=\frac{1}{2}\ddot{q}^2+\frac{A}{2}\dot{q}^2+\frac{B}{2}q^2-\dot{q}q^{(3)} \; . \tag{2.19}$$

The Euler-Lagrange equation reads

$$0=q^{(4)}-A\ddot{q}-Bq \; . \tag{2.20}$$

The momenta are

$$P_1=A\dot{q}-q^{(3)}, \quad P_2=\ddot{q} \; .$$

For $A=B=0$, one gets the positive energy solution $q=t^2$ and the negative energy solution $q=t^3+t$.

For $B=0$, $A=\pm 1$, the solution

$$q=\alpha t+\beta s(t)+\gamma c(t) \; ,$$

where $s(t)=\sinh t$ for $A=1$, $s(t)=\sin t$ for $A=-1$, analogously $c(t)$, has the energy

$$H=\pm\tfrac{1}{2}(\alpha^2-\beta^2)+\tfrac{1}{2}\gamma^2 \; .$$

For $A=0$, $B=1$, the general solution of Eq. (2.20) reads

$$q=\alpha\sin t+\beta\cos t+\gamma\sinh t+\delta\cosh t \; .$$

One gets

$$H=\alpha^2+\beta^2-\gamma^2+\delta^2 \; .$$

The general case shows similarly that both signs of the energy appear. For nonlinear equations, of course, the solutions do not simply add, but the behavior of the signs of the energy is similar.

III. MINIMAL, NOT ONLY STATIONARY ACTION

Instability may occur if the action is not minimal, but only stationary. We shall check whether this type of instability can occur in the second-order Lagrangian discussed. More on this topic, especially applied to fourth-order gravity, can be found in Ref. [6]; however, the point here is only to convince the reader that requiring minimality of the action does not trivially rule out fourth-order theories. To do so, we develop the action $S[q+\epsilon h]$ defined by

$$S[q]=\int_0^T L\,dt \tag{3.1}$$

into powers of ϵ, where L is the same as in Eq. (2.1) and $T>0$. Without loss of generality, the initial point of time was put $t=0$. Let h be any differentiable function satisfying $h(0)=\dot{h}(0)=h(T)=\dot{h}(T)=0$. After partial integration and use of the notation Eq. (2.6), we get

$$S[q+\epsilon h]=S[q]+\epsilon\int_0^T h\frac{\delta L}{\delta q}dt+\frac{\epsilon^2}{2}V[q,h]+O(\epsilon^3) \; , \tag{3.2}$$

where

$$V[q,h]=\int_0^T\left[h^2\frac{\partial^2 L}{\partial q^2}+\dot{h}^2\frac{\partial^2 L}{\partial\dot{q}^2}+\ddot{h}^2\frac{\partial^2 L}{\partial\ddot{q}^2}+2h\dot{h}\frac{\partial^2 L}{\partial q\,\partial\dot{q}}+2h\ddot{h}\frac{\partial^2 L}{\partial q\,\partial\ddot{q}}+2\dot{h}\ddot{h}\frac{\partial^2 L}{\partial\dot{q}\,\partial\ddot{q}}\right]dt \; .$$

In dealing with V, partial integration does not help. So one should discuss it directly.

Remark: Before discussing the fourth-order case, let us repeat the behavior for the harmonic oscillator $L=\frac{1}{2}\dot{q}^2-\frac{1}{2}q^2$. One gets

$$V=\int_0^T(\dot{h}^2-h^2)dt \; .$$

One only needs the boundary conditions $h(0)=h(T)=0$. From the first glance, V seems to possess only a saddle point, but a Fourier analysis with

$$h=\sum_{n=1}^{\infty}a_n\sin(n\pi t/T)$$

gives

$$V=\frac{T}{2}\sum_{n=1}^{\infty}\left[\frac{\pi^2n^2}{T^2}-1\right]a_n^2 \; .$$

For $T<\pi$, this is positive definite, for $T=\pi$, it is positive semidefinite, and only for $T>\pi$ does it become a saddle point. The harmonic oscillator is *the* standard example of a stable model, and so one should require the action to be locally minimal, i.e., minimal if considered over sufficiently short but finite time intervals.

Let us now come to the Lagrangian Eq. (2.18). Inserting it into Eqs. (3.1) and (3.2), we get

$$V=\int_0^T(\ddot{h}^2+A\dot{h}^2-Bh^2)dt \; . \tag{3.3}$$

We perform the same Fourier analysis as in the previous example and get the following. The maximally allowed value T depends on A and B, but it is always positive, so that one has the same kind of stability here: If the time interval considered is sufficiently short, then each stationary point of the action represents a minimum.

IV. ANOTHER HAMILTONIAN FORMALISM

In addition to Ostrogradski's approach discussed in Sec. II, there exists another possibility to get a Hamiltonian from a higher-order Lagrangian. It has the advantage that the relation from classical mechanics,

$$p_i = \frac{\partial L}{\partial \dot{q}^i} , \tag{4.1}$$

remains valid, whereas Ostrogradski changed it to Eq. (2.9).

Further possibilities to get a Hamiltonian are discussed in Ref. [7] (see also the references cited there); the difference is that in [7] there is always a constraint, whereas we look for a method where no additional constraint must be introduced.

Again, we start from Eq. (2.1) and concentrate on Lagrangians of the type (2.3). The difference is now that the new coordinates are chosen to be

$$q^1 = q, \quad q^2 = \ddot{q} . \tag{4.2}$$

So the dependence of $L(q^1, \dot{q}^1, q^2, \dot{q}^2)$ is unique, whereas by use of Eq. (2.8), there is an ambiguity between $\dot{Q}^1$ and Q^2.

To show up the procedure, we find it more appropriate to concentrate on one single Lagrangian; the general procedure might become clear from it. Moreover, it is just that Lagrangian which will appear in Sec. V. So we take

$$L = (\ddot{q})^2 e^{3q} . \tag{4.3}$$

To prevent an identical vanishing of p_2 according to Eqs. (4.1)–(4.3), we add a suitable total derivative to the Lagrangian (4.3). Let us first take

$$\tilde{L} = L + \frac{4}{3}\frac{d}{dt}[(\dot{q})^3 e^{3q}] , \tag{4.4}$$

i.e., an equation to be used in Sec. V A:

$$\tilde{L} = [(\ddot{q})^2 + 4\ddot{q}(\dot{q})^2 + 4(\dot{q})^4] e^{3q} . \tag{4.5}$$

In a second step we take

$$\hat{L} = L - 2\frac{d}{dt}(\dot{q}\ddot{q}e^{3q}) , \tag{4.6}$$

i.e.,

$$\hat{L} = -[(\ddot{q})^2 + 2\dot{q}q^{(3)} + 6(\dot{q})^2\ddot{q}] e^{3q} . \tag{4.7}$$

We insert Eq. (4.2) into Eq. (4.7) and get

$$\hat{L} = -[(q^2)^2 + 2\dot{q}^1\dot{q}^2 + 6(\dot{q}^1)^2 q^2] \exp(3q^1) , \tag{4.8}$$

where we use $\dot{q}^n \equiv (d/dt)(q^n)$.

Applying (4.1), we get

$$\begin{aligned} p_1 &= -[2\dot{q}^2 + 12\dot{q}^1 q^2] \exp(3q^1) , \\ p_2 &= -2\dot{q}^1 \exp(3q^1) . \end{aligned} \tag{4.9}$$

The Jacobian is

$$\frac{\partial(p_1, p_2)}{\partial(\dot{q}^1, \dot{q}^2)} = -4\exp(3q^1) . \tag{4.10}$$

This differs from zero, and so we can invert Eq. (4.9) to

$$\begin{aligned} \dot{q}^1 &= -\tfrac{1}{2} p_2 \exp(-3q^1) , \\ \dot{q}^2 &= [3p_2 q^2 - \tfrac{1}{2} p_1] \exp(-3q^1) . \end{aligned} \tag{4.11}$$

We get from Eqs. (4.8) and (2.10) with the help of Eqs. (4.11) the Hamiltonian

$$H = -\tfrac{1}{2}(p_1 p_2 - 3p_2^2 q^2)\exp(-3q^1) + (q^2)^2 \exp(3q^1) . \tag{4.12}$$

It is essential to observe that the equation $\ddot{q}^1 = q^2$ follows from the canonical equations of H [Eq. (4.12)] without imposing it as an additional constraint. So Eqs. (4.2) become automatically compatible.

To give some feeling for Hamiltonians with negative kinetic energy, we give six typical examples, hoping that this gives better insight than general formulations do.

Example 1. Let k be a parameter satisfying $0 < k < 1$. We consider the Lagrangian

$$L = (\ddot{q})^2 - 2(\dot{q})^2 + kq^2 \tag{4.13}$$

for a one-dimensional point particle $q(t)$. The Euler-Lagrange equation reads

$$0 = q^{(4)} + 2\ddot{q} + kq . \tag{4.14}$$

We insert the ansatz

$$q = e^{\lambda t}, \quad 0 = \lambda^4 + 2\lambda^2 + k \tag{4.15}$$

into Eq. (4.14), which leads to

$$\lambda = \pm i(1 \pm \sqrt{1-k}\,)^{1/2} , \tag{4.16}$$

representing four different purely imaginary numbers. Therefore the general solution of Eq. (4.14) can be written as

$$q(t) = \sum_{n=1}^{2} c_n \sin\{t_n + t[1 + (-1)^n \sqrt{1-k}\,]^{1/2}\} , \tag{4.17}$$

where c_1, c_2, t_1, t_2 are the four integration constants. Each solution is bounded in time. In the limiting case $k = 0$, the unbounded function $q(t) = t$ is a solution. In the other limiting case $k = 1$, $q(t) = t \sin t$ is also an unbounded solution.

Example 2. Let ϵ be a parameter satisfying $0 < \epsilon < 1$. Let

$$H = \tfrac{1}{2}p^2 + \tfrac{1}{2}q^2 + \tfrac{1}{2}P^2 + \tfrac{1}{2}Q^2 + \epsilon qQ \tag{4.18}$$

be a Hamiltonian for two one-dimensional point particles q, Q [or, equivalently, one two-dimensional particle with coordinates (q, Q)]; p is the momentum corresponding to q, P to Q. Because of the restriction put on ϵ, H is positive definite in all its arguments. For $\epsilon = 0$, this is nothing but two independent harmonic oscillators of frequency 1. Only the term $\sim \epsilon$ introduces some interaction.

The canonical equations following from Eq. (4.18) are

$$\frac{\partial H}{\partial p}=\dot{q}=p, \quad \frac{\partial H}{\partial q}=-\dot{p}=q+\epsilon Q=-\ddot{q} \ , \tag{4.19}$$

$$\frac{\partial H}{\partial P}=\dot{Q}=P, \quad \frac{\partial H}{\partial Q}=-\dot{P}=Q+\epsilon q=-\ddot{Q} \ . \tag{4.20}$$

To integrate the system, it proves useful to eliminate Q by use of Eq. (4.19) as follows:

$$Q=-\frac{1}{\epsilon}(\ddot{q}+q) \ . \tag{4.21}$$

This leads with Eq. (4.20) to

$$0=q^{(4)}+2\ddot{q}+(1-\epsilon^2)q \ . \tag{4.22}$$

With $k=1-\epsilon^2$, we meet exactly the system (4.14) from example 1. The result of example 1 is in agreement with the Kolmogorov-Arnowitt-Moser (KAM) theorem, which applies to the system considered here and states that there exists an interval of positive ϵ values such that the corresponding system is solved by toruslike (i.e., periodic) solutions. Each arbitrarily small value ϵ gives rise to a bifurcation of the frequency according to $|\lambda|$ in example 1. Each solution is periodic and, hence, bounded. In the limiting case $\epsilon\to 0$ corresponding to $k\to 1$, the equivalence of example 2 to example 1 breaks, because for $\epsilon=0$ all solutions remain bounded here. That this equivalence breaks as $\epsilon=0$ becomes also plausible from the relation (4.21) between Q and q.

H can be considered to be the energy of the system. Let us express it as function of q and its derivatives alone (calculations have been done by REDUCE 3.41):

$$2H\epsilon^2=q[q+2\ddot{q}](1-\epsilon^2)+[q^{(3)}]^2+2\dot{q}q^{(3)} +(\ddot{q})^2+(\dot{q})^2(1+\epsilon^2) \ . \tag{4.23}$$

A direct calculation leads to

$$\epsilon^2\frac{dH}{dt}=\dot{Q}[q^{(4)}+2\ddot{q}+(1-\epsilon^2)q] \ ; \tag{4.24}$$

hence, $H=\text{const}$ follows from the q equation (4.22), but $H=\text{const}$ implies a solution for $\dot{Q}\neq 0$ only.

Let us mention that the fourth-order equation of example 1 is equivalent to the positive-definite Hamiltonian of example 2.

Example 3. Let us start again from the system of example 1, but now we apply the Ostrogradski approach to make a Hamiltonian from it. New coordinates are $Q^1=q$, $Q^2=\dot{q}$ [see Eq. (2.8)]; new momenta are

$$P_1=\frac{\partial L}{\partial \dot{q}}-\frac{d}{dt}\frac{\partial L}{\partial \ddot{q}}=-4\dot{q}-2q^{(3)}, \quad P_2=\frac{\partial L}{\partial \ddot{q}}=2\ddot{q} \tag{4.25}$$

[see Eqs. (2.9) and (2.2)]. The Hamiltonian is

$$\tilde{H}=-L+\sum_{n=1}^{2}P_n\dot{Q}^n=P_1Q^2+\tfrac{1}{4}P_2^2+2(Q^2)^2-k(Q^1)^2 \ . \tag{4.26}$$

Now let us forget about the origin of $\tilde{H}$ and calculate the canonical equations. Inserting $Q^1=q$, we get after some calculus $Q^2=\dot{q}$, $P_1=-4\dot{q}-2q^{(3)}$, $P_2=2\ddot{q}$, and finally

$$0=q^{(4)}+2\ddot{q}+kq \ , \tag{4.27}$$

i.e., as it should be, just Eq. (4.14) of example 1.

We insert the values for P_n, Q^n into $\tilde{H}$ and get

$$\tilde{H}=(\ddot{q})^2-2\dot{q}q^{(3)}-2(\dot{q})^2-kq^2 \ . \tag{4.28}$$

This is also a conserved quantity:

$$\frac{d\tilde{H}}{dt}=-2\dot{q}[q^{(4)}+2\ddot{q}+kq] \ , \tag{4.29}$$

but $\tilde{H}$ is not evidently bounded. Clearly, each solution remains bounded, but the set of solutions for a fixed value $\tilde{H}$ need not to be bounded: Let us insert the solution (4.17) of example 1 with $t_1=t_2=0$; then, one gets

$$\tilde{H}=2\epsilon[c_2^2(1+\epsilon)-c_1^2(1-\epsilon)] \ . \tag{4.30}$$

$\tilde{H}$ can take each real value, and with arbitrarily fixed value $\tilde{H}$, we can find an unbounded set of functions solving for just this $\tilde{H}$.

Let us compare this result with the analogous calculations in example 2: There one gets, from the same initial conditions,

$$H=c_2^2(1+\epsilon)+c_1^2(1-\epsilon) \ . \tag{4.31}$$

Up to the inessential prefactor 2ϵ, it is just the other sign in front of c_1^2 which makes the difference. Here $H\geq 0$ and the set of solutions for a fixed value H forms a compact set of bounded functions; moreover, it is a uniformly bounded set of functions. So the conserved quantity $\tilde{H}$ should not be considered as the energy of the system, because H better meets the point. This is another argument against using the Ostrogradski approach. Let us further mention that the Poisson brackets of these two conserved quantities $H,\tilde{H}$ identically vanish, and so it does not give rise to a further conserved quantity.

Example 4. We take now the same $\tilde{H}$ as in example 3, but we interchange coordinates and momenta ($P_1\to q$, $Q^1\to -p$, $P_2\to Q$, $Q^2\to -P$):

$$\hat{H}=2P^2-kp^2-qP+\tfrac{1}{4}Q^2 \ . \tag{4.32}$$

Here the kinetic energy is indefinite, but the solutions are, of course, also only the periodic ones of example 1. Here we see that two different Hamiltonians may describe the same system: One has indefinite kinetic energy, and the other has definite kinetic energy. On the other hand, a regular Hamiltonian must have a nonvanishing Jacobian

$$J=\frac{\partial(\dot{q},\dot{Q})}{\partial(p,P)}=\frac{\partial^2 H}{\partial p^2}\frac{\partial^2 H}{\partial P^2}-\left[\frac{\partial^2 H}{\partial p\,\partial P}\right]^2 \ . \tag{4.33}$$

Here $J<0$. [This J is the inverse of the Jacobian used in Eq. (4.10); this does not matter since only the sign of J is essential here.]

A one-parameter family of regular canonical transformations connected with the identity transformation cannot change the sign of J. Therefore the Hamiltonians H (example 2) and $\hat{H}$ (example 4) cannot be continuously deformed into each other by such a transformation, because, in example 2, H is positive definite and $J>0$. Nevertheless, they describe the same system.

Example 5. Let us take the Hamiltonian

$$\bar{H}=Pp+\frac{\epsilon}{2}(q^2+Q^2)+qQ,\quad 0<\epsilon<1\ . \tag{4.34}$$

Both the kinetic and potential parts are indefinite. The canonical equations give

$$P=\dot{q},\quad p=\dot{Q},\quad \dot{P}=-q-\epsilon Q,\quad \dot{p}=-Q-\epsilon q\ . \tag{4.35}$$

After some calculus we get

$$0=q^{(4)}+2\ddot{q}+(1-\epsilon^2)q\ , \tag{4.36}$$

which is again the previously discussed system.

Example 6. Now we start from a Lagrangian which differs from example 2 only in two changes of a sign:

$$H^*=\tfrac{1}{2}p^2+\tfrac{1}{2}q^2-\tfrac{1}{2}P^2-\tfrac{1}{2}Q^2+\epsilon qQ,\quad 0<\epsilon<1\ . \tag{4.37}$$

So both kinetic and potential energies are indefinite. But as was seen in example 5, this does not exclude the equivalence.

Let us first consider the limiting case $\epsilon=0$. Here, again, it is fully equivalent to the second example: There is no interaction between the two oscillators, and there is no *a priori* sign preferred for the energy. H and H^* are two conserved quantities, whose Poisson brackets vanish.

The situation drastically changes if we come back to $\epsilon>0$. One of the assumptions of the KAM theorem is no longer valid, and so we expect qualitatively different solutions for arbitrarily small values ϵ. In the particle picture one can imagine the following: The spontaneous creation of pairs of particles, one with positive energy, the other with negative energy, is energetically allowed, and it *should* take place with a typical doubling time $\sim 1/\epsilon$.

For fixed energy, arbitrarily large momenta are possible. We perform the calculations analogous to the previous ones. We can prevent any calculations if we look at H and H^*: Multiplying P and Q by i and multiplying ϵ by $(-i)$, one is changed into the other. So, clearly, the other formulas are valid if ϵ^2 is replaced by $(-\epsilon^2)$. Then the dynamics follows from

$$0=q^{(4)}+2\ddot{q}+kq,\quad k=1+\epsilon^2\ , \tag{4.38}$$

and it is example 1 with $k>1$. The corresponding fourth-order polynomial for λ is then solved by

$$\lambda=\pm i\sqrt{1\pm i\epsilon}=\pm\frac{\epsilon}{2}\pm i\left[1+\frac{\epsilon^2}{8}\right]+O(\epsilon^3)\ . \tag{4.39}$$

The four solutions correspond to the four combinations of the signs $\pm$. Therefore the general solution can be written as

$$q(t)=\sum_{n=1}^{2} c_n \exp\left[t(-1)^n\left[\frac{\epsilon}{2}+O(\epsilon^3)\right]\right] \times\sin\left[t_n+t\left[1+\frac{\epsilon^2}{8}+O(\epsilon^4)\right]\right]\ , \tag{4.40}$$

where c_1,c_2,t_1,t_2 are the four integration constants. (By the way, the $O(\epsilon^3)=[\frac{9}{64}+O(\epsilon^2)]\epsilon^3$ for this formula.) $q(t)\equiv 0$ is the only bounded solution, and for $c_2\neq 0$, one gets an exponential increase as expected. This is, of course, a resonance effect. If, on the other hand, H^* is altered by a suitable positive factor in front of p^2, then for small ϵ, the general solution remains periodic, but the periods are mixed.

V. STAROBINSKY MODEL

In Ref. [5], Starobinsky proposed to use

$$L=\left[\frac{R}{2}-\frac{l^2}{12}R^2\right]\sqrt{-g} \tag{5.1}$$

as gravitational Lagrangian. Here R is the curvature scalar, g the determinant of the metric of space-time, and l is a length being somehow in the region $l=10^{-28}$ cm.

A. High-curvature limit

Let us first consider the high-curvature limit

$$\tilde{L}=\tfrac{1}{36}R^2\sqrt{-g}\ . \tag{5.2}$$

For the metric of a spatially flat Friedmann model,

$$ds^2=dt^2-e^{2q(t)}(dx^2+dy^2+dz^2)\ , \tag{5.3}$$

we get

$$R=-6\ddot{q}-12(\dot{q})^2,\quad g=-e^{6q}\ , \tag{5.4}$$

and the Lagrangian becomes

$$\tilde{L}=[\ddot{q}+2(\dot{q})^2]^2e^{3q}\ , \tag{5.5}$$

which coincides with Eq. (4.5).

Now we could apply both the Ostrogradski approach Sec. II as well as the approach of Sec. IV. We prefer to use the latter one because of the validity of Eq. (4.1), but for comparison we write down both of them.

Let us first apply the Ostrogradski approach to Eq. (5.5). Looking at Eqs. (4.3)–(4.5), one can see that (up to a divergence) we have to consider

$$L=(\ddot{q})^2e^{3q}\ . \tag{5.6}$$

Applying the formalism of Sec. II, we get

$$Q^1=q,\quad Q^2=\dot{q},\quad P_1=-2\frac{d}{dt}(\ddot{q}e^{3q}),\quad P_2=2\ddot{q}e^{3q}\ ,$$

and then

$$H=P_1Q^2+\tfrac{1}{4}(P_2)^2\exp(-3Q^1)\ , \tag{5.7}$$

i.e.,

$$H=e^{3q}[(\ddot{q})^2-2\dot{q}(q^{(3)}+3\dot{q}\ddot{q})]\ . \tag{5.8}$$

As is expected, the canonical equations to the Hamiltonian H [Eq. (5.7)] give again the original system, where H [Eq. (5.8)] represents a conserved quantity. Moreover, one knows that the gravitational field equation (here especially its zero-zero component) forces H to vanish. This can easiest be shown by making the ansatz $dt^2=N^2(\tau)d\tau^2$ and putting $N=1$ only after the variation (and not before as we did).

Let us now come to the Wheeler-DeWitt equation (i.e., the zero-energy Schrödinger equation of a cosmological model) for this system. In units where $\hbar=1$, it is obtained via substituting P_n by $i\partial_n \equiv i\partial/\partial Q^n$. Applying this to Eq. (5.7), one can see that the fact that P_1 is contained linearly and not quadratically gives as a consequence that one of the coefficients of the Wheeler-DeWitt equation fails to be real. A third reason against this approach is the fact that $Q^2=\dot{q}$ [which is just the Hubble parameter of the cosmological model Eq. (5.3)] is not invariantly defined; it changes its sign by a change of the time direction.

Now we try the same with the formulas of Sec. IV. In fact, we can work directly with Eqs. (4.3) and (4.6)–(4.12). If we reinsert Eqs. (4.2) and (4.9) into Eq. (4.12), we get exactly Eq. (5.8). So no second conserved quantity appears here, and the Wheeler-DeWitt equation is derived as follows.

The material from Eq. (5.9) to (5.13) is taken from Ref. [8], and from Eq. (5.14) to (5.15) is taken from Ref. [9].

The Lagrangian $\hat{L}$ [Eq. (4.8)] has the structure

$$\hat{L}=\tfrac{1}{2}g_{ij}\dot{q}^i\dot{q}^j-V(q^i)\ , \tag{5.9}$$

where g_{ij} depends on the q^i only and the Einstein sum convention is applied. One gets

$$p_i=\frac{\partial L}{\partial \dot{q}^i}=g_{ij}\dot{q}^j\ . \tag{5.10}$$

To apply the Hamiltonian formalism, it is necessary to invert Eq. (5.10) such that the velocities are written with dependence on coordinates and momenta. This is possible if and only if g_{ij} is an invertible matrix. This takes place for the case considered here: A comparison of (4.8) with (5.9) gives

$$g_{11}=-6q^2\exp(3q^1),\quad g_{12}=-\exp(3q^1),\quad g_{22}=0\ .$$

Let g^{ij} be the inverse matrix to g_{ij}. Then Eq. (5.10) can be inverted to

$$\dot{q}^i=g^{ij}p_j\ . \tag{5.11}$$

In the interesting case (4.8), this gives

$$g^{22}=6q^2\exp(-3q^1),\quad g^{12}=-\exp(-3q^1),\quad g^{11}=0\ .$$

The Hamiltonian becomes

$$H=p_i\dot{q}^i-L=\tfrac{1}{2}g^{ij}p_ip_j+V(q^i)\ , \tag{5.12}$$

and here $V=(q^2)^2\exp(3q^1)$. If we quantize now by substituting p_n with $i\partial/\partial q^n$, then the procedure is no longer covariant, and the factor-ordering problem appears. In classical mechanics this problem is absent, because g_{ij} is a constant matrix. We circumvent the problem by substituting p_n by $i\nabla_n$ where ∇_n denotes the covariant derivative into the q^n direction with respect to the metric g_{ij}. Then the Wheeler-DeWitt equation reads

$$0=(\Box-2V)\psi(q^i)\ , \tag{5.13}$$

where ψ is the world function and

$$\Box=\nabla_i\nabla^i=\frac{1}{\sqrt{-g}}\partial_i\sqrt{-g}\,g^{ij}\partial_j$$

is the D'Alembertian. For our example we get

$$0=[3\partial_2 q^2\partial_2-\partial_1\partial_2-2(q^2)^2\exp(6q^1)]\psi\ . \tag{5.14}$$

To simplify, let us apply the transformation

$$\sigma=\exp(-3q^1),\quad \tau=\tfrac{3}{2}q^2\exp(3q^1)\ .$$

This transformation explicitly brings g_{ij} to the flat form $g_{\sigma\tau}=1$, $g_{\sigma\sigma}=0$, $g_{\tau\tau}=0$. The Lagrangian becomes

$$L=\frac{1}{\sigma}\dot{\sigma}\dot{\tau}-\tau^2\sigma\ .$$

The Hamiltonian is correspondingly

$$H=\sigma[\pi_\tau\pi_\sigma+\tau^2]\ ,$$

and the Wheeler-DeWitt equation reads

$$0=[\partial_\tau\partial_\sigma-\tau^2]\psi(\sigma,\tau)\ . \tag{5.15}$$

This linear differential equation can be solved in closed form by

$$\psi=\int_{-\infty}^{\infty}a(\lambda)\exp\left[\lambda\sigma+\frac{\tau^3}{3\lambda}\right]d\lambda\ , \tag{5.16}$$

where the amplitude function a, $a(0)=0$, can be arbitrarily chosen both as continuous as well as a sum of δ functions.

B. Superfluous degrees of freedom

We look at higher (higher than second) order gravity theories under the point of view that the higher order yields more degrees of freedom than is to be expected. Further results on higher-order gravity and inflationary phase of cosmic evolution can be found in Refs. [10–24] (a list which is not intended to be representative, but essentially contains the papers we refer to in the subsequent text). Let us start with some historical comments taken from Ref. [14].

In [25], Weyl proposes a new theory which is intended to unify gravitation with electromagnetism. He generalizes both the Riemannian geometry underlying the general relativity theory of Einstein to a nonintegrable theory (introduction of the Weyl vector) as well as the Einstein-Hilbert Lagrangian (which is linear in curvature) to a Lagrangian quadratic in curvature. At that point Weyl already realized [25, p. 477] that "This has the consequence, though our theory leads to Maxwell's electromagnetic equations, it fails to lead to Einstein's gravitational ones; instead of them, 4th order differential equations appear." We cite this sentence to show that already in 1918 the possibility of fourth-order gravitational field equations had been discussed as an alternative to general relativity. One year later, Pauli [26] calculated the static spherically symmetric solutions of Weyl's theory; he got the result that the Schwarzschild solution is a solution for all the equations following from one of the three Lagrangians:

$$R^2,\quad R_{ik}R^{ik},\quad R_{ijkl}R^{ijkl} \tag{5.17}$$

(R is the curvature scalar, R_{ik} the Ricci tensor, and R_{ijkl} the Riemann tensor).

Pauli assumed the Weyl vector to be zero, so that he had a Riemannian structure of Lorentz signature as the underlying geometry. He concluded that measurements of the Mercury perihelion advance and light deflection in the field of the Sun which are in agreement with general relativity are also in agreement with all the variants of Weyl's theory, but the fourth-order theory has too many ambiguities (both in finding the correct Lagrangian as well as choosing the correct solution); more explicitly, this is done in [27]. Today one can say, more generally, each vacuum solution of Einstein's equation is also a vacuum solution of each of the variants of fourth-order gravity [where Pauli [26] believed this to be the case for the first two expressions in Eq. (5.17) only].

Pauli [26] wrote about the superfluous degrees of freedom that they are a consequence of the fact that he only considered the vacuum equations and that it should be possible to cancel them by finding the correct interior solution at the source. The latter is only a mathematical problem; we proceed along this line in Sec. VI A. Further, he assumes that the far field of a mass m can be developed into powers of m/r, where r is the distance from the center of the source.

The last point we want to repeat from the old papers is the following: R has dimension $\langle \text{length} \rangle^{-2}$, and, therefore, the action

$$\int R^m \sqrt{-g}\, d^n x \tag{5.18}$$

(where g is the determinant of the metric in the n-dimensional space-time) is scale invariant (i.e., does not change by a change of the used length unit) if and only if $n=2m$ holds. For the usual case $n=4$, this gives $m=2$, an argument which was already used by Weyl in 1918.

Now let us come back to Simon's argument [2] that the superfluous degrees of freedom have to be canceled: Surely, he has found one possibility, but that one is *a priori* not better than the following ones. In units where $8\pi G = c = 1$, we use the Lagrangian

$$L = \left(\tfrac{1}{2}R + \tfrac{1}{4}k^2 C^2 - \tfrac{1}{12}l^2 R^2\right)\sqrt{-g}\ , \tag{5.19}$$

where

$$C^2 = C_{ijkl}C^{ijkl} \tag{5.20}$$

is the square of the Weyl tensor and can be written as linear combination of the terms in Eq. (5.17). The Lagrangian (5.19) gives rise to a tachyonic-free theory if and only if both $k^2 \geq 0$ and $l^2 \geq 0$ hold. (In principle, k^2 and l^2 may have both signs; we prefer the nontachyonic case.)

It holds (see Ref. [28], and see Ref. [29] for the presentation used here) that if we redefine the original metric g_{ij} to G_{ij} via

$$G_{ij} = g_{ij} - 2k^2 R_{ij} + \tfrac{1}{3}(k^2 - l^2) R g_{ij}\ , \tag{5.21}$$

then the linearized Einstein tensor of G_{ij} vanishes if and only if g_{ij} solves the linearized fourth-order equation following from Eq. (5.19). For microscopically small lengths k and l, both metrics cannot be distinguished by experiment, and so one is free to use G_{ij} as a physical metric possessing the required second-order dynamics, at least on the linearized level.

The second possibility is the following: For the spatially flat Friedmann model, the typical solution of fourth-order gravity is composed of (cf., e.g., Refs. [19] and [20]) damped oscillations around the expansion law of the Einstein–de Sitter model. This is not only due to the high symmetry: In Ref. [30] there is considered the general anisotropic Bianchi type-I model, and the result was the same. The general behavior with inhomogeneous models is not known, but there exist reasons (by the conformal transformation of fourth-order gravity to Einstein's theory with a minimally coupled scalar field; see, e.g., Ref. [31], where flat space is related to a local minimum of the potential) to believe that there are similar typical solutions. We interpret them as follows. The superfluous degrees of freedom are just the phases of the oscillations, and by the damping of the amplitudes they simply disappear.

C. Starobinsky inflation as a power series

In this section we consider, in more detail than can be found in the literature, in which sense the Starobinsky inflationary solution can be developed in a power series. To this end we make the following consideration [which makes more explicit what has been done in Ref. [20], Sec. 5, especially Eq. (18)]. It is essential to note that there is neither a cosmological term nor an additional inflaton or scalar field—all inflation comes from the R^2 term in the Lagrangian (5.19). The exact inflationary de Sitter space-time is defined by Eq. (5.3) with $q(t)=ht$, h having a constant positive value. In general, $h=\dot{q}$ is the Hubble parameter. The quasi–de Sitter stage is that period where $|dh/dt| \ll h^2$.

Now the field equation following from the Lagrangian (5.19) is considered. The Friedmann model is conformally flat, and so the term with the Weyl tensor identically vanishes. So without loss of generality we put $k=0$. Further, we consider the nontachyonic case $l^2>0$ only. Suen [17,18] showed that flat space is unstable and that the Starobinsky solution is stable. A partial stability of flat space with respect to initially expanding perturbations can be found; however, in [20] it was shown that *all* vacuum solutions representing an expanding spatially flat Friedmann model of the field equation following from the Lagrangian (5.19) with $l^2>0$ can be integrated up to infinity; they all have the same asymptotic behavior: damped oscillations with frequency $1/l$ about the Einstein–de Sitter model $q=t^{2/3}$. The spatially flat Friedmann model has the metric (5.3). We take the Lagrangian (5.19) and get a fourth-order differential equation for the metric. However, the 00 component of the differential equation is a constraint and gives a second-order equation for h which reads

$$2h\frac{d^2h}{dt^2} - \left[\frac{dh}{dt}\right]^2 + 6h^2\frac{dh}{dt} = -h^2 l^{-2}\ . \tag{5.22}$$

Linearization of this equation gives $0=0$, so that it is clear that the linearized equation gives no information about the full one. It holds that each solution of (5.22) is also a solution of the other nine components of the field equation; however, each function h solves the linearized equation, but in general not the linearized trace equation, which reads simply

$$\frac{d^3h}{dt^3}+l^{-2}\frac{dh}{dt}=0 . \tag{5.23}$$

Next, it is clear that Eq. (5.22) has a singular point at $h=0$, and so the numerical integration has to be done with care. The best method to integrate the system numerically is the following. One uses the constraint only at the initial moment, and then one integrates the trace equation; the trace is regular even for $h=0$. Also, the existence of oscillations with frequency $1/l$ becomes clear from (5.23), and the sign of the right-hand side (RHS) of Eq. (5.22) decides whether or not the oscillations are damped. But the result of [20], that for $l^2>0$ and initial value $h>0$ Eq. (5.22) can be integrated up to infinite time t, is strong and does not depend on the numerics, and it does not change if we include classical matter like dust or radiation.

In Eq. (5.22) the inflationary period can be found by requiring that the first two items be negligible in comparison with the third one. (Afterwards, it will turn out that the first two terms remain finite, whereas the third one tends to infinity as $h\to\infty$; so this approximation is consistent.) We get the first step of the approximation by removing the first two terms of Eq. (5.22); this leads to the equation

$$\frac{dh}{dt}=-\frac{1}{6l^2} . \tag{5.24}$$

The larger the value h, the better the approximation (5.24). This justifies using a Laurent sequence in h^2 as a general ansatz as follows:

$$\frac{dh}{dt}=-\frac{1}{l^2}\sum_{i=0}^{\infty}(-1)^i g_i (hl)^{-2i} . \tag{5.25}$$

We included such powers of the length l as factors that the coefficients g_i become real numbers. Comparing (5.24) with (5.25), one gets $g_0=\frac{1}{6}$. The motivation of the factor $(-1)^i$ will become clear afterwards: All numbers g_i will turn out to be positive. Just for the same reason we did not write odd powers of $1/h$ in (5.25), because all their coefficients automatically vanish if we insert the sequence into Eq. (5.22). This is very satisfactory, because even powers of $1/h$ correspond to powers of $\hbar$, whereas the odd powers would correspond to $\sqrt{\hbar}$, which is a less natural quantity. The coefficients g_i can be obtained as follows. h times the derivative of Eq. (5.25) gives

$$h\frac{d^2h}{dt^2}=\frac{dh/dt}{l^2}\sum_{i=0}^{\infty}(-1)^i 2ig_i (hl)^{-2i} . \tag{5.26}$$

Now we insert Eqs. (5.25) and (5.26) into (5.22), multiply by l^2, and get, step by step,

$$\frac{dh}{dt}\left[6h^2l^2+\sum_{i=0}^{\infty}(-1)^i(4i+1)g_i(hl)^{-2i}\right]=-h^2 . \tag{5.27}$$

After division by $(-h^2)$ and some rearrangement, we get

$$\sum_{k=0}^{\infty}(-1)^k 6g_k(hl)^{-2k} - \sum_{k=1}^{\infty}(-1)^k(hl)^{-2k}\sum_{i=0}^{k-1}(4i+1)g_i g_{k-i-1}=1 . \tag{5.28}$$

The absolute value of Eq. (5.28) gives again $g_0=\frac{1}{6}$, and for each $k>0$ we get

$$g_k=\frac{1}{6}\sum_{i=0}^{k-1}(4i+1)g_i g_{k-i-1} , \tag{5.29}$$

e.g., $g_1=1/6^3=\frac{1}{216}$, $g_2=1/6^4$, and $g_3=65/6^7$. The next natural step seems to be the insertion of (5.29) into the ansatz (5.25). But it turns out that one gets the result more quickly by integrating that equation which is obtained from (5.27) after division by h^2 (time translation is only a coordinate transformation, and so we get no essential constant of integration):

$$6l^2h+\sum_{i=0}^{\infty}(-1)^{i+1}g_i\frac{4i+1}{2i+1}h^{-1}(hl)^{-2i}=-t . \tag{5.30}$$

This equation can be inverted as

$$h=-\frac{t}{6l^2}-\frac{1}{6t}+\sum_{i=1}^{\infty}f_i(l/t)^{2i}t^{-1} , \tag{5.31}$$

with certain dimensionless constants f_i. With Eq. (5.31) we solve Eq. (5.22) and insert the result into the metric (5.3). To simplify the expressions, we perform the coordinate transformation $t=l\tau$. Then the metric describing the Starobinsky inflation reads

$$ds^2=l^2\left[d\tau^2-\exp(-\tau^2/6)|\tau|^{-1/3}\times\sum_{i=0}^{\infty}q_i\tau^{-2i}(dx^2+dy^2+dz^2)\right] , \tag{5.32}$$

where $q_0=1$ and the other q_i are certain real constants. The metric (5.32) gives a good presentation in the region $-\infty<\tau\ll-1$. One can see that it would not have been found by a simple guess, e.g., by a Fourier or Laurent sequence in t or so. How to come from (5.29) to the analogous expressions for the f_i and q_i is straightforward analysis, and the convergence of the sequences can be proved; from the line after Eq. (5.29), it becomes at least quite plausible. In the presentation (5.32) it is not immediately clear that this is inflation; for $\tau\ll-1$, the parabola $-\tau^2/6$ is an almost linearly increasing function, so that the cosmic scale factor is almost exponentially increasing, because the other terms do not essentially change the picture.

Let us now come to the main question here: What happens for $l\to0$? In the metric (5.32), all metric coefficients can be developed into powers of l^2; moreover, they are quadratic polynomials in l. However, for $l\to0$, the metric degenerates. This is essentially the argument of Simon [2], that Starobinsky inflation is not self-consistent in semiclassical gravity. One should look at whether this effect depends on the special coordinates chosen. To this end we go back to the synchronized coordinates $t=\tau l$. Then the factor $\exp(-t^2/6l^2)$ makes the problem (besides the third root of l in the next factor), whereas the further sequence is a sequence in l^2. This is in agreement with the fact that for $l=0$ the corresponding field equation has only the flat Minkowski space-time as a solution.

VI. SIXTH- AND HIGHER-ORDER EQUATIONS

In this section we consider gravitational field equations of order higher than fourth; this is mainly done to show how the fourth-order Starobinsky model is situated between the Einstein theory and the sixth- and higher-order ones.

A. Newtonian limit

The Newtonian limit is the slow-motion approximation of the linearized field equation. In this limit the fourth-order field equation following from (5.19) becomes tractable. For a δ source of mass m, one gets

$$ds^2=(1-2\Phi)dt^2-(1+2\theta)(dr^2+r^2d\Omega^2) , \tag{6.1}$$

where $d\Omega^2$ denotes the metric of the unit S^2,

$$\Phi=\frac{m}{r}[1-\tfrac{4}{3}\exp(-r/k)+\tfrac{1}{3}\exp(-r/l)] \tag{6.2}$$

(see [32]), and

$$\theta=\frac{m}{r}[1-\tfrac{2}{3}\exp(-r/k)-\tfrac{1}{3}\exp(-r/l)] \tag{6.3}$$

(see [15]). It is essential to observe that the solutions (6.2) and (6.3) are unique. One should note that in spite of the higher order of the differential equation one needs the same restriction (namely, the vanishing of Φ and θ as r tends to infinity) to get a unique Newtonian limit. We have considered the same question for a class of gravitational field equations of arbitrary high order and got the same result [21] for the tachyonic-free case. We used

$$L=\frac{R}{2}-\frac{R}{12}\sum_{i=0}^{p}\sum_{0\le j_0<j_1<\cdots<j_i\le p}\prod_{m=0}^{i}l_{j_m}^2\Box^i R , \tag{6.4}$$

where $p\ge 0$, $0<l_0<l_1<\cdots<l_p$ are characteristic lengths, and $\Box$ denotes the D'Alembertian. [For comparison, Eq. (5.19) with $k=0$ and Eq. (6.4) with $p=0$ coincide.] For Eq. (6.4) the Newtonian limit gives (6.1) with

$$\Phi=\frac{m}{r}\left[1+\frac{1}{3}\sum_{i=0}^{p}(-1)^{i+p}\prod_{j\neq 1}\left|\frac{l_j^2}{l_i^2}-1\right|^{-1}\exp(-r/l_i)\right] \tag{6.5}$$

and $\Phi+\theta=2m/r$. For an extended mass distribution, the result is the same because of the linearity, and for the full nonlinear equations one can conjecture that at least in the vicinity of flat space the result remains the same. There is an essential point of departure from the Pauli approach mentioned before and the calculations here: For $L=R^2$ [33] and also for the other purely quadratic Lagrangians [i.e., linear combinations of the terms in Eq. (5.17); see Refs. [34,35]], one does not get the correct Newtonian limit unless one adds the Einstein-Hilbert Lagrangian to the action. [In [16] and [36] the same problem is considered with the same Lagrangian but another variation (Palatini's one, which gives the same theory for the Einstein-Hilbert Lagrangian only), i.e., independent variation with respect to metric and affinity; the result agrees not only with respect to the fact that the Einstein-Hilbert Lagrangian must be added, but also with respect to the general Newtonian plus Yukawa-type potential.]

Let us end this section with a further point of departure from Pauli's approach [26]: He (and the authors of [37] and [38] too) required the outer solution to be developable in powers of m/r. But then *only* the Schwarzschild solution appears, which is definitely *not* the outer solution for a point mass. And neither Eq. (6.2) nor (6.3) can be developed in powers of $1/r$.

B. Generalization of Simon's approach to higher-order gravity

In the units chosen here $(8\pi G=c=1)$, the Planck length $l_{\rm Pl}$ is related to Planck's constant via $\hbar=8\pi l_{\rm Pl}^2$. So Simon's expansion [2,11] into powers of $\hbar$ is equivalent to an expansion into powers of l^2, where l is a fixed length. Let us take as an example the Lagrangian (6.4), which was already considered in [22] and [39] for $p=1$ and in [23] for general p. Equation (6.4) can be written as

$$L=\frac{R}{2}-\frac{R}{12}\sum_{i=0}^{p}c_i l^{2i+2}\Box^i R , \tag{6.6}$$

with numerical constants c_i. We suppose $c_p\neq 0$, and (6.4) leads to a field equation of order $2p+4$. [We do not need it in its entirety here; one can find it in [23], Eq. (8).] For simplicity, we consider the vacuum equations only. We show that by the method of Simon the order can be reduced to $2p+2$ as follows. We suppose (6.4) to be the truncation of an infinite sequence, and so it is valid only up to corrections of order $O(l^{2p+4})$. Then we multiply the field equation following from (6.4) by l^{2p+4} with the result that only the term from the Einstein-Hilbert part of the Lagrangian survives; all other contributions can be subsumed to another $O(l^{2p+4})$. So we get

$$0=R_{ij}l^{2p+2}+O(l^{2p+4}) . \tag{6.7}$$

For $p=0$, this is just Eq. (5.3) of Ref. [11]. We can form the covariant derivatives of (6.7), multiply it by R, and form traces. Then it is possible to add such a linear combination of these equations to the field equation that, up to terms of the order $O(l^{2p+4})$, all terms stemming from $c_pR\Box^pR$ in Eq. (6.6) are compensated and the field equation reduces to the order $2p+2$. For $p=0$ this coincides with Simon's approach.

If the higher-order terms in the Lagrangian do not contain derivatives of the curvature, then the field equation is of fourth order in each step; this has been analyzed in [12]; there, it is also mentioned that Starobinsky inflation remains a consistent solution if one interprets fourth-order gravity as a classical theory and not as a semiclassical one. This point of view (see also [13]) is compatible with Stelle's result [40] that fourth-order gravity, if taken as a classical theory, becomes, in contrast with Einstein's theory, renormalizable.

But here we have chosen an example where the order of the differential equation is increased step by step. The next question which is interesting for $p>0$ is whether the procedure can be repeated such that the order can be reduced from $2p+2$ to even lower order; one should expect that it must be order 2 at the end. The simplest nontrivial example is $p=1$, where, after the first step described above, the following fourth-order equation appears:

$$0=R_{ij}-\frac{R}{2}g_{ij}-\frac{l^2}{3}\left[\Box R g_{ij}-R_{;ij}+RR_{ij}-\frac{R^2}{4}g_{ij}\right] +O(l^6)\ , \tag{6.8}$$

where the semicolon denotes the covariant derivative. The essential difference between Eqs. (6.8) and (1.1) of [11] is now the power [here $O(l^6)$, there $O(l^4)$] of the remainder. The trace of (6.8) reads

$$0=R+l^2\Box R+O(l^6)\ . \tag{6.9}$$

We apply $l^2\Box$ to Eq. (6.9) and get

$$0=l^2\Box R+l^4\Box\Box R+O(l^6)\ . \tag{6.10}$$

The same done with (6.10) yields

$$0=l^4\Box\Box R+O(l^6)\ . \tag{6.11}$$

The sum of Eqs. (6.9) and (6.11) minus Eq. (6.10) yields

$$0=R+O(l^6)\ . \tag{6.12}$$

Similarly, one can handle the trace-free part of Eq. (6.8). So we have shown that (at least this type of) sixth-order gravity can be brought to second order by Simon's approach.

But one should mention that we have, as Simon did, made such assumptions that the application of $l^2\Box$ does not change the power of the general remainder. This is a consistent assumption because $l^2\Box$ is a dimensionless operator.

By inclusion of matter, one gets then covariant derivatives up to the fourth one of the energy-momentum tensor [instead of second derivatives found in [11], Eq. (5.4)]. The corresponding calculation is straightforwardly done, and so we do not write out the formulas. (Also, they are not so essential here, because in regions where the higher-order terms are dominant, one usually believes that matter is not yet essential for the dynamics.) Let us sketch them: the LHS of Eq. (6.7) becomes, in analogy to Eq. (5.3) of [11], $\kappa l^{2p+2}(T_{ij}-\frac{1}{2}Tg_{ij})$, where T_{ij} is the energy-momentum tensor and T its trace. The LHS of Eq. (6.8) gets the form $\kappa[T_{ij}+l^2(\Box T+\text{similar terms})]$, and then two further covariant derivatives to the LHS appear similar to those in Eqs. (5.4) and (5.5) of Ref. [11].

VII. DISCUSSION

The recent review of the Starobinsky model can be found, e.g., in Refs. [41,42] and a more geometrically oriented review in [43].

The Starobinsky model goes back to early ideas of Zeldovich and Sakharov (see, e.g., Ref. [44]), where the addition of higher-curvature terms to the Einstein-Hilbert action was intended to mimic quantum gravitational effects; it was hoped that these terms can prevent the initial singularity.

Another approach can be found in Ref. [45], where the stress tensor renormalization of quantized matter fields in a classical background metric lead to curvature-squared terms in the effective action with spin-dependent calculable coefficients in front of them.

A third approach was performed by Stelle [40], who showed that the Lagrangian (5.19) leads to a renormalizable theory of gravity; the coefficients in front of the curvature-squared terms are not calculable, but should be measured.

We distinguished these three approaches explicitly, because they are often mixed.

The Starobinsky model follows from the Lagrangian Eq. (5.1), which coincides with Eq. (5.19) if $k=0$. This is not a renormalizable theory of gravity, and it shares this property with Einstein's general relativity theory (GRT).

One instability of the theory following from Eq. (5.19) comes from the fact that, for $k^2<0$, tachyons appear and for $k^2>0$ ghosts appear (the latter are particles with negative kinetic energy). The Starobinsky model, however, contains neither tachyons nor ghosts.

A further instability can appear if there is no minimum of the total energy of a given local system. In Einstein's GRT this is prevented by the well-known positive energy theorem, whereas Eq. (5.19) with $k\neq 0$ allows an analogous theorem only in a very restricted sense; cf. Ref. [46]. For the theory following from Eq. (5.1), however, a positive energy theorem, analogous to that one in GRT, is valid; see [47]. It needs only the additional assumption that $R<3l^{-2}$. This represents no practical restriction because l is microscopically small and the inflationary period of cosmic evolution is connected with negative values of R. Connected with this fact is the point discussed in Sec. III: Requiring minimality of the action does not rule out fourth-order gravity.

A third instability could occur if one looks at Eq. (5.1) as a perturbation of Einstein's GRT, l playing the role of the smallness parameter. This is, of course, a singular perturbation, and usually one would expect quickly increasing solutions to appear. In general, this takes place, but under the special circumstances met here, this does not happen. This has its origin in the special kind of nonlinearity of the singular differential equation (5.22): It has the property that for each initial condition with $h>0$ (i.e., initially, the Universe expands), the system can be integrated up to infinite time, and there it tends to the corresponding solution of Einstein's GRT. (We made this explicit here in Sec. V C because of statements found in Refs. [17,18] which seem to contradict this, but in fact only use another notion of instability.) This regular behavior of the solutions can also be seen in the Newtonian limit; see Sec. VI A. If one looks at the solutions Eqs. (6.2) and (6.3), one can see that they converge to the corresponding Newtonian potential as $k, l \rightarrow 0$, but they cannot be developed into powers of k and l. So the problem of the superfluous degrees of freedom can be solved by stating that in the weak-field region, the coefficients of these terms are unobservably small. Another way to deal with the superfluous degrees of freedom is carried out by Simon in [2,10].

Also, in [1, p. 408] it is pointed out that the Starobinsky model is not more unstable than Einstein's theory itself.

These stability statements are all compatible. To see this, one has to remember that for initially contracting perturbations, both Einstein's theory and fourth-order gravity yield a big-bang-type instability after finite time.

The instability appearing from the fact that a fourth-order equation can be brought to a Hamiltonian with indefinite kinetic energy (see Refs. [1,4]) was analyzed in detail in Sec. II. We showed by some typical examples that this can lead to instabilities, but it need not do so. We proposed another general approach to bring a fourth-order theory to Hamiltonian form in Sec. IV. The advantages of our approach are also listed there. The Hamiltonian form of the theory is needed to deduce the Wheeler-DeWitt equation of the system. In Sec. V A we made it for the high-curvature limit, and there the Wheeler-DeWitt equation could be solved in closed form [Eq. (5.16)]. It is planned to make the analogous calculations for less symmetric space-times and the theory including the R term, i.e., for the Lagrangian (5.1). For the interpretation of them, cf., e.g., Ref. [48]. But for the problem discussed here one only needs the sign of the kinetic energy in the Hamiltonian formulation; it is the same as the signature of the superspace metric, as is clear from Eq. (5.12). In Ref. [49] the following was shown: The signature S (= number of negative eigenvalues) of the superspace metric leading to the Wheeler-DeWitt equation following from Einstein's GRT equals $S = 1 + s(n - s)$, where n is the dimension of the spatial part (usually = 3) of the space-time metric and s its signature [$s = 0$ both for the Lorentzian as well as for the Euclidean signature of the underlying $(n+1)$-dimensional manifold]. So $S = 1$ for Einstein's GRT and usual signature, which has the consequence that the Wheeler-DeWitt equation is a normal hyperbolic wave equation. What is essential here is the fact that for fourth-order gravity the Hamiltonian formulation leads to an indefinite kinetic energy (superspace metric signature equals 1) and this is a property which it has in common with GRT.

A discussion of the $R + R^2$ theory in connection with topological defects can be found in Ref. [50].

Let us finally make some remarks about what happens if one adds some higher-order terms, e.g., those discussed in Sec. VI, especially the Lagrangian (6.4) with $p \geq 1$ leading to the order of the differential equation ≥ 6. Then the problems become more serious. The superspace metric gets the signature ≥ 2, and so the Wheeler-DeWitt equation is no longer normally hyperbolic. The conformal transformation to Einstein's theory with several scalar fields (see [22] for $p = 1$ and [23] for $p > 1$) leads always to ghosts. The hope that sixth-order gravity naturally leads to models with double inflation is not fulfilled; see Ref. [39]. It is unclear yet whether eighth-order gravity (partial results can be found in Ref. [51]; further work is in progress) can solve these problems. So the proposal by Simon [2] to reduce fourth-order gravity seems practicable to be generalized (as we did in Sec. VI) to reduce sixth- and higher-order models to second order.

ACKNOWLEDGMENTS

Discussions with W.-M. Suen, A. Berkin (during the MG 6 Conference, Kyoto), and U. Kasper (after his lectures at Potsdam University) helped me to understand the problem. Further, I thank V. Müller, A. A. Starobinsky, S. Reuter, M. Rainer, S. Kluske, J. Audretsch, A. Economou, J. Kurths, and the members of the Rome group headed by F. Occhionero for valuable comments. I am grateful to M. Mohazzab for his lecture at Potsdam University, especially for his hint to Ref. [1]. Financial support from KAI e.V. Berlin, Contract No. 015373/E, and from DFG Bonn, Contract No. Schm 911/5-1, are gratefully acknowledged.

[1] D. Eliezer and R. Woodard, Nucl. Phys. **B325**, 389 (1989).
[2] J. Simon, Phys. Rev. D **45**, 1953 (1992).
[3] M. Ostrogradski, Mem. Acad. St. Petersbourg Series VI **4**, 385 (1850), especially pp. 395 and 514.
[4] T. Damour and G. Schäfer, J. Math. Phys. **32**, 127 (1991); R. Hill, *ibid.* **8**, 1756 (1967); E. Kerner, *ibid.* **6**, 1218 (1965).
[5] A. A. Starobinsky, Phys. Lett. **91B**, 99 (1980).
[6] H.-J. Schmidt, "Minimal, not only stationary action for $R + R^2$ theores of gravitation," Potsdam Report No. PRE-ZIAP 85-01, 1985 (unpublished).
[7] U. Kasper, Class. Quantum Grav. **10**, 869 (1993).
[8] H.-J. Schmidt, "Classical mechanics with lapse," Potsdam Report No. 93/10, 1993 (unpublished).
[9] S. Reuter and H.-J. Schmidt, in Differential Geometry and Applications, Proceedings of the Conference, Opava, Czechoslovakia, 1993, edited by O. Kowalski and D. Krupka (unpublished), p. 243.
[10] J. Simon, in *Sixth Marcel Grossmann Meeting,* Proceedings, Kyoto, Japan, 1991, edited by H. Sato (World Scientific, Singapore, 1992), p. 1246.
[11] L. Parker and J. Simon, Phys. Rev. D **47**, 1339 (1993).
[12] A. Liddle and F. Mellor, Gen. Relativ. Gravit. **24**, 897 (1992).
[13] P. Steinhardt, in *Sixth Marcel Grossmann Meeting* [10], p. 269.
[14] R. Schimming and H.-J. Schmidt, NTM-Schriftenr. Gesch. Naturwiss. Tech. Med. **27**, 41 (1990).
[15] H.-J. Schmidt, Astron. Nachr. **307**, 339 (1986).
[16] V. Hamity and D. Barraco, Gen. Relativ. Gravit. **25**, 461 (1993).
[17] W.-M. Suen, Phys. Rev. Lett. **62**, 2217 (1989).

[18] W.-M. Suen, Phys. Rev. D **40**, 315 (1989).
[19] W.-M. Suen and P. Anderson, Phys. Rev. D **35**, 2940 (1987).
[20] V. Müller and H.-J. Schmidt, Gen. Relativ. Gravit. **17**, 769 (1985).
[21] I. Quandt and H.-J. Schmidt, Astron. Nachr. **312**, 97 (1991).
[22] S. Gottlöber, H.-J. Schmidt, and A. Starobinsky, Class. Quantum Grav. **7**, 893 (1990).
[23] H.-J. Schmidt, Class. Quantum Grav. **7**, 1023 (1990).
[24] A. Vilenkin, Phys. Rev. D **32**, 2511 (1985).
[25] H. Weyl, Sitzungsber. Preuss. Akad. Wiss. Phys. Math. Kl. **1**, 465 (1918).
[26] W. Pauli, Ber. Dtsch. Phys. Ges. **21**, 742 (1919).
[27] C. Lanczos, Z. Phys. **73**, 147 (1932).
[28] G. 't Hooft and M. Veltman, Ann. Inst. Henri Poincaré **20**, 69 (1974).
[29] H.-J. Schmidt, Europhys. Lett. **12**, 667 (1990).
[30] H.-J. Schmidt, Astron. Nachr. **309**, 307 (1988).
[31] H.-J. Schmidt, Astron. Nachr. **308**, 183 (1987).
[32] K. Stelle, Gen. Relativ. Gravit. **9**, 353 (1978).
[33] E. Pechlaner and R. Sexl, Commun. Math. Phys. **2**, 165 (1966).
[34] H.-J. Treder, Ann. Phys. (Leipzig) **32**, 383 (1975).
[35] P. Havas, Gen. Relativ. Gravit. **8**, 631 (1977); see also H. Goenner and P. Havas, J. Math. Phys. **21**, 1159 (1980).
[36] G. Stephenson, J. Phys. A **10**, 181 (1977).
[37] B. Fiedler and M. Günther, Astron. Nachr. **307**, 129 (1986).
[38] R. Pavelle, Phys. Rev. Lett. **40**, 267 (1978).
[39] L. Amendola, A. Battaglia Mayer, S. Capozziello, S. Gottlöber, V. Müller, F. Occhionero, and H.-J. Schmidt, Class. Quantum Grav. **10**, L43 (1993).
[40] K. Stelle, Phys. Rev. D **16**, 953 (1977).
[41] S. Gottlöber, V. Müller, H.-J. Schmidt, and A. A. Starobinsky, Int. J. Mod. Phys. D **1**, 257 (1992/93).
[42] S. Capozziello, F. Occhionero, and L. Amendola, Int. J. Mod. Phys. D **1**, 615 (1992/93).
[43] H.-J. Schmidt, Fortschr. Phys. **41**, 179 (1993).
[44] A. D. Sakharov, Dokl. Akad. Nauk SSSR **177**, 70 (1967) [Sov. Phys. Dokl. **11**, 1040 (1968)].
[45] N. Birrell and P. Davies, *Quantum Fields in Curved Space* (Cambridge University Press, Cambridge, England, 1982).
[46] H.-J. Schmidt, Ann. Phys. (Leipzig) **44**, 361 (1987).
[47] A. Strominger, Phys. Rev. D **30**, 2257 (1984).
[48] C. Kiefer, Phys. Rev. D **47**, 5414 (1993).
[49] H.-J. Schmidt, in *Differential Geometry and Applications,* Proceedings of the Conference, Brno, Czechoslovakia, edited by D. Krupka (World Scientific, Singapore, 1990), p. 405.
[50] J. Audretsch, A. Economou, and C. Lousto, Phys. Rev. D **47**, 3303 (1993).
[51] A. Battaglia Mayer and H.-J. Schmidt, Class. Quantum Grav. **10**, 2441 (1993).