Higher-derivative Lagrangians, nonlocality, problems, and solutions

Jonathan Z. Simon*
Department of Physics, University of California, Santa Barbara, California 93106
(Received 15 November 1989)

Higher-derivative theories are frequently avoided because of undesirable properties, yet they occur naturally as corrections to general relativity and cosmic strings. We discuss some of their more interesting and disturbing problems, with examples. A natural method of removing all the problems of higher derivatives is reviewed. This method of "perturbative constraints" is required for at least one class of higher-derivative theories—those which are associated with nonlocality. Nonlocality often appears in low-energy theories described by effective actions. The method may also be applied to a wide class of other higher-derivative theories. An example system is solved, exactly and perturbatively, for which the perturbative solutions approximate the exact solutions only when the method of "perturbative constraints" is employed. Ramifications for corrections to general relativity, cosmic strings with rigidity terms, and other higher-derivative theories are explored.

I. INTRODUCTION

Theories with higher derivatives (third derivative or higher in time in the equations of motion, second derivative or higher in the Lagrangian) occur naturally for various reasons in different areas of physics. Quite often the higher-derivative terms are added to a more standard (lower-derivative) theory as a correction. This occurs in general relativity, for instance, where quantum corrections naturally contain higher derivatives of the metric (see, e.g., Birrell and Davies[1]), or where nonlinear σ models of string theory predict terms of order R^2 and higher (see, e.g., de Alwis[2]). It occurs in the case of cosmic strings where higher-order corrections, dependent on the "rigidity" of the string, contain higher derivatives,[3,4] and in Dirac's relativistic model of the classical radiating electron.[5] Unlike lower-derivative corrections, however, it is false to assume that adding a higher-derivative correction term with a small coefficient will only perturb the original theory. The presence of an unconstrained higher-derivative term, no matter how small it may naively appear, makes the new theory dramatically different from the original.

Unconstrained higher-derivative theories have very distinctive features. As will be shown below, they have more degrees of freedom than lower-derivative theories, and they lack a lower-energy bound. There is nothing mathematically inconsistent with these features, but they make two almost identical-looking theories, one a lower-derivative theory and the other the same theory with a higher-derivative correction added, very different. The lack of a lowest-energy state for the higher-derivative theory is probably the most dramatic change. This *always* occurs when higher-derivative terms are present (assuming no degeneracy or constraints), independently of how small their coefficients are. The addition of more degrees of freedom might be physically more accurate, but then it means that the original lower-derivative theory was incomplete and missing (the most interesting) new families of solutions. It is particularly disturbing if there is a progression of higher-order, higher-derivative corrections, each system of which has more and more degrees of freedom. Classically, more degrees of freedom means that more initial data are required to specify motion. Quantum mechanically this means that, for a particle, x and $\dot{x}$ now commute since they are freely specifiable, and it becomes possible to measure the position and velocity at the same time. The momentum conjugate to x, π_x, still does not commute with x; $[x,\pi_x]=i\hbar$, but $\pi_x \neq m\dot{x}$. From the path-integral point of view, the paths which dominate the functional integral are of a different class: where once they were nowhere differentiable, now they are everywhere once differentiable. Examples of all these types of behavior are presented below. No familiarity with any of the properties of higher-derivative theories is assumed.

There is a large class of theories naturally containing higher derivatives that do not suffer the above problems. Nonlocal theories, where the nonlocality is regulated by a naturally small parameter, have perturbation expansions with higher derivatives. They avoid the above problems because they are constrained systems. They contain implicit constraints which keep the number of degrees of freedom constant and maintain a lower-energy bound. Higher-derivative theories that are truncated expansions of a nonlocal theory also avoid these problems, once the proper constraints are imposed. Any theory for which the higher-derivative terms have been added as small corrections can be treated in the same manner, also avoiding the above problems.

Nonlocality naturally appears in effective theories, valid only in a low-energy limit and derived from a larger theory with some degrees of freedom frozen out. A good example is Wheeler-Feynman electrodynamics,[6] in which the degrees of freedom of the electromagnetic field are frozen out. For two particles of mass m,

$$S=-\sum_i mc\int ds\left[\frac{dx_i^\mu}{ds}\frac{dx_{i\mu}}{ds}\right]^{1/2}+\frac{e_1e_2}{c}\int ds\,ds'\frac{dx_1^\mu}{ds}\frac{dx_{2\mu}}{ds'}\delta(\Delta x^\nu\Delta x_\nu)\,, \tag{1}$$

where $\Delta x^\nu=x_1^\nu-x_2^\nu$. The only degrees of freedom remaining are of the charged particles. This is nonlocal because the particle-particle interaction is not instantaneous and pointlike, but occurs in retarded time (action at a distance with finite propagation speed). The nonlocal Wheeler-Feynman theory is not valid for large v/c (e.g., particle creation and annihilation is not allowed for), so there is a natural perturbative expansion in powers of v/c. Higher derivatives occur directly as a result of the nonlocality. The action can be naturally expanded as[7]

$$S=\int dt\left[-\sum_i mc\left[1-\frac{\mathbf{v}_i^2}{c^2}\right]^{1/2}-e_1e_2\sum_{p=0}^{\infty}\frac{(-D_1D_2)^p}{2p!c^{2p}}\left[1-\frac{\mathbf{v}_1\cdot\mathbf{v}_2}{c^2}\right]r^{2p-1}\right], \tag{2}$$

where D_i signifies differentiation with respect to t of $\mathbf{x}_i$ only, and $r=|\mathbf{x}_1-\mathbf{x}_2|$. To achieve the same solutions as the original Wheeler-Feynman theory, however, particular constraints must be imposed. Without the constraints, the expansion would have all the problems associated with higher-derivative theories, which are not present in the Wheeler-Feynman theory. The effect of the constraints is to throw away "runaway" solutions. This is accomplished by only allowing solutions that can be Taylor expanded in powers of c^{-1} about $c^{-1}=0$ (corresponding to infinite propagation speed).

These constraints allow the series expansion to be considered as a legitimate perturbative expansion. Without the constraints, higher order does not correspond to higher powers of v/c, but instead all terms contribute equally. With the constraints imposed, each term in the series contributes commensurately less as its order increases. For this reason the constraints will be referred to as "perturbative constraints."

The need for perturbative constraints was first pointed out by Bhabha[8] in the context of Dirac's classical theory of the radiating electron[5] (and its higher-order generalizations), although Dirac realized that runaway solutions should be excluded. The use of perturbative constraints as a method to remove the problems of higher-derivative theories in general was discovered independently by Jaén, Llosa, and Molina[9] (JLM) and Eliezer and Woodard[10] (EW). An explicit method for finding the perturbative constraints for any system expanded in a higher-derivative series about some small expansion parameter was found also by JLM (Ref. 9). Given the perturbative constraints, a method of implementing them in a canonical fashion which greatly simplifies the calculation was found by EW (Ref. 10). Perturbative constraints can be implemented for either infinite or finite series expansions, though for the infinite case the perturbative constraints are already implicitly present if it is demanded that the equations of motion converge. For finite series expansions, where convergence of the series is not an issue, the perturbative constraints play an extremely important role. The finite series expansion, with perturbative constraints imposed, describes a system with the same solutions as those of the full nonlocal series (up to the appropriate order). The finite series expansion without the perturbative constraints describes a system with solutions most of which are nothing like the solutions to the original nonlocal system.

Finding the perturbative constraints does not depend on knowledge of the full nonlocal theory. It can be done just as easily if only a finite number of perturbative terms are known. For this reason it can equally be applied to any higher-order system (with a small expansion coefficient) without knowing whether or not the theory is part of an infinite expansion. This is where the application of perturbative constraints is most powerful and most underutilized.

In general relativity typical corrections take the form of curvature-squared terms in the Lagrangian.[11–14] Even for small coefficients these terms can easily dominate the evolution of the system (as in Starobinsky inflation[15]). Applying the appropriate perturbative constraints describes a (different) system, in which the number of degrees of freedom is the same as in Einstein gravity, and which has no runaway solutions or ghostlike particles. It is a perturbative correction to Einstein gravity, which we know to be a very good approximation of nature.

Applying the perturbative constraints is not just an *ad hoc* procedure. It is completely natural and necessary in cases where the higher-derivative theory is a truncated perturbative expansion of some larger, nonlocal (but otherwise well-behaved) theory. The nonlocal theory itself may be the low-energy effective limit of some even larger theory for which fields have been integrated out. It will be shown that the case of cosmic strings with higher-derivative "rigidity" terms falls in this category.

There is sometimes a small cost to the use of perturbative constraints with higher-derivative theories. Even for finite series expansions, locality can be lost under the influence of explicitly time-dependent sources. This is well known in the case of the self-interacting electron, where the nonlocal phenomenon of preacceleration (acceleration in response to a force that has yet to be applied) occurs. An example is demonstrated below. In the case of the electron, this is considered unimportant, since it takes place only on the scale of the time light travels across the classical electron radius. At any rate, acausality only arises at scales for which the approximation of the electron as a classical particle breaks down. If a similar effect were to occur in a theory of gravity the nonlocality would be at the Planck scale. Most physicists, though, would agree that at the Planck scale the usual notions of geometry probably break down (e.g., the appearance of spacetime foam), and so the possible presence of nonlocality (and the accompanying loss of causality) should not be worrisome.

The structure of the paper is as follows. First is a review of the behavior of unconstrained higher-derivative

theories in general, both classical and quantum, with simple examples of all interesting properties. Next is a discussion of various higher-derivative theories that have been studied in the literature, including Dirac's classical electrodynamics, corrections to general relativity, and corrections to cosmic strings. This is followed by a discussion of the higher-derivative theories that do not suffer from the above problems, and how the problems are avoided by the use of perturbative constraints.

Nonlocal systems, when cast into their higher-derivative expansion, demand the use of perturbative constraints to reproduce the results of the original equations of motion. For any finite expansion approximation to the nonlocal theory, perturbative constraints are still required, and they must be imposed explicitly. The same finite expansion without the perturbative constraints would be a very different theory, completely unrelated to the original nonlocal theory in terms of its available solutions. The constraints must also be applied to systems where the purpose of the higher-derivative terms is to provide small corrections to the original theory. Any theory which is intended (by construction or by physical motivation) to provide perturbative corrections to known solutions, but does not do so, is either incorrect or is being applied beyond its domain of applicability. The method of perturbative constraints is the only means by which a theory with higher-derivative corrections can self-consistently avoid these problems.

Finally, the specific effects of applying the perturbative constraints are calculated for the cases of higher-derivative extensions to general relativity and cosmic strings.

II. A REVIEW OF HIGHER-DERIVATIVE THEORIES

(Many of the ideas in this review section are also covered by EW in a particularly lucid presentation.[10] All the equations presented here apply to one-particle, one-dimensional systems, but the generalization is trivial.) The Lagrangian formalism is straightforwardly applied to higher-derivative theories. For a Lagrangian

$$L = L(q,\dot{q},\ldots,q^{(N)}) , \tag{3}$$

applying the variational principle gives

$$\delta S = \int_{t_i}^{t_f} dt\, \delta q \left[-\sum_{i=0}^{N} \left[-\frac{d}{dt} \right]^i \frac{\partial L}{\partial q^{(i)}} \right] + \sum_{i=0}^{N-1} p_{q^{(i)}} \delta q^{(i)} \Bigg|_{t_i}^{t_f} , \tag{4}$$

where the $p_{q^{(i)}}$ are given by

$$p_{q^{(i)}} \equiv \sum_{k=i+1}^{N} \left[-\frac{d}{dt} \right]^{k-i-1} \frac{\partial L}{\partial q^{(k)}} . \tag{5}$$

Assuming that the $\delta q^{(i)}$ are all held fixed at the boundary, the Euler-Lagrange equation is

$$0 = \sum_{i=0}^{N} \left[-\frac{d}{dt} \right]^i \frac{\partial L}{\partial q^{(i)}} . \tag{6}$$

The canonical formalism for higher-derivative theories was developed by Ostrogradski.[16] The canonical momenta are defined by (5), which shows the generality of the Hamiltonian-Jacobi formalism. The Hamiltonian, as expected, is given by

$$H = \sum_{n=0}^{N-1} p_{q^{(n)}} \dot{q}^{(n)} - L = \sum_{n=0}^{N-1} p_{q^{(n)}} q^{(n+1)} - L . \tag{7}$$

It is conserved and generates evolution in time, and so is equal to the energy of the system. Note that $q^{(N)} = q^{(N)}(q,\dot{q},\ldots,q^{(N-1)},p_{q^{(N-1)}})$ (assuming no degeneracy), but all the remaining $q^{(n)}$ are independent generalized coordinates and so are not inverted. For this reason, $L = L(q,\dot{q},\ldots,q^{(N-1)},p_{q^{(N-1)}})$ as well. The first-order equations of motion are

$$\begin{aligned} &\frac{\partial H}{\partial p_{q^{(n)}}} = \dot{q}^{(n)} = q^{(n+1)} , \\ &\frac{\partial H}{\partial q^{(n)}} = -\dot{p}_{q^{(n)}}, \quad n = 0,1,\ldots,N-1 , \end{aligned} \tag{8}$$

which reproduce the Euler-Lagrange equation. Note that it is entirely self-consistent from within this formalism to consider q and all its derivatives up to N completely independent. The dependence is regained from the equations of motion by the first relation of (8), which states

$$\frac{d}{dt} q^{(n)} = q^{(n+1)} \quad \text{for } n = 0,1,\ldots,N-2 . \tag{9}$$

To demonstrate how higher-derivative theories differ from their lower-derivative counterparts, I will use the simple example

$$L = \tfrac{1}{2}(1+\epsilon^2\omega^2)\dot{x}^2 - \tfrac{1}{2}\omega^2 x^2 - \tfrac{1}{2}\epsilon^2 \ddot{x}^2 \tag{10}$$

which is a simple harmonic oscillator with the mass term slightly modified, and an acceleration-squared piece. It may be helpful to think of $\epsilon\omega \ll 1$, but this is never assumed in our calculations. The kinetic term has been modified only to make the calculations easier; it has no qualitative effect whatsoever, and all quantitative effects are small: $O(\epsilon^2\omega^2)$. This contrasts strongly with the effects of the last term. It is tempting at first to view the last term as a small correction, but we shall see that this is false, independently of how small ϵ is. (The analogous example in scalar field theory has been examined by Hawking.[17])

That the number of degrees of freedom of a higher-derivative theory is more than the lower-derivative theory can be seen by examining Eqs. (4) and (8). For the unconstrained system, there are $2N$ constants that determine the motion, corresponding to the $2N$ initial and final $q^{(n)}$'s, or to the N initial (or final) $q^{(n)}$'s and the N initial (or final) $p_{q^{(n)}}$'s. This is a major qualitative difference from the lower-derivative theory, which needs only two constants to specify the motion. This is also reflected in the quantum theory. The wave function has N arguments,

and the commutation relations reflect the Ostrogradski canonical structure:

$$[q^{(n)},p_{q^{(m)}}]=i\hbar\delta_{nm}\ ,$$
$$[q^{(n)},q^{(m)}]=0=[p_{q^{(n)}},p_{q^{(m)}}]\ . \tag{11}$$

The second of these equations looks especially odd: the position and velocity of a particle commute. The wave function of the system will typically be functions of all the $q^{(n)}$, although one may, of course, Fourier transform any of the generalized coordinates and obtain it in terms of any of the conjugate momenta in their place.

Next we examine how these properties are exhibited in the example. The equation of motion is

$$\epsilon^2 D^4 x+(1+\epsilon^2\omega^2)D^2 x+\omega^2 x=0,\quad \text{where } D=\frac{d}{dt}\ . \tag{12}$$

Being fourth order in time, it requires twice as many initial conditions as the $\epsilon=0$ case, independent of the size of ϵ. This is also reflected in the Ostrogradski canonical formalism, where the independent generalized coordinates are x and $\dot{x}$, and their respective generalized momenta are

$$\pi_x=\frac{\partial L}{\partial\dot{x}}-D\left[\frac{\partial L}{\partial\ddot{x}}\right]=(1+\epsilon^2\omega^2)\dot{x}+\epsilon^2\dddot{x}\ ,$$
$$\pi_{\dot{x}}=\frac{\partial L}{\partial\ddot{x}}=-\epsilon^2\ddot{x}\ . \tag{13}$$

The Hamiltonian is

$$H=\tfrac{1}{2}[2\pi_x\dot{x}-\epsilon^{-2}\pi_{\dot{x}}^2-(1+\epsilon^2\omega^2)\dot{x}^2+\omega^2x^2]\ . \tag{14}$$

Note the impossibility of taking the $\epsilon\to0$ limit in this case. The general solution is

$$x=A_+\cos(\omega t+\phi_+)+A_-\cos(\epsilon^{-1}t+\phi_-)\ . \tag{15}$$

For $\epsilon\omega\ll1$, the second mode oscillates extremely rapidly. The modes separate exactly because the Lagrangian is quadratic in all terms; nonquadratic terms would couple the modes. The oscillatory nature of the second term is not related to the fact that the $\epsilon=0$ case is a simple harmonic oscillator: in the case $\omega=0$, the solution is $x=x_0+v_0t+A_-\cos(\epsilon^{-1}t+\phi_-)$. In the case $\epsilon^2<0$, the solution is

$$x=A_+\cos(\omega t+\phi_+)+A_c\cosh(|\epsilon|^{-1}t)+A_s\sinh(|\epsilon|^{-1}t)\ . \tag{15a}$$

Quantum mechanically, since x and $\dot{x}$ are independent coordinates, the wave function will be a function of both: $\psi=\psi(x,\dot{x})$ [though we could also use $\psi=\psi(x,\pi_x)$, $\psi=\psi(\pi_x,\dot{x})$, or $\psi=\psi(\pi_x,\pi_{\dot{x}})$]. Note that $[x,\dot{x}]=0$, allowing the position and velocity to be measured in the same experiment to arbitrary accuracy. This is also independent of the size of ϵ, so long as it is nonvanishing.

The quantum-mechanical system is solved exactly in Appendix A. The energy eigenstates are labeled by two non-negative integers:

$$E=(n+\tfrac{1}{2})\omega-(m+\tfrac{1}{2})\epsilon^{-1}\quad\text{for } n,m=0,1,2,\ldots\ . \tag{16}$$

The simplest wave function to calculate is

$$\psi_{00}(x,\dot{x})=\left[\frac{\omega}{\epsilon\pi^2}\right]^{1/4}\exp\left[-\frac{\omega^2(1-\epsilon^2\omega^2)x^2-4i\epsilon\omega^2x\dot{x}-(1-\epsilon^2\omega^2)\dot{x}^2}{2\epsilon^{-1}(1+\epsilon^2\omega^2)}\right]\ . \tag{17}$$

As expected, the limit $\epsilon\to0$ does not approach the purely simple harmonic-oscillator ground-state wave function.

Strongly related to the fact that $[x,\dot{x}]=0$ is that the class of paths that dominate the Feynman path-integral changes. The path-integral sums over all possible paths, but a particular class of paths dominates the sum, which can be seen by examining the expectation values in transition amplitudes.[18] First we examine the properties of these paths for a lower-derivative theory. For a path-integral skeletonized into time slices of duration δ, the expectation value of the distance crossed in that time is approximately

$$\langle\Delta x\rangle\sim\delta^{1/2}\ . \tag{18}$$

So, as $\delta\to0$, the typical paths (averaged with a complex weighting) are continuous. But the expectation value of the particle's velocity diverges:

$$\langle\dot{x}_{\rm cl}\rangle\sim\left\langle\frac{\Delta x}{\delta}\right\rangle\sim\delta^{-1/2}\ . \tag{19}$$

For a higher-derivative theory, this is not true. The typical paths for acceleration-dependent Lagrangians have finite velocities, but their acceleration diverges; i.e., the paths are continuous in $(x,\dot{x})$ space. The higher the derivatives in the Lagrangian, the smoother the paths become. An infinite number of higher derivatives would have, in some sense, only perfectly smooth paths contributing. (In fact, because the path-integral formulation can be used to derive Schrödinger's equation, one can read off expectation values from the Hamiltonian, as done by Feynman,[18] and as shown in Appendix B.)

To illustrate the path-integral properties of higher-derivative Lagrangian, we will use the simpler case $\omega=0$: a free particle with a (seemingly small) quadratic acceleration term (the $\omega\neq0$ case is conceptually no more difficult but requires enormously more calculation):

$$L=\tfrac{1}{2}(\dot{x}^2-\epsilon^2\ddot{x}^2),\quad H=\tfrac{1}{2}(-\epsilon^{-2}\pi_{\dot{x}}^2+2\pi_x\dot{x}-\dot{x}^2)\ . \tag{20}$$

We calculate the transition expectations

$$\langle \Delta x \rangle \sim \dot{x}\delta \to 0, \quad \langle \Delta \dot{x} \rangle \sim \epsilon^{-1}\delta^{1/2} \to 0 ,$$
$$\langle x_{\rm cl} \rangle \sim x \quad \langle \dot{x}_{\rm cl} \rangle \sim \dot{x} + O(\langle \Delta x \rangle / \delta) \sim \dot{x} , \tag{21}$$
$$\langle \ddot{x}_{\rm cl} \rangle \sim \frac{\dot{x}}{\delta} + O(\langle \Delta \dot{x} \rangle / \delta) + O(\langle \Delta x \rangle / \delta^2) \to \infty .$$

The dominant paths are now once differentiable. The exact propagator has also been calculated for this system using modes, as shown in Appendix C.

Another extremely important property of higher-derivative theories, both classical and quantum, is the lack of any lower-energy bound. This can be seen most easily through (7). The only dependence on the $p_{q^{(n)}}$ for $n < N-1$ is linear, permitting the Hamiltonian to take on arbitrarily negative values. This carries over into the quantized system as well.[10] This property is easily demonstrated by our example system: the Hamiltonian is manifestly indefinite in Eq. (A2). The energy for the general solution given in Eq. (15) is

$$E = \tfrac{1}{2}(1-\epsilon^2\omega^2)(\omega^2 A_+^2 - \epsilon^{-2} A_-^2) , \tag{22}$$

which is also manifestly indefinite. The effect of even a small amplitude for the negative mode leads to enormously negative energies (for $\epsilon\omega \ll 1$). Even though exciting the negative-energy modes leads only to oscillatory behavior (for the $\epsilon^2 > 0$ case), it is nevertheless unstable since even small excitations of those modes lowers the energy dramatically. Any coupling present in a not purely quadratic Lagrangian system would make the problem worse. The quantized system has the same negative-energy problems, as seen in (16) and Appendix A. Attempts have been made within quantum mechanics to change the minus sign in (16) into a plus by giving half of the quantum states negative norm.[19,17] This merely shifts the problem from the lack of a ground state to the lack of unitarity (arising from the now possible zero-norm modes), but it is really the same problem transformed.

Higher-derivative field theories have the related problem of ghosts: excitations (particles) of negative energy (mass) (see, e.g, Hawking[17]). They behave analogously to the oscillatory excitations of negative-energy states in our example. Creation of ghost particles not only costs no energy, it produces excess energy, causing them to be spontaneously produced in infinite numbers.

In short, the distinct features of higher-derivative theories fall into two major categories, either deriving from the more numerous degrees of freedom than the lower-derivative case, or from the loss of a lowest-energy state. It should be noted that there is nothing fundamentally contradictory or mathematically inconsistent with higher-derivative systems. A good example of this kind of theory is the pure R^2 theory of Horowitz[20] (although there are still problems with the negative-energy modes, as pointed out by Eliezer and Woodard[10]).

These features do become serious problems in most cases, however. Except in a purely cosmological context, the lack of a ground state is very unphysical. It is also unphysical when there is a sequence of higher-order theories for which the higher-order terms are supposed to provide small corrections, but instead introduce new degrees of freedom and new behavior at every step. In our example above, the problems become manifest when the system is compared to a simple harmonic oscillator ($\epsilon = 0$). These problems cannot be avoided in unconstrained higher-derivative theories, whether oscillating particles, flexing cosmic strings, for R^2 gravity.

III. NATURALLY OCCURRING HIGHER-DERIVATIVE THEORIES

As stated in the Introduction, higher-derivative theories appear naturally in at least two contexts. The first is as corrections to a lower-derivative theory. The oldest example of this is the Abraham-Lorentz model of a nonrelativistic, classical, radiating, charged particle (see, e.g., Jackson[21]) and the relativistic generalization due to Dirac.[5] In attempting to take into account the loss of energy due to radiation, a third-derivative term is introduced into the equation of motion. The higher-derivative term has a small coefficient, $\tau = \frac{2}{3}(e^2/mc^3) \sim 10^{-23}$ s, yet there are now twice as many solutions as for the nonradiating electron, and half the solutions are runaways: solutions qualitatively different from solutions of the nonradiating electron. (This is a dissipative system due to the radiation, so the lack of a lower-energy bound is not manifest.) As an example, Dirac's equation of motion in the absence of external forces

$$\dot{v}_\mu = \tau(\ddot{v}_\mu - \dot{v}_\nu \dot{v}^\nu v_\mu) \quad \text{where } \eta_{\mu\nu} = (-,+,+,+) , \tag{23}$$

has the general solution (for motion in the x direction)

$$v_x = \sinh(e^{s/\tau} + b), \quad v_t = \cosh(e^{s/\tau} + b) , \tag{24}$$

where s is proper time and b is an integration constant, or

$$v_\mu = \text{const} . \tag{25}$$

For the first solution, the free electron accelerates to near the speed of light in a time comparable to τ. For the second, the electron remains unaccelerated, which is the expected answer for zero external force.

Another example of a naturally occurring higher-derivative theory is the case of cosmic strings. If treated as an unconstrained higher-derivative theory, as is always done in the literature, it suffers from all the above problems. The number of degrees of freedom is dependent on which order the higher-order expansion is stopped. The excitations of the newly available modes contain negative energy, just as in the case of the oscillator above. Note that this is *independent of the sign* of the coefficient of the higher-derivative term (corresponding to rigidity). It is completely analogous to the acceleration-dependent harmonic-oscillator example above. For $\epsilon^2 < 0$ the negative-energy modes are exponential in time and obviously unstable, but even for $\epsilon^2 > 0$, exciting the negative-energy modes allows arbitrary amounts of energy to be extracted. Even a small kick ($A_- \ll A_+$ in the example) can extract large amounts of energy since the negative energies are inversely proportional to the small parameter, in this case the width of the string. We shall see below that the perturbative constraints must be applied for con-

sistency.

Higher-order corrections to general relativity itself can arise either quantum mechanically or classically. Quantum mechanically, conformal anomalies give rise to an effective action with higher-derivative terms (see, e.g., Birrell and Davies[1]), which can be local ($\propto R^2$) or nonlocal (not expressible in local quantities such as the metric and Riemann tensor). Renormalizability arguments demand the presence of terms $\propto R^2$ and $\propto R_{ab}R^{ab}$ (Ref. 13). The effect of these terms is to give additional families of solutions, some of which are "runaways." The negative-energy problem manifests itself when coupled to matter; there is in general no positive-energy theorem[12] (with the exception of special initial conditions for certain higher-derivative terms[22]). There are also, in general, problems with ghost fields and local instabilities due to the presence of negative-energy modes.[11] The concept of a runaway solution on a cosmological scale is somewhat unclear in the case of gravity where, with an ordinary cosmological constant, exponential inflation is a physical solution. Nevertheless, the smaller the coefficient of the higher-derivative terms, the faster the rate of inflation it can induce. These extra solutions are non-Taylor expandable in powers of the small coefficient. Lovelock gravities[23,24] have similar problems, despite the fact that they are, strictly speaking, not higher-derivative theories. Lovelock theories contain higher-order terms in the Lagrangian which are dimensionally extended Euler densities. They allow extra solutions to the field equations, though the number of new solutions is finite, not a continuous family. Nevertheless, some of the new solutions are dramatically different from the original and can be considered runaways.

Classically, string theory gives higher-order (local) corrections to the action in higher power of curvature and its derivatives, as shown below in (50). If left unconstrained, in addition to all the problems of the preceding paragraph, each theory obtained by truncating the expansion at a given order has a different number of degrees of freedom than the previous one.

IV. HIGHER-DERIVATIVE THEORIES WITHOUT THE PROBLEMS

When higher-derivative theories occur as a result of truncating a perturbative expansion of a nonlocal theory, the usual problems of higher-derivative theories do not occur because the perturbative constraints must be applied. They guarantee that of all possible solutions to the unconstrained higher-derivative equation of motion, only solutions that are Taylor expandable in the small expansion parameter are permitted. All other solutions are considered runaways, solutions that do not exist in the limit of zero-expansion parameter. This corresponds to the limit of infinite propagation speed of instantaneous interactions in the Wheeler-Feynman model. The extra solutions have extremely rapid behavior for small expansion coefficients and are always associated with negative-energy behavior. The remaining solutions form a two-parameter family.[9] The first use of the exclusion of runaway solutions was in the context of removing obviously unphysical solutions such as (24) from the Dirac model (and its nonrelativistic variations). It was suggested that runaway solutions be isolated and defined by their late time behavior, that the acceleration should be finite in the far future for finite forces. Imposing future boundary conditions at large scales is undesirable, since, if there is acausality present in the Universe, it is likely to be only at the smallest (i.e., Planck) scales. Furthermore, the finite acceleration criterion does not generalize well to other classical theories, such as general relativity, which has extremely varied cosmological solutions. Bhabha pointed out that for Dirac's theory (and higher-order extensions) all nonrunaway solutions are Taylor expandable in the natural small expansion parameter of the theory τ. Imposing the perturbative constraints is equivalent to throwing away runaway solutions, but relieves us of the obligation to specify future conditions.[8]

Nonlocal systems, such as the Wheeler-Feynman theory, when cast into their higher-derivative expansion (first done for the Wheeler-Feynman theory by Kerner[7]), demand the use of perturbative constraints to reproduce the results of the original equations of motion. The perturbative constraints are implicit in the Lagrangian by demanding convergence. This will be demonstrated below with a simple model. For any finite expansion approximation to the nonlocal theory, perturbative constraints are still required, but must be imposed explicitly, since they are no longer demanded by convergence of the series. The same finite expansion without the perturbative constraints is a very different theory, completely unrelated to the original nonlocal theory in terms of its available solutions.

It is also perfectly self-consistent and valid to impose perturbative constraints on a higher-derivative system even without the sure knowledge that it is a truncated expansion of some nonlocal theory. This was done by Dirac when he threw away the undesirable runaway solutions. It is not unreasonable to apply it to any higher-derivative theory for which the method is applicable and examine the consequences. It is absolutely necessary if the higher-derivative theory is to resemble at all the original lower-order theory in its behavior.

The JLM procedure for finding the perturbative constraints strictly imposes the condition that all solutions must be Taylor expandable in the perturbative expansion parameter.[9] It is not possible to invert all of the canonical momenta within the limits of Taylor expandability, which signals the presence of a primary constraint. Secondary constraints are obtained by taking time derivatives of the primary constraints. Linear combinations of the constraints can always be put in the form

$$x^{(i)}-f_i(x,\dot{x})=0, \quad i=2,\ldots,N \ . \tag{26}$$

Add the constraints in this form to the Lagrangian with Lagrange multipliers, and proceed either to the Euler-Lagrange equations or the Hamilton-Dirac equations (whose equivalence for constrained higher-derivative systems has been shown by Pons[25]).

There is no issue of whether or not to quantize on a large phase space and apply weak constraints afterwards, because there is no larger phase space. The perturbative

constraints are second-class constraints, and hold strongly. One must use the Dirac procedure, with Dirac brackets instead of Poisson brackets, calculate the Hamiltonian, and use the Hamilton-Dirac equations to quantize the system.[26,27]

There is an easier method to put the system in canonical form, however, based on the fact that we know what form the final answer must take, due to EW (Ref. 10) and quickly reviewed here. There exists a local, lower-order theory equivalent to the higher-derivative theory plus the constraints.[27] By knowing the energy, reduced to a function of x and $\dot{x}$ only, $E_r(x,\dot{x})$ (once the constraints have been applied), and knowing that the Hamiltonian is equal in value to the energy, the value of p, the canonical momentum to x, for a reduced version of the same system can be inferred. In order that

$$\dot{x}=\{x,H_r(x,p)\} \tag{27}$$

hold true, we must have

$$p(x,\dot{x})=\int_0^{\dot{x}}\frac{dv}{v}\frac{\partial E_r(x,v)}{\partial\dot{x}}+p(x,0)\ . \tag{28}$$

[There may be some uncertainty in the choice of $p(x,0)$, but this corresponds to the uncertainty in the initial Lagrangian of whether to add total derivatives of the form $\Delta L=dF(x(t))/dt$, which one always has the freedom to do. This addition corresponds to the making the canonical transformation

$$x'=x,\quad p'=p+\partial F/\partial x\ .] \tag{29}$$

We can then invert (28) to get $\dot{x}(x,p)$ and arrive at $H_r(x,p)=E_r(x,\dot{x}(x,p))$ and $L_r(x,\dot{x})=p(x,\dot{x})\dot{x}-E_r(x,\dot{x})$. Using the new x and p, where $\{x,p\}=1$, is equivalent in every way to using all the $x^{(n)},p_{x^{(n)}}$, and constraints, with Dirac brackets.[27]

In quantum cosmology there is not, in general, a special physical quantity that takes the role of time which is so essential to Hamiltonian-based quantum mechanics. In these cases the action is taken to be the fundamental basis and quantization can proceed using the Feynman sum-over-histories approach. The technique is to Euclideanize the action and sum over all paths with a specified boundary condition, which produces a specific quantum state. This does not require any canonical formalism. If there is a special time parameter or symmetry that allows a useful canonical formalism, the EW method can be used for ease of calculation of the action, but is not required.

Let us observe all the above properties and characteristics of a nonlocal theory with a simple model. The model is solved exactly, both classically and quantum mechanically, and can be expanded in an infinite or arbitrarily truncated series in higher derivatives. The system is a one-dimensional harmonic oscillator but for which the potential depends not only on the position at a given time, but at all times past and future. Distant times, however, contribute exponentially weakly. We begin with the equation of motion

$$\begin{aligned}0&=\ddot{x}+\omega^2\int_0^{\infty}ds\,e^{-|s|}\tfrac{1}{2}[x(t+\epsilon s)+x(t-\epsilon s)]\\&=\ddot{x}+\omega^2\int_{-\infty}^{+\infty}\frac{ds}{2}e^{-|s|}x(t+\epsilon s)\ .\end{aligned} \tag{30}$$

When $\epsilon\omega\ll 1$ we might expect this system's behavior to be very similar to a simple harmonic oscillator, and indeed this is the case. Using the ansatz $x=Ae^{i\gamma t}$ we find two roots:

$$\gamma^2=\begin{cases}\omega^2\chi^2\ ,\\ -\epsilon^{-2}\chi^{-2}\ ,\end{cases}\quad\text{where}\quad \chi=(\tfrac{1}{2}+\tfrac{1}{2}\sqrt{1+4\epsilon^2\omega^2})^{-1/2}=1-\tfrac{1}{2}\epsilon^2\omega^2+\cdots\ . \tag{31}$$

The second root, when reinserted into the equation of motion, fails to converge, and so is not a solution. The remaining root corresponds to the solution we expect: harmonic motion with frequency close to the original, $\gamma=\omega[1+O(\epsilon^2\omega^2)]$. The general real solution is

$$x=A\cos(\gamma t+\phi)\ . \tag{32}$$

To put the system into a Lagrangian form, we expand out the equation of motion into an infinite series

$$0=\ddot{x}+\omega^2\sum_{n=0}^{\infty}(\epsilon D)^{2n}x \tag{33}$$

and we can construct a Lagrangian that will give us this:

$$\begin{aligned}0&=-\sum_{n=0}^{\infty}(-D)^n\frac{\partial L}{\partial(D^n x)}\ ,\\ L&=\frac{1}{2}\left[\dot{x}^2-\omega^2x^2+\omega^2\sum_{n=0}^{\infty}\sum_{m=0}^{\infty}[(\epsilon D)^{2n+1}x(\epsilon D)^{2m+1}x+(\epsilon D)^{2n+2}x(\epsilon D)^{2m+2}x]\right]\\&=\frac{1}{2}\left\{\dot{x}^2-\omega^2x^2+\omega^2\left[\left[x-\int_{-\infty}^{+\infty}\frac{ds}{2}e^{-|s|}x(t+\epsilon s)\right]^2+\left[\int_{-\infty}^{+\infty}\frac{ds}{2}e^{-|s|}\epsilon\dot{x}(t+\epsilon s)\right]^2\right]\right\}.\end{aligned} \tag{34}$$

There are, of course, other Lagrangians that give us the same equations of motion. For instance, we can always add the total derivative $(d/dt)f(x)$ without changing the classical equation of motion. When adding total derivatives with higher derivatives, however, one must exercise caution, or the variational principle necessary for quantization will be lost. Details of the need for a valid variational formulation are discussed in Appendix D.

Since there is only one true degree of freedom (i.e., the evolution is specified completely by x_i and $\dot{x}_i$), not infinitely many as implied by the infinite expansion, there must be an infinite number of constraints implicit in the expansion Lagrangian. They must express all higher derivatives in terms of x and $\dot{x}$, and can be put in the form of (26). By inspection of the known solution (32) they must be

$$\begin{aligned} D^{2n+2}x &= (-1)^{n+1}\gamma^{2n+2}x \ , \\ D^{2n+3}x &= (-1)^{n+1}\gamma^{2n+2}\dot{x} \ , \end{aligned} \qquad n=0,1,\ldots , \tag{35}$$

or some linear combination. The JLM procedure is unnecessary here because we have the general solution (32), and finding constraints to enforce it can be done by inspection. These constraints may be put explicitly into the Lagrangian with Lagrange multipliers. It is not necessary, since the convergence of the series enforces the constraints implicitly, but it is helpful to acknowledge them explicitly as well.

The energy is calculated using the Ostrogradski Hamiltonian:

$$\begin{aligned} E &= \sum_{n=1}^{\infty}\sum_{m=0}^{n-1}(D^{n-m}x)\left[(-D)^m\frac{\partial L}{\partial(D^n x)}\right]-L \\ &= \tfrac{1}{2}\dot{x}^2+\omega^2 x\int_{-\infty}^{+\infty}\frac{ds}{2}e^{-|s|}x(t+\epsilon s) \\ &\quad -\tfrac{1}{2}\omega^2\int_0^{\infty}ds\,e^{-s}x(t+\epsilon s)\int_0^{\infty}ds'e^{-s}x(t-\epsilon s') \ . \end{aligned} \tag{36}$$

Note that $\lim_{\epsilon\to 0}E=\frac{1}{2}(\dot{x}^2+\omega^2x^2)$, as expected.

The next step is to quantize the system. As usual, finding the ground state of the system (which would not exist if it really were an unconstrained higher-derivative theory) can be done without reference to the canonical formalism or Poisson and/or Dirac brackets. The ground state can be calculated using the Euclidean sum over paths:

$$\psi_0(x_f)=\int \mathcal{D}\tilde{x}\,e^{-I[\tilde{x}(\tau)]} \ , \tag{37}$$

where the Euclidean action $I=\iota S, t=-\iota\tau$, the sum is over all paths of finite Euclidean action ending at x_f, and $\tilde{x}(\tau)=x(t)$. We may take the final Euclidean time to be $\tau_f=0$, without loss of generality:

$$I=\tfrac{1}{2}\int_{-\infty}^{0}d\tau\left\{\dot{\tilde{x}}^2+\omega^2\tilde{x}^2-\left[\left[\tilde{x}-\int_{-\infty}^{+\infty}\frac{ds}{2}e^{-|s|}\tilde{x}(t+\iota\epsilon s)\right]^2-\left[\int_{-\infty}^{+\infty}\frac{ds}{2}e^{-|s|}\epsilon\dot{\tilde{x}}(t+\iota\epsilon s)\right]^2\right]\right\} . \tag{38}$$

The action is still nonlocal in real time, even though the paths are in Euclidean time. The path integral can be done exactly since the action is quadratic in x (Ref. 18). There is only one classical solution with finite Euclidean action: $x_{\rm cl}=x_0e^{\gamma\tau}$. Let $\tilde{x}=x_{\rm cl}+q$. Then

$$\begin{aligned} \psi_0(x_0) &= \underbrace{\int \mathcal{D}q e^{-I[q]}}_{\text{const}}e^{-I[x_{\rm cl}]} \\ &= \text{const}\times e^{-\gamma\eta x_0^2/2} \ , \end{aligned} \tag{39}$$

where

$$\eta=2-\chi^2=1+O(\epsilon^2\omega^2) \ , \tag{39a}$$

so

$$\psi_0(x)\propto e^{-\omega x^2[1+O(\epsilon^2\omega^2)]/2} \tag{39a}$$

as expected.

This system can be put into canonical form in two ways: expanding the nonlocal integrals into infinite sums, defining the momenta by Eq. (5), applying the (second-class) constraints of Eq. (35) with Lagrange multipliers, and using Dirac brackets, or by the calculationally much simpler EW method discussed above. The reduced relevant quantities given by this method are

$$\begin{aligned} E_r &= \tfrac{1}{2}\eta(\dot{x}^2+\gamma^2x^2), \quad p=\eta\dot{x} \ , \\ H_r &= \tfrac{1}{2}(\eta^{-1}p^2+\eta\gamma^2x^2) \ , \\ L_r &= \tfrac{1}{2}\eta(\dot{x}^2-\gamma^2x^2) \ . \end{aligned} \tag{40}$$

The function $p(x,0)$ is determined in this case by demanding $L(x_{\rm cl})=L_r(x_{\rm cl})$, which gives $p(x,0)=0$. Now we know the whole system is canonically equivalent to a simple harmonic oscillator, so, in particular,

$$\begin{aligned} H\psi_n &= \tfrac{1}{2}\gamma(n+1)\psi_n \quad \text{where} \quad n=0,1,\ldots , \\ \psi_0 &= ce^{-\eta\gamma x^2/2} \ , \end{aligned} \tag{41}$$

which agrees with the Euclideanized sum-over-histories calculation of the ground-state wave function.

Now suppose we are not given the full theory, but only the first N terms. We may not even know where they came from. But we do know that the zeroth-order term is a good approximation when $\epsilon\omega\ll 1$. Or perhaps we do not have the tools to solve the full theory, but only for the first N terms:

$$L_N=\frac{1}{2}\left[\dot{x}^2-\omega^2x\sum_{n=0}^{N}(\epsilon D)^{2N}x\right]+\text{total derivatives} \ . \tag{42}$$

The cases $N=0,1$ are trivial, and the solutions are the same as the solution to the full nonlocal theory, to order $(\epsilon\omega)^0$ and $(\epsilon\omega)^2$, respectively. If left unconstrained, however, the case $N=2$ has all the quirks and problems of the acceleration-dependent oscillator above: twice as many solutions as the zeroth-order case, negatively unbounded energy, etc. For $N=3,4,\ldots$, the number of solutions continues to increase and all associated problems get worse.

The perturbative constraints are needed explicitly here (they are implicit in the full theory). The new "solutions" to the unconstrained finite series approximations do not converge when put in the equation of motion for the full theory. The appropriate constraints can be obtained by the JLM procedure, which is necessary when the full theory is not known, and gives

$$\ddot{x}+\gamma_N^2 x=0\ , \tag{43}$$

where γ_N depends on the expansion order N, and $\gamma_N^2-\gamma^2=O((\epsilon\omega)^{2N+2})$. Higher-derivative constraints are obtained by differentiating and substituting as necessary. Solving this system, for finite N, gives the correct solutions of the full infinite system (to the appropriate order), whether classical or quantum, whether via the Lagrangian or Hamiltonian. Solving the system without the perturbative constraints, while describing a well-defined system, does not approximate the full nonlocal system in any sense.

The above nonlocal oscillator is an example of a perfectly well-behaved system that appears sick when expanded naively in a perturbation series. But when expanded properly, with the knowledge that the only contributing solutions are those close to the zeroth-order solutions, the expansion is useful, perturbative, physical, and agrees with the full theory to the appropriate order.

The one important aspect of higher-derivative and nonlocal systems that does not appear in the above example is the appearance of acausal solutions, i.e., preacceleration types of effects. These appear when the system is coupled to explicitly time-dependent terms. The best known example is Dirac's classical electron. In the case of a nonzero force, the electron experiences a preacceleration on times of the order of τ (Refs. 21 and 5). For a one-dimensional delta-function impulse the equation of motion (23) in the nonrelativistic limit becomes

$$\tau\ddot{x}-\dddot{x}=\kappa\delta(t) \tag{44}$$

which has the general solution $\dot{x}=c_1e^{t/\tau}+c_2$ for $t\neq 0$ with an instantaneous change in $\ddot{x}$ of κ across $t=0$. For $t>0$, $c_1=0$ by requiring finite acceleration in the infinite future. For $t<0$, $c_2=0$ if we desire zero velocity in the far past. The solution is

$$\dot{x}=\begin{cases}\kappa\tau e^{1/\tau}, & t<0\ ,\\ \kappa\tau, & t>0\ ,\end{cases} \tag{45}$$

which has nonzero acceleration before the force is applied, but the time scale it occurs on is τ. In general, noncausality only appears at the scale of the small expansion parameter, and as shown by Wheeler and Feynman for similar theories, the noncausality decreases as the number of particles rises.[28] At any rate, for theories in which the nonlocality appears only as a low-energy effective theory, the theory itself is only an approximation at that scale, and its results at that scale will reflect this. If the theory is truly nonlocal, as the true theory of everything might be (we have no experimental evidence whether or not nature is local near the Planck scale), then the nonlocality will be manifest in the solutions.

The JLM procedure fails in the presence of external sources, though a related procedure has been proposed.[29] Note that (45) is not analytic in τ, and so whatever method is needed to remove ill-behaved solutions will most likely not allow a solution of this form. The noncausality may still manifest itself, though probably in a different way. The form of the source term might be restricted due to (nonlocal) back reaction, on the order of the expansion parameter.

It is important to mention and put to rest a common fallacy that the extra degrees of freedom of higher-derivative theories are somehow related to the degrees of freedom frozen out in creating the effective nonlocal theory; for instance, that the higher-derivative degrees of freedom in a cosmic string arise from the lost degrees of freedom of the scalar and gauge fields, or that the higher-derivative degrees of freedom from curvature squared corrections to general relativity arise from the frozen-out massive string modes. The field degrees of freedom frozen out to create the nonlocal effective theory are gone and cannot be regained. The apparent higher-derivative degrees of freedom, as in the model above, are artifacts from trying to perturbatively expand a nonlocal theory without the necessary perturbative constraints. This is also particularly relevant in the case of cosmic strings, where negative-energy modes are available in the unconstrained higher-derivative case, yet exact traveling wave solutions along a string in the full field theory have only positive energy.[30]

V. EFFECTS ON GENERAL RELATIVITY, COSMIC STRINGS, AND OTHER THEORIES

Under what circumstances should perturbative constraints be applied? They must always be applied in the case where the theory in question is known to be a truncated expansion (with a small coupling constant) of a nonlocal theory. The nonlocal theory may itself be a low-energy limit of some larger local theory. Specific cases of this are Wheeler-Feynman electrodynamics and cosmic strings (as will be shown below). If the expansion is to be perturbative, then the perturbative constraints are the only means of enforcing it. To verify whether the perturbative expansion itself is appropriate, check the behavior of the zeroth-order approximation (e.g., a slowly moving electron for which radiation effects are ignored, or a very straight slow cosmic string). If it is appropriate to approximate the system with only the zeroth-order term, then it is appropriate to use higher-order perturbative corrections, and hence the perturbative constraints. This is certainly the case for Wheeler-Feynman electrodynamics and cosmic strings without too much curvature

(i.e., no kinks or cusps). Where the zeroth-order approximate theory is inappropriate (e.g., electrons with speeds near c and intersecting cosmic strings), the perturbative expansion is inappropriate as well. The expansion without perturbative constraints is never appropriate.

Consider the case of cosmic strings. The usual derivation of the string action[4,3] begins with the full gauge theory

$$S=\int d^4x\sqrt{-g}\,\mathcal{L}(\phi,\partial_\mu\phi,A_\mu,F_{\mu\nu})\ . \tag{46}$$

Let $\phi=\phi_0(x^\mu)$, $A_\nu=A_{\nu 0}(x^\mu)$ be a field configuration that describes a string and let $x^\mu=X^\mu$ be the location in spacetime of the string. We want to write down an effective action based only on the movement of the string, i.e., formed only from functions and operators acting on X^μ. First pick a coordinate system such that two coordinates ξ^a are in the world sheet of the string and two coordinates ρ^A are Gaussian normal coordinates perpendicular to the world sheet

$$x^\mu=X^\mu(\xi^a)+\rho^A n_A^\mu\ , \tag{47}$$

where n_A^μ are two (arbitrary) unit vectors normal to the world sheet, $a=1,2$, and $A=1,2$. The action, still exact, now reads

$$S=\int d^2\xi\,d^2\rho[\mathcal{L}(\phi_0,A_{\mu 0})+\cdots]\times\left[\sqrt{-\gamma}\left[1+\rho^A K_A+\frac{1}{2!}\rho^A\rho^B K_A K_B+\cdots\right]\right], \tag{48}$$

where $\gamma_{ab}(\xi^a)$ is the metric on the world sheet and $K_A(\xi^a)$ are the traces of the two extrinsic curvatures. We can make an effective, nonlocal theory by integrating out all degrees of freedom off the string world sheet. It is nonlocal because the string has finite thickness, so the energy of a piece of the string propagates not only along the string but also over and around it, but those extra degrees of freedom are now frozen out of the picture. Once we have done this, the lost degrees of freedom cannot be recovered; i.e., we cannot reconstruct the original system from the effective theory. This effective, nonlocal theory is not usually examined *per se*, but is itself perturbatively expanded, effectively in powers of the string width multiplied by the extrinsic curvatures

$$S_{\rm pert}=-\mu\int d^2\xi\sqrt{-\gamma}\left[1+\frac{s_0}{\mu}K^A K_A+\cdots\right], \tag{49}$$

where μ is the string tension, s_0 is the "rigidity," etc. The zeroth-order term is just the Nambu-Goto action. Higher-order expansions contain higher derivatives via the extrinsic curvatures and their derivatives. If left unconstrained, these higher-derivative terms would have the usual disastrous effect, making the so-called perturbative theory totally different from the full nonlocal theory. Instead, enforcing the perturbative constraints produces solutions consistent with the full gauge field theory, allows the expansion to be truly perturbative, and removes all the problems of extra degrees of freedom and negative energy.

For cosmic strings the perturbative constraints ensure that all higher-order solutions remain close to the solutions of the old, zeroth-order, Nambu-Goto action. Since it has been shown that solutions of the Nambu-Goto action are also solutions when the first-order "rigidity" corrections are present,[3] the appropriate constraint, for a noninteracting string, is that the original equations of motion remain unchanged. The result is that, for an isolated string, the rigidity term has no effect at all on the motion, *independently of the sign of its coefficient.* The first contributing corrections to the Nambu-Goto action must come from higher-order terms, e.g., torsion or derivatives of extrinsic curvature. As in the case of Dirac's theory, however, the rigidity term will certainly play a role once external forces are considered.

Similar to the cases of nonradiating electrons and cosmic strings, Newtonian gravity and its post-Newtonian corrections can be derived from a perturbation expansion of an effective non-local theory ultimately derived from Einstein gravity. One might expect the same phenomenon to occur, since the effects of gravitons have been integrated out. Acceleration-dependent terms do occur in the post-Newtonian and post-post-Newtonian approximations, but only linearly, which is a degenerate case (though perfectly suited to the Ostrogradski canonicalization procedure with second-class constraints[31]). The higher-derivative terms at these low orders arise only from coordinate and gauge choices,[32] but there is little reason to doubt that nondegenerate higher-derivative terms will appear at higher order.

Another important case where perturbative constraints have not been considered, but should be, is the case of gravity as a low-energy limit of string theory. Einstein's equations and the corrections to arbitrarily high orders (in the slope parameter) can be obtained from nonlinear σ models (e.g., de Alwis[2]). The first-order corrected gravitational action in D dimensions is

$$S\propto\int d^Dx\sqrt{-g}\left[R-\frac{\alpha'}{4}(R_{\mu\nu\rho\sigma}R^{\mu\nu\rho\sigma}-\nabla^2 R)+(\text{matter terms})+O(\alpha'^2)\right], \tag{50}$$

where α' is the slope parameter. As in the preceding case, to the extent that the zeroth-order theory (Einstein gravity) is a good approximation of nature, higher-order terms produced by this method should only be perturbative corrections; they should not completely alter the dynamics of the system. The perturbative constraints must be applied for consistency.

It is important to note in this case that the large nonlocal theory which Einstein gravity and the constrained higher-order terms approximates well is not string theory itself. It is a nonlocal, low-energy effective theory that is derivable from string theory and is appropriate in cases where Einstein gravity is also a good approximation (though not as good as the nonlocal low-energy effective theory). Neither is appropriate in regions of very high curvature or near singularities. The analog of the intermediate nonlocal theory for the case of electrodynamics

would be the Wheeler-Feynman theory, which falls between full field-theoretic electrodynamics and slowly moving, nonradiating, Lorentz-force-law motion. Just as the Wheeler-Feynman theory is not accurate at high energies (comparable to the electron mass), neither general relativity nor general relativity plus string corrections will be accurate at high curvatures.

When the higher-derivative corrections arise from quantum effects, then the above argument does not hold. One may or may not choose to apply the perturbative constraints. But one should be aware that the higher-derivative theory without perturbative constraints is dramatically different from Einstein gravity, while the same higher-derivative theory with perturbative constraints is a true perturbative correction. If there is any reason to believe that the quantum corrections will not radically alter the behavior of the system, then the perturbative constraints must be applied. (The same holds for Lovelock gravities, which, while strictly speaking not higher-derivative theories, have some solutions that are close to Einstein gravity and others that are far from Einstein gravity. One must throw away the dramatically different solutions, i.e., impose the perturbative constraints, if the corrections to Einstein gravity are intended to be small.)

The perturbatively constrained system has qualitatively different properties than the unconstrained higher-derivative theory. The renormalizability gained from the higher-derivatives is lost once the constraints are applied. The extra particles (degrees of freedom) present in the unconstrained theory do not exist in the constrained theory, since any solution containing them is nonanalytic in the expansion parameters.

For higher-order terms $\propto R^2$, all vacuum solutions to the Einstein action are still vacuum solutions, and so just as for the case of cosmic strings and the unforced Dirac electron, the perturbative constraint is just the old equation of motion. This has the effect that the new equations of motion ignore the R^2 piece entirely (though again, when coupled to matter, this could easily change). Dramatically different solutions, such as those offered in Starobinsky-type inflation,[15] are excluded from the realm of acceptable solutions. For the additional nonlocal, higher-derivative terms arising from quantum corrections, the first-order terms contribute nontrivially even in vacuum, but do not dominate the evolution.[33] In the case of Lovelock gravities, the de Sitter-like and anti–de Sitter-like solutions[24] are disallowed as acceptable spherically symmetric solutions.

VI. SUMMARY

Higher-derivative theories occur in various places throughout theoretical physics, usually as a correction term to a standard, lower derivative theory. In particular, they arise in the context of theories of gravity and of cosmic strings. Though the higher-derivative theories are mathematically self-consistent, there are distinctive features of unconstrained higher-derivative theories that set them apart from similar lower-derivative theories. There are more degrees of freedom, associated with new solutions called "runaways," qualitatively different from those of a related lower-derivative theory. There is no lower-energy bound. In the case of field theories, these features can cause the problems of "ghost" fields and loss of unitarity.

There is a natural way to constrain many higher-derivative theories and save them from the above problems. It applies in cases where the higher-derivative terms are associated with a small, perturbative, expansion parameter. The method, called the method of perturbative constraints, is to exclude solutions that have no Taylor expansion in that small expansion parameter. The effect is to throw away all the runaway, negative-energy solutions. Without the perturbative constraints, higher-order terms in the expansion contribute as much as the lower-order terms, not commensurately less.

Higher-derivative theories that are expansions of a nonlocal theory require these perturbative constraints to give the same results as the full nonlocal theory. The perturbative constraints are actually present implicitly in the full theory, but they must be included explicitly for any finite expansion. Important examples of this case, where the perturbative constraints should be (but have not been) applied include higher-order corrections to general relativity from string theory, and to cosmic strings from the original gauge theory from which they arise. Nonlocality is a common feature in low-energy effective theories, and is not at all necessarily present in the full theory from which they are derived.

Higher-derivative theories which are not necessarily a truncated version of an infinite series, but can still be viewed as corrections to a valid lower-derivative theory, can also reap the benefits of the method of perturbative constraints. The constrained theory will resemble the original, lower-order theory in its solutions and number of degrees of freedom, and will have a lower-energy bound, all of which one would hope for in a perturbative expansion. The constraints are necessary if the perturbative higher-derivative corrections are to produce perturbative solutions. The unconstrained version would have all of the associated problems of higher-derivative theories, and the higher-derivative "correction" would completely dominate the behavior of the solutions, complete with negative-energy modes.

ACKNOWLEDGMENTS

The author would like to especially thank the Department of Applied Mathematics and Theoretical Physics at the University of Cambridge, where most of this work was done, for its warm hospitality and for the many important conversations held with its members. Special thanks go to Tanmay Vachaspati and Nico Giulini for many helpful conversations, and to Jim Hartle for his help, guidance, and support. This work was supported in part by NSF Grant No. PHY85-06686.

APPENDIX A

The quantized version of the system presented in Eq. (10) can be solved exactly by algebraic methods. We

define the new canonical variables

$$
\begin{aligned}
q_+ &= \frac{1}{\omega}\frac{1}{\sqrt{1-\epsilon^2\omega^2}}(\epsilon^2\omega^2\dot{x}-\pi_x)\ , \\
p_+ &= \frac{\omega}{\sqrt{1-\epsilon^2\omega^2}}(x-\pi_{\dot{x}})\ , \\
q_- &= \frac{\epsilon}{\sqrt{1-\epsilon^2\omega^2}}(\dot{x}-\pi_x)\ , \\
p_- &= \frac{1}{\epsilon}\frac{1}{\sqrt{1-\epsilon^2\omega^2}}(\epsilon^2\omega^2 x-\pi_{\dot{x}})\ ,
\end{aligned}
\tag{A1}
$$

such that the Hamiltonian is in the form of the difference of two harmonic oscillators:

$$
H = p_+^2+\omega^2 q_+^2-(p_-^2+\epsilon^{-2}q_-^2)\ . \tag{A2}
$$

The energy spectrum is then given by

$$
E=(n+\tfrac{1}{2})\omega-(m+\tfrac{1}{2})\epsilon^{-1} \quad \text{for} \quad n,m=0,1,2,\ldots\ . \tag{A3}
$$

The wave function can then be put in the form

$$
\psi_{nm}(\pi_x,\dot{x})=\chi_{n,\omega}(q_+)\chi_{m,\epsilon^{-1}}(q_-)\ , \tag{A4}
$$

where $\chi_{n',\omega'}$ is the standard simple harmonic-oscillator wave function with energy level n' and frequency ω', and $q_\pm$ are defined above. $\psi_{nm}(x,\dot{x})$ is given by the Fourier transform: e.g.,

$$
\begin{aligned}
\psi_{00}(x,\dot{x}) &= \frac{1}{\sqrt{2\pi}}\int d\pi_x e^{ix\pi_i}\psi_{00}(\pi_x,\dot{x}) \\
&= \left[\frac{\omega}{\epsilon\pi^2}\right]^{1/4}\exp\left[-\frac{\omega^2(1-\epsilon^2\omega^2)x^2-4i\epsilon\omega^2x\dot{x}-(1-\epsilon^2\omega^2)\dot{x}^2}{2\epsilon^{-1}(1+\epsilon^2\omega^2)}\right]\ .
\end{aligned}
\tag{A5}
$$

APPENDIX B

The rules for reading off the expectation values from the Hamiltonian are simple:

$$
\begin{aligned}
&\langle q^{(n)}\rangle = i\hbar\delta\times\text{coefficient of } (p_{q^{(n)}})\ , \\
&\langle (q^{(n)})^2\rangle = i\hbar\delta\times\text{coefficient of } ((p_{q^{(n)}})^2)\ , \quad \text{etc.}\ ,
\end{aligned}
\tag{B1}
$$

but they only apply when the Hamiltonian and/or Schrödinger formulation is equivalent to the Feynman path-integral formulation of quantum mechanics (for more details on when this is true, see, e.g, Popov[34]).

To calculate meaningful quantities, take the expectation value of classical expressions (since path integrals are semiclassical approximations at the smallest scale). For example, when calculating the expectation value of velocity, take the transition expectation value of

$$
v_{\rm cl}\sim\dot{x}+O(\Delta x/\delta)\ . \tag{B2}
$$

For the example, in Eq. (20), $v_{\rm cl}$ is given exactly by the derivative of (C1).

APPENDIX C

The general solution for the system described by (20) is

$$
\begin{aligned}
x(t)=\{&(\dot{x}_i-\dot{x}_f)\sin[\epsilon^{-1}(T-t)]+(\epsilon^{-2}tx_i-\epsilon^{-2}tx_f-\epsilon^{-2}Tx_i-\dot{x}_i+\dot{x}_f)\sin(\epsilon^{-1}T) \\
&+(\dot{x}_i-\dot{x}_f)\sin(\epsilon^{-1}t)-\epsilon^{-1}(T\dot{x}_i+x_i-x_f)\cos[\epsilon^{-1}(T-t)] \\
&+\epsilon^{-1}(T\dot{x}_f+x_i-x_f)\cos(\epsilon^{-1}t)+\epsilon^{-1}(-t\dot{x}_i-t\dot{x}_f+T\dot{x}_i-x_i-x_f)\cos(\epsilon^{-1}T) \\
&+\epsilon^{-1}(t\dot{x}_i+t\dot{x}_f+T\dot{x}_f+x_i+x_f)\}\epsilon[-2\cos(\epsilon^{-1}t)-\epsilon^{-1}T\sin(\epsilon^{-1}t)+2]^{-1}
\end{aligned}
\tag{C1}
$$

for a particle beginning its motion at x_i at time $t=0$ and ending at x_f at time $t=T$. From this we can compute the classical action

$$
\begin{aligned}
S=\frac{\epsilon}{2}E^{-1}[&A(\dot{x}_i^2+\dot{x}_f^2)-2B\dot{x}_i\dot{x}_f \\
&-2\epsilon^{-1}C(\dot{x}_i+\dot{x}_f)(x_i-x_f) \\
&+\epsilon^{-2}D(x_i-x_f)^2]\ ,
\end{aligned}
\tag{C2}
$$

where

$$
\begin{aligned}
A &= -4\epsilon^{-1}T\cos\epsilon^{-1}T+3\epsilon^{-1}T\cos2\epsilon^{-1}T \\
&\quad +4\sin\epsilon^{-1}T+(\epsilon^{-2}T^2-2)\sin2\epsilon^{-1}T+\epsilon^{-1}T\ , \\
B &= 4\epsilon^{-1}T\cos\epsilon^{-1}T+\epsilon^{-1}T\cos2\epsilon^{-1}T \\
&\quad +2(\epsilon^{-2}T^2+2)\sin\epsilon^{-1}T-2\sin2\epsilon^{-1}T-5\epsilon^{-1}T\ , \\
C &= 8\cos\epsilon^{-1}T-2\cos2\epsilon^{-1}T+2\epsilon^{-1}T\sin\epsilon^{-1}T \\
&\quad -\epsilon^{-1}T\sin2\epsilon^{-1}T-6\ , \\
D &= \epsilon^{-1}T\cos2\epsilon^{-1}T+4\sin\epsilon^{-1}T-2\sin2\epsilon^{-1}T-\epsilon^{-1}T\ , \\
E &= 16\cos\epsilon^{-1}T+(\epsilon^{-2}T^2-4)\cos2\epsilon^{-1}T \\
&\quad +8\epsilon^{-1}T\sin\epsilon^{-1}T-4\sin2\epsilon^{-1}T-\epsilon^{-2}T^2-12\ .
\end{aligned}
\tag{C3}
$$

Because the Lagrangian is quadratic, the quantum transition amplitude given by the path-integral formulation is exactly

$$K(x_f,\dot{x}_f,t_i+T;x_i,\dot{x}_i,t_i)=F(T)e^{\imath S}, \tag{C4}$$

where $F(T)$ can be derived from the classical action alone (see, e.g., Marinov[35] for details):

$$F(T)=\frac{1}{2\pi i}\left[\frac{-BD-C^2}{E^2}\right]^{1/2}. \tag{C5}$$

Schrödinger's equation for this system can be obtained directly from (C4) without the use of canonical formalism or the Hamiltonian.[18]

APPENDIX D

One may always add any total derivative to a Lagrangian without affecting the equations of motion, but not every Lagrangian had a valid variational formulation associated with it, and without one, the associated quantum-mechanical wave function will not fold. A simple example of a Lagrangian without a valid variational formulation is

$$L_\alpha=\tfrac{1}{2}(\dot{x}^2-\omega^2x^2)+\alpha\frac{d}{dt}(x\dot{x}) \quad \text{for } \alpha\neq 0,-1 \tag{D1}$$

for which the variational principle is

$$\delta S_\alpha=-\int_1^2(\ddot{x}+\omega^2x)\delta x\,dt+(1+\alpha)\dot{x}\,\delta x\,|_1^2+\alpha x\,\delta\dot{x}\,|_1^2 . \tag{D2}$$

This tells us to fix four boundary conditions for a second-order equation, which is an overdetermined system. Any Lagrangian lacking a valid variational formulation can regain it by adding a total derivative. In this case, the most obvious total derivative to add is $-\alpha(d/dt)(x\dot{x})$. Note that for $\alpha=-1$ there is a valid variational formulation (even though there are also constraints); this is the theory obtained by choosing the canonical momentum as the generalized coordinate. (A valid variational formulation is not needed to put the system into canonical form, it is only necessary for quantum mechanics.)

From Eq. (4) and from calculating the $p_{x^{(n)}}$ for our nonlocal oscillator [expressed as the infinite sum in the second line of (34)], we can see that the model Lagrangian presented above does have a valid variational formulation, once the implied constraints of (35) have been taken into account. The implied constraints tell us that holding all even derivatives fixed on the boundaries holds x fixed, and holding all odd derivatives fixed holds $\dot{x}$ fixed. Because the $p_{x^{(n)}}$ vanish for odd n,

$$\delta S=\int_1^2(\text{equations of motion})+(\text{some function})\delta x\,|_1^2 \tag{D3}$$

once the constraints are used. It is correct and, in fact, necessary to use the constraints to determine whether or not the system is over determined, as they are just part of the equations of motion (cf. the case of L_α). In general, adding an arbitrary total derivative to the Lagrangian, if it contains higher derivatives, corresponds to a canonical change in variable, which would destroy the variational formulation.

*Electronic address: JSimon@Voodoo.

[1]N. D. Birrell and P. C. W. Davies, *Quantum Fields in Curved Space* (Cambridge University Press, Cambridge, England, 1982).
[2]S. P. de Alwis, Phys. Rev. D **34**, 3760 (1986).
[3]T. L. Curtright, G. I. Ghandour, and C. K. Zachos, Phys. Rev. D **34**, 3811 (1986).
[4]K. Maeda and N. Turok, Phys. Lett. B **202**, 376 (1988); R. Gregory, *ibid.* **199**, 206 (1988).
[5]P. A. M. Dirac, Proc. R. Soc. London **A167**, 148 (1938).
[6]J. A. Wheeler and R. P. Feynman, Rev. Mod. Phys. **21**, 425 (1949).
[7]E. J. Kerner, J. Math. Phys. **3**, 35 (1962).
[8]H. J. Bhabha, Phys. Rev. **70**, 759 (1946).
[9]X. Jaén, J. Llosa, and A. Molina, Phys. Rev. D **34**, 2302 (1986).
[10]D. A. Eliezer and R. P. Woodard, Nucl. Phys. **B325**, 389 (1989).
[11]K. S. Stelle, Gen. Relativ. Gravit. **9**, 353 (1978).
[12]David G. Boulware, S. Deser, and K. S. Stelle, in *Quantum Field Theory and Quantum Statistics: Essays in Honor of the Sixtieth Birthday of E. S. Fradkin*, edited by I. A. Batalin, C. J. Isham, and C. A. Vilkovisky (Hilger, Bristol, England, 1987).
[13]K. S. Stelle, Phys. Rev. D **16**, 953 (1977).
[14]David G. Boulware, in *Quantum Field Theory and Quantum Statistics: Essays in Honor of The Sixtieth Birthday of E. S. Fradkin* (Ref. 12).
[15]A. A. Starobinsky, Phys. Lett. **91B**, 99 (1980).
[16]M. Ostrogradski, Mem. Act. St. Petersbourg **VI 4**, 385 (1850).
[17]S. W. Hawking, in *Quantum Field Theory and Quantum Statistics: Essays in Honor of the Sixtieth Birthday of E. S. Fradkin* (Ref. 12).
[18]R. P. Feynman, Rev. Mod. Phys. **20**, 267 (1948); R. P. Feynman and A. R. Hibbs, *Quantum Mechanics and Path Integrals* (McGraw-Hill, New York, 1965).
[19]T. D. Lee and G. C. Wick, Nucl. Phys. **B9**, 209 (1969).
[20]G. T. Horowitz, Phys. Rev. D **31**, 1169 (1985).
[21]J. D. Jackson, *Classical Electrodynamics* (Wiley, New York, 1975).
[22]A. Strominger, Phys. Rev. D **30**, 2257 (1984).
[23]D. Lovelock, J. Math. Phys. **12**, 498 (1971); **13**, 874 (1972).
[24]D. G. Boulware and S. Deser, Phys. Rev. Lett. **55**, 2656 (1985); J. T. Wheeler, Nucl. Phys. **B268**, 737 (1986); **B273**, 732 (1986); R. C. Myers and J. Z. Simon, Phys. Rev. D **38**, 2434 (1988).
[25]J. M. Pons, Lett. Math. Phys. **17**, 181 (1989).
[26]P. A. M. Dirac, *Lectures on Quantum Mechanics* (Yeshiva University, New York, 1964).
[27]A. Hanson, T. Regge and C. Teitelboim, *Constrained Hamiltonian Systems* (Acad. Naz. dei Lincei, Rome, 1976).

[28]J. A. Wheeler and R. P. Feynman, Rev. Mod. Phys. **17**, 157 (1945).
[29]L. Bel and H. Sirousse-Zia, Phys. Rev. D **32**, 3128 (1985).
[30]Vachaspati and T. Vachaspati, Phys. Lett. B **238**, 41 (1990).
[31]Y. Saito, R. Sugano, T. Ohta, and T. Kimura, J. Math. Phys. **30**, 1122 (1989).
[32]G. Schäfer, Phys. Lett. **100A**, 128 (1984).
[33]J. Z. Simon (in preparation).
[34]V. N. Popov, *Functional Integrals in Quantum Field Theory and Statistical Physics* (Reidel, Boston, 1983).
[35]M. S. Marinov, Phys. Rep. **60**, 1 (1980).