Lagrangian vs Hamiltonian: The best approach to relativistic orbits

Richard H. Price
Department of Physics, MIT, 77 Massachusetts Avenue, Cambridge, Massachusetts 02139

Kip S. Thorne
California Institute of Technology, Pasadena, California 91125

In introductory general relativity courses, free particle trajectories, such as astronomical orbits, are generally developed via a Lagrangian and variational calculus, so that physical examples can precede the mathematics of tensor calculus. The use of a Hamiltonian is viewed as more advanced and typically comes later if at all. We suggest here that this might not be the optimal order in a first course in general relativity, especially if orbits are to be solved with numerical methods. We discuss some of the issues that arise in both the Lagrangian and Hamiltonian approaches. © *2018 American Association of Physics Teachers.*

I. INTRODUCTION AND SUMMARY

In a recent paper,[1] one of us (KST) described the numerical computation of photon orbits in a wormhole metric to produce images for the science fiction movie *Interstellar*. That description emphasized the convenience of basing the equations of motion on the Hamiltonian rather than the more traditional Lagrangian-based approach. This relative convenience was rooted in computational difficulties inherent in the traditional approach.

Here, we will compare the two approaches somewhat more broadly, but not too broadly. We will avoid some of the interesting questions lurking in the background of the Hamiltonian-vs.-Lagrangian choice such as the availability of the advanced techniques of canonical transformations, adiabatic invariants, the connection to quantum mechanics, and thermodynamics, etc. We will be concerned only with the Hamiltonian-vs.-Lagrangian choice in computing particle and photon orbits in a curved spacetime.

Below, within that limited scope, we will deal with the following claims and will present coupled counterclaims: (i) Claim: The Lagrangian formulation allows orbital equations to be presented with a minimum of mathematical preliminaries; this explains why the Hamiltonian approach is not presented in most texts. Counterclaim: we will argue that the Hamiltonian approach *could* be presented with minimal mathematical foundation; it is simply not the tradition. (ii) Claim: The Hamiltonian approach is superior because it leads to first-order equations of motion that are better for numerical integration, not the second-order equations of the Lagrangian approach. Counterclaim: The second order ODE *could* always be recast as two coupled first-order ODEs; we shall demonstrate this explicitly in one of the examples. (iii) Claim: The Hamiltonian approach makes constants of motion more apparent. Counterclaim: Without the advanced methods of Hamiltonian mechanics this is not true (as will be shown in an example in one of the appendices). (iv) Claim: The equations of motion of the Hamiltonian formulation lack the singularities that are a difficulty in working with the Lagrangian equations. (This was the primary motive in Ref. 1 for extolling the Hamiltonian approach.) Counterclaim: As we shall show, the singularities are a consequence of seeking a "solution by quadrature," and even in the solution by quadrature, it is possible to overcome this difficulty with techniques detailed in another of the Appendixes.

This short paper is organized as follows: In Sec. II, we briefly review the traditional Lagrangian approach, and in Sec. III we outline the Hamiltonian approach. The two methods are compared in Sec. IV for orbits in a spherically symmetric spacetime. We conclude and summarize in Sec. V. In three appendices, we present some details of techniques for dealing with singular integrals, and two examples of the Hamiltonian approach.

Throughout this paper we use the notation of the text *Gravitation*.[2] In particular, we use the sign convention $-+++$ and units in which $c = G = 1$.

II. TRADITIONAL APPROACH: CONSTANTS OF MOTION AND QUADRATURE

Though a geodesic is best described as a locally straight spacetime curve, the implementation of that description requires a substantial background in the mathematics of curved spacetime: covariant differentiation, Christoffel symbols, etc. Introductory texts, in a hurry to get to the physics of orbits, defer the mathematics of curved spacetime by introducing a geodesic as a curve $x^\mu(\lambda)$ that is the extremum of the action integral $\int \mathcal{L} d\lambda$ along the curve, where[3,4]

$$\mathcal{L} = \frac{1}{2} g_{\mu\nu} \dot{x}^\mu \dot{x}^\nu. \tag{1}$$

Here, $\dot{x}^\mu \equiv dx^\mu/d\lambda$, and λ is an affine parameter that is typically taken to be proper time τ for a particle with nonzero mass m so that $\dot{x}^\mu = U^\mu$, a component of the particle 4-velocity. Equally well, λ can be taken to be τ/m, so that $\dot{x}^\mu = P^\mu$, a component of the particle 4-momentum. For a photon, the proper time does not exist; λ is typically defined so that $\dot{x}^\mu = P^\mu$, the photon 4-momentum. For uniformity, and simplicity, we will let $\dot{x}^\mu = P^\mu$ below, for both massive and massless particles.

The equations of motion follow from the Euler-Lagrange (EL) equations $\delta \mathcal{L}/\delta x^\mu = 0$, that is,

$$\frac{d}{d\lambda}\left(\frac{\partial \mathcal{L}}{\partial \dot{x}^\mu}\right) = \frac{\partial \mathcal{L}}{\partial x^\mu}. \tag{2}$$

This shortcut to the equations of motion largely explains the focus in introductory texts on the Lagrangian-based equations of motion.

The most typical example used in texts is orbits in a spherically symmetric static spacetime; the wormhole metric in Ref. 1 is of this type. We will use this typical case to illustrate the traditional approach. That approach uses standard spherical polar coordinates[5] t, r, θ, ϕ. Since neither t nor ϕ appears in the metric coefficients, the EL equations tell us that $\partial\mathcal{L}/\partial\dot{t} = P_t$ and $\partial\mathcal{L}/\partial\dot{\phi} = P_\phi$ are constants. Following tradition, we call these, respectively—E and L_z; they are constants of motion and they show up just as directly in the Lagrangian approach as in the Hamiltonian.

In the traditional (Lagrangian) approach, we combine these two constants with a third constant of motion, the value of $\mathcal{L}$ itself. It is constant since, by Eq. (1), $\mathcal{L}$ is just half the square of the mass of the particle whose trajectory is being computed. We exploit symmetry to take the orbital plane to be the equatorial plane, so that $\theta = \pi/2$ is, in a sense, another constant of motion. With this set of four constants we find, as closed form functions of r, the first derivatives $dr/d\lambda$, $dt/d\lambda$, $d\phi/d\lambda$ with the fourth coordinate θ constant at $\pi/2$. These first integrals are then solved for the trajectory, i.e., for t, r, θ and ϕ as functions of λ, by integrating. One says the orbit has been solved by, or reduced to "quadrature," a somewhat old-fashioned name for integration.

The reduction of a problem to an integral has many advantages. Often an integral is easier to evaluate numerically than the solution of coupled differential equations. The integral also has the advantage of showing the whole solution "at once," rather than showing a step-by-step process. This can be valuable for seeing the dependence on parameters, the location of turning points, etc.

What are the disadvantages of this traditional method, disadvantages that might push us to investigate the Hamiltonian alternative? A consideration—not a disadvantage *per se*—is that the numerical solution of differential equations is no longer a barrier; computation is part of a physics curriculum. A real disadvantage, specific to orbits, is that the quadrature integrals have singularities. These singularities are integrable, but lead to sign changes at orbital turning points that are cumbersome in numerical work. For tasks such as the ray tracing in Ref. 1 one also wants photon orbits, in many different planes, and one does not want to have to transform back and forth between the working coordinate frame and the frame in which the equatorial plane is the orbital plane for each photon being considered.

Appendix A gives some details of these inconveniences. It also shows how they can be overcome. The attraction of the Hamiltonian approach then is not so much that it is needed, but that it is simple and has no such disadvantages.

III. THE HAMILTONIAN APPROACH: COVARIANT MOMENTA

The standard path from a Lagrangian formalism to a Hamiltonian formalism starts with the Lagrangian viewed as a function of generalized coordinates q_j, and their derivatives $\dot{q}_j$ with respect to a trajectory parameter, and with the introduction of the Hamiltonian

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

The second step is to consider the spacetime coordinates to be the generalized coordinates, to take the affine parameter λ

as the trajectory parameter, and to define the "canonical momenta" by $P_\mu \equiv \partial\mathcal{L}/\partial\dot{x}^\mu$.

When we apply this procedure to the expression in Eq. (1), changing notation *mutatis mutandis*, we find that the canonical momenta are simply the covariant components of the particle 4-momentum, and the Hamiltonian is

$$\mathcal{H} = \dot{x}^\alpha \frac{\partial\mathcal{L}}{\partial\dot{x}^\alpha} - \mathcal{L}, \tag{4}$$

where now, as usual, summation over repeated indices is assumed. With $\mathcal{L}$ given in Eq. (1) and $\dot{x}^\alpha = P^\alpha = g^{\alpha\beta}P_\beta$, we can write the Hamiltonian in the useful form

$$\mathcal{H} = \frac{1}{2} g^{\mu\nu} P_\mu P_\nu, \tag{5}$$

and we have what we need for Hamilton's equations of motion

$$\frac{dP_\mu}{d\lambda} = -\frac{\partial\mathcal{H}}{\partial x^\mu} \quad \frac{dx^\mu}{d\lambda} = \frac{\partial\mathcal{H}}{\partial P_\mu}. \tag{6}$$

One of the points that we most want to emphasize is that the Hamiltonian (5) does not have to be introduced via the Lagrangian, as above. The right hand side of Eq. (5), aside from an irrelevant factor of $-1/2$, is simply the square of the total rest-mass-energy of a particle. Students are almost as likely to have been exposed to the basic Hamilton equations of motion as to EL variation. Introducing Eq. (5) as the relativistic extension of the Newtonian particle energy "$T + V$" is therefore about as palatable as the variational principle for orbital equations of motion.

Let us now compare the relativistic mathematical preliminaries necessary for introducing equations of motion in the Hamiltonian vs. Lagrangian formalism in a beginning course. For the Lagrangian approach, the student needs to have only the concept of a line element written in a coordinate basis, i.e., the metric expression; cf. Eq. (1). For the Hamiltonian approach, the student also needs the contravariant components $g^{\mu\nu}$ of the metric tensor $g_{\mu\nu}$, and how to raise and lower indices with these objects; cf. Eq. (5). We believe that the student cannot help but learn these elements early on in any case, so that the argument is weak that the Lagrangian approach is favored in an introductory course. Although some might make that argument, there can be no question that *neither approach requires the mathematics of curved spacetime*.

IV. COMPARISON OF EQUATIONS OF MOTION

To understand how the two methods differ in actual usage, let us take, again, the paradigmatic example of motion in a static spherical spacetime. With the radial coordinate taken to be the usual "areal" radius, we will write the metric as

$$ds^2 = -A(r)dt^2 + B(r)dr^2 + r^2(d\theta^2 + \sin^2\theta d\phi^2). \tag{7}$$

The Lagrangian approach starts with

$$\mathcal{L} = \frac{1}{2}\left(-A\dot{t}^2 + B\dot{r}^2 + r^2\dot{\theta}^2 + r^2\sin^2\theta\dot{\phi}^2\right). \tag{8}$$

Following the traditional approach and specifying that our particle is a photon (so that $\mathcal{L} = 0$), we arrive at

$$0 = -\frac{1}{A}E^2 + \frac{1}{B}\left(\frac{dr}{d\lambda}\right)^2 + \frac{1}{r^2}L_z^2, \qquad (9)$$

at which point, or soon after, we exclaim "solved by quadrature."

The Hamiltonian approach starts by identifying the non-zero elements of the inverse metric

$$g^{tt} = -\frac{1}{A} \quad g^{rr} = \frac{1}{B} \quad g^{\theta\theta} = \frac{1}{r^2} \quad g^{\phi\phi} = \frac{1}{r^2 \sin^2 \theta}, \quad (10)$$

with which the Hamiltonian [Eq. (5)] becomes

$$\mathcal{H} = \frac{1}{2}\left(-\frac{1}{A}P_t^2 + \frac{1}{B}P_r^2 + \frac{1}{r^2}P_\theta^2 + \frac{1}{r^2 \sin^2\theta}P_\phi^2\right). \qquad (11)$$

With θ fixed at $\pi/2$ the equations of motion for this Hamiltonian are

$$\frac{dP_t}{d\lambda} = 0 \Rightarrow P_t \equiv -E, \quad \frac{dt}{d\lambda} = -\frac{P_t}{A}, \qquad (12)$$

$$\frac{dP_\phi}{d\lambda} = 0 \Rightarrow P_\phi \equiv L_z, \quad \frac{d\phi}{d\lambda} = \frac{P_\phi}{r^2}, \qquad (13)$$

$$\frac{dP_r}{d\lambda} = -\frac{1}{2}\left[-\frac{d}{dr}\left(\frac{1}{A}\right)E^2 + \frac{d}{dr}\left(\frac{1}{A}\right)P_r^2 - \frac{2}{r^3}L_z^2\right],$$
$$\frac{dr}{d\lambda} = \frac{P_r}{B}. \qquad (14)$$

The central difference of the two methods lies in the comparison of Eqs. (14) and (9) for $r(\lambda)$. The first is a pair of coupled first order differential equations with no obvious singularities; the second is a single first order equation with singularities at turning points. The first pair of equations is very well suited to solution for $P_r(\lambda)$ and $r(\lambda)$ by straightforward numerical methods for ordinary differential equations. The second method leads to an integral that can be evaluated numerically but not with straightforward methods (see Appendix A).

It is important to notice that Eq. (14) can be extracted from the Lagrangian method as well as from the Hamiltonian method. The EL equation from Eq. (8), for the r variable, is

$$\frac{d}{d\lambda}(B\dot{r}) = \frac{1}{2}\left[-\frac{dA}{dr}\dot{t}^2 + \frac{dB}{dr}\dot{r}^2 + \frac{2}{r^3}L_z^2\right]. \qquad (15)$$

If we add to this the definition in the second part of Eq. (14), the result (with $\dot{t}$ and $\dot{\phi}$ eliminated in favor of E and L_z) is just the first part of Eq. (14). Notice that we can define $\psi \equiv B dr/d\lambda$, without any guidance from the fact that it is P_r, and write

$$\frac{d\psi}{d\lambda} = \frac{1}{2}\left[-\frac{A_{,r}}{A^2}E^2 + \frac{B_{,r}}{B^2}\psi^2 + \frac{2}{r^3}L_z^2\right] \quad \frac{dr}{d\lambda} = \frac{\psi}{B}. \qquad (16)$$

We have arrived at the Hamilton equations without the Hamiltonian approach.

An important point to note here is that the solution of these Lagrangian-based equations does not involve dealing with any singularities. The singularities appear only when the value of the Lagrangian is added as an additional constraint in order to arrive at a "solution" by quadrature. It seems counterintuitive that adding information[6] (the constant

value of the Lagrangian) makes the solution more difficult, but the difficulty is only a technical one associated with a quest for quadrature.

The Hamiltonian approach does have an advantage in that it works from the outset with the canonical momenta; they do not have to be constructed (as $\partial\mathcal{L}/\partial\dot{x}^\mu$) in the Lagrangian process. This difference is trivial for diagonal metrics. For metrics that are not diagonal, the Hamiltonian approach can result in considerable simplification. We give such an example in Appendix C. The use of canonical momenta at the outset, however, should not be misinterpreted as aiding the identification of constants of motion. Appendix B presents an example in which the symmetry is hidden in both approaches.

V. CONCLUSION

The traditional (Lagrangian) approach to the equations of motion in general relativity is founded on the path of least required mathematics; the student needs to have experience only with standard variational calculus. We have shown that this method is adequate, but a perhaps better (Hamiltonian) method requires perhaps less mathematics of a student; a student must only accept that it is the covariant components of particle 4-momentum that are the canonical momenta for the equations of motion. The only remaining step is to take the Hamiltonian to be half the square of the 4-momentum, the negative of the squared particle mass, a constant which is an acceptable, perhaps compelling, if not obvious, generalization of the constant "energy" for a nonrelativistic system. The resulting Hamiltonian approach simplifies both numerical and analytical work with the equations of motion.

APPENDIX A: DIFFICULTIES AND REMEDIES IN THE TRADITIONAL APPROACH

Because the first derivative $dr/d\lambda$ occurs quadratically in Eq. (1), the expression for $dr/d\lambda$ contains a square root, and the square root changes sign at "turning points," values of r at which $dr/d\lambda = 0$. The quadrature for solving for $r(\lambda)$ [more precisely for $\lambda(r)$] will therefore have the form

$$\lambda = \pm\int^r \frac{dr'}{\sqrt{E - V(r')}}. \qquad (A1)$$

Typically one divides $dr/d\lambda$ by $d\phi/d\lambda$ to eliminate the unnecessary reference to λ and solves for the orbital shape $r(\phi)$ [more accurately for $\phi(r)$]. The resulting integral for ϕ has the same form as in Eq. (A1).

The integral in Eq. (A1) may have one or two turning points. (That is, the expression inside the square root may have one or two first-order zeros.) The two-turning-points case occurs for bound orbits. Let us call these turning points r_1 and $r_2 > r_1$. We can then write the quadrature as

$$\lambda = \pm\int^r \frac{dr'}{\sqrt{(r_2 - r')(r' - r_1)S(r')}}. \qquad (A2)$$

Here, the function S is taken to be positive on the interval $r_1 \leq r \leq r_2$. Though the integral in Eq. (A2) is finite, evaluation of the integral with a standard method, e.g., Simpson's rule, will generally involve evaluating the integrand at points at which the denominator vanishes. A student encountering such an integrand might try to take very small integration

intervals $\Delta r'$, in the neighborhood of the singularities, but any such approach involves unnecessarily large truncation errors for a given $\Delta r'$, and requires dealing with the sign changes in Eqs. (A1) and (A2).

There are much more efficient ways to handle the singularities.[7] The simplest is to make the substitution

$$r' = \frac{r_2 - r_1}{2} \cos \xi + \frac{r_2 + r_1}{2}, \tag{A3}$$

$$\lambda = \pm \int^{\xi(r)} \frac{d\xi}{\sqrt{S(\xi)}}, \tag{A4}$$

which has no singularities. [In the case of a single zero, say at $r = r_1$, in the square root, as there would be for a hyperbolic orbit, the transformation $r = r_1 + \xi^2$ converts the problem to one with the form of Eq. (A4).] The singularity is thus eliminated, costing only the additional computational overhead of inverting Eq. (A3).

In addition to the problem with turning points, the second inconvenience of the traditional approach to orbits is the need to have the photon orbit always in the equatorial plane. This is simple to eliminate. One removes the choice $\theta = \pi/2$ and replaces it with the constant of motion

$$L^2 = p_\theta^2 + L_z / \sin^2 \theta. \tag{A5}$$

With this fourth general constant,[8] motion in any plane is reduced to quadrature.

APPENDIX B: EXAMPLE 2: HIDDEN SYMMETRY

We demonstrate here that constants of motion may not be obvious, in either the Lagrangian or the Hamiltonian formalism. Let us consider the spacetime metric

$$ds^2 = -e^{2\Phi} dt^2 + a^2(\sinh^2 u + \sin^2 v)(du^2 + dv^2) + dz^2, \tag{B1}$$

where a is a constant and Φ is a function only of z. The Hamiltonian for this spacetime is

$$H = \frac{1}{2} \left[-e^{-2\Phi} P_t^2 + \frac{P_u^2 + P_v^2}{a^2(\sinh^2 u + \sin^2 v)} + P_z^2 \right]. \tag{B2}$$

Two of the equations of motion are simple

$$P_t = \text{constant} \quad \frac{dP_z}{ds} = -\Phi_{,z} e^{-2\Phi} P_t^2. \tag{B3}$$

The equations for P_u, P_v are not simple

$$\frac{dP_u}{ds} = -\frac{1}{a^2} \left(P_u^2 + P_v^2 \right) \frac{\partial}{\partial u} \left(\frac{1}{\sinh^2 u + \sin^2 v} \right) \tag{B4}$$

$$\frac{dP_v}{ds} = -\frac{1}{a^2} \left(P_u^2 + P_v^2 \right) \frac{\partial}{\partial v} \left(\frac{1}{\sinh^2 u + \sin^2 v} \right). \tag{B5}$$

It turns out, however, that the combination

$$P_{\text{cons}} \equiv \frac{\sin v \cos v \, P_u + \sinh u \cosh u \, P_v}{\sinh^2 u + \sin^2 v} \tag{B6}$$

is a constant of motion, i.e., $dP_{\text{cons}}/ds = 0$. Thus this constant can be used to eliminate (say) P_v in Eq. (B4) and in the equations for du/ds and dv/ds.

It is not a coincidence that $dP_{\text{cons}}/ds = 0$ is a constant. In terms of new coordinates

$$x = a \cosh u \cos v \quad y = a \sinh u \sin v \tag{B7}$$

the u, v part of the metric is transformed according to

$$a^2(\sinh^2 u + \sin^2 v)(du^2 + dv^2) \rightarrow dx^2 + dy^2, \tag{B8}$$

which in turn can be transformed to the two-dimensional flat metric in plane polar coordinates r, ϕ. The constant of motion P_{cons} is in fact just P_ϕ, which we know to be a constant by symmetry.

APPENDIX C: EXAMPLE 3: HAMILTONIAN APPROACH WITH A NON-DIAGONAL METRIC

As an example, we consider trajectories $t(\lambda)$, $x(\lambda)$ in the yz plane of the spacetime with a metric

$$ds^2 = \frac{-dt^2 - 2x \, dx \, dt + dx^2}{1 + x^2} + dy^2 + dz^2, \tag{C1}$$

for which the Lagrangian is

$$\mathcal{L} = \frac{1}{2} \left(\frac{-\dot{t}^2 - 2x\dot{t}\dot{x} + \dot{x}^2}{1 + x^2} \right). \tag{C2}$$

(This metric has no physical importance that we know of, but is contrived to lead to a closed form solution.)

Since t is an ignorable coordinate we know that there is a conserved momentum P_t. This turns out to be

$$P_t = \frac{-\dot{t} + x\dot{x}}{1 + x^2}, \tag{C3}$$

where a dot signifies $d/d\lambda$. The EL equation $d(\partial \mathcal{L}/\partial \dot{x})/d\lambda = \partial \mathcal{L}/\partial x$ gives us

$$\frac{d}{d\lambda} \left(\frac{-x\dot{t} + \dot{x}}{1 + x^2} \right) = -\frac{\dot{x}\dot{t}}{1 + x^2}$$
$$- \frac{x}{(1 + x^2)^2} \left(-\dot{t}^2 - 2x\dot{t}\dot{x} + \dot{x}^2 \right). \tag{C4}$$

Equation (C4) can be converted to a differential equation for $x(t)$, by using $d/d\lambda = \dot{t} d/dt$, and by eliminating $\dot{t}$ with the use of Eq. (C3).

While this is possible in principle, it is quite tedious, so it is interesting to see how this compares with the tedium of the Hamiltonian approach. The extra effort[9] in that approach goes into finding the components of the inverse metric, the $g^{\mu\nu}$, which, for the metric of Eq. (C1), are $g^{xx} = g^{yy} = g^{zz} = -g^{tt} = 1$ and $g^{xt} = -x$. The Hamiltonian for motion in the yz plane is therefore

$$\mathcal{H} = \frac{1}{2} \left(-P_t^2 - 2xP_xP_t + P_x^2 \right). \tag{C5}$$

We rather easily get the following Hamilton equations in which, of course, P_t is a constant:

$$\frac{dP_x}{d\lambda} = P_t P_x, \quad \frac{dx}{d\lambda} = P_x - xP_t, \quad \frac{dt}{d\lambda} = -P_t - xP_x. \quad \text{(C6)}$$

The explicit solutions follow immediately:

$$P_x = \alpha e^{P_t \lambda}, \quad x = \frac{\alpha}{2P_t} e^{P_t \lambda} + \beta e^{-P_t \lambda},$$

$$t = -(P_t + \alpha\beta)\lambda + t_0 - \frac{\alpha^2}{4P_t^2} e^{-2P_t \lambda}. \quad \text{(C7)}$$

Here, α, β, and t_0 are constants, and are constrained by

$$\mathcal{H} = -\frac{1}{2} P_t(P_t + 2\alpha\beta) = 0 \quad \text{(C8)}$$

if the particle is a photon.

[1]O. James, E. von Tunzelmann, P. Franklin, and K. S. Thorne, "Visualizing interstellar's wormhole," Am. J. Phys. **83**(6), 486–499 (2015).
[2]C. W. Misner, K. S. Thorne, and J. A. Wheeler, *Gravitation* (W. H. Freeman, San Francisco, 1973); reprinted by Princeton University Press, 2017.
[3]The factor of 1/2 is included so that the quantities $\partial \mathcal{L}/\partial \dot{x}^\mu$ agree with the usual meaning of the component P^μ of the 4-momentum.
[4]It can be argued (quite correctly) that the right form of the Lagrangian should be the square root of the negative of the expression in Eq. (1). In the interest of simplicity we are cutting a corner here. It can be shown without too much difficulty that either form of the Lagrangian gives the same equations of motion if the Lagrangian itself has a constant value for the solution. In the case of an affine parameter, which is the case that we consider, the Lagrangian *is* a constant for the solution curve.
[5]For the wormhole of Ref. 1 the radial coordinate is ℓ, not r.
[6]When we add the information that the Lagrangian is constant, we are actually modifying the problem. Without this the curve parameter λ is unconstrained; when we specify that the Lagrangian we are constraining λ to be an affine parameter.
[7]The integral can be evaluated with good accuracy through the use of Gauss-Chebyshev integration, but this is not familiar to most undergraduates and has no real advantage over the substitution presented here.
[8]The "constant of motion" L^2 is a true integral of motion, so relativistic orbits in a spherical spacetime have a full complement—four—of first integrals. Interestingly, L^2 can be generalized to "the Carter constant," for the Kerr geometry, which is not spherically symmetric. See B. Carter, "Global structure of the Kerr family of gravitational fields," Phys. Rev. **174**(5), 1559–1571 (1968); Ref. 2, Sec. 33.5.
[9]With modern symbolic manipulation packages this effort is much easier than it had formerly been, adding another consideration in favor of the Hamiltonian approach.