Getting the most action out of least action: A proposal

Thomas A. Moore[a]

Department of Physics and Astronomy, Pomona College, 610 N. College Avenue, Claremont, California 91711

Lagrangian methods lie at the foundation of contemporary theoretical physics. Several recent articles have explored the possibility of making the principle of least action and Lagrangian methods a part of the first-year physics curriculum. I examine some of this proposal's implications for subsequent courses in the undergraduate physics major, and focus on the influence that this proposal might have on the selection of topics and the opportunities this proposal presents for teaching these courses in a more contemporary way. Many of these ideas are relevant even if students first learn Lagrangian methods in a sophomore mechanics course. © *2004 American Association of Physics Teachers.*
[DOI: 10.1119/1.1646133]

I. INTRODUCTION

Hamilton's principle,[1] more generally known as the principle of least action (particularly since the publication of Feynman's lectures[2]) has played a seminal role in the development of theoretical physics in the latter part of the 20th century. Lagrangian methods that extend this principle lie at the heart of general relativity, quantum field theory, and the standard model of particle physics, and such methods play a crucial role in conceptually framing and expressing these theories.

Edwin Taylor has recently argued that this principle provides a simple but powerful framework for unifying Newtonian mechanics, relativity, and quantum mechanics,[3] and he and his collaborators have begun to lay the foundations for teaching the principle in the introductory course.[4–8] If we presume that this proposal is possible and desirable, it has implications for subsequent courses in the physics major. In this article, I will examine some of these implications, focusing on new opportunities that teaching least action in the introductory course makes possible, as well as on what changes in upper-level courses might best support these opportunities in subsequent courses.

My purpose is not to describe a new upper-level curriculum in detail. Instead, I hope that by presenting an overview of the issues and providing references to some available resources, I will provide some guidance to those who might develop such curricula. This article also might be interesting to those seeking to modernize the upper-level courses that follow an intermediate mechanics course which discusses Lagrangian methods.

II. THE MODERN PHYSICS COURSE

Most upper-level physics curricula open with a course in "modern physics," which for the sake of argument in what follows, I will assume to be a sophomore-level class that at least discusses special relativity, some basic quantum theory, atomic and nuclear physics, and perhaps some particle physics. In a curriculum where the classical principle of least action is taught in introductory physics, the modern physics course might be reworked somewhat to address two important goals: connect the classical principle of least action with quantum mechanics and relativity, and build a solid foundation for using the principle in subsequent courses. I will discuss the link to quantum mechanics first (for reasons that will become clearer as we go).

Taylor, Vokos, O'Meara, and Thornber have recently published a curricular plan that connects quantum mechanics with the principle of least action at a level that seems appropriate for sophomores.[9] This plan starts with the students working through the first half of Feynman's popular book *QED*.[10] Feynman's book demonstrates that it is possible to explain the results of classical optics in a variety of practical situations using the following simple model: a photon explores all possible paths between emission and detection, we imagine the photon traveling along each possible path to carry an arrow that rotates a number of times that is proportional to the action along that path, and the probability that the photon will be observed at the detection event is proportional to the squared length of the vector sum of the final arrows for all the paths that the photon explores. Sophomore-level majors (unlike Feynman's intended audience) should be able to understand that the arrows are visual representations of complex numbers, but this visualization is powerful and useful even when students can do the calculations with complex numbers.

The fundamental problem with the "explore all paths" model is that actually summing the arrows over all possible paths is a daunting task. Taylor and his collaborators make this task simpler by providing computer programs that compute the sums for various simple paths so that students can explore the implications of the model. Building on this foundation, Taylor and his collaborators (aided by more programs) then extend Feynman's description to help students discover methods for handling free electrons and then electrons with potential energy, the concept of a wave function, the concept of the free-particle propagator, and ultimately the concept of a bound-state wave function, all with very little mathematics.

The method Taylor and his collaborators use to develop the free-particle propagator illustrates their general approach to making difficult ideas more accessible. The key to making the "explore many paths" approach practical is to get rid of the summation over all possible paths. The world line through spacetime between a given starting event *a* and a given ending event *b* that has the least action is by definition the world line along which the particle's arrow undergoes the fewest turns from start to finish. With the help of the computer programs, a student can find that the only paths that contribute significantly to the arrow representing the final sum at event *b* are those contributing final arrows that make an angle of less than π with the arrow contributed by the world line of least action; neither the length nor the direction
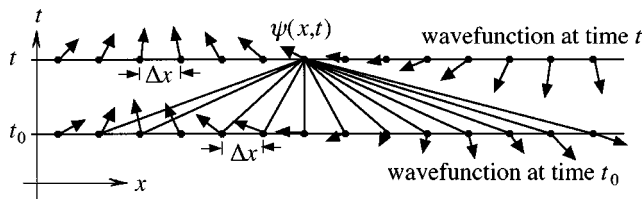
Fig. 1. We can calculate the wave function amplitude $\psi(x,t)$ at position x at time t by using Eq. (1) to calculate the contribution of the wave function amplitude $\psi(x_i,t_0)$ at a position x_i at an earlier time t_0 and then summing over all x_i. The diagonal lines show the direct paths that connect the various points x_i with the final point x.

of the sum is much affected if one ignores all other paths. Indeed, one finds that for a free particle, the direction (in the complex plane) of the arrow representing the sum at b is always rotated by 45° relative to the direction of the arrow contributed by the least-action world line at b (which in turn is simply a rotated version of the arrow at the initial event a), and the sum's magnitude depends on how far a path must deviate from the least-action world line to yield a contributed arrow that makes an angle of π with the least-action arrow.

Therefore it should be possible in principle to forego the sum entirely and calculate the arrow representing the sum over all paths by rotating the direction of the arrow contributed by the single least-action path by 45° and multiplying by a factor that specifies the degree to which small deviations from this path affect the angle of the path's contributed arrow. For a free particle, this factor can only be a function of the particle's mass m, Planck's constant h, the time interval between the initial and final events, and the spatial separation of those events. Taylor and his collaborators[9] argue that we can determine the correct expression for this factor by assuming that a free-particle wave function which is uniform over space at a certain time must remain uniform as time passes (a result required by symmetry). We can consider any wave function at a given time to be a set of arrows (that is, complex numbers) distributed over space. Assume that we know the wave function arrows $\psi(x_i,t_0)$ at various positions x_i at some initial time t_0. The arrow $\psi(x,t)$ at a different position x and later time t is determined by determining the sum of the arrows contributed by all paths starting from the arrow $\psi(x_i,t_0)$ at a given x_i at time t_0 and arriving at position x and time t, and then summing over all x_i (see Fig. 1). For the free particle, we can do the sum over all paths by calculating the arrow contributed by the least-action path from x_i,t_0 to x, t (which is a straight world line for a free particle) and use a formula (rule) involving h, m, $t-t_0$, and $x-x_i$ to convert this arrow to an arrow representing the sum over all paths. By using a program constructed for this purpose, students can experiment with different rules until they find one that preserves the uniform wave function. The process is quite intuitive and requires very little mathematics.

Once we know how to generate a future wave function from a past one, we can generalize to particles that are not free and begin to explore both stationary and dynamic states of bound particles.[9] After developing the general concept of a stationary state, we might introduce the Schrödinger equation and explore bound states of other systems in a more conventional manner.

The approach in Ref. 9 is plausibly accessible to sophomore-level physics majors, and has the advantage of giving these students a deeper, more intuitive, and perhaps more engaging understanding of quantum mechanics than one typically gets in a modern physics course. Moreover, this approach is the only way that I know in which we might plausibly link the classical principle of least action and quantum mechanics at this level. This approach, however, will take a fair amount of class time, and thus will probably displace some other topics usually covered in such a course.[11]

Next I would like to discuss the treatment of special relativity in the modern physics course. The argument about the propagator assumes that the reader understands what events, world lines, and spacetime diagrams are (Fig. 1 is essentially a spacetime diagram). Therefore, a careful treatment of these concepts in the relativity portion of the course is essential for the success of the quantum section. My experience is that taking the time to teach students to use spacetime diagrams and the geometric analogy to relativity before teaching the Lorentz transformation equations greatly improves their understanding. Students understand much better the meaning of the Lorentz transformation equations after they have seen a spacetime diagram that shows the axes for two different reference frames, and after they have understood the crucial differences between coordinate measurements and the invariant spacetime interval.

The other topic that needs to be explored is the concept of a four-vector. This concept not only makes the relationship between energy, momentum, and mass much easier to understand, but it provides an essential foundation for any future application of Lagrangian methods to special relativity, general relativity, or electricity and magnetism. This course is not where we should introduce index notation and the Einstein summation convention, but most students at this level understand column vectors and matrix multiplication, and we can go a long way with these tools and explore the most crucial characteristics of four-vectors (such as their transformation properties, the invariance of a four-vector's magnitude, the invariance of the dot product of four-vectors, and the frame-independence of four-vector equations).

This part of the course also should link the classical principle of least action with the principle that a straight world line is the world line of longest proper time between two given events (the latter is easily proved using an elementary argument[12,13] and should be a part of the development of the concept of proper time). The action S for a relativistic free particle for a given world line can be written as

$$S = -mc^2 \int d\tau, \tag{1}$$

where c is the speed of light, and the integral yields the total proper time measured along the path. The minus sign ensures that the action is a minimum for whatever path has maximal proper time, and the factor mc^2 gives the action the appropriate units and the correct linear dependence on the particle's mass.

We can write Eq. (1) in the form of a coordinate-time integration over a Lagrangian as follows:

$$S = -mc^2 \int \frac{d\tau}{dt}\, dt = -mc^2 \int \sqrt{1 - \frac{v^2}{c^2}}\, dt, \tag{2a}$$

which implies that

$$L(v_x, v_y, v_z) = -mc^2 \sqrt{1 - \frac{v^2}{c^2}}. \tag{2b}$$

A simple application of the Euler–Lagrange equations and some basic calculus establishes that the particle's velocity components must be constant. We see, therefore, that we can develop a relativistic principle of least-action for a free particle and obtain the constant-velocity result that we know must be true from other arguments. This result supports the idea (used in the quantum section) that the world line of least action for even a relativistic free particle is indeed a straight world line, and Eq. (2b) is an essential first step in developing an electromagnetic Lagrangian.

Such a discussion would imply a relativity section that is three to four weeks long, which is more time than is usually spent on the topic. In what follows, however, I will show that this discussion would open up significant opportunities for subsequent courses. Because applications of relativity are increasingly important in modern technology, a solid understanding of relativity is more important to physicists and engineers now than it was even two decades ago.[14]

III. THE INTERMEDIATE MECHANICS COURSE

The next course a typical physics major might encounter would be one in intermediate classical mechanics, which typically discusses subjects such as orbital motion, damped and driven harmonic oscillators, rotation of rigid bodies, and perhaps even some chaos and non-linear dynamics. Texts for this course commonly include a discussion of the principle of least action and Lagrangian methods.[15] If these ideas are thoroughly discussed in the introductory course, then some time would become available in this course. My recommendation is that at least some of this extra time be spent exploring the application of Lagrangian methods to continuous media. This application is important because the same methods apply to fields, so this discussion of continuous media would provide essential background for any subsequent application of Lagrangian methods to the electromagnetic field. Reference 16 presents a very nice discussion of continuous media.

IV. QUANTUM MECHANICS

Most undergraduate major programs include a quantum mechanics course in the junior or senior year. I will assume that students in this course are familiar with partial derivatives, complex numbers, looking up integrals, and Taylor-series expansions.

A crucial first step in this course would be to firmly and formally connect the explore all paths model presented in the sophomore course with the time-dependent Schrödinger equation. Once this connection has been made, the rest of the course can be taught in the standard way. In what follows, I will briefly sketch the logic of the argument: more details can be found in Ref. 17.

In the sophomore-level course, students should have discovered that for a free particle, the propagator function that species the contribution to the quantum amplitude (arrow) $\psi(x,t)$ made by arrows of the particle's wave function within a sufficiently small range Δx_i around the position x_i at an earlier time t_0 is given by

$$K(x,t,x_i,t_0)\psi(x_i,t_0)\Delta x_i = \sqrt{\frac{m}{h(t-t_0)i}}e^{iS_{\text{direct}}/\hbar}$$
$$\times \psi(x_i,t_0)\Delta x_i, \qquad (3)$$

where S_{direct} is the action measured along the straight world-line from x_i,t_0 to x, t. For a free particle moving in one dimension with a constant potential energy V, the value of S_{direct} is simply

$$S_{\text{direct}} = (T-V)\Delta t = \left[\frac{1}{2}m\left(\frac{x-x_i}{\Delta t}\right)^2 - V\right]\Delta t = \frac{mu^2}{2\Delta t} - V\Delta t, \qquad (4)$$

where $\Delta t \equiv t - t_0$ is the (coordinate) time difference between the events and $u \equiv x_i - x$. So in this case, we have

$$K(x,t,x_i,t_0) = \sqrt{\frac{m}{h(t-t_0)i}}\exp\left(\frac{imu^2}{2\hbar\Delta t}\right)\exp\left(-\frac{i\Delta t}{\hbar}V\right). \qquad (5)$$

To find the complete wave function amplitude $\psi(x,t)$, we must sum $K\psi(x_i,t_0)\Delta x_i$ over all possible initial positions x_i, as schematically shown in Fig. 1. Note that the middle factor in Eq. (5) is the only thing that varies as x_i varies, because it will cause u to vary, and this term rotates the phase angle of the resulting complex amplitude. As discussed, arrows rotated by an angle greater than π relative to the arrow for $u=0$ do not contribute significantly to the result, and we really only need to be concerned about the contributions from the initial positions x_i close enough to the final position x so that

$$\frac{mu^2}{2\hbar\Delta t} < \pi \quad \text{or} \quad u^2 < \frac{h}{m}\Delta t. \qquad (6)$$

Equation (6) proves to be the key to using the explore all paths approach to derive the Schrödinger equation. Note that if we choose the time step $\Delta t = t - t_0$ between the initial and final wave functions to be infinitesimal, then u also must be infinitesimal, which means that the positions of points along all the paths in Fig. 1 that contribute significantly will not be much different from x. Therefore, even if the particle's potential energy varies with position, its value over the range of interest for calculating $\psi(x,t)$ will be essentially equal to $V(x)$, its value at x, so Eqs. (3)–(5) apply even to the case of nonuniform $V(x)$ in the limit $\Delta t \to 0$. The sum over all x_i in this limit therefore becomes

$$\psi(x,t) = \sqrt{\frac{m}{ih\Delta t}}\int_{-\infty}^{\infty}\exp\left(\frac{imu^2}{2\hbar\Delta t}\right)$$
$$\times \exp\left(-\frac{i\Delta t}{\hbar}V(x)\right)\psi(x+u,t_0)du, \qquad (7)$$

because $x_i = x + u$ and $dx_i = du$. If we expand the exponential involving V to order Δt, $\psi(x+u,t_0)$ to order u^2, and do some integrals of the form $\int_{-\infty}^{\infty}u^n e^{-au^2}du$,[18] we find that

$$\psi(x,t) = \psi(x,t_0) + 0 - \frac{\hbar\Delta t}{2im}\frac{\partial^2\psi}{\partial x^2} - \frac{i\Delta t}{\hbar}V(x)\psi(x,t_0). \qquad (8)$$

If we subtract $\psi(x,t_0)$ from both sides, multiply through by $i\hbar/\Delta t$, and take the limit $\Delta t \to 0$, we find the time-dependent Schrödinger equation for one dimension. (It is not very difficult to generalize this derivation to three dimensions, but it does not yield any deeper understanding.)

V. ELECTRICITY AND MAGNETISM

The undergraduate curriculum also typically includes a course in electricity and magnetism offered at the sophomore, junior, or senior level. I will assume that this course is offered for juniors and/or seniors and that students have taken a modern physics course and intermediate mechanics course of the type already described.

The first task in this course would be to discuss index notation and the Einstein summation convention, the Lorentz transformation properties of scalars, vectors, and covectors, and the four-gradient. My experience is that juniors and seniors can become comfortable with this material within four to five class sessions if the material is taught carefully.[19] The relativistic Lorentz force law provides a good physical context for practicing the notation. In appropriate units,[20] this law can be written as

$$\frac{dp^\mu}{d\tau} = qF^{\mu\nu}u_\nu, \tag{9a}$$

where

$$F^{\mu\nu} = \begin{bmatrix} 0 & -E_x & -E_y & -E_z \\ E_x & 0 & -B_z & B_y \\ E_y & B_z & 0 & -B_x \\ E_z & -B_y & B_x & 0 \end{bmatrix}, \tag{9b}$$

and u^ν is the charged particle's four-velocity with components $u^t = [1 - v^2/c^2]^{-1/2} \equiv \gamma$, $u^i = \gamma v^i/c$, $p^\mu = mc\, u^\mu$ is the particle's four-momentum, q is its charge, τ is the proper time measured along its world line and I am using a metric with a timelike signature $(+ - - -)$. Equation (9) involves scalars, vectors, covectors, and tensors and yet when the sums are written out explicitly, the three spatial components reduce to the Lorentz law taught in introductory physics and the time component reduces to conservation of energy. By examining the transformation properties of all the pieces, students can demonstrate that Eq. (9) must have the same form in all reference frames. It also is a good exercise for students to show that the antisymmetric nature of $F^{\mu\nu}$ ensures that $d(p^\mu p_\mu)/d\tau = 0$, meaning that the particle's rest mass $m = p^\mu p_\mu$ is fixed.

To fully connect electricity and magnetism with the principle of least action, we also must develop the concept of the magnetic potential **A**. Textbooks at this level avoid or marginalize the magnetic potential, partly because when it is presented in the usual way, it can be a tricky and abstract concept. However, there are ways to make the magnetic potential more accessible,[21] and there are some good reasons to discuss it fully even if we ignore the principle of least action.[22]

One possible story line for introducing the four-potential is made possible by the principle of least action. The action for a non-relativistic particle moving in a static electric field is

$$S = \int (T - V)dt = \int \left(\frac{1}{2}mv^2 - q\phi \right)dt. \tag{10}$$

Our goal is to see if we can guess the appropriate relativistic action for this case. We already know how to generalize the kinetic energy part; the action for a free particle is given in Eq. (2a). Like this part, whatever we add to the action to account for the field must be a relativistic scalar. But is the electric potential ϕ a relativistic scalar or something else? By

considering the field between the plates of a parallel-plate capacitor when viewed in a frame moving parallel to the plates, it can be quickly argued that ϕ must transform like the time component of a four-vector. So in a fully relativistic expression for the action, the electromagnetic field must appear in the form of a four-vector that we will call A^μ. However, the term we add to the Lagrangian must be a relativistic scalar, so the term must be the dot product of A^μ and some other four-vector. The only available four-vector in the case of a point particle is the particle's own four-velocity u^μ. So we propose a relativistic action of the form

$$S = -\int (mc^2 + qu_\mu A^\mu)d\tau = -\int (mc^2 + qu_\mu A^\mu)\frac{d\tau}{dt}dt \tag{11a}$$

$$= \int \left(-mc^2\sqrt{1 - \frac{v^2}{c^2}} - q\phi + \frac{q}{c}\mathbf{v}\cdot\mathbf{A} \right)dt, \tag{11b}$$

where the components of **A** are the spatial components of A^μ. We can easily show that S in Eq. (11b) reduces to Eq. (10) in the non-relativistic limit (except for an extra rest energy term that does not affect the motion).

What kind of motion does this principle imply? Although we can quickly give the result in index notation, let me demonstrate the argument in a form that might be more accessible to a junior physics major. Consider the x component of the Euler–Lagrange equation. The partial derivatives of the Lagrangian in this case are

$$\frac{\partial L}{\partial x} = -q\frac{\partial\phi}{\partial x} + \frac{q}{c}\left(v^x\frac{\partial A^x}{\partial x} + v^y\frac{\partial A^y}{\partial x} + v^z\frac{\partial A^z}{\partial x} \right), \tag{12a}$$

$$\frac{\partial L}{\partial v^x} = \frac{mv^x}{\sqrt{1 - v^2/c^2}} + \frac{q}{c}A^x = p^x + \frac{q}{c}A^x, \tag{12b}$$

where p^x is the relativistic momentum. The Euler–Lagrange equations in this case therefore imply that

$$\frac{dp^x}{dt} + \frac{q}{c}\left(\frac{\partial A^x}{\partial t} + \frac{\partial A^x}{\partial x}v^x + \frac{\partial A^x}{\partial y}v^y + \frac{\partial A^x}{\partial z}v^z \right)$$

$$= -q\frac{\partial\phi}{\partial x} + \frac{q}{c}\left(v^x\frac{\partial A^x}{\partial x} + v^y\frac{\partial A^y}{\partial x} + v^z\frac{\partial A^z}{\partial x} \right), \tag{13a}$$

which implies that

$$\frac{dp^x}{dt} = -q\frac{\partial\phi}{\partial x} - \frac{q}{c}\frac{\partial A^x}{\partial t} + \frac{q}{c}v^y\left(\frac{\partial A^y}{\partial x} - \frac{\partial A^x}{\partial y} \right)$$

$$- \frac{q}{c}v^z\left(\frac{\partial A^x}{\partial z} - \frac{\partial A^z}{\partial x} \right). \tag{13b}$$

The usual definition of the electric field is the force per unit charge on a test charge at rest, so we have

$$E^x = -\frac{1}{c}\frac{\partial A^x}{\partial t} - \frac{\partial\phi}{\partial x}. \tag{14a}$$

If we identify

$$\mathbf{B} = \nabla\times\mathbf{A}, \tag{14b}$$

we can easily see that Eq. (13b) is equivalent to the x component of the Lorentz force law given by Eq. (9a). We also can see quite generally that

$$F^{\mu\nu} = \partial^\mu A^\nu - \partial^\nu A^\mu, \tag{15}$$

and that Faraday's law and div $\mathbf{B}=0$ are identities implied by Eq. (15).

Once we have gone this far, we can derive the source-dependent Maxwell equations from a plausible principle of least action.[23] Students should know from the treatment of continuous media in the intermediate mechanics course that a least-action principle for the electromagnetic field will involve integrating a Lagrangian density over all space and time. This Lagrangian density must be a relativistic scalar and must involve a term that is quadratic in the field quantities. These requirements imply that the resulting Euler–Lagrange equations will produce linear differential equations in the field, which is required for the field to obey the superposition principle. The only plausible candidates for such terms are $A_\mu A^\mu$ and $F_{\mu\nu}F^{\mu\nu}$. The first of these leads to absurd results, for example, the resulting field equations in the electrostatic case involve ϕ directly, not the derivatives of ϕ, which does not match Gauss' law. For the second case, we can argue that the sign of the integral has to be negative for the quantity to have a plausible minimum,[24] and that we must have a factor of $1/k$ (where k is Coulomb's constant) to make the units come out right. The Lagrangian density also must involve a term that is linear in the four-current $J^\mu = [\rho, \mathbf{j}/c]$, where $\mathbf{j}$ is the ordinary current density, so that the sources will appear linearly in the field equation. The only plausible term with the right units in this case is $A_\mu J^\mu$. Therefore, the least-action principle for the electromagnetic field must be something like

$$S = \int \left(-\frac{1}{k} F_{\mu\nu}F^{\mu\nu} + b A_\mu J^\mu \right) dt\, dx\, dy\, dz$$

$$= \int \left(-\frac{1}{k} [g^{\mu\alpha}g^{\nu\beta}(\partial_\mu A_\nu - \partial_\nu A_\mu)(\partial_\alpha A_\beta - \partial_\beta A_\alpha)] \right.$$

$$\left. + b A_\mu J^\mu \right) dt\, dx\, dy\, dz, \tag{16}$$

where $g^{\mu\alpha}$ is the inverse flat-space metric and b is some unitless constant that specifies the relative magnitude and sign of the two terms. The field quantities A_μ play the role of coordinates and the gradients $\partial_\mu A_\nu$ play the role of "velocities." With only a bit of work,[25] the Euler–Lagrange equations yield

$$\partial_\mu(\partial^\mu A^\nu - \partial^\nu A^\mu) = \partial_\mu F^{\mu\nu} = -\frac{bk}{4}J^\nu. \tag{17}$$

If we choose $b = -16\pi$, the time component of Eq. (17) matches Gauss' law. By writing them out, students can discover that the other components spell out the Ampere–Maxwell relation.

VI. OTHER INTERESTING APPLICATIONS OF LEAST ACTION

Because many electromagnetic circuits have direct mechanical analogs, we often can use Lagrangian methods to find equations of motion for such circuits, even for very complicated electromechanical circuits. It turns out that we can even handle realistic resistors by treating them as generalized external forces. These issues (along with many other applications of Lagrangian techniques) are beautifully discussed by Wells.[26]

Another interesting source of applications of the principle of least action to fields at a fairly advanced level is a book written some time ago by Soper.[27] This book even includes a discussion of dissipative effects that might be appropriate in an upper-level course.

Once students are used to the principle of least action, other variational calculations become conceptually simpler. Several years ago, Van Baak discussed a variational technique that enables one to solve complicated steady-state circuits without invoking Kirchoff's loop rule.[28] Because applying the loop rule requires careful attention to signs, it is a common source of student errors. Van Baak's approach avoids this problem.

Finally, I point out that if students have studied special relativity in some depth and have seen index notation and know about four-vectors, covectors, and tensors, they have a background that provides a great springboard for studying general relativity. The geodesic equations of motion can be treated as a least-action principle. One can even use a Lagrangian to find equations of motion for the gravitational field,[29] a method widely used by researchers in the field (particularly those doing numerical simulations).

VII. CONCLUSIONS

My goal has been to reflect on what kinds of changes to the upper-level curriculum might help students take full advantage of an introductory-level exploration of the principle of least action. I have only provided a broad sketch; there is much work to be done before these suggestions can become anything approaching a practical curriculum. The proposed changes would in some cases mean shifting priorities to allow sufficient time for the development of some of the techniques, and I have no doubt that some of the changes would present problems that would have to be worked out.

However, the proposed changes could create a very exciting upper-level curriculum that could more clearly display the deep underlying connections between mechanics, relativity, electrodynamics, and quantum mechanics. These changes would give us a thoroughly 21st-century physics curriculum that teaches viewpoints and techniques currently used by researchers. The principle of least action is among the most beautiful and powerful physical principles ever envisioned. With some vision and effort, the least action principle could become a greater part of the common background of physics undergraduates.

ACKNOWLEDGMENTS

I would like to thank E. F. Taylor, the editors of this special issue, and the reviewers for making valuable suggestions about how to improve this article.

[a] Electronic mail: tmoore@pomona.edu

[1] Herbert Goldstein, Charles P. Poole, Jr., and John L. Safko, *Classical Mechanics* (Addison–Wesley, San Francisco, 2002), 3rd ed., Vol. 1, Chap. 2, pp. 34ff.

[2] Richard P. Feynman, Robert B. Leighton, and Matthew Sands, *The Feynman Lectures on Physics* (Addison–Wesley, Reading, MA, 1964), Vol. 2, Chap. 19, pp. 19–1ff.

[3] Edwin F. Taylor, "A call to action," Am. J. Phys. **71**, 423–425 (2003).

[4] Jozef Hanc, Slavomir Tuleja, and Martina Hancova, "Simple derivation of Newtonian mechanics from the principle of least action," Am. J. Phys. **71**, 386–391 (2003).

[5] Jozef Hanc, Edwin F. Taylor, and Slavomir Tuleja, "Deriving Lagrange's equations using elementary calculus," Am. J. Phys. (submitted). See ⟨www.eftaylor.com/leastaction.html⟩.

[6] Edwin F. Taylor and Jozef Hanc, "From conservation of energy to the principle of least action: A story line," Am. J. Phys. (submitted). See ⟨www.eftaylor.com/leastaction.html⟩.

[7] Jozef Hanc, Slavomir Tuleja, and Martina Hancova, "Symmetries and conservation laws: Consequences of Noether's theorem," Am. J. Phys. (submitted). See ⟨www.eftaylor.com/leastaction.html⟩.

[8] Jozef Hanc, "The original Euler's calculus-of-variations method: Key to Lagrangian mechanics for beginners," Am. J. Phys. (submitted). See ⟨www.eftaylor.com/leastaction.html⟩.

[9] Edwin F. Taylor, Stamatis Vokos, John M. O'Meara, and Nora S. Thornber, "Teaching Feynman's sum-over-paths quantum theory," Comput. Phys. **12**, 190–199 (1998). Current versions of the draft teaching materials and computer programs discussed in this article are available online at ⟨www.eftaylor.com/download.html#quantum⟩.

[10] Richard S. Feynman, *QED: The Strange Theory of Light and Matter* (Princeton U.P., Princeton, 1985).

[11] Such modern physics courses often include a discussion of the historical development of quantum mechanics that would be less relevant to this approach. Cutting much of this material will help make some room.

[12] Edwin F. Taylor and John Archibald Wheeler, *Spacetime Physics* (Freeman, New York, 1992), 2nd ed., p. 149ff.

[13] Thomas A. Moore, *A Traveler's Guide to Spacetime* (McGraw–Hill, New York, 1995), pp. 86–87. The same argument also appears on pp. 83–84 of Moore's introductory textbook, *Six Ideas That Shaped Physics, Unit R: The Laws of Physics are Frame-Independent* (McGraw–Hill, New York, 2003), 2nd ed.

[14] The relativity of simultaneity has become a very practical engineering problem for the designers of the global positioning system. Students can see the delay imposed by light travel time when satellite communications are used on television. Experimental general relativity has mushroomed in recent years, and gravitational waves will likely be discovered in the coming decade. Moreover, aspects of relativistic cosmology previously considered esoteric are likely to have a large impact on physics in the next couple of decades.

[15] Examples include Jerry B. Marion and Stephen T. Thornton, *Classical Dynamics of Particles and Systems* (Saunders, Fort Worth, 1995), 4th ed.; Ralph Baierlein, *Newtonian Dynamics* (McGraw–Hill, New York, 1983); and Grant R. Fowles, *Analytical Mechanics* (Saunders, Philadelphia, 1986), 4th ed.

[16] Herbert Goldstein, Charles P. Poole, Jr., and John L. Safko, *Classical Mechanics* (Addison–Wesley, San Francisco, 2002), 3rd ed. Secs. 13.1 and 13.2 (up to the middle of p. 563) are at a level suitable for sophomores or juniors. One would probably not need to derive the Euler–Lagrange equations the way that they do, but rather state the equations (appealing to analogy) and show that they work for a simple case (as the authors do at the top of p. 563).

[17] Ramamurti Shankar, *Principles of Quantum Mechanics* (Plenum, New York, 1980), Sec. 8.5, pp. 240–241.

[18] The results for these definite integrals given in standard integral tables assume (usually implicitly) that a is real. However, the same results apply even if a is complex, as long as the real part of $a > 0$. See, for example, Milton Abramowitz and Irene A. Stegun, *Handbook of Mathematical Functions* (Dover, New York, 1964), p. 302, where no such assumption is made. I am not sure that it is necessary to have students worry about this issue unless they ask.

[19] I regularly teach a junior-level course in general relativity where students are required to master this material. I have found that there are some tricks for teaching index notation at this level that are beyond the scope of this article to discuss in detail, but it helps greatly if students are explicitly taught to recognize the difference between free and summed indices, and if they write out expanded versions of the equations when necessary. Students also should be required to calculate the time derivative of a product involving an implied sum and do other exercises where the correct answer depends on correctly recognizing the implied sums. J. B. Hartle's *Gravity* (Addison–Wesley, San Francisco, 2003) is better than most general relativity books in teaching the notation (and in presenting the entire subject of relativity to undergraduates).

[20] We can conveniently combine the advantages of Gaussian and SI units by defining $\mathbf{B} \equiv c\mathbf{B}_{conv}$, where $\mathbf{B}_{conv}$ is the conventional magnetic field measured in teslas. The redefined $\mathbf{B}$ has units of N/C, just like the electric field (with 300 MN/C corresponding to 1 T.) All electromagnetic equations then take the same mathematical form as they would in Gaussian units, except that factors of 4π become $4\pi k$, where k is the Coulomb constant. However, the units for all quantities other than the magnetic field are in SI. This system makes the symmetries between the electric and magnetic fields apparent (and the equations much more beautiful) without having to deal with Gaussian units. This unit system also has the advantage of making it easy to show the connections between electromagnetic field theory and gravitational field theory (where the gravitational constant G is not typically suppressed as is the corresponding Coulomb constant k in Gaussian units).

[21] For example, $\mathbf{A}$ can be given a more physical meaning than often is supposed. In a static situation where $\phi = 0$ and a particle moves perpendicular to $\mathbf{A}$, the Euler–Lagrange equations implied by Eq. (11) imply that the quantity $\mathbf{p} + (q/c)\mathbf{A}$ is constant in time. Just as the scalar potential ϕ at a point in space near a static charge distribution is the total work per unit charge that one would have to do on a charged test particle to move it from infinity to that point, the quantity $\mathbf{A}/c$ at a point in space near a static (and neutral) current distribution is the total momentum per unit charge that one would have to supply to a charged test particle to keep it moving from infinity to that position along a path that is always perpendicular to $\mathbf{A}$. Therefore, if ϕ represents potential energy per unit charge, $\mathbf{A}$ represents "potential momentum" per unit charge.

[22] For example, the Aharonov–Bohm effect suggests that the magnetic potential is more fundamental than $\mathbf{E}$ and $\mathbf{B}$, and is certainly more directly connected to quantum mechanics. See J. J. Sakurai, *Modern Quantum Mechanics*, edited by San Fu Tuan (Addison–Wesley, Redwood City, CA, 1985), pp. 136–139, or John S. Townsend, *A Modern Approach to Quantum Mechanics* (McGraw–Hill, New York, 1992), pp. 399–404 for good discussions of this effect. The four-potential also provides significant advantages for calculating electromagnetic fields: indeed, R. L. Coren of Drexel University once told me that computer programs used by electrical engineers almost always calculate the scalar and magnetic potentials instead of calculating $\mathbf{E}$ and $\mathbf{B}$ directly.

[23] The general argument for the least-action derivation of the field equations comes from L. D. Landau and E. M. Lifschitz, *The Classical Theory of Fields* (Pergamon, Oxford, 1975), 4th ed., pp. 67–74, and from John David Jackson, *Classical Electrodynamics* (Wiley, New York, 1999), 3rd ed., Sec. 12.7.

[24] Reference 23, Landau and Lifschitz, p. 68.

[25] With students who are still becoming familiar with the index notation, the easiest way to have them work out the implications of the electromagnetic Lagrangian is for them to write out the implied sums in the two terms (because the metric is diagonal, there are not that many terms to write) and then calculate the Euler–Lagrange equation for a specific field coordinate (say A^x) to see how the calculation goes.

[26] Dare A. Wells, *Shaum's Outline of Theory and Problems of Lagrangian Dynamics* (McGraw–Hill, New York, 1967). The section on electrical and electromechanical systems is Chap. 15.

[27] Davison E. Soper, *Classical Field Theory* (Wiley, New York, 1976).

[28] D. A. Van Baak, "Variational alternatives to Kirchoff's loop theorem," Am. J. Phys. **67**, 36–44 (1999).

[29] Charles W. Misner, Kip S. Thorne, and John Archibald Wheeler, *Gravitation* (Freeman, San Francisco, 1973), Chap. 21.