Action: Forcing Energy to Predict Motion

Dwight E. Neuenschwander, Southern Nazarene University, Bethany, OK
Edwin F. Taylor, Massachusetts Institute of Technology, Cambridge, MA
Slavomir Tuleja, Gymnazium arm. gen. L. Svobodu, Humenne, Slovakia

In this paper we use scalar energy, rather than vector force and momentum, to predict how a particle will move. The result is a quantity called *action*. Action and its relatives undergird Newton's laws and transcend them, also predicting motion in the quantum world and in the curved spacetime of general relativity. An example exhibits action in action.

Most introductory physics courses begin with mechanics, as physics itself did historically. We recall everyday experiences with toy wagons, balls, and automobiles and refine descriptions of their motion using vectors: force, momentum, and acceleration. Newton tells us that $\mathbf{F} = d\mathbf{p}/dt$. Beyond equations, we learn to represent motion graphically with a *worldline,* a plot of displacement versus time, which provides a complete description of the path a particle takes through space and time.

Energy, which is mathematically simpler than force because it is *not* a vector, wafts in as a breath of fresh air with forms kinetic K and potential U. But potential energy leads back once again to a vector formulation $\mathbf{F} = -\,\text{grad}\ U$, telling us that the sharper the incline on which you stand the more difficult it is to resist rolling downhill in the steepest direction.

In spite of its awesome power, conservation of energy cannot predict the motion of even a single particle. Why not? Surprise! Because energy is a scalar without direction while displacement of a particle is a vector. Knowing a particle's kinetic energy, we know its speed, and thus the *distance* ds it will move during the next clock tick dt, but not the *direction* of that motion, especially in two and three dimensions (Fig. 1).

In this paper we force energy to predict how a particle will move. The result is a quantity called *action,* the invention of a string of geniuses that lived after Newton. To start toward action, think of the simplest possible motion, that of a free particle—a particle subject to no forces. Newton tells us that with respect to an inertial reference frame, a free particle moves in a straight line at constant speed. So choose our space dimension x to lie along the direction of motion of this free particle and plot its worldline (Fig. 2). (Note that the axes are t and x in Fig. 2, not x and y as in Fig. 1.) Constant speed means constant slope of the worldline in the spacetime diagram; that is, a free particle follows a straight worldline. That is what Newton tells us.

Predicting Motion Using Kinetic Energy

Now we go over Newton's head and appeal directly to Nature herself, respectfully requesting that she justify the straight worldline of a free particle in terms of our sweet scalar energy. Can she give us an *energy reason* why the particle follows the straight worldline direct from P to Q in Fig. 2? For example, why doesn't the particle take the alternative worldline PRQ composed of segments A and B? To take this alternative path, the particle would move with higher kinetic energy along the first segment A between P and R, arriving at the final x-value in half the time. Then the particle would relax at rest—with zero kinetic energy—letting time carry it along the horizontal segment B from R to Q. Tell us, oh Nature, is there an *energy* reason why this alternative worldline PRQ is not acceptable?

DOI: 10.1119/1.2173320

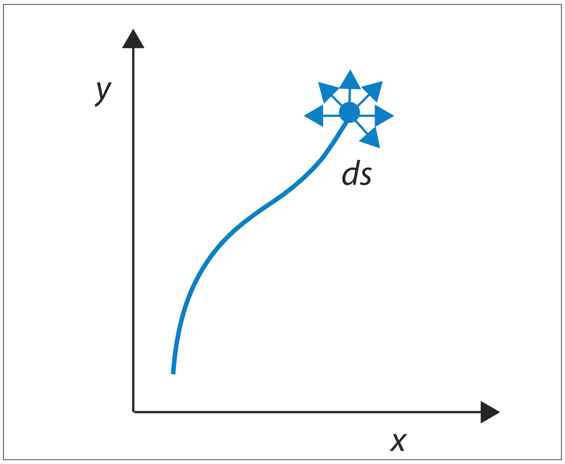

Fig. 1. Path of a particle in two dimensions *x, y*. From the value of the potential energy *U* at a given location, the conservation of total energy *E* tells us the value of kinetic energy *K* at that location. From the scalar kinetic energy we find the scalar *distance ds* of the next incremental step along the path during the next time step *dt*. But this procedure does not tell us the *direction* of that next step. The result is a pincushion of possible next steps as shown above.

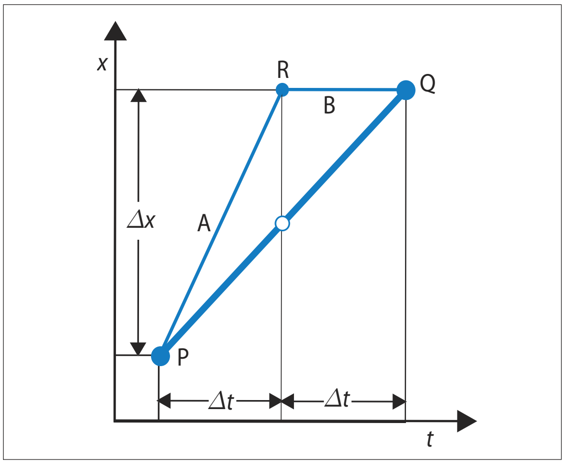

Fig. 2. Alternative worldlines between events P and Q. Can we force energy to tell us why the particle follows the straight worldline from P to Q rather than the alternative worldline along segments A and B from P to R to Q? (For simplicity we choose R at the midpoint in time between P and Q; any other fixed fraction of the time would yield the same results as the analysis in the text.)

Let's guess: Nature wants the particle to have the lowest possible value of kinetic energy needed to move from P to Q. If the straight worldline satisfies Nature's desire, then the value of the kinetic energy along this path PQ should be less than the kinetic energy along the alternative path PRQ. Check this out by calculating the kinetic energy along the different worldlines using the symbols in Fig. 2. Along the direct worldline the kinetic energy is

$$K_{\text{direct}} = \frac{1}{2}m\left(\frac{\Delta x}{2\Delta t}\right)^2 = \frac{1}{4}\left[\frac{1}{2}m\left(\frac{\Delta x}{\Delta t}\right)^2\right]. \qquad (1)$$

Kinetic energy along each of the two alternative segments A and B is

$$K_{\text{A}} = \frac{1}{2}m\left(\frac{\Delta x}{\Delta t}\right)^2 \quad \text{and} \quad K_{\text{B}} = 0. \qquad (2)$$

We guessed that Nature wants "the lowest possible value of particle kinetic energy from P to Q." But what do we (and Nature) *mean* by "lowest possible value"? For the direct worldline there is only one value of kinetic energy. Along the alternative worldline PRQ, however, there are two different values of kinetic energy given in Eq. (2). How do we combine these two values to give "*the* value of kinetic energy along AB"? Again we have to guess: We predict that Nature wants the smallest possible *average* value of the kinetic energy between P and Q, and she chooses an average over time. Call this average value $\langle K\rangle_{\text{direct}}$ for the direct line PQ and simply $\langle K\rangle$ for the indirect line PRQ. Along the direct path the kinetic energy is constant, so that value is its average:

$$\langle K\rangle_{\text{direct}} = K_{\text{direct}} = \frac{1}{4}\left[\frac{1}{2}m\left(\frac{\Delta x}{\Delta t}\right)^2\right]. \qquad (3)$$

Here's how we compute the average, $\langle K\rangle$, when the worldline is broken into segments A and B:

$$\langle K\rangle = \begin{pmatrix}\text{fraction of time}\\ \text{spent on segment A}\end{pmatrix} \times K_A + \begin{pmatrix}\text{fraction of time}\\ \text{spent on segment B}\end{pmatrix} \times K_B\,. \qquad (4)$$

For segments A and B in Fig. 2, each of these fractions is one-half. Therefore, from Eq. (2):

$$\langle K\rangle = \frac{1}{2}\times\frac{1}{2}m\left(\frac{\Delta x}{\Delta t}\right)^2 + \frac{1}{2}\times 0 = \frac{1}{2}\left[\frac{1}{2}m\left(\frac{\Delta x}{\Delta t}\right)^2\right]. \qquad (5)$$

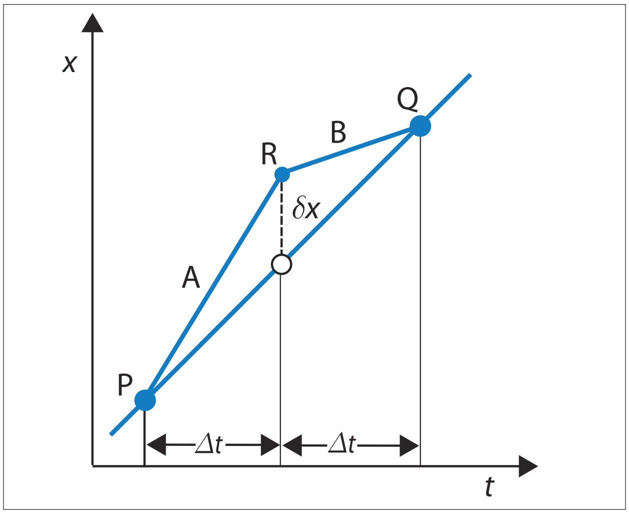

Fig. 3. Small portion of a longer worldline with two alternative nearby worldlines connecting P and Q, the whole thing enormously magnified.

Comparing the right-hand sides of Eqs. (5) and (3), we see that the value of the average kinetic energy along the alternative worldline PRQ is twice that along the direct worldline PQ. This result conforms to our guess: Nature prefers the direct worldline because the time-average kinetic energy is less. This is a start toward predicting motion with energy alone, but we need to check it more carefully.

Let's try to squeeze Nature's use of kinetic energy to predict motion of a free particle by picking an alternative worldline PRQ that is close to the direct one, as shown in Fig. 3. To be specific, we examine a small portion PQ of a longer worldline, a portion divided into two equal times Δt, and call δx the small displacement of the center point that changes direct worldline PQ into alternative worldline PRQ. How do we prove that average kinetic energy is a minimum along the straight worldline? *Minimum* implies zero change in $\langle K \rangle$ for small values of δx. We will calculate $\delta\langle K \rangle$ along PRQ and find the consequences if this change is zero for small values of δx.

As the center point R in Fig. 3 is moved up by δx, the slope of the first segment of the worldline increases, while the slope of the second segment decreases. The slope of the worldline is the velocity v; change of slope means change of velocity.

$$\delta v_A = \frac{\delta x}{\Delta t} \quad \text{and} \quad \delta v_B = -\frac{\delta x}{\Delta t}. \tag{6}$$

The average kinetic energy $\langle K \rangle$ along segments A and B is (remember each value is multiplied by the fraction of time it spends at that value, namely one-half)

$$\langle K \rangle = \frac{1}{2} K_A + \frac{1}{2} K_B. \tag{7}$$

Kinetic energy is $K = (1/2)mv^2$; the *change* in kinetic energy is approximated by its differential:

$$\delta K \approx mv\delta v = p\delta v. \tag{8}$$

Apply Eq. (8) to segments A and B, noting that in Eqs. (6) the signs are opposite (because raising the midpoint in Fig. 3 tilts segment B *down* while it tilts segment A *up*):

$$\begin{aligned} \delta K_A &\approx p_A \delta v_A = p_A \frac{\delta x}{\Delta t} \quad \text{and} \\ \delta K_B &\approx p_B \delta v_B = -p_B \frac{\delta x}{\Delta t}. \end{aligned} \tag{9}$$

The *change* in the average kinetic energy is

$$\begin{aligned} \delta\langle K \rangle &= \frac{1}{2}\delta K_A + \frac{1}{2}\delta K_B = -\frac{1}{2}(p_B - p_A)\frac{\delta x}{\Delta t} \\ &= -\frac{1}{2}\frac{\Delta p}{\Delta t}\delta x \approx -\frac{1}{2}\frac{dp}{dt}\delta x. \end{aligned} \tag{10}$$

Is $\langle K \rangle$ a minimum for the straight worldline? When a minimum occurs, a small deviation δx must not change the value of $\langle K \rangle$, thus making $\delta\langle K \rangle = 0$. Equation (10) tells us that for small δx we have $\delta\langle K \rangle = 0$ provided $dp/dt = 0$. No change in momentum between segments A and B means no change in velocity, which means no change in slope of the worldline. The worldline is straight, which we derived by asking that the average kinetic energy of the particle be a minimum.

Victory! Nature agrees with our guess that minimizing the average kinetic energy of a free particle leads to a straight worldline. This is a big step away from Newton's vector equation and toward using energy to predict motion. The next step is to add potential energy to the minimization process.

Predicting Motion Using Kinetic and Potential Energy

A particle with potential energy $U(x)$ accelerates under force $F = -dU/dx$, resulting in a curved worldline, approximated by a change of direction at the joints of our segmented worldline. Assume that segments A and B in Fig. 4 are very short, so the potential energy U varies a small amount along each segment. Then we approximate the average potential energy along any segment by its value in the middle of that segment. Let δU be the change in the potential energy at the center point R in Fig. 4 when it moves up by δx on the spacetime diagram. Then Fig. 4 shows us that the midpoint of each segment moves up half as much. For short segments, we assume that half the displacement means half the change in potential energy:

$$\delta U_A = \frac{\delta U}{2} \quad \text{and} \quad \delta U_B = \frac{\delta U}{2}. \tag{11}$$

The *change* in average potential energy over the two segments is

$$\delta\langle U\rangle = \frac{1}{2}\delta U_A + \frac{1}{2}\delta U_B = \frac{\delta U}{2} \approx \frac{1}{2}\frac{dU}{dx}\delta x. \tag{12}$$

When potential energy is combined with kinetic energy, what is Nature's preference about their average values? The simplest idea is that Nature wants the average of the *total* energy $E = K + U$ to be a minimum along every segment of the worldline. Again, when it is a minimum this average cannot change for small values of δx. So ask what is required for $\delta\langle E\rangle$ to be zero for small δx. From Eqs. (10) and (12):

$$\begin{aligned}\delta\langle E\rangle &= \delta\langle K+U\rangle = \delta\langle K\rangle + \delta\langle U\rangle \\ &= \frac{1}{2}\left(-\frac{dp}{dt} + \frac{dU}{dx}\right)\delta x = 0. \qquad \text{UNPHYSICAL}\end{aligned} \tag{13}$$

Equation (13) tells us that when $\delta\langle E\rangle$ is zero for small δx, the quantity in the parenthesis equals zero, or

$$\frac{dp}{dt} = \frac{dU}{dx}. \qquad \text{UNPHYSICAL} \tag{14}$$

This is WRONG. Equation (14) says that if the potential energy U increases with x, the momentum will increase as the particle rises. Objects would race uphill faster and faster; both kinetic and potential energy would increase, violating conservation of energy. Climbing Mount Everest would be a pushover (actually a push-up); getting *down* again would be the problem. Indeed, everything on Earth would fly into space!

How can we modify Eq. (13) to avoid the flying-into-space catastrophe? The next-simplest assumption is that Nature wants to minimize the *difference* between kinetic and potential energy. Try that:

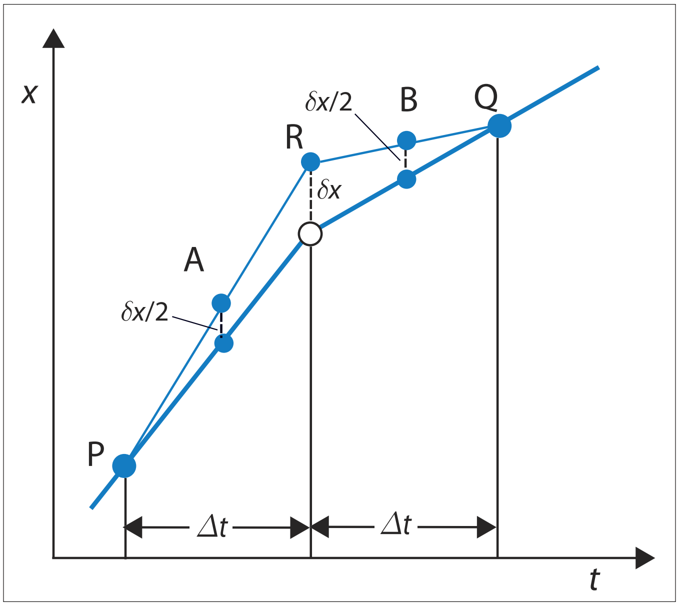

Fig. 4. For a particle moving in a potential, its worldline may be bent (heavy line, using straight segments to approximate the worldline). Two alternative worldlines connect P and Q. Approximate the change in potential along each segment as the change in the potential in the middle of the segment. Each midpoint of segments A and B moves up half as much as the displacement of the connecting point R.

$$\delta\langle K-U\rangle = \delta\langle K\rangle - \delta\langle U\rangle = \frac{1}{2}\left(-\frac{dp}{dt} - \frac{dU}{dx}\right)\delta x = 0. \tag{15}$$

For small values of δx, this equation is satisfied provided

$$\frac{dp}{dt} = -\frac{dU}{dx} = F. \tag{16}$$

This is Newton's second law of motion. When we satisfy Nature's demand that the average $\langle K - U\rangle$ be a minimum along a short segment of worldline, Newton's second law of motion automatically appears. And look where this victory leads! Equation (15) applies to *every* short segment of a worldline. Therefore it applies to the *entire* worldline; so this must be *the* worldline followed by the particle.

We have found that minimizing the time-average of a particle's kinetic minus potential energy leads to Newton's laws of motion. Next we show how to find the entire worldline of a particle by minimizing $\langle K - U\rangle$.

Predicting Motion Using Action

Minimizing $\langle K - U \rangle$ on every small portion of a worldline gives us one way to construct the actual worldline. Start with a trial worldline, our best guess of the path the particle will take through spacetime between some fixed initial event P at the beginning and a fixed final event Q at the end occurring a (fixed) total travel time T later. We divide that trial worldline into tiny segments, each spanning an incremental time dt. The fraction of time the particle spends on each segment is dt/T. Now we temporarily freeze all intermediate events except one, and move that event up and down in the spacetime diagram (as in Figs. 3 and 4) to minimize $\langle K - U \rangle$ along the adjacent pair of segments, then do this in turn for every intermediate event along the worldline. For incremental dt the resulting summation becomes an integral:

$$\langle K - U \rangle = \int_P^Q (K - U)\left(\frac{dt}{T}\right). \tag{17}$$

Total travel time T is fixed by the events P and Q that anchor the two ends of the worldline. This leads to what is called the *action*, given the symbol S:

$$\text{ACTION} \equiv S \equiv \langle K - U \rangle T = \int_P^Q (K - U)dt. \tag{18}$$

Moving one intermediate event at a time while freezing the others typically does not allow each event to move far enough to reach that event's final location on the worldline with minimum action along every segment. We usually need repeated sweeps through all intermediate events for every one of them to reach a position in which action is a minimum along the pair of adjacent segments, verifying that the worldline is a correct one between P and Q. When we reach this condition, further small adjustment of any intermediate event does not change the total action:

$$\delta S = 0. \tag{19}$$

The technical term for the condition $\delta S = 0$ is that the action is *stationary*. *Stationary* means that if you vary the correct worldline just a little, the change in action is negligibly small. Our result is called the *principle of stationary action*:

When the action is a minimum along every segment of a worldline between fixed end-events (resulting in stationary action S), that worldline is one that a particle will follow between those events.

Nature has granted our request; we have gone over Newton's head to predict motion using scalar energy. Finding stationary action is often simpler than applying Newton's mechanics with its vector forces and momenta because the principle of stationary action involves simple addition and subtraction of scalar energies. An added benefit of scalar addition of energy is that Eqs. (18) and (19) work just as well for three-dimensional motion of a particle as they do for one-dimensional motion. Indeed, the scalar nature of action allows us to generalize these equations to predict motion of more than one particle, particles connected with springs and rods, rigid bodies suspended from pivots, and a wealth of simple and complicated systems.

The principle of stationary action is especially powerful when you want to control in advance the initial and final conditions in space and time. For example, firing a space probe from Earth orbit to the Moon requires that the probe arrive at the location of the Moon *when the Moon is there*; we specify initial and final events in advance. The principle of stationary action is perfect for planning this Moon shot.

The principle of stationary action has some drawbacks. Often we do not know ahead of time where a particle is going; rather we want to use the laws of motion to *find out* where it is going. For example, will an incoming asteroid miss the Earth? Moreover, action does not easily predict motion when friction is present. In these cases, Newton can often help us more than action.

For technical reasons the specific process of finding a worldline by multiple sweeps through intermediate events does not always work; *sometimes* it does *not* move the initial guessed trial worldline toward the correct worldline. But that is another story for another day; we can always find *some* calculating method that finds a worldline with minimum action along every small portion, which therefore satisfies the principle of stationary action. An example is shown in Box 1. All successful methods have this in common: They are repetitive and boring—a perfect task for a computer. One of us (ST) has created computer programs that allow you to use stationary action to hunt for world-

Box 1.

EXAMPLE: Using Action to Find the Worldline

See how the action principle can help construct the worldline of an apple thrown vertically upward in a uniform gravitational field. We use an approximation that can be generalized to an automatic computer program. Divide a portion of the worldline into four adjacent straight line segments, each of duration Δt, as shown in Fig. 5.

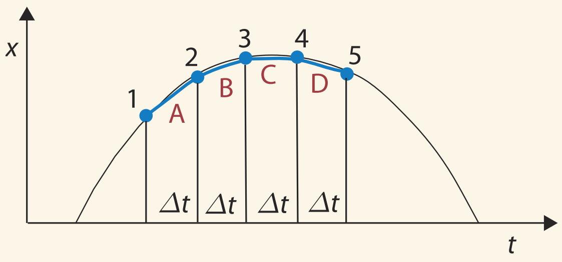

Fig. 5. Four worldline segments used in computer solution

Take the potential energy on each segment to be the value at its center. Then the straight-segment approximation yields the action along segment A:

$$S_A = \langle K - U \rangle_A \Delta t = \frac{1}{2} m \frac{(x_2 - x_1)^2}{(\Delta t)^2} \Delta t - mg \frac{x_1 + x_2}{2} \Delta t \quad ,$$

with similar expressions for segments B, C, and D. The total action along all four segments is

$$S_{\text{total}} = \frac{1}{2} m \frac{(x_2 - x_1)^2 + (x_3 - x_2)^2 + (x_4 - x_3)^2 + (x_5 - x_4)^2}{\Delta t} - mg \frac{(x_1 + x_2) + (x_2 + x_3) + (x_3 + x_4) + (x_4 + x_5)}{2} \Delta t.$$

We fix initial and final events 1 and 5 and ask: What should be the positions of events 2, 3, and 4 in order to minimize the value of S_{total} in our straight-segment approximation? Answer: The derivatives of S_{total} with respect to each of the three intermediate coordinates should vanish. Take the derivative with respect to x_2 and set the result equal to zero to obtain

$$\frac{dS_{\text{total}}}{dx_2} = 0 = m \frac{(x_2 - x_1) - (x_3 - x_2)}{\Delta t} - mg\Delta t.$$

Similar expressions result from setting equal to zero derivatives of S_{total} with respect to x_3 and x_4. The result is three equations in the three unknowns x_2, x_3, and x_4. Divide each equation through by m and multiply through by Δt to obtain

$$-x_1 + 2x_2 - x_3 = g(\Delta t)^2$$
$$-x_2 + 2x_3 - x_4 = g(\Delta t)^2$$
$$-x_3 + 2x_4 - x_5 = g(\Delta t)^2 .$$

As an example, fix initial and final positions x_1 = 2 m and x_5 = 3 m and elapsed time = 1 s, so that Δt = 0.25 s. Setting g = 9.8 m/s^2 yields three equations that anyone can solve (no need for a computer!) to give x_2 = 3.17 m, x_3 = 3.72 m, and x_4 = 3.67 m. The result fixes this portion of a worldline so that action is a minimum along each pair of adjacent segments. Instead of just four segments, we can use N segments that span and approximate the entire worldline from initial fixed event to final fixed event; the larger the value of N the better the approximation. The result is $N-1$ linear equations in $N-1$ unknowns. You can easily use the four-segment, three-equation example above to write down these $N-1$ equations directly. There are well-established computer solutions for such equations, one of them used by the interactive programs in Ref. 1. By such a method the action principle minimizes action along every section of a worldline, leading to stationary action, Eq. (19), and the construction of the entire worldline with a single procedure.

lines for various kinds of motion.[1]

The principle of stationary action is a powerful tool that helps us to get from here-now to there-then. Adding the action principle to our toolbox multiplies our understanding of Nature and our ability to influence her. But there are deeper reasons to have respect, even passion, for the action principle. Unlike force, action roots the world of Newton deep in the quantum world, as Richard Feynman showed us decades ago.[2] As particle mass increases from electron to tennis ball, Feynman's description of quantum motion goes over seamlessly into the principle of stationary action.[3] On the other end of the cosmic scale, the motion of a particle in Einstein's curved spacetime is elegantly predicted using a version of the principle of stationary action.[4] Hook onto the action principle and hitch your understanding to a star!

Acknowledgment

Suggestions by Jon Ogborn greatly simplified and augmented the argument of this paper.

References

1. Interactive plotting of worldlines using Java software is available at http://www.eftaylor.com/software/Action ClockTicks and also at http://www.eftaylor.com/ software/ActionApplets/LeastAction.html.
2. Richard P. Feynman, *QED: The Strange Theory of Light and Matter* (Princeton University Press, Princeton, NJ, 1986).
3. Jon Ogborn and Edwin F. Taylor, "Quantum physics explains Newton's laws of motion," *Phys. Educ.* **40**, 26–34 (Jan. 2005). Available at http://stacks.iop.org/0031-9120/ 40/26.
4. Edwin F. Taylor, "A call to action," *Am. J. Phys.* **71**, 423–425 (May 2003). Available to subscribers at http:// scitation.aip.org/journals/doc/AJPIAS-ft/vol_71/iss_ 5/423_1.html.

PACS codes: 01.40.gb , 01.50.ht, 01.55.+b, 02.70.-c , 45.20.dg, 45.40.Aa

Further Reading

Richard Feynman has an excellent introduction to the action principle in Chapter 19 of Volume II of *The Feynman Lectures on Physics* (Addison-Wesley, Reading, MA, 1964).

A selection of papers on the action principle is available at http://www.eftaylor/action.

Edwin F. Taylor *is co-author of textbooks on quantum mechanics, special relativity, and general relativity. In 1998 he received the Oersted Medal, the highest award of the American Association of Physics Teachers for "notable contributions to the teaching of physics." He is senior research scientist emeritus at the Massachusetts Institute of Technology.*

Massachusetts Institute of Technology, 77 Massachusetts Ave. Cambridge, MA 02139-4307; eftaylor@mit.edu

Slavomir Tuleja *is currently a high school teacher of physics and mathematics in Humenne, Slovakia. His interests are in developing interactive Java computer programs for physics education and working with talented high school students in a physics club.*

Gymnazium arm. gen L. Svobodu, Humenne 06651, Slovakia; stuleja@gmail.com

Dwight E. Neuenschwander *is professor of physics at Southern Nazarene University, editor of the Society of Physics Students publications for the American Institute of Physics, and enjoyed teaching introductory physics for 25 years.*

Dept of Physics, Southern Nazarene University, Bethany, OK 73008; DNeuensc@snu.edu