Action on Stage: Historical Introduction

Jon Ogborn, Institute of Education, University of London, UK
Jozef Hanc[1], Institute of Physics, P.J. Safarik University in Kosice, Slovakia,
Edwin F. Taylor[2], Department of Physics, Massachusetts Institute of Technology, USA

Abstract

The action principle is a powerful tool for understanding, applying, and building bridges among fields of physics, from quantum theory through relativity to current research. We dramatize those who devised the action principle and its precursors – Fermat, Huygens, Maupertuis, Euler, Hamilton, Einstein and Feynman – with the authors performing the roles of these great physicists and mathematicians. We accept no responsibility for the accuracy of the words of our characters! This is an effort to introduce fundamental physical principles, not to reconstruct the actual historical development of these principles.

Action on Stage

Animateur:

This symposium is about building bridges between things students would like to learn– relativity, quantum theory, particle physics – and things they have to learn – notably classical mechanics. We are interested in simplicity and unity in physics, as well as with exciting students about physics.

The idea that links these topics is the concept of *stationary action*.

Now we have a problem. Either you know nothing at all about the physical quantity called *action*, or you learned it in a difficult course of theoretical mechanics. Both will make you hostile to our proposals. Either, "I never studied it, so it can't be important", or "Anything I don't understand *must* be too hard for students".

So – we have to tell you about these ideas, starting from zero. How better than to ask the people who invented them to explain what they were doing? First up is the Frenchman Pierre de Fermat.

Pierre Fermat (1601-1665):

Pierre de Fermat

Although I never published a scientific paper, my reputation as one a leading mathematician came from my correspondence with other scientists and from them publishing my ideas and methods in their work. I am known primarily for my work in number theory. I also developed analytic geometry independent of Descartes and worked in many other mathematical fields – completely as an amateur.

I had a terrible fight with Descartes. He thought that light is transmitted instantaneously from point to point "like the cane of a blind man," so I had to express my optical theory in terms of "resistance" of different media through which light passes. You have no such difficulty, and Fermat's principle of least time is the oldest variational principle; one that you still use. The idea is simple: the path that light takes is just the one that takes the least time. Among all possible paths, the minimum total time picks out the unique path between fixed initial and final points. What could be easier?

Animateur:

Monsieur Fermat, there's an obvious objection to your idea. How does the light know in advance which path will be the quickest?

[1] email: <jozef.hanc@upjs.sk>
[2] email: <eftaylor@mit.edu>, website: <http://www.eftaylor.com>

Fermat:

When I was alive I could not answer your question. The objection was not overcome until long after my death, when you came to see that every point on an advancing wave acts as a source of little wavelets. Then between point source and point detector the wavelets add up with coherent phase along *the path of stationary time*. I hope you will tell us how some Dutchman figured it out.

Animateur:

We had hoped that, as a Dutchman, Christiaan Huygens could join us here in Amsterdam, but unfortunately he is away at the Royal Court in France. His big idea was that light is a wave, and that where the wave goes next can be predicted by supposing that each point on the wave front acts as a source of little wavelets. The many wavelets all superpose, adding up in constructive interference to generate the new wave front, but canceling in destructive interference everywhere else. Centuries later, Richard Feynman was to adapt the same idea to build a new formulation of quantum mechanics: the "many paths" approach.

Christiaan Huygens

Now we jump a hundred years, and our next guest is another Frenchman, Pierre Louis de Maupertuis.

Maupertuis (1698-1759):

Pierre de Maupertuis

I am Pierre-Louis Moreau de Maupertuis. My father, a wealthy pirate, gave me every advantage. I led an expedition to Lapland to measure the length of a degree along the Earth's meridian, proving that the Earth is hamburger-shaped. Its fame led to my becoming president of the Prussian Academy and a favorite in the court of Louis the fifteenth.

I conceived the principle of least action, that in all events of Nature there is a certain quantity, called *action*, which is always a minimum; that collisions of bodies or refraction of light occur in such a way that the amount of the quantity *mvs* is as little as possible. My original definition of action as the product of mass, speed and distance traveled by a moving object was later restated by my friend Leonhard Euler [see eq. (1)].

My paper was titled, "The laws of motion and rest deduced from the attributes of God" and stated: "Here then is this principle, so wise, so worthy of the Supreme Being: Whenever any change takes place in Nature, the amount of action expended in this change is always the smallest possible." I am horrified to hear that people think that action can sometimes be a saddle point; I reject this idea entirely because the perfection of God is incompatible with anything other than utter simplicity and minimum expenditure of action.

Animateur:

Ignorant people say that nothing of intellectual distinction greater than the cuckoo-clock ever came out of Switzerland. To give them the lie, we now hear from the great Swiss mathematician Leonhard Euler, who supported and developed Maupertuis' idea.

Leonhard Euler(1707-1783):

It is no boast to say that I am the most prolific mathematician of all time, producing about 900 papers and books in my lifetime. I spent my years largely in the courts of the Tzars of Russia and in the court of Frederick the Great.

Maupertuis is a great buddy of mine, but sloppy in formulating his action principle. I realized that without the law of conservation of energy the action quantity of Maupertuis loses all significance. So I

Leonard Euler

cleaned it up, formulating the principle of least action as an exact dynamical theorem and giving his action a correct mathematical form:

$$W = \int_{\substack{\text{initial}\\ \text{position}}}^{\substack{\text{final}\\ \text{position}}} mv\,\mathrm{d}s \qquad \text{(assume energy conserved)}, \qquad (1)$$

(The integral is calculated along a path of a moving particle.)

My statement “since the plan of the universe is the most perfect possible and the work of the wisest possible creator, nothing happens which has not some maximal or minimal property!” was my acknowledgement of Maupertuis as originator of the action principle.

I also developed a simple, intuitive, geometrically understandable way of finding the minimum or stationary action path (see figure 1)

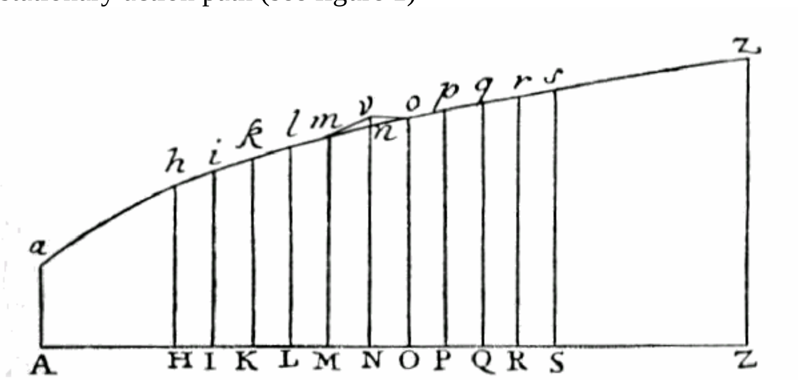


Figure 1. Euler realized that if the action integral is minimal along the entire path, it must also be minimal for every subsection of the path: triplets of nearby points on the path, e.g. *mno* in my figure. Minimal action means that any change in the path, e.g. point *n* varied slightly to point *v*, leads to zero first order change in action. If this condition is be satisfied for each triplet and we go to the limit in which lengths of segments tend to zero, we get a differential equation (the Euler-Lagrange equation), whose solution is the stationary action path.

Later, I helped the career of the young Joseph Louis Lagrange (1736-1813), who wrote to me about his elegant mathematical way to express conditions of minimum action. His ideas led me to drop my intuitive graphical approach and coin the phrase "calculus of variations". Great for mathematics and theoretical physics, but a disaster for physics education! Lagrange's abstract method has dominated your advanced mechanics classes. Too bad, because my graphical method is perfect for modern computers (fig. 2)

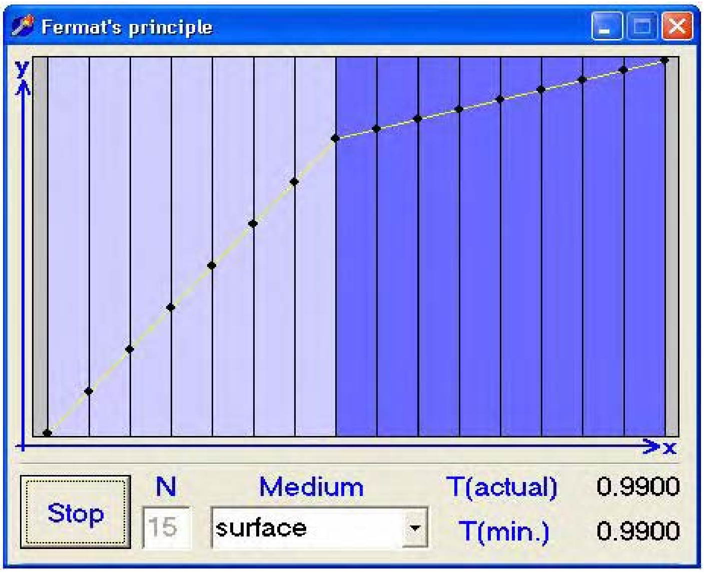


Figure 2. The universality of Euler's graphical approach is demonstrated by this computer display used in modeling Fermat's principle. Click the computer mouse to select an arbitrary moveable intermediate point on the path, then drag the point up and down, looking at the value of the total time, to find the minimal (stationary) time of that point. Then do the same for other points, cycling through them until the time for each results in the least (stationary) value of the total action or time. This *method of successive displacements* or *hunting for the least time path* is straightforward but tedious. However the task can be done quickly by computer.

Animateur:

We jump a hundred years again. Here is the Irishman William Rowan Hamilton to tell you about his new version of action, more powerful than ever.

Hamilton (1805-1865):

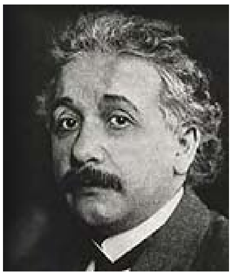
My name is William Rowan Hamilton. I wanted to develop a common mathematical language for particles and waves, so starting from Fermat's least time principle and using the Lagrangian, I found what you call Hamilton's action S:

$$S = \int_{\substack{\text{initial}\\ \text{event}}}^{\substack{\text{final}\\ \text{event}}} L\,\mathrm{d}t = \int_{\substack{\text{initial}\\ \text{event}}}^{\substack{\text{final}\\ \text{event}}} (K - U)\,\mathrm{d}t \qquad (2)$$

William R. Hamilton

Maupertuis' action, remember, determines *trajectories in space* between fixed initial and final locations and requires that energy be conserved. In contrast, my action principle determines *worldlines in spacetime* between initial and final events and is true even if the potential energy is a function of time as well as position, in which case the energy of the particle may not be a constant. I understood that action along an actual worldline is not necessarily a minimum but is always stationary compared with action along adjacent alternative worldlines between the same fixed initial and final events. Actually the word "worldline" is a stranger to me; it took Albert Einstein to make the term important. For me it was just the path between fixed places and times – between what Einstein called fixed *events*.

Einstein:



Albert Einstein

I, Albert Einstein, am Swiss by nationality – another blow to cuckoo-clock theory! It was my idea that the physical world has to be structured as space-time events. I emphasized the fact that such events are connected by worldlines in space-time. I could show that the natural, unforced, path from one event to another was that for which ageing – wristwatch time – is a maximum.

In relativity, the Hamilton action S for a free particle is just:

$$S = -mc^2 \int_{\substack{\text{initial}\\ \text{event}}}^{\substack{\text{final}\\ \text{event}}} \mathrm{d}\tau \qquad (3)$$

Minimal (or stationary) action along a real worldline makes the total proper time $\tau = \int d\tau$ maximal (or stationary). Therefore, because of the minus sign in front of the integral, the relativistic principle of least action is the same as the *principle of maximal proper time*, called by Dr. Taylor the *principle of maximal aging*.

Moreover it is not difficult to show that for small velocities Eq. (3) gives the same results as classical nonrelativistic Hamilton action.

Animateur:

You will have noticed that these contributions all came from Europe. But in the last century, American physics blossomed, and one of its finest products was Richard Feynman, who completes our story.

Richard Feynman:

Richard Feynman

When I was in high school my physics teacher Mr Bader told me that Newton's laws could be stated not only in the form $F = ma$, but also in the form: "average kinetic energy minus average potential energy is as little as possible for the path of an object going from one point to another." I next got involved with least action as a PhD student with John Archibald Wheeler. This led to the development of the "many paths" version of quantum mechanics, a third formulation mathematically equivalent to the Schrödinger and Heisenberg versions. It also helped me develop my "*Feynman diagrams*" for doing calculations in quantum electrodynamics. For this work I shared the Nobel Prize with Schwinger and Tomonaga.

So how does a mindless particle recognize the least action path or worldline? Does it smell neighboring paths to find out whether or not they have increased action? According to my formulation, yes! The electron explores all worldlines between source and detector. For each possibility there exists a little rotating stopwatch whose hand, or *arrow*, makes a total number of turns equal to Hamilton's action S divided by Planck constant h (see Fig. 3). The Lagrangian L, divided by h, is nothing other than the rate of arrow rotation.

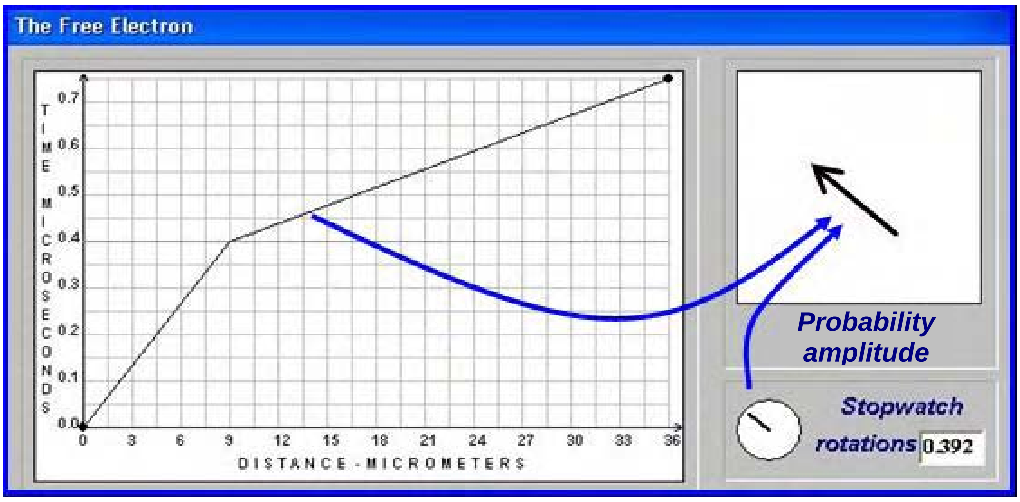


Figure 3. In the "many paths" version of quantum mechanics the electron explores all possible worldlines from initial emission event to final detection event. The figure shows a single one of these worldlines. Along this path a little stopwatch hand rotates at the rate *L/h*, leading to a contribution to the final amplitude at the detection event.

All these quantum arrows (probability amplitudes) add up constructively (*line up*) if they have similar phases. This is so for worldlines close to the stationary action path (the blue pencil of paths in Fig.4). The arrows cancel out or *curl up* for other sets of worldliness, as you can see in a piece of Dr. Hanc's program. The bigger the mass of an object the narrower is the pencil of nearby worldlines that significantly contributes to the resulting amplitude.

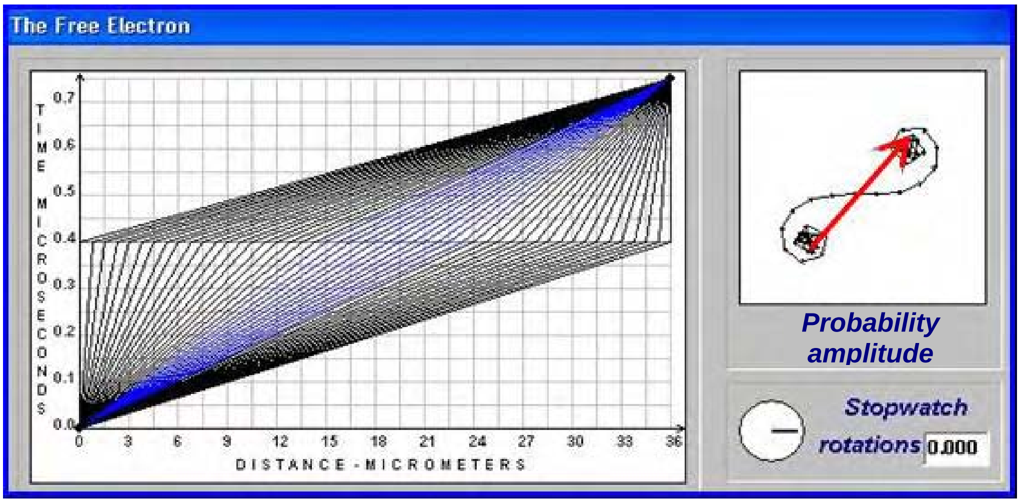


Figure 4. Some of the infinite number of possible worldlines connecting fixed initial and final events. The squared magnitude of the resultant arrow is proportional to the probability of detecting the particle at the final event. As the particle mass increases, the pencil of worldlines contributing significantly to the resultant arrow (shown in blue) becomes narrower and narrower, approaching the single path of classical mechanics.

Animateur:

Christiaan Huygens has sent me a letter claiming priority for your “many paths” idea. He claims that it is just his idea of wavelets, in modern clothing. What do you say to that?

Richard Feynman:

It’s true (I acknowledged Huygens in my PhD thesis). However the use of Planck’s constant, and the fact that the idea also works for particles like electrons or atoms, goes beyond Huygens. The idea for the number of quantum stopwatch rotations came from Paul Dirac.

Animateur:

OK, that’s it. We will be describing in the papers that follow, how the scalar quantity *action* adds to the physicist’s toolkit for analyzing and predicting motion. It looks like this:

1. Use Maupertuis action W when we fix in advance the initial and final POSITIONS, and energy is conserved.
2. Use Hamilton action S when we fix initial and final EVENTS and energy is may or may not be conserved.
3. Use Newton or Lagrange when we do NOT know where the motion is going from its initial conditions.
4. Use Newton when friction is significant, so vectors are inevitable.

We hope that we have started to break up some of your thought-glaciers about action.

References

Background papers with historical references are available at the web-site:
http://www.eftaylor.com/leastaction.html

Photos

What is the action model?
Introducing and modeling principles of least action.

Jozef Hanc[1], Institute of Physics, P. J. Safarik University in Kosice, Slovakia,

Abstract

The Action Principle predicts motion using the scalars energy and time, entirely avoiding vectors and differential equations of motion. Action is the tool of choice when we want to specify both initial and final conditions. Maupertuis-Euler action finds the trajectory when initial and final positions are prescribed in advance, but requires that energy be a constant of the motion. Hamilton action finds the worldline when initial and final events are prescribed in advance and easily describes motion when potential energy is a function of time as well as position. A simple toolkit of motion tells us when to use action, when to use Lagrange's equations, and when we must return to the vector methods of Newton. The original Euler method of handling action also provides a basis for computer modeling. Interactive software allows students to employ basic concepts of the principle of least action and increase conceptual understanding.

Least action approach in teaching

The least action principle approach, so important for modern physics, is widely considered to be a difficult topic and is usually only used in advanced mechanics textbooks and courses. Why does it seem a peculiar way to introduce and teach classical mechanics? Why is action as a physical quantity understood as being very abstract and unsuitable for introductory physics, despite the fact that it is a scalar very similar to energy – one of the central concepts of introductory courses?

The reasons are that in the majority of standard advanced texts [like Landau & Lifschitz 1976, Goldstein et al 2002, Marion & Thornton 2003, Hand & Finch 1998], — (1) the mathematics used is *the calculus of variations,* which is not part of the common mathematical toolkit acquired at introductory college level; (2) action is usually introduced extremely briefly and is only used for a quick variational derivation of Lagrange’s equations; (3) action is not immediately illustrated by examples; texts typically include no or a very few examples; (4) there is also no study of the properties of action after introducing it, since they are taught at the end of courses and texts, (5) and finally you find no computer modeling, which means that students cannot obtain direct experience and intuition.

So the important question is how to introduce least action principles? Our experience says that it is possible in the frame of introductory courses provided that we concentrate on the following crucial issues:

- *Starting with one dimensional cases and using a powerful graphical language*. More general cases only bring in more complicated mathematical expressions, but essentially nothing new in physics ideas
- *Using concrete, but easily generalizable examples.* We have chosen Newton’s falling apple, but the arguments will work for any reasonable potential energy function.
- *Building a clear connection to Newton’s laws* in terms of comparison.
- *Using* ingeniously *simple original arguments* of the greatest physicists and mathematicians of all times: (1) *Newton’s* argument from his celebrated *Mathematical Principles of Natural Philosophy* (1687) and (2) *Euler’s* argument from his pioneering work on the variational calculus *The method of finding curved lines enjoying properties of maximum or minimum* (1744). As result we will not need advanced mathematics (all arguments require only high school algebra such as expressions $(a \pm b)^2$ and basic properties of parabola). Moreover we also obtain an excellent foundation for computer modeling, which is important in getting good intuition and experience.

[1]email: <jozef.hanc@upjs.sk>

So how does classical mechanics explain the motion of a falling apple? We will show three different approaches – tools for answering this question: *Newton's laws, Hamilton's* and *Maupertuis'* principle of least action. As we will see below, together they form a simple tool-kit of mechanics in which the question being asked about any system determines directly which tool should be used to predict the motion of that system.

Motion of a falling apple from different points of views

Newton's laws of motion

We start with the well-known Newton's laws of motion, which are already taught at high schools. Firstly Newton says: "Give me the initial state of the apple, which means the initial position and velocity of the apple".

However giving the initial velocity and position means experimentally measuring two nearby positions at very close instants. In this case the initial (and indeed any) velocity is graphically nothing else than the slope of the *position vs. time graph,* or in the language of spacetime physics the slope of the apple's *worldline*.

Then Newton offers us his laws of motion and answers the questions: *What happens next with the apple? That is, what is the position of the apple at the next instant, if there is Earth's gravity or in general some force F (see fig.1)?*

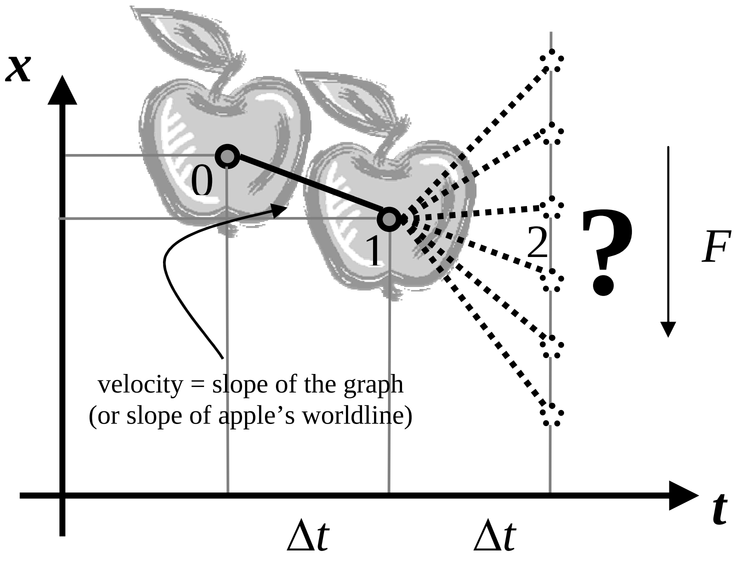


Figure 1. Newton's laws answer the question *what is the position of the apple at the next instant*, if we know the initial velocity and position or in other words, we know two very nearby positions of the apple.

If there were no acting force, then according to the first law of motion, the principle of inertia, the apple would continue in motion at the same velocity, so graphically it would follow a straight-line worldline (see fig.2). Instead Earth's gravity (or generally some force F) causes a component of motion in the direction of the applied force, as described by the second law of motion, the momentum principle $m\Delta v = F\Delta t$. Competition between these two tendencies results in the parallelogram of which the diagonal represents the worldline of the actual motion (fig.2). This process of constructing worldlines (which is conceptually the same for trajectories) is simple, repetitive and universally applicable, so it provides a first-rate foundation for computer modeling. Since today's computers are very fast there is really no need for fancy algorithms in introductory physics teaching. To get a better approximation to the actual motion we simply take smaller time steps.

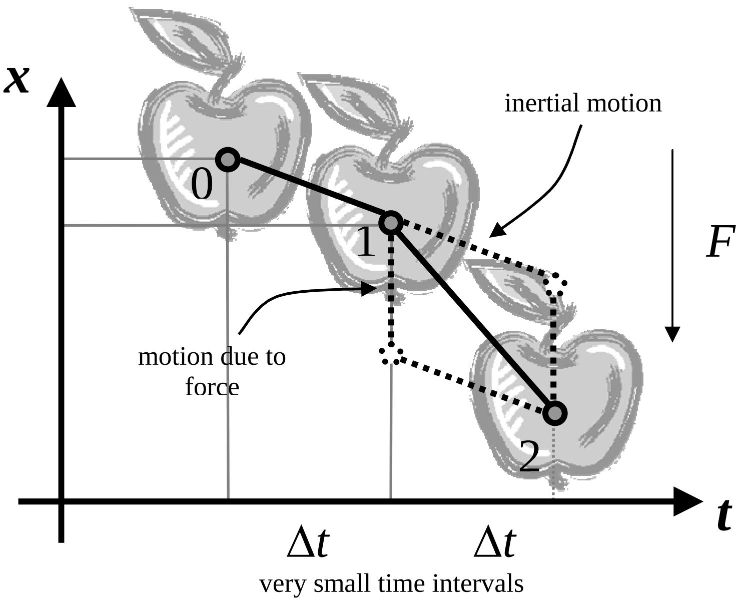


Figure 2. According to Newton's laws the motion of the apple is produced by two "effects": the apple's inertial motion at constant velocity, if no forces act upon it and the motion due to an acting force *F*. The net motion is then given by *the diagonal of the parallelogram* of the separate motions that would have occurred.

This numerical method appears in some classics physics texts notably in Chapter 9 in Vol.1 of Feynman's lectures on physics (1964). But it is also very effectively applied in the modern introductory physics curriculum, e.g. in *Modern Mechanics* of Chabay and Sherwood (2002) or in *Unit N* of Moore's Six Ideas That Shaped Physics (2003).

Hamilton's principle of least action

Now we apply a first action model to our falling apple based on energy concepts and Hamilton's least action principle. Hamilton tells us: "Give me both initial and final positions and times of the apple," called in spacetime physics *events*. If we specify the initial and final events in advance, then Hamilton's principle can successfully answer the following question: *What is the middle event for the apple? Or which worldline is followed by the falling apple between the initial and final events, if the apple has potential energy U(x) (see fig.3)?*

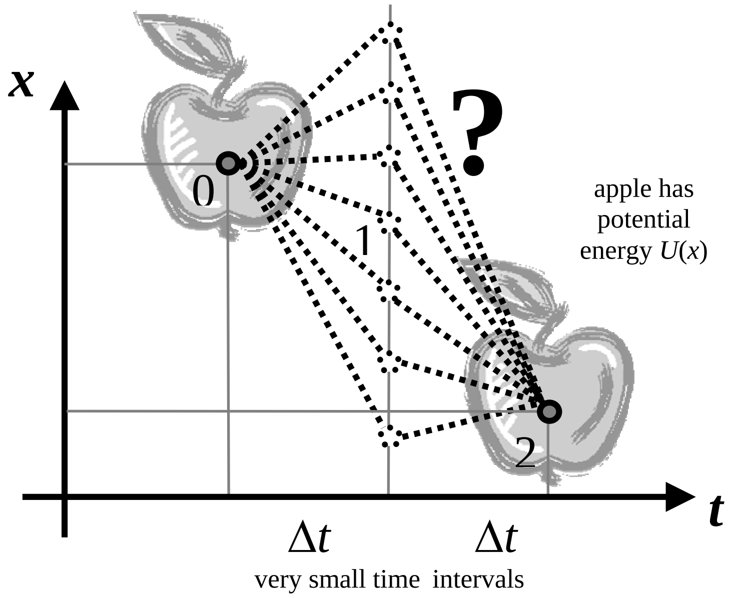


Figure 3. Hamilton's principle answers the question *what is the middle position* or more generally *the middle event* for the apple, if we know the apple's initial and final events.

Now what special property does the actual worldline obey? The principle of least action discovered by Hamilton says that the apple follows *the worldline for which the average kinetic energy minus the average potential energy is as little as possible* or put more briefly worldline has *the least action*, because Hamilton's action S is defined as

$$S \equiv \begin{pmatrix} \text{difference bewteen} \\ \text{average kinetic and potential} \\ \text{energy along the worldline} \end{pmatrix} \cdot \begin{pmatrix} \text{time duration} \\ \text{of motion} \end{pmatrix} \quad \text{or} \quad S = \left(\langle K \rangle - \langle U \rangle \right) \cdot (t_{final} - t_{initial}) \qquad (1)$$

Using integral calculus the definition (1) has the form

$$S = \int_{t_{initial}}^{t_{final}} (K - U)dt = \int_{t_{initial}}^{t_{final}} L dt\,, \qquad (2)$$

where the difference $K - U$ is called the *Lagrangian L*, the quantity that appears in Lagrange's equations of motion.

In the case of our falling apple (and also in general case for small Δt) it is easy to calculate all terms in the expression (1) for action S along any worldline 012 (fig. 3). The time duration $t_{final} - t_{initial}$ equals $2\Delta t$. The average kinetic energy $\langle K \rangle$ is given by $(K_A + K_B)/2$, the average of kinetic energies for the first and second segment of the worldline, that is, by $(1/2)(mv_A^2/2 + mv_B^2/2)$, where $v_A = (x_1 - x_0)/\Delta t$ and $v_B = (x_2 - x_1)/\Delta t$.

We now have to pay attention to the potential energy $U(x)$. The shortness of Δt allows us to approximate $U(x)$ by a linear function Cx in the region near point 1. (An additive constant is not important, because it is always zero after an appropriate choice of a reference point.) For the apple constant C is positive and equals mg. Generally we will consider it here as some positive constant. From the viewpoint of the force concept used previously in Newtonian analysis it represents a force $F = -\Delta U/\Delta x = -C\Delta x/\Delta x = -C$, a force in the downward direction. Then $\langle U \rangle$ equals $(U_A + U_B)/2 = (1/2)[C(x_0 + x_1)/2 + C(x_1 + x_2)/2]$. Since the events 0 and 2 are fixed and only position of the middle event 1 is variable, the apple's action S must be only a function of x_1, in which case it is a quadratic function.

To find a worldline with the least action therefore means that we must vary and find a position x_1 which makes the action a minimum. There are two natural ways to do this. One is the trial-and-error method, perfectly suited for a computer which can quickly calculate and compare the action (1) for millions of worldlines. The detailed description of computer modeling based on the so-called Euler variational method is described in our symposium contributions *Action on Stage* (see fig. 1, 2) and *Use, Abuse, and Unjustified Neglect of the Action principle* (see fig. 1).

The second way to find the least action worldline is the use of mathematical methods. According to Hamilton's principle the action has to become larger, if we change the position x_1 of the middle event 1 of the actual worldline by any small displacement δx. Using only high school algebra one can obtain the following expression for the corresponding change in action:

$$\delta S = S(x_1 + \delta x) - S(x_1) = (m\Delta v - F\Delta t)\delta x + 2\frac{m}{\Delta t}(\delta x)^2 \qquad (3)$$

Mathematically equation (3) represents a simple quadratic function with respect to δx whose graph is a parabola. The graphical method proves that the least-action worldline is identical with the worldline predicted by Newton's laws (fig. 4). The method gives students an intuitive and visual understanding of the meaning of the least action principle, as does the computer modeling described earlier.

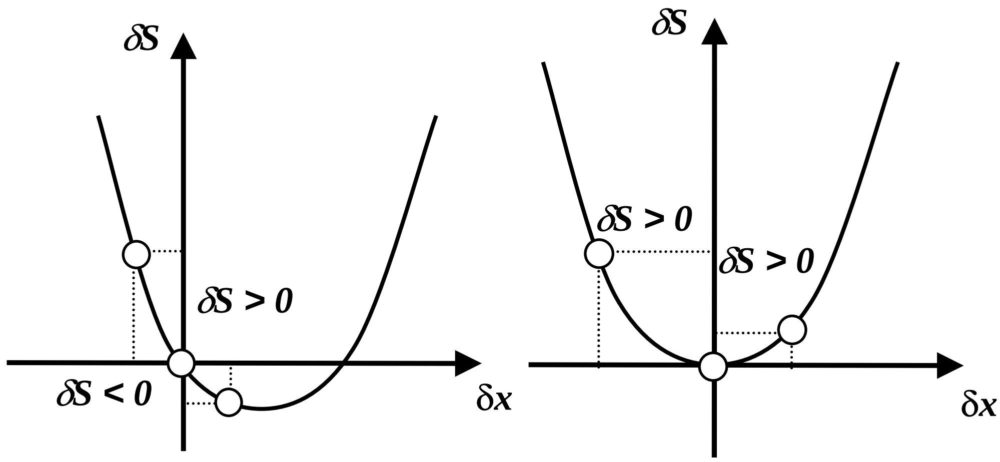


Figure 4. Both parts of the figure display changes in action with respect to displacement δx. In the left part the linear term in eq. (3) is not zero, i.e. $m\Delta v - F\Delta t \neq 0$. The action demonstrates both negative and positive changes, so the chosen worldline does not yield a minimal action. In the right part the condition that the linear term be zero, $m\Delta v - F\Delta t = 0$, gives a required minimum.

Maupertuis' principle of least action

Finally we will analyze the apple's motion from the viewpoint of the second least action principle called Maupertuis' principle of least action. Maupertuis requires: "Consider a conservative system. Give initial and final position and total energy of the system."

The total energy and its conservation (we again assume knowledge of the potential energy), lead to knowing the apple's initial speed, which implies two possibilities of motion – with the upward or downward direction. But in the case of the falling apple we are interested only in the downward motion. Then according to Maupertuis we are able to answer the question: "*What is the apple's final event, if we know its initial and final positions (see fig. 5)?*"

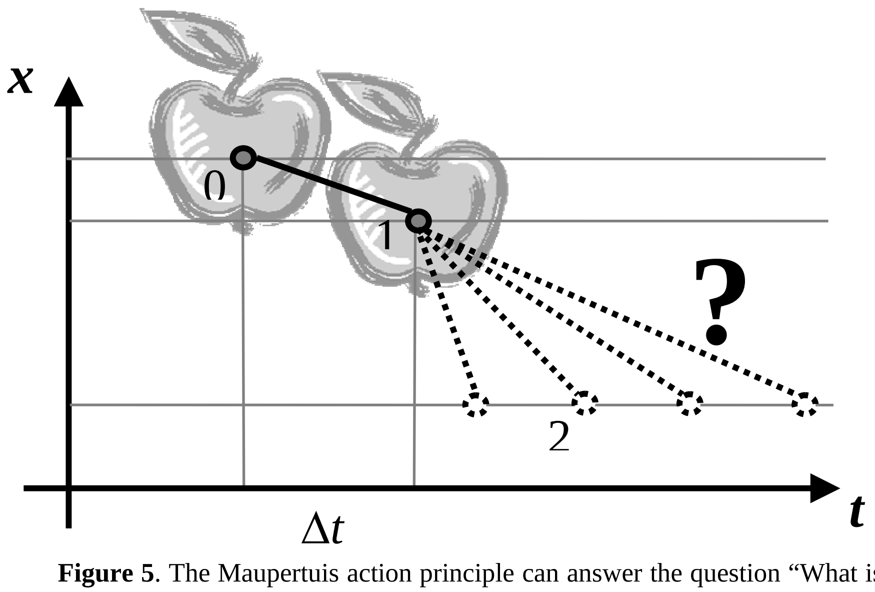


Figure 5. The Maupertuis action principle can answer the question "What is the apple's final event, if we know its initial and final positions?"

Since we consider a conservative system the actual motion of the apple must satisfy energy conservation. From Newtonian mechanics we know that it is the same motion as predicted by Newton's laws. Everything seems to be good. So the natural question arises: where is the action principle? But we now see that we did not realize that energy conservation alone actually allows other worldlines, strange and unrealistic with respect to Newton's laws. One example is shown in fig. 6. How to recognize a motion as actual or unrealistic?

The criterion is just the Maupertuis action. It can be shown in a very similar way as in Hanc et al. 2005 that a useful graphical tool in the case is the velocity vs. position diagram called in mechanics *the phase diagram*, which says that for unrealistic worldlines the area under *the phase curve* is always bigger than for the actual one. This geometric idea provides a foundation for the definition of the second version of action, Maupertuis action *W*:

$$W \equiv \begin{pmatrix} \text{area under} \\ \text{phase curve} \end{pmatrix} \cdot \begin{pmatrix} \text{object's} \\ \text{mass} \end{pmatrix} \quad \text{or} \quad W = \int_{\substack{\text{initial}\\\text{position}}}^{\substack{\text{final}\\\text{position}}} mv\,ds \qquad (4)$$

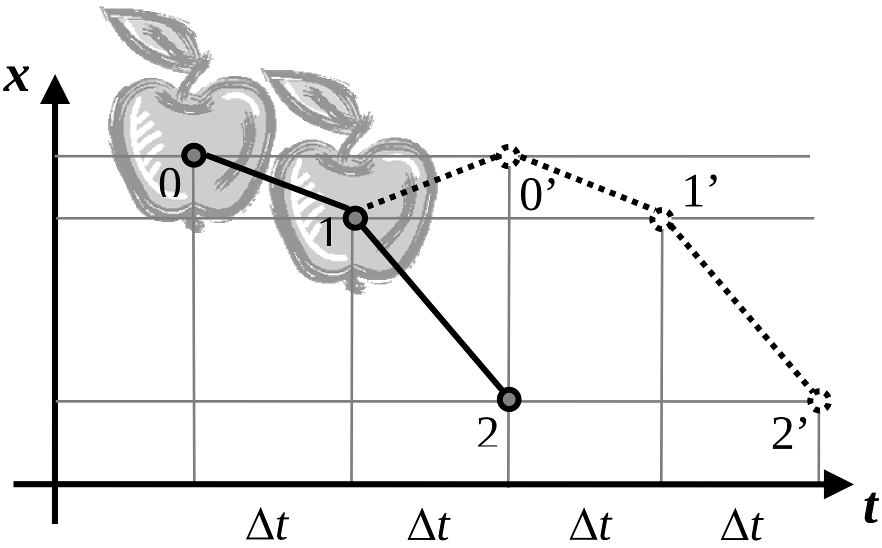


Figure 6. Both depicted worldlines 010 and 010'1'2' satisfy energy conservation, but only one describes the real motion.

Summarizing we can say that Maupertuis' principle of least action tells the falling apple to move so that the product of mass and area under the phase curve has the smallest possible value (subject to energy conservation). So far as computer modeling is concerned, it is the problem as before.

Acknowledgments. The work was partly supported from the grant of Slovakian cultural and educational agency (KEGA), Project No. 3/3005/05.

List of references

Chabay, R.W., Sherwood, B. A. (2002) *Matter and Interactions: I. Modern Mechanics*. John Wiley & Sons, New York. <http://www4.ncsu.edu/~rwchabay/mi/>

Euler, L. (1744). *Methodus Inveniendi Lineas Curvas Maximi Minimive Proprietate Gaudentes sive…* Laussanne and Geneva; see also chap. 2 in Goldstine H. H. (1980), *A History of the Calculus of Variations from the 17th Through the 19th Century*. Springer-Verlag, New York.

Feynman, R.P., Leighton, R.B., Sands, M. (1964) *The Feynman Lectures on Physics.* Addison-Wesley, Reading, MA, Vol. I

Goldstein, H., Poole, Ch., Safko, J. (2002). *Classical Mechanics*. (3rd ed.) Addison–Wesley, Reading, MA

Hanc, J., Taylor, E.F., Tuleja, S. (2005). Variational mechanics in one and two dimensions. *Am. J. Phys.* (73), 603-610 (and references in the article) <http://www.eftaylor.com/leastaction.html>

Hand, L.N., Finch, J. D. (1998). *Analytical Mechanics*. Cambridge University Press, Cambridge, UK.

Landau, L. D., Lifshitz, E. M. (1976). *Mechanics* (3rd ed.). Butterworth-Heinemann, Oxford, UK.

Moore, T. A. (2003) *Six ideas that Shaped Physics: Unit N - The Laws of Physics are Universal.* McGraw Hill, New York. <http://www.physics.pomona.edu/sixideas/>

Marion, J.B., Thornton, S.T. (2003). *Classical Dynamics of Particles and Systems* (5th ed.). Thomson Brooks/Cole.

Newton, I. (1687) *Philosophiae Naturalis Principia Mathematica*, London; see also chap. 3 in Goodstein, D. L. et al (1999). *Feynman's Lost Lecture* …W.W. Norton, New York.

A First Introduction to Quantum Behavior

Jon Ogborn Institute of Education University of London

Abstract

The physics curriculum in England and Wales has a requirement to introduce quantum phenomena to students in the first year of the two-year pre-university physics course in schools. Usually this means discussing the photoelectric effect, with a few words about "waves or particles". In the innovative course "Advancing Physics" we take a more fundamental approach, following Feynman's "many paths" formulation of quantum physics. The experiences of teaching this material for the past five years are discussed, together with difficulties it has thrown up.

Introducing quantum behavior: a national requirement

In England and Wales, all "A-level" Physics courses – that is, high school physics courses leading to University entrance are required to introduce the quantum behavior of photons and electrons. For the innovative course *Advancing Physics* (Ogborn & Whitehouse 2000) we decided to base our approach on Richard Feynman's remarkable small book "QED" (Feynman 1985), in which he describes in the simplest possible terms his 'sum over histories' or 'many paths' approach to quantum mechanics. Others have tried something similar (see Hanc et al 2005).

The Feynman approach in essence

> '…Dick Feynman told me about his … version of quantum mechanics. "The electron does anything it likes," he said. "It just goes in any direction at any speed, forward or backward in time, however it likes, and then you add up …." I said to him, "You're crazy." But he wasn't.'
> *Freeman Dyson*

Feynman's big idea, starting with his 1942 doctoral thesis (Brown 2005), was to track all the space-time paths available to photons or electrons in a given situation. Every possible path is associated with a quantum amplitude. To find the probability of an event, you add up (taking account of phase) the amplitudes for all possible paths leading to that event. The probability is then given by the square of the amplitude, suitably normalized. The phase of the amplitude is given by a result first noted by Dirac, namely that the number of rotations of the phase 'arrow' along a path is just the classical action S along the path, divided by the Planck constant h.

This alternative way of setting up quantum mechanics led to enormous simplification of calculations in quantum field theory, and is still today an essential tool for theoretical physics. Our concern, however, is with the simplification and clarification it can bring to a first introduction to the quantum world.

Six steps in introducing quantum behavior

We think of our teaching program in six steps:

1. random arrival
2. photon energy in lumps of size $E = hf$
3. superposition of amplitudes
4. what is quantum behavior?
5. quantum behavior can explain ....
6. electrons do it too.

Step 1 Random arrival

Perhaps the most important first experience of quantum behavior is to listen to a Geiger counter detecting gamma photons: "click....click. click........click". The key point is that the gamma photons arrive *at random*. The time of arrival of a photon is not predictable: the only thing we can know is the *probability* of arrival. Here is the first cornerstone of an understanding of quantum behavior: only the probability of events is predictable.

A well known set of photographs (Figure 1) illustrates the idea beautifully.

A more careful treatment would want to show that the arrival of photons follows a Poisson distribution, but this level of discussion is not available to us in this course, though we are working on it (Ogborn, Collins & Brown 2003a,b).

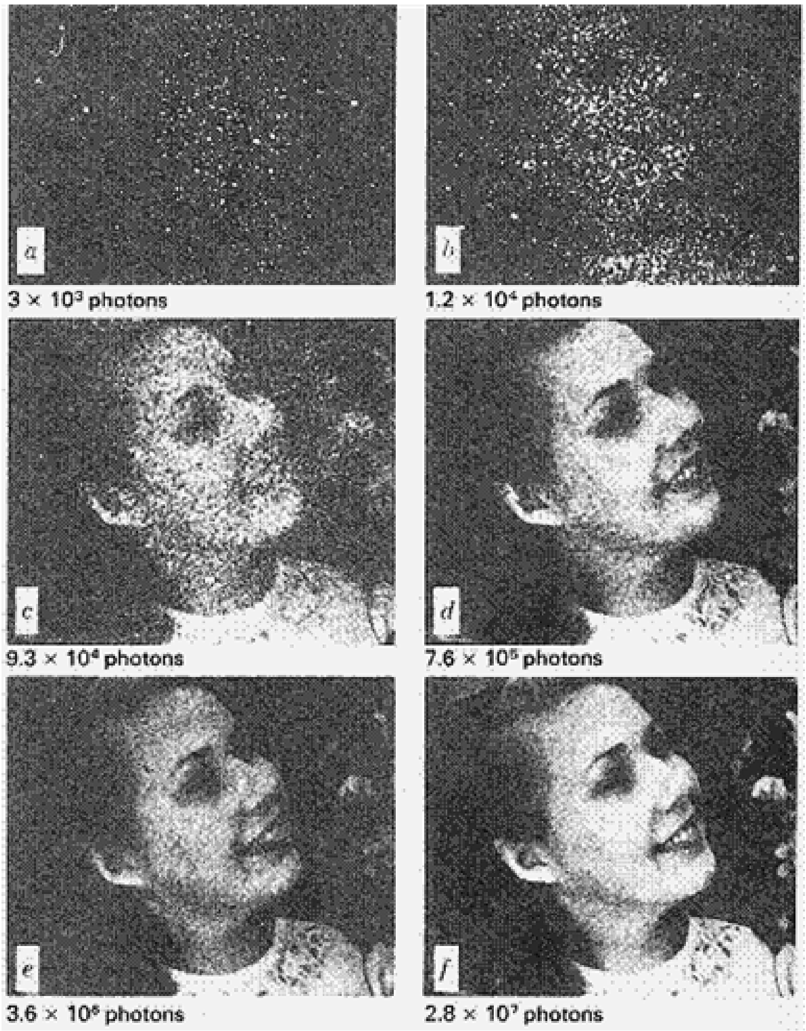


Figure 1 The same picture taken with progressively more and more photons. The photons arrive randomly, but with probabilities such that the picture gradually builds up

Step 2 Photon energy in lumps of size E = hf

The next step is to measure the energy and frequency to arrive at an estimate of the Planck constant h. We suggest the use of a set of light-emitting diodes (LEDs), measuring the wavelength of the light they emit, and the minimum potential difference needed for light just to be emitted.

The point is to get across, as simply and directly as possible, that whatever happens in between emission and arrival, light is always emitted and absorbed in discrete amounts $E = hf$.. Now we have to think about “what happens in between”.

Step 3 Superposition of amplitudes

In *Advancing Physics*, the study of quantum behavior follows a study of the nature of light. There, the classical story of the development of the wave picture is told, from Huygens and

Fermat through to Young and Fresnel. Thus students know that interference effects arise when there are alternative paths between emission and absorption events. Now we marry up this wave picture of superposition with the story of quantum behavior.

The big idea of quantum theory can be put very simply. It is just: “steal the wave calculation but forget about the waves”. That is, associate with each path a phase ‘arrow’, and add up the ‘arrows’ to get the resultant amplitude from all paths.

This is just Huygens’ wavelet principle. But something new and essential is added. It is tht the rate of rotation of the quantum arrow along a path is given by $f = E/h$. In this way, the results of the wave calculations of interference and diffraction patterns are all taken over. The novelty is to start with the photon energy E as given and fundamental, and to use h, the quantum of action, to translate it into a rate of rotation of phase.

Here we arrive at the heart of quantum behavior. In wave theory, the existence of a phase is a consequence of the nature of wave motion. In quantum thinking, the existence of a phase is rock-bottom fundamental. It is to be thought of as a given, not as a consequence.

Step 4 Quantum behavior

We can now describe the behavior of quantum objects. For each possible path between initial and final discrete events, there is an ‘arrow’ (a phasor). For photon paths, the rate of rotation of the arrow between start and finish is E/h. Add up the arrows for all the possible paths, tip to tail, to get the resultant ‘arrow’ for the pair of events. The square of the resultant ‘arrow’ is proportional to the probability of the pair of events.

Although initial and final events are localized, there is no reason to think of a photon as localized “in between”. Lumpiness in energy does not imply lumpiness in space. To say as Feynman does that the particles “go everywhere” is not to picture them as trying out all the possibilities one at a time. Part of the essence of “being a particle”, namely its continuing existence at a succession of places and times, has been taken away. Waves, of course, do “go everywhere”, but part of their essence has gone too. Their energy can no longer be divided into smaller and smaller amounts, without limit.

Not surprisingly, this account gives students difficulties. Like most of us, they naturally try to form as concrete a picture as possible. So, thinking of photons as particles, and imagining tracking each along a given path, they wrongly imagine them trying out all possible paths one by one. They tend to think of the rotating phasor, not as associated with the *path*, but as ‘riding on the back of a photon’ as it travels.

However, these difficulties are simply the difficulties of getting used to quantum thinking. A possible merit of the approach is that it brings them out so clearly, by denying that photons are some mixed-up approximation to waves and particles. Thus we present quantum behavior as itself, not as like something else. Its essence is that all possibilities contribute, superposing taking account of phase.

Step 5 Quantum behavior explains...

We conclude by offering students some comfort, by showing how quantum behavior explains some familiar things, in particular, the laws of reflection and refraction.

Figure 2 shows a diagram from Feynman’s book, illustrating how the law of reflection at a plane surface arises directly out of the quantum behavior of photons.

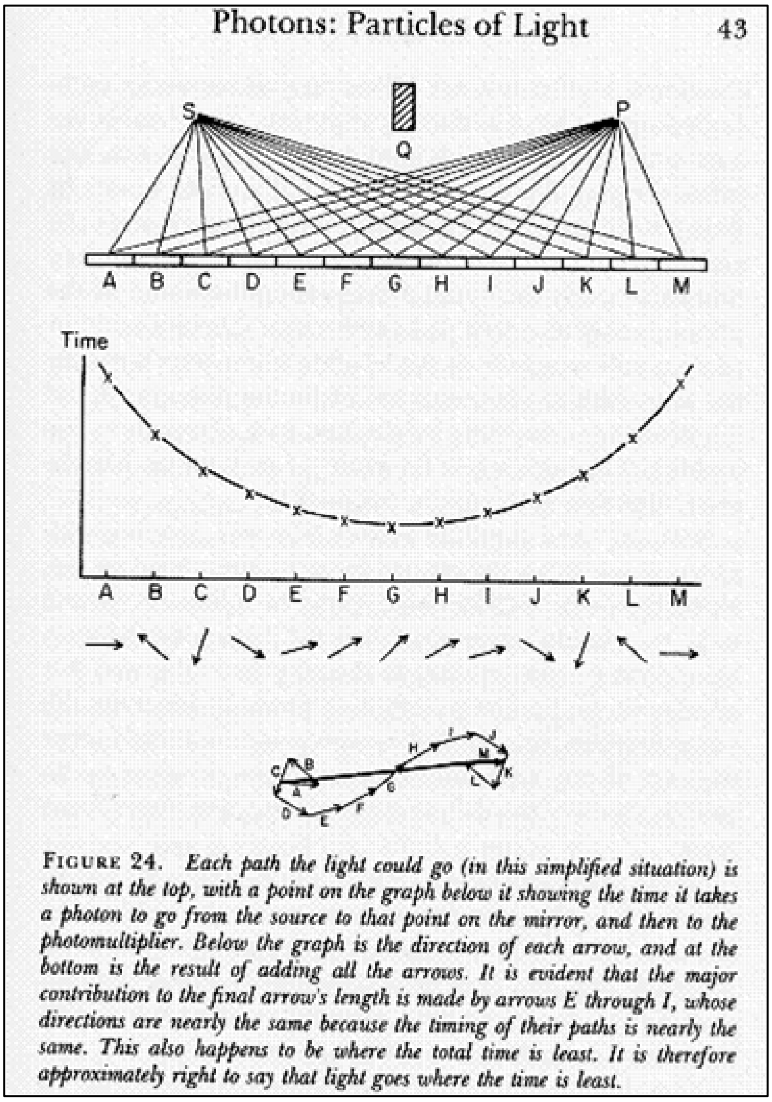


FIGURE 24. *Each path the light could go (in this simplified situation) is shown at the top, with a point on the graph below it showing the time it takes a photon to go from the source to that point on the mirror, and then to the photomultiplier. Below the graph is the direction of each arrow, and at the bottom is the result of adding all the arrows. It is evident that the major contribution to the final arrow's length is made by arrows E through I, whose directions are nearly the same because the timing of their paths is nearly the same. This also happens to be where the total time is least. It is therefore approximately right to say that light goes where the time is least.*

Figure 2 Reflection of photons at a plane mirror

Every possible path from source to detector counts. But only in the middle are nearby paths closely similar in length, so that the ‘arrows’ associated with them are nearly in phase and thus ‘line up’. Further out, the phase changes rapidly as the path changes, so that the arrows from these paths ‘curl up’, giving a very small contribution to the final resultant arrow. Only the paths very close to the path prescribed by the law of reflection make an important contribution.

This explains Fermat’s principle of least time. The graph shows how the travel-time varies as the path changes. At the minimum the time does not vary as the path changes. So the phase associated with these paths does not change either. However, quantum thinking gives us something more. By thinking about *how much* the phase changes between nearby paths we can estimate the limits to the description given by geometrical optics.

The traditional “wave-particle duality” story makes it seem as if wave and particle behavior are partial explanations of quantum behavior. Something more like the reverse is true. Particle behavior, in particular all of classical Newtonian mechanics *is explained by* quantum behavior (see Ogborn & Taylor 2005). Equally, the wave behavior of light *follows from* the quantum behavior of photons, in the limit of low photon energy and many photons. Quantum physics helps to explain why – and when – classical ideas work.

Step 6 Electrons do it too

The description of quantum behavior is appropriate not just for photons, but for all particles, including electrons. For a free (non-relativistic) electron, the rate of rotation of the quantum arrow is just *K/h*, where *K* is the kinetic energy. Demonstration apparatus available for school laboratories makes it straightforward to show electron diffraction. Even better is to see electrons arrive one by one, building up gradually into a two-slit interference pattern. This has been achieved by a Hitachi team led by Akira Tonomura, who have produced a beautiful film clip of this most fundamental experiment (Tonomura et al 1999).

The rate of rotation of the phase of an electron is given by *L/h*, where *L* is the Lagrangian, and the total phase rotation is *S/h*, where *S* is the action along the path. We don't mention this to students in *Advancing Physics*, but we mention it here to show how the simple approach can be extended in later work.

Does it work?

The only evaluation we have, so far, of the value of this approach is the experience of it being taught in about 25% of UK schools teaching physics at this level. The teachers have available to them an email network, to discuss whatever they please. Every year, when this topic is taught, there is a flurry of discussion, about what exactly the ideas mean and about how to respond to students' questions.

These discussions show that teachers new to the approach are understandably nervous, and need a good deal of re-assuring. It also shows that they, like their students, are prone to giving the ideas an over-concrete interpretation, particularly in wanting to associate a 'traveling phasor' with a 'traveling photon', rather than associating a phasor with a possible path.

Some ways of teaching the ideas lead very strongly to over-concrete interpretation. One popular idea is to push a rotating wheel along a number of paths and to note the total rotation for each path. This very effectively shows how, for example, the results of Figure 2 arise. But students inevitably want to know what role the wheel, and its diameter and speed, play in the theory.

However, it remains true that there is, in evaluation of the whole *Advancing Physics* course, no demand to remove or change this topic. Nor do students complain about the examination questions set on it. We can say, therefore, that this rather radical innovation has succeeded at least to the extent of having survived for six years so far, in a course taught on a national scale in schools of widely differing kinds. And at least a proportion of both teachers and students find it very interesting and thought-provoking.

References

Ogborn, J & Whitehouse, M. (Editors) (2000) *Advancing Physics AS.* Bristol: Institute of Physics Publishing

Feynman, R. (1985) *QED: The strange theory of light and matter*. London: Penguin Books

Brown, L. M. (Editor) (2005) *Feynman's Thesis: A new approach to quantum theory*. Singapore: World Scientific.

Hanc, J., Tuleja, S., (2005) 'The Feynman Quantum Mechanics with the help of Java applets and physlets in Slovakia', Proceedings of MPTL 10, Freie Universität Berlin, 5 - 7 October, <pen.physik.uni-kl.de/w_jodl/MPTL/MPTL10/Proceedings.html>

Ogborn, J., Collins, S. & Brown, M. (2003a) 'Randomness at the root of things 1: random walks'. *Physics Education* volume 38, number 5, pages 391-397

Ogborn, J., Collins, S. & Brown, M. (2003b) 'Randomness at the root of things 2: Poisson sequences'. *Physics Education* volume 38, number 5, pages 398-405

Ogborn, J. & Taylor, E.F. (2005) 'Quantum physics explains Newton's laws of motion'. *Physics Education* Volume 40, number 1, pages 26-34

Tonomura, A. et al (1999) 'Demonstration of single-electron buildup of an interference pattern'. American Journal of Physics Volume 57, No 2, pages 117-120 (movie clip downloadable from www.pnas.org. Search for article under Author: Tonomura, A; Year 2005; Volume 102; First Page 14952.)

Use, Abuse, and Unjustified Neglect of the Action Principle

Edwin F. Taylor, Massachusetts Institute of Technology, eftaylor@mit.edu

Abstract

Traditionally, differential equations dominated physics education; the action principle was used primarily to derive differential equations such as Lagrange's equations. Now the computer allows the action principle to be applied directly from first principles, often bypassing analytic solutions entirely. The action principle can illuminate and unify physics education and physics research from quantum field theory to cosmology.

Computer solutions without solving equations

In classical mechanics it is not true that action along a particle worldline is always minimum. Often it is a minimum, and it is never a maximum. But sometimes action can be a saddle point, which means that, compared with the true worldline, some adjacent curves have a greater value of the action, while some have a smaller value of the action. Seems complicated. However, it is always true that the action along a sufficiently short worldline -- or a sufficiently short segment of a worldline -- is always a minimum. Therefore we can correctly say:

A particle follows a worldline such that the action along every small segment is a minimum compared with that along every nearby segment with the same endpoints.

Looking at this universal principle tells you immediately that computers are perfect for applying action. Computers do increments beautifully. If a computer finds, by whatever means, a worldline along every segment of which action is a minimum, then that is a true worldline, one that a particle can follow.

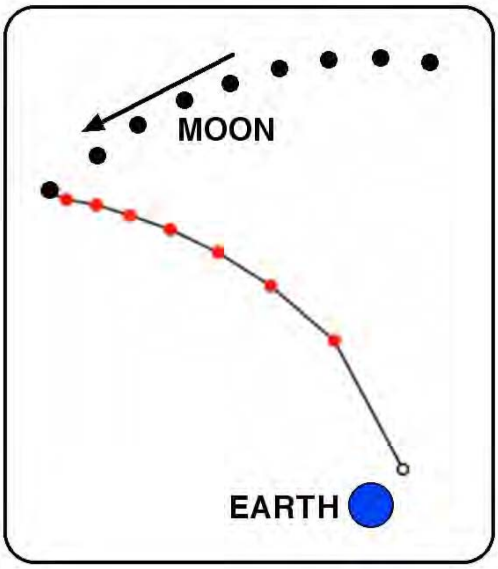


Figure 1. Transfer trajectory between parking orbits around earth and moon. Dots on the trajectory are the events of clock ticks that completely determine the worldline. The operator can add intermediate clock ticks, drag individual ticks to minimize the action, or have the computer minimize action automatically.

Figure 1 shows a frame of an interactive program by Slavomir Tuleja (2005). The goal is to transfer a probe in an unpowered trajectory from a parking orbit around earth to a parking orbit around the moon. We use Hamilton's action S because the fixed initial and final points must be events; the probe must get to the location of the moon *when the moon is there*. The dots along the trajectory are ticks on the clock carried by the probe, so the representation completely determines the worldline: position vs time. The operator can drag the clock ticks back and forth one at a time to minimize the total action, can add many intermediate ticks to increase the accuracy, and can ask the computer to minimize the action automatically, which it does in a split second. Notice that this analysis moves directly from the action principle to a visualized solution with no intermediate mathematical analysis. And automatically generated spread sheet data, the time and location of every clock tick, can be analyzed to any desired level of detail.

A critic might object that the solution is now too easy; all the student does is push a few buttons. On the contrary, the student has now been freed to investigate a hugely expanded world of possible problems. For example, by changing the time lapse between initial and final events, the student can search for the worldline that minimizes the total rocket impulse required for the transfer from earth orbit to moon orbit. She can try the same for different parking orbits around earth and moon. She can apply a similar program to transfers between earth and mars. The software offers analysis of motion in other potentials as well.

The action principle for special and general relativity and the cosmos

When I was editor of the *American Journal of Physics*, I despaired about the twin paradox, which seemed to poison the literature. Every month some engineer or retired doctor would submit a paper disproving the obviously ridiculous predictions of the twin paradox, thus invalidating special relativity.

In fact the twin paradox is central to the use of action to describe high speed motion: In flat spacetime a true particle worldline yields the longest proper time (wristwatch time, aging) between fixed initial and final events. A general expression for the action uses the Lagrangian L, which for low speeds is the difference between the kinetic and potential energies.

$$S \equiv \int L dt$$

A particle moving at any speed in an electromagnetic field has the Lagrangian

$$L = \frac{1}{2} m \mathbf{v} \bullet \mathbf{v} - \frac{q}{c} \mathbf{v} \bullet \mathbf{A}$$

where ϕ is the scalar potential and $\mathbf{A}$ is the vector potential. Two equations determine all possible worldlines under electromagnetic influence in flat spacetime!

Now, general relativity is weird, but has the following gorgeous simplification: When there are no singularities or gravitational waves, then at every event on a particle worldline you can always find a local inertial frame in which special relativity holds. Since the worldline is the sum of segments in these local frames, and since proper time is an invariant, the same for all observers, therefore in general relativity the worldline of a particle between fixed end events is the one with maximum aging along each small segment. We call this result *the principle of maximal aging*. For low speeds and small curvature of spacetime, the principle of maximal aging reduces to the principle of least action.

How do we find the value $d\tau$ of the proper time (aging) along a segment of a worldline in curved spacetime? From the *metric*, the solution to the field equations. On the left of the metric equation is the increment $d\tau$ of proper time (wristwatch time, aging) between a nearby pair of events on the worldline. On the right side of the metric equation are the corresponding increments of the (arbitrarily chosen) coordinates between that pair of events.

Now, the metric is expressed in increments; manipulating the metric requires only calculus. This means that if, instead of starting with the field equation, we start with the metric solutions, we can introduce general relativity to sophomores using only calculus and the principle of maximal aging. John Archibald Wheeler and I did this in our text *Exploring Black Holes, Introduction to General Relativity,* Fig. 2 (Taylor & Wheeler 2000)

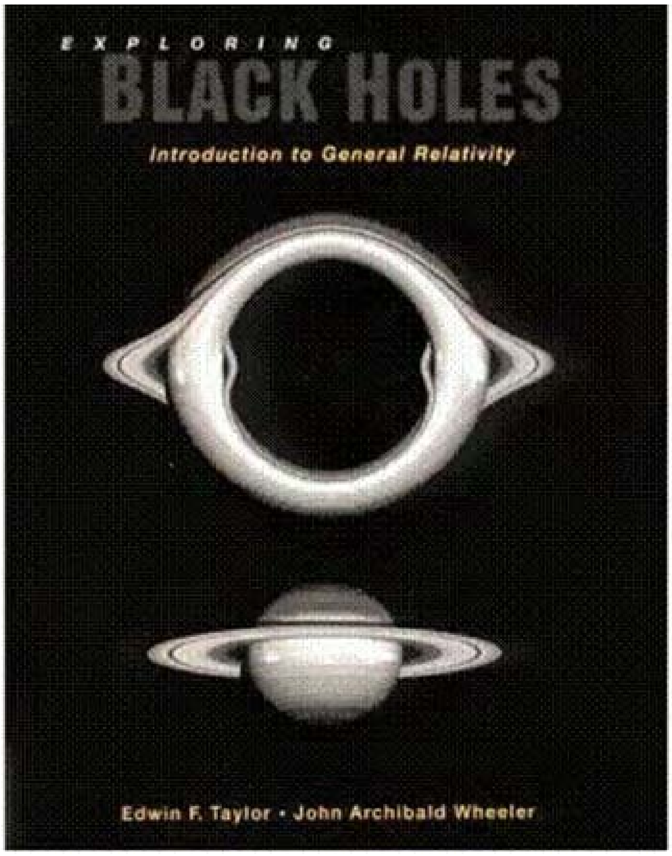


Figure 2. Text *Exploring Black Holes* by Edwin F. Taylor and John Archibald Wheeler that introduces general relativity to sophomores using the metric and the principle of maximal aging.

The action principle not only tracks particles in curved spacetime. It is also at the root of Einstein's field equations themselves. Hilbert derived the field equations from an action principle, some say before Einstein completed his theory. Landau and Lifshitz do the same for advanced physics students. Thus one can say that action describes the fundamental non-quantum laws of the cosmos.

Quantum mechanics from the bottom

Recall Feynman's formulation of nonrelativistic quantum mechanics, exemplified by the electron:

- The electron explores all possible worldlines between fixed end events that we choose.
- Along each trial worldline the total rotation of the quantum phasor is *S/h = (Action)/h*.
- At the end event, add up phasors for all worldlines to give the resultant quantum amplitude.
- The probability of detecting the electron at the final event is proportional to the squared magnitude of the resultant quantum amplitude.

This is not a new idea! Sixty four years ago it was summarized in the introduction to Feynman's 1942 Ph.D. thesis under John Wheeler: *A generalization of quantum mechanics is given in which the central mathematical concept is the analogue of the action in classical mechanics . . . It is only required that some form of least action principle be available . . . if a*

Lagrangian exists . . . the generalization reduces to the usual form of quantum mechanics. In the classical limit, the quantum equations go over into the corresponding classical ones, with the same action function. (Brown 2003)

Feynman was a co-recipient of a Nobel Prize for expanding these ideas to quantum field theory, which one can take to be the most fundamental current theory of the very small (a status that string theory has not yet achieved). And quantum field theories can be derived from action. A quote from the Web: "Of all possible fields with a given boundary condition the one that provides an extremum . . . of the action is The Solution."

Thus action carries us seamlessly from the smallest that we know to the largest, the universe as a whole.

Variational principles: parents of action

. . . *once the laws of physical theory are expressed as differential equations, the possibility of their reduction to a variational principle is evident from purely mathematical reasoning* . . . (Yourgrau & Mandelstam 1968)

Traditionally differential equations have been the analytic tool of choice in physics, both for education and research. As we have seen, the computer can apply variational principles directly, often eliminating intermediate mathematical analysis. The use of variational principles in current physics text is, at best, spotty. A few scattered examples:

Landau and Lifschitz develop the first two Maxwell Equations from experiment and hand-waving. The last two equations they derive from a variational principle, which they call action.

A standard method for finding the ground state wave function of an atom is to minimize the electromagnetic energy.

The relaxation method is a powerful one for determining the electrostatic field resulting from an array of fixed charges.

Textbooks often miss powerful and simplifying applications of variational principles. For example, Van Baak (1999) replaces Kirchoff's circuit theorems by requirements that (1) current is conserved, and (2) the rate of dissipation of energy is minimized. By using this method, he says, the usual "extravagance of equations is wholly avoided."

Strategies for introducing action

Variational principles -- and action -- are tools, like differential equations, and not fundamental physical principles (though they are related to conservation laws through Noether's theorem; see e.g. Hanc et al. 2004). Still, they are unifying tools that the computer can implement throughout physics education and research. Here are some notes on strategies for introducing them to physics instruction:

1. Sneak bits of action and variational principles into secondary classes and introductory undergraduate courses (see particular examples in Hanc et al. 2003 or Hanc & Taylor 2004).

2. Do NOT make action the primary tool in introductory physics; it is not concrete enough. Introductory students need to feel the pushes and pulls of forces.

3. Use action and variational principles as unifying tools in the remainder of undergraduate and graduate programs.

References

Hanc, J., Taylor E.F., (2004) ‘From Conservation of Energy to the Principle of Least Action: A Story Line’, *Am. J. Phys.* 72 (4), 514 - 521

Hanc, J., Tuleja, S., Hancova, M. (2004). ‘Symmetries and Conservation Laws: Consequences of Noether's Theorem’, *Am. J. Phys.* 72 (4), 428 - 435;

Hanc, J., Tuleja, S., Hancova, M., (2003), ‘Simple derivation of Newtonian mechanics from the principle of least action’, *Am. J. Phys.* 71 (4), 386-391.

Brown, L. M. (Editor) (2005) *Feynman's Thesis -- A New Approach to Quantum Theory*, Singapore: World Scientific Publishing Co. page 1

Taylor, E. F. & Wheeler, J. A. (2000) *Exploring Black Holes: Introduction to General Relativity*, San Francisco: Addison Wesley.

Tuleja, S. et al (2005). Use of a computer in advanced mechanics - Principle of least action, Proceedings of MPTL 10, Freie Universität Berlin, 5 - 7 October, 2005, <pen.physik.uni-kl.de/w_jodl/MPTL/MPTL10/Proceedings.html>

Van Baak, D. A. (1999) ‘Variational alternatives to Kirchoff's loop theorem in dc circuits’, *Am. J. Phys.* **67**, 36-44.

Yourgrau, W. and Mandelstam, S. (1968), *Variational Principles in Dynamics and Quantum Theory*, New York: Dover Publications, page 175.

Software and background publications are available at www.eftaylor.com/leastaction.html;