ORIGIN OF THE PRINCIPLE OF LEAST ACTION

Link to: physicspages home page.

To leave a comment or report an error, please use the auxiliary blog and include the title or URL of this post in your comment.

Post date: 19 Jul 2023.

This post is based on the book by Leonard Susskind and George Hrabovsky (2013), *Classical Mechanics: The Theoretical Minimum,* Lecture 6.

We've used the Lagrangian formulation to analyze numerous problems in both classical and quantum physics, but in all these applications, it was just assumed that the principle of least action, that is, the principle that a particle always travels along a trajectory that minimizes the integral of the Lagrangian between two points, is obeyed. The origin of this principle is never mentioned in any of the books I've looked at, which is why Susskind's treatment proved quite interesting and valuable.

Incidentally, there are a large collection of Susskind's lectures available on the Theoretical Minimum web site. The lectures are aimed at people who have some background in college-level mathematics (familiarity with the basics of calculus, linear algebra and so on is needed), but they tend not to delve too deeply into any given topic. As such, you won't find detailed mathematical derivations, but the lectures are valuable for the physical insight they give into the principles behind most of the modern theories in physics. Certainly worth a look.

Anyway, back to least action. Where does it come from? It seems that the idea was formulated by people such as Lagrange and Hamilton by observing that when travelling through various transparent media, a light ray always takes the path that requires the least time. For example, when you shine a light beam from a source in air into another medium such as glass or water, the ray bends because the speed of light in the denser medium is less. This is known as Fermat's principle of least time.

It appears that physicists at the time wondered if this principle could be generalized to say that the trajectory of any system between a starting time t_1 and ending time t_2 also follows a path where some quantity is minimized. The philosophical basis of this idea is that nature always chooses the most efficient way of doing anything (like water finding the shortest path downhill) so there should be *something* that is being minimized in any physical process.

This idea becomes formalized when we propose a function (the Lagrangian) whose integral between two points is minimized (strictly speaking, the integral is *stationary* rather than minimized, but in practice, it is always a minimum). The first question to be answered is: what is this function a function *of*? In order to solve Newton's equations for the motion of a particle (or collection of particles) we need to specify their initial positions and velocities (along with the force law), so it seems natural to make L a function of positions and velocities. Thus we usually write

$$L = L(q_i, \dot{q}_i) \tag{1}$$

where the q_is are generalized coordinates and the $\dot{q}_i$s are velocities. At this point, there is no restriction on L apart from this assumption. The principle of least action then says that we should require the integral

$$S = \int_{t_1}^{t_2} L(q_i, \dot{q}_i)\, dt \tag{2}$$

to be a minimum over all possible choices of L. At this point, we can apply the usual derivation via the calculus of variations to get the Euler-Lagrange equations

$$\frac{\partial L}{\partial q_i} - \frac{d}{dt}\frac{\partial L}{\partial \dot{q}_i} = 0 \tag{3}$$

The derivation in Susskind's book is actually done without using the calculus of variations explicitly, and is a lot more intuitive. It goes like this.

Consider a general function L between the two times t_1 and t_2. The action (the thing we want to minimize), is given by the integral 2. In order to find the minimum of this integral, we need to vary L at every point along the curve and find the minimum of all these varied curves. Since there are an infinite number of points along the curve, we're effectively minimizing a function with respect to an infinite number of variables, which seems to be a hopeless task.

To approach this problem, we approximate the integral by dividing the time interval into discrete time steps, each of size Δt. Then the integral becomes (approximately)

$$S = \sum_n L_n(q, \dot{q})\, \Delta t \tag{4}$$

where now the subscript n refers to the time increment at which we're evaluating L and we've dropped the subscript i from the q_is so that we're considering a single coordinate for simplicity (the derivation is the same for more coordinates). However, since we're discretizing the intervals, we need to know the values of q and $\dot{q}$ at each point. A reasonable approximation would seem to be

$$q(t) = \frac{q_n + q_{n+1}}{2} \tag{5}$$

$$\dot{q}(t) = \frac{q_{n+1} - q_n}{\Delta t} \tag{6}$$

That is, for each position, we're taking the average of the positions at times t_n and t_{n+1} and for each velocity, we're taking the slope of the line connecting times t_n and t_{n+1}. Note that now the subscripts n and $n+1$ on the q_n terms denote the time at which the measurement is made (we're still dealing with only one coordinate). Thus our approximation for the action is now

$$S = \sum_n L\left(\frac{q_n + q_{n+1}}{2}, \frac{q_{n+1} - q_n}{\Delta t}\right)\Delta t \tag{7}$$

Now suppose we vary L at one particular point along the curve, say, point b. How much does this change the action? Since we're varying L at only one point, we need to consider only those terms in the sum that depend on this point. Since L depends only on nearest neighbour points, we'll need to consider only the points $b-1$, b and $b+1$. There are only two terms in the sum that depend on the time point b, so the terms to be examined (which we'll call A) are

$$A = L\left(\frac{q_b + q_{b+1}}{2}, \frac{q_{b+1} - q_b}{\Delta t}\right)\Delta t + L\left(\frac{q_{b-1} + q_b}{2}, \frac{q_b - q_{b-1}}{\Delta t}\right)\Delta t \tag{8}$$

To minimize the action over all possible variations of L at the one point b, we can take the derivative of this term with respect to q_b and set it to zero. To do this, note that the first L term is effectively the value of L at time t_{b+1} and the second L term is the value at t_b. Thus the derivative consists of two terms, one at t_{b+1} and the other at t_b. That is

$$\frac{\partial A}{\partial q_b} = \frac{\Delta t}{2}\left[\left.\frac{\partial L}{\partial q_n}\right|_{n=b+1} + \left.\frac{\partial L}{\partial q_n}\right|_{n=b}\right] + \tag{9}$$

$$\frac{\Delta t}{\Delta t}\left[-\left.\frac{\partial L}{\partial \dot{q}_n}\right|_{n=b+1} + \left.\frac{\partial L}{\partial \dot{q}_n}\right|_{n=b}\right] \tag{10}$$

To get these derivatives, we've used the chain rule and the shorthand notation

$$\dot{q}_n = \frac{q_{b+1} - q_b}{\Delta t} \tag{11}$$

so that, for example,

$$\frac{\partial}{\partial q_n}L\left(\frac{q_n+q_{n+1}}{2},\frac{q_{n+1}-q_n}{\Delta t}\right)=\frac{\partial}{\partial\frac{q_n+q_{n+1}}{2}}L\left(\frac{q_n+q_{n+1}}{2},\frac{q_{n+1}-q_n}{\Delta t}\right)\frac{\partial}{\partial q_n}\left(\frac{q_n+q_{n+1}}{2}\right)+$$
$$\frac{\partial}{\partial\frac{q_{n+1}-q_n}{\Delta t}}L\left(\frac{q_n+q_{n+1}}{2},\frac{q_{n+1}-q_n}{\Delta t}\right)\frac{\partial}{\partial q_n}\left(\frac{q_{n+1}-q_n}{\Delta t}\right) \tag{12}$$
$$=\frac{\partial L}{\partial q_n}\left(\frac{1}{2}\right)+\frac{\partial L}{\partial \dot{q}_n}\left(\frac{-1}{\Delta t}\right) \tag{13}$$

Setting 10 to zero and dividing through by Δt we get

$$\frac{1}{2}\left[\left.\frac{\partial L}{\partial q_n}\right|_{n=b+1}+\left.\frac{\partial L}{\partial q_n}\right|_{n=b}\right]+\frac{1}{\Delta t}\left[-\left.\frac{\partial L}{\partial \dot{q}_n}\right|_{n=b+1}+\left.\frac{\partial L}{\partial \dot{q}_n}\right|_{n=b}\right]=0 \tag{14}$$

Taking the limit $\Delta t\to 0$, we see that the first bracket becomes $\frac{\partial L}{\partial q}$ and the second bracket becomes $-\frac{d}{dt}\frac{\partial L}{\partial \dot{q}}$, so we regain the Euler-Lagrange equation 3. This derivation also makes clear that 3 is valid at every point along the trajectory.

As to where Lagrangians come from, Susskind confirms my suspicion that there really isn't any standard way of deriving a Lagrangian. It's a combination of inspired guesswork and considerations from experimental results and symmetries. The relation between L and the kinetic and potential energies $L=T-V$ is valid only for systems where the force is the gradient of the potential. For other cases, such as the magnetic force $\boldsymbol{F}=q\boldsymbol{v}\times\boldsymbol{B}$, the Lagrangian is *not* just $T-V$. When Susskind gives the electromagnetic Lagrangian (in Lecture 11), he does it mainly by inspired guesswork, and by verifying that it does give the correct force law.