PRINCIPLE OF LEAST ACTION

This principle, first formulated by W.R. Hamilton, can be treated as the starting point for all of classical mechanics. In modern physics it is the starting point for everything - every physical system in Nature, whether it be quantum mechanical, or described by general relativity, will have an action and an associated Lagrangian.

In ordinary classical mechanics we have

$$S[2,1] = \int_{t_1}^{t_2} dt \, L[q, \dot{q}; t] \qquad (1)$$

where $L[q,\dot{q};t]$ is the Lagrangian, a function of a coordinate $q(t)$, a velocity $\dot{q}(t) \equiv dq/dt$, and of time t.

Normally we will treat "closed systems" not interacting with any outside agent; then

$$L(q,\dot{q};t) \rightarrow L(q,\dot{q}) \qquad (\text{closed system}) \qquad (2)$$

The principle of least action applies to any system described by a Lagrangian function

$$L(\varphi,\dot{\varphi}) \equiv L(q_1, q_2, \ldots q_N; \dot{q}_1, \dot{q}_2 \cdots \dot{q}_N) \qquad (3)$$

where $Q = (\{q_j\}) \equiv (q_1, q_2, \ldots q_N)$ is a set of generalized coordinates, and where $\dot{\varphi}$ is the associated generalized velocity. If we deal with a set of P particles moving in 3 dimensions, then $N = 3P$, and there are $6P$ degrees of freedom in the system.

Even if we don't know what $L(\varphi,\dot{\varphi})$ is, the principle of least action still tells us what differential eqtn is satisfies. The way this works is as follows:

(1) The principle of least action says that of <u>all</u> <u>possible</u> <u>paths</u> that the system might take, it will only take <u>one</u> of them, the path that minimizes the total action $S[2,1]$ accumulated between the initial and final state.

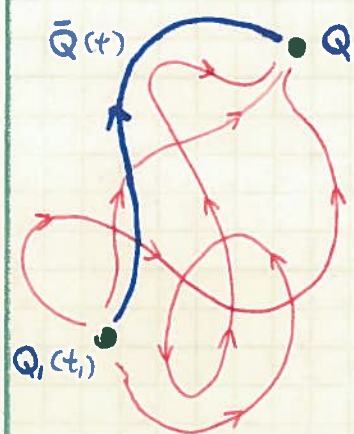
Notice that this is quite different from how Newton's laws work. Newton tells us that in order to determine the dynamics of a system, we need to know some "initial conditions" (eg. the position

$q(t_0)$ and the velocity $\dot{q}(t_0)$ at some time t_0), and then we use the differential eqtn

$$\underline{F}(t) = m\ddot{\underline{q}}(t) \qquad (4)$$

to find the motion at any other time – we must therefore solve (4). The big difficulty with this, in practise, is in determining $F(t)$. For many problems, particularly those involving an assembly of solid bodies of odd shapes which may be rolling, sliding, etc., across each other, with or without friction, this is can be quite difficult.

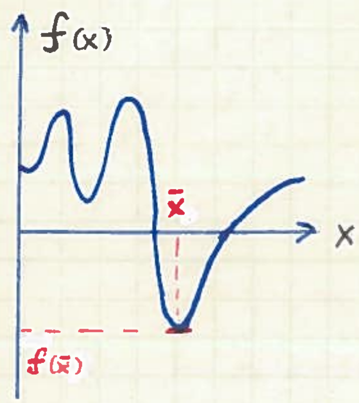
Now to express the principle of least action mathematically we imagine the situation as shown schematically at left. At time t_1 and t_2 we assume the system is at coordinates Q_1 and Q_2. We also assume that there is one path of minimum action, between $Q_1(t_1)$ and $Q_2(t_2)$, which we call $\bar{Q}(t)$. All other possible paths (some are shown in red at left) have action $S(2,1)$ greater than that for the path $\bar{Q}(t)$, ie



$$S[Q(t)] \geq S[\bar{Q}(t)] \qquad (5)$$

for all paths $Q(t) \neq \bar{Q}(t)$.

However, recall that if we try to vary a function $f(x)$ around a value $f(\bar{x})$, where $f(x)$ has a minimum at $\bar{x}$, by varying the variable x, we will find that



$$\left.\frac{df}{dx}\right|_{x=\bar{x}} = 0 \qquad (6)$$

ie.,

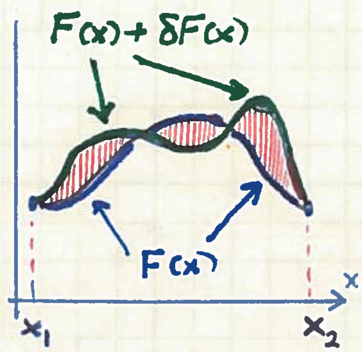
$$df = \left.\frac{df}{dx}\right|_{x=\bar{x}} dx = 0 \qquad (7)$$

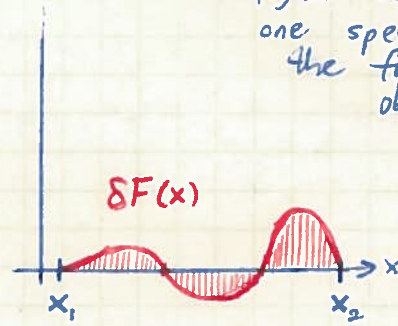
In other words, even though we have a minimum of $f(x)$ at $x=\bar{x}$, any infinitesimal variation of x around $\bar{x}$ will not change $f(x)$ (we deal with other extrema of $f(x)$ below).

We can do the same thing with $S[Q(t)]$, but now by varying the path $Q(t)$ around $\bar{Q}(t)$. This is a much more complicated operation, for there are many different ways to vary the path, all of them infinitesimal, whereas there is only one way to vary the 1-dimensional variable x (by adding dx). Here it is useful to invent some notation to cover the more general case where we vary a function $F(x)$ instead of varying a simple variable x. We write, in analogy with $x \to x + dx$, the change

$$F(x) \longrightarrow F(x) + \delta F(x) \tag{8}$$

which signifies that some change $\delta F(x)$ is being made to $F(x)$ for <u>all values</u> of x. We see that this change $\delta F(x)$ is itself a <u>function</u>. To see how this works, we look at the example in the figure below, which shows just one specific small change made to the function $F(x)$. But there are obviously many ways to do this – as many ways as there are different functions of x. To make the change in $F(x)$ infinitesimal, we can imagine taking some finite function $\delta F(x)$, then multiplying it by an infinitesimal ϵ, and letting $\epsilon \to 0$.





Notice one other thing. Since we have fixed the end-points of $F(x)$ to be $F(x_1) = F_1$ and $F(x_2) = F_2$, then when we vary $F(x)$ to $F(x) + \delta F(x)$, we cannot change F_1 and F_2. This then means that

$$\delta F(x)\Big|_{x=x_1} = \delta F(x)\Big|_{x=x_2} = 0 \tag{9}$$

Let's now go back to our action, for which the function S is given by (1). This is a little more complicated than (8), because we can vary both $\varphi(t)$ and $\dot{\varphi}(t)$, with the condition

$$\delta\varphi(t)\Big|_{t=t_1} = \delta\varphi(t)\Big|_{t=t_2} = 0 \tag{10}$$

(but no condition on the velocities $\dot{\varphi}(t=t_1)$ or $\dot{\varphi}(t=t_2)$; these can be anything we like). It then follows that

$$\delta S = \frac{\delta S[\varphi,\dot{\varphi}]}{\delta\varphi(t)}\Bigg|_{\delta\dot{\varphi}=0} \delta\varphi(t) + \frac{\delta S[\varphi,\dot{\varphi}]}{\delta\dot{\varphi}}\Bigg|_{\delta\varphi=0} \delta\dot{\varphi} \tag{11}$$

in analogy with the eqtn.

$$df(x,y) = \frac{\partial f}{\partial x}\Big|_{dy=0} dx + \frac{\partial f}{\partial y}\Big|_{dx=0} dy \tag{12}$$

which we have in ordinary calculus for 2 variables.

We can rewrite (11) using (1) as

$$\delta S = \int_{t_1}^{t_2} dt \left[\frac{\partial L(\varphi,\dot\varphi)}{\partial \varphi}\bigg|_{\delta\dot\varphi=0} \delta\varphi + \frac{\partial L(\varphi,\dot\varphi)}{\partial \dot\varphi}\bigg|_{\delta\varphi=0} \delta\dot\varphi \right] \qquad (13)$$

and then, reducing the clutter a bit, and using the requirement that we want our path to be a minimum of S, we write

$$\delta S = \int_{t_1}^{t_2} dt \left[\frac{\partial L}{\partial \varphi} \delta\varphi + \frac{\partial L}{\partial \dot\varphi} \delta\dot\varphi \right] = 0 \qquad (14)$$

Just one technical note here. Just as $df/dx = 0$ is the eqtn for an extremum of $f(x)$, ie., either a maximum or minimum, in the same way (14) only tells us when $S[\varphi,\dot\varphi]$ will be an extremum. However, as we will see, in any realistic case, $S[\varphi,\dot\varphi]$ has no maximum - it can become infinitely large - and although there can be several minima, they will all have the same eqtn. of motion, which is what we are interested.

2 Now let's proceed to the solution of (14). We will see that this gives us a differential eqtn., known as the "Euler-Lagrange eqtn". The derivation is pretty simple - it just requires an integration by parts.

We would like to write (14) entirely in terms of $\delta\varphi(t)$. To do this lets take the 2nd term, and integrate by parts:

$$\int_{t_1}^{t_2} dt \frac{\partial L}{\partial \dot\varphi} \delta\dot\varphi = \left[\frac{\partial L}{\partial \dot\varphi} \delta\varphi \right]\bigg|_{\varphi_1}^{\varphi_2} - \int \frac{d}{dt}\left(\frac{\partial L}{\partial \dot\varphi} \right) \delta\varphi \qquad (15)$$

where in the first term, we have $\int dt\, \delta\dot\varphi = \int dt\, \delta(d\varphi/dt) = \int d\,\delta\varphi$.

We now use the constraint in (10), which tells us that $\delta\varphi = 0$ when $\varphi = \varphi_1$ or $\varphi = \varphi_2$. Thus the 1st term on the right-hand side of (15) is zero, and we can take the 2nd term and substitute it into (14), to get

$$\delta S = \int_{t_1}^{t_2} dt \left[\frac{\partial L}{\partial \varphi} - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{\varphi}}\right)\right] \delta\varphi(t) = 0 \qquad (16)$$

Now recall that this is supposed to be true for <u>any</u> variation $\delta\Phi(t)$ around the path of least action. The only way this can be true is if the expression in square brackets is zero. We thus get the Euler-Lagrange eqtn:

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{\varphi}}\right) - \frac{\partial L}{\partial \varphi} = 0 \qquad (17)$$

Note that if $Q \equiv (q_1, q_2, .. q_N)$, then this implies, from (17), that

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_j}\right) - \frac{\partial L}{\partial q_j} = 0 \qquad (\forall_j = 1,2, .. N) \qquad (18)$$

ie., we have a set of N different differential eqtns for N degrees of freedom.

It will take a while for us to see the great advantages that this kind of approach will give us. However our immediate problem is – what is L? This is a physical question, which is discussed in the document on the form of simple Lagrangians.

Finally, let's return to the distinction between extrema and minima. It is perfectly possible for us to have more than one solution to the eqtns of motion. This typically happens when the system has to "choose" between taking 2 or more different paths, and the initial condition allows more than one of them (eg., if we have a rolling ball at the top of a hill, it can begin rolling in any direction). However in almost any realistic situation, there is never exact symmetry between different paths – in the case of the ball at the top of the hill, it will always be slightly displaced from the exact top – and so which one is chosen will usually be clear.