# The Action Functional

September 9, 2014

## 1 Functionals

Informed discussion of Lagrangian methods is helped by introducing the idea of a *functional*. To understand it, think of a function, $f(x)$, as a mapping from the reals to the reals,

$$f : R \longrightarrow R$$

that is, given one real number, $x$, the functions hands us another real number, $f(x)$. This generalizes readily to functions of several variables, for example, $f(\mathbf{x})$ is a map from $R^3$ to $R$ while the electric field $\mathbf{E}(\mathbf{x}, t)$ maps $E : R^4 \longrightarrow R^3$.

In integral expressions, we meet a different sort of mapping. Consider

$$F[x(t)] = \int_{t_1}^{t_2} x(t)\, dt$$

where we introduce square brackets, [ ], to indicate that $F$ is a *functional*. Given any function $x(t)$, the integral will give us a definite real number, but now we require the *entire function*, $x(t)$, to compute it. Define $\mathcal{F}$ to be a function space, in this case the set of all integrable functions $x(t)$ on the interval $[t_1, t_2]$. Then $F$ is a mapping from this function space to the reals,

$$F : \mathcal{F} \longrightarrow R$$

Over the course of the twentieth century, functionals have played an increasingly important role. Introduced by P. J. Daniell in 1919, functionals were used by N. Weiner over the next two years to describe Brownian motion. Their real importance to physics emerged with R. Feynman's path integral formulation of quantum mechanics in 1948 based on Dirac's 1933 use of the Weiner integral.

We will be interested in one particular functional, called the *action* or *action functional*, given for the Newtonian mechanics of a single particle by

$$S[\mathbf{x}(t)] = \int_{t_1}^{t_2} L(\mathbf{x}, \dot{\mathbf{x}}, t)\, dt$$

where the *Lagrangian*, $L(\mathbf{x}, \dot{\mathbf{x}}, t)$, is the difference between the kinetic and potential energies,

$$L(\mathbf{x}, \dot{\mathbf{x}}, t) = T - V$$

## 2 Some historical observations

At the time of the development of Lagrangian and Hamiltonian mechanics, and even into the $20^{th}$ century, the idea of a uniquely determined classical path was deeply entrenched in physicists' thinking about motion. The

great deterministic power of the idea underlay the industrial age and explained the motions of planets. It is not surprising that the probabilistic preditions of quantum mechanics were strongly resisted[1] but experiment - the ultimate arbiter - decrees in favor of quantum mechanics.

This strong belief in determinism made it difficult to understand the variation of the path of motion required by the new approaches to classical mechanics. The idea of "varying a path" a little bit away from the classical solution simply seemed unphysical. The notion of a "virtual displacement" dodges the dilemma by insisting that the change in path is virtual, not real.

The situation is vastly different now. Mathematically, the development of functional calculus, including integration and differentiation of functionals, gives a language in which variations of a curve are an integral part. Physically, the path integral formulation of quantum mechanics tells us that one consistent way of understanding quantum mechanics is to think of the quantum system as evolving over all paths simultaneously, with a certain weighting applied to each and the classical path emerging as the expected average. Classical mechanics is then seen to emerge as this distribution of paths becomes sharply peaked around the classical path, and therefore the overwhelmingly most probably result of measurement.

Bearing these observations in mind, we will take the more modern route and ignore such notions as "virtual work". Instead, we seek the extremum of the action functional $S[\mathbf{x}(t)]$. Just as the extrema of a function $f(x)$ are given by the vanishing of its first derivative, $\frac{df}{dx} = 0$, we ask for the fanishing of the first *functional derivative*,

$$\frac{\delta S[x(t)]}{\delta x(t)} = 0$$

Then, just as the most probable value of a function is near where it changes most slowly, i.e., near extrema, the most probable path is the one giving the extremum of the action. In the classical limit, this is the *only* path the system can follow.

# 3 Variation of the action and the functional derivative

For classical mechanics, we do not need the formal definition of the functional derivative, which is given in a separate Note for anyone interested. Instead, we make use of the extremum condition above and use our intuition about derivatives. The derivative of a function is given by

$$\frac{df}{dx} = \lim_{\epsilon \to 0} \frac{f(x+\epsilon) - f(x)}{\epsilon}$$

Notice that the limit removes all but the linear part of the numerator, $f(x+\epsilon) - f(x) = \left(f(x) + \epsilon \frac{df}{dx} + \frac{1}{2}\epsilon^2 \frac{d^2 f}{dx^2} + \cdots\right) - f(x) \Rightarrow \left(f(x) + \epsilon \frac{df}{dx}\right) - f(x)$. The function at $x$ cancels and we are left with the derivative. If the derivative vanishes, we do not need the $dx$ part, but only

$$df = (f(x+dx) - f(x))|_{linear\, order} = 0$$

where we have set $\epsilon = dx$. Applying the same logic to the vanishing functional derivative, we require

$$\delta S[x(t)] \equiv (S[x(t) + \delta x(t)] - S[x(t)])|_{linear\, order} = 0$$

$\delta S$ is called the variation of the action, and $\delta x(t)$ is an arbitrary variation of the path. Thus, if $x(t)$ one path in the $xt$-plane, $x(t) + \delta x(t)$ is another path in the plane that differs slightly from the first. The variation $\delta x$ is required to vanish at the endpoints, $\delta x(t_1) = \delta x(t_2) = 0$ so that the two paths both start and finish in the same place at the same time.

[1] Einstein's frustration is captured in his assertion to Cornelius Lanczos, "... dass [Herrgott] würfelt ... kann ich keinen Augenblick glauben." (I cannot believe for an instant that God plays dice [with the world]). He later abbreviated this in conversations with Niels Bohr, "Gott würfelt nicht...", God does not play dice. Bohr replied that it is not for us to say how God chooses to run the universe. See http://de.wikipedia.org/wiki/Gott_würfelt_nicht

In defining the variation in this way, we avoid certain subtleties arising from places where the paths cross and $\delta x(t) = 0$, and also the formal need to allow $\delta x$ to be completely arbitrary rather than always small. The variation is sufficient for our purpose.

Now consider the actual form of the variation when the action is given by

$$S[\mathbf{x}(t)] = \int_{t_1}^{t_2} L(\mathbf{x}, \dot{\mathbf{x}}, t)\, dt$$

with $L(\mathbf{x}, \dot{\mathbf{x}}, t) = T - V$. For a single particle in a position-dependent potential $V(\mathbf{x})$,

$$S[\mathbf{x}(t)] = \int_{t_1}^{t_2} \left(\frac{1}{2} m\dot{\mathbf{x}}^2 - V(\mathbf{x})\right) dt$$

and setting $\delta\mathbf{x}(t) = \mathbf{h}(t)$, the vanishing variation gives

$$\begin{aligned}
0 &= \delta S[\mathbf{x}(t)] \\
&= (S[\mathbf{x}+\mathbf{h}] - S[\mathbf{x}])|_{linear\ order} \\
&= \int_{t_1}^{t_2} dt \left(\left(\frac{1}{2} m\left(\dot{\mathbf{x}} + \dot{\mathbf{h}}\right)^2 - V(\mathbf{x}+\mathbf{h})\right) - \left(\frac{1}{2} m\dot{\mathbf{x}}^2 - V(\mathbf{x})\right)\right)\Bigg|_{linear\ order} \\
&= \int_{t_1}^{t_2} dt \left(\left(\frac{1}{2} m\left(\dot{\mathbf{x}}^2 + 2\dot{\mathbf{x}}\cdot\dot{\mathbf{h}} + \dot{\mathbf{h}}^2\right) - V(\mathbf{x}+\mathbf{h})\right) - \left(\frac{1}{2} m\dot{\mathbf{x}}^2 - V(\mathbf{x})\right)\right)\Bigg|_{linear\ order}
\end{aligned}$$

Now we drop the small quadratic term, $\dot{\mathbf{h}}^2$, cancel the kinetic energy $\frac{1}{2}m\dot{\mathbf{x}}^2$ along the original path $\mathbf{x}(t)$, and expand the potential in a Taylor series,

$$\begin{aligned}
0 &= \int_{t_1}^{t_2} dt \left(m\dot{\mathbf{x}}\cdot\dot{\mathbf{h}} - V(\mathbf{x}+\mathbf{h}) + V(\mathbf{x})\right)\Big|_{linear\ order} \\
&= \int_{t_1}^{t_2} dt \left(m\dot{\mathbf{x}}\cdot\dot{\mathbf{h}} - \left(V(\mathbf{x}) + \mathbf{h}\cdot\boldsymbol{\nabla}V(\mathbf{x}) + \mathcal{O}\left(\mathbf{h}^2\right) + \ldots\right) + V(\mathbf{x})\right)\Big|_{linear\ order} \\
&= \int_{t_1}^{t_2} dt \left(m\dot{\mathbf{x}}\cdot\dot{\mathbf{h}} - \mathbf{h}\cdot\boldsymbol{\nabla}V(\mathbf{x})\right)
\end{aligned}$$

Our next goal is to rearrange this so that only the arbitrary vector $\mathbf{h}$ appears as a linear factor, and not its derivative. We integrate by parts. Using the product rule to write

$$\frac{d}{dt}(m\dot{\mathbf{x}}\cdot\mathbf{h}) = m\ddot{\mathbf{x}}\cdot\mathbf{h} + m\dot{\mathbf{x}}\cdot\dot{\mathbf{h}}$$

and solving for the term we have, $m\dot{\mathbf{x}}\cdot\dot{\mathbf{h}} = \frac{d}{dt}(m\dot{\mathbf{x}}\cdot\mathbf{h}) - m\ddot{\mathbf{x}}\cdot\mathbf{h}$, the vanishing variation of the action implies

$$\begin{aligned}
0 &= \int_{t_1}^{t_2} dt \left(\frac{d}{dt}(m\dot{\mathbf{x}}\cdot\mathbf{h}) - m\ddot{\mathbf{x}}\cdot\mathbf{h} - \mathbf{h}\cdot\boldsymbol{\nabla}V(\mathbf{x})\right) \\
&= m\dot{\mathbf{x}}(t_2)\cdot\mathbf{h}(t_2) - m\dot{\mathbf{x}}(t_1)\cdot\mathbf{h}(t_1) - \int_{t_1}^{t_2} dt\,(m\ddot{\mathbf{x}} + \boldsymbol{\nabla}V(\mathbf{x}))\cdot\mathbf{h}
\end{aligned}$$

$$= -\int_{t_1}^{t_2} dt\,(m\ddot{\mathbf{x}} + \boldsymbol{\nabla} V(\mathbf{x})) \cdot \mathbf{h}$$

since $\mathbf{h}(t_2) = \mathbf{h}(t_1) = 0$. Finally, since $\mathbf{h}(t)$ is arbitrary in both direction and magnitude, the only way the integral can vanish[2] is if

$$m\ddot{\mathbf{x}} = -\boldsymbol{\nabla} V(\mathbf{x})$$

and this is Newton's second law where the force is derived from the potential $V$.

# 4 The Euler-Lagrange equation

For many particle systems, we may write the action as a sum over all of the particles. However, there are vast simplifications that occur. In a rigid body containing many times Avogadro's number of particles, the rigidity constraint reduces the number of degrees of freedom to just six - three to specify the position of the center of mass, and three more to specify the direction and magnitude of rotation about this center. Moreover, the use of generalized coordinates may give expressions only vaguely reminiscent of the single particle kinetic and potential energies. Therefore, it is useful to take a general approach, supposing the Lagrangian to depend on $N$ generalized coordinates $q_i$, their velocities, $\dot{q}_i$, and time. We take the potential to depend only on the positions, not the velocities or time, so that

$$L(q_i, \dot{q}_i, t) = T(q_i, \dot{q}_i, t) - V(q_i)$$

Despite the generality of this form, we may find the equations of motion.

Carrying out the variation as before, the $i^{th}$ position coordinate may change by an amount $h_i(t)$, which vanishes at $t_1$ and $t_2$. Following the same steps as for the single particle, vanishing variation gives,

$$\begin{aligned}
0 &= \delta S[q_1, q_2, \ldots, q_N] \\
&= (S[q_i + h_i] - S[q_i])|_{linear\,order} \\
&= \int_{t_1}^{t_2} dt \left( T\left(q_i + h_i, \dot{q}_i + \dot{h}_i, t\right) - V(q_i + h_i) - (T(q_i, \dot{q}_i, t) - V(q_i)) \right)\Big|_{linear\,order} \\
&= \int_{t_1}^{t_2} dt \left( \left( T(q_i, \dot{q}_i, t) + \sum_{i=1}^{N} h_i \frac{\partial T}{\partial q_i} + \sum_{i=1}^{N} \dot{h}_i \frac{\partial T}{\partial \dot{q}_i} - V(q_i) - \sum_{i=1}^{N} h_i \frac{\partial V}{\partial q_i} \right) - (T(q_i, \dot{q}_i, t) - V(q_i)) \right)\Big|_{linear\,order} \\
&= \sum_{i=1}^{N} \int_{t_1}^{t_2} dt \left( h_i \frac{\partial T}{\partial q_i} + \dot{h}_i \frac{\partial T}{\partial \dot{q}_i} - h_i \frac{\partial V}{\partial q_i} \right)
\end{aligned}$$

where the Taylor series to first order of a function of more than one variable contains the linear term for each, $f(x + \epsilon, y + \delta) = f(x, y) + \epsilon \frac{\partial f}{\partial x} + \delta \frac{\partial f}{\partial y} + higher\,order\,terms$. The center term contains the change in velocities, so we integrate by parts,

$$\begin{aligned}
\sum_{i=1}^{N} \int_{t_1}^{t_2} dt\, \dot{h}_i \frac{\partial T}{\partial \dot{q}_i} &= \sum_{i=1}^{N} \int_{t_1}^{t_2} dt \left( \frac{d}{dt} \left( h_i \frac{\partial T}{\partial \dot{q}_i} \right) - h_i \frac{d}{dt} \left( \frac{\partial T}{\partial \dot{q}_i} \right) \right) \\
&= \sum_{i=1}^{N} \left( h_i(t_2) \frac{\partial T}{\partial \dot{q}_i}(t_2) - h_i(t_1) \frac{\partial T}{\partial \dot{q}_i}(t_1) \right) - \sum_{i=1}^{N} \int_{t_1}^{t_2} dt\, h_i \frac{d}{dt} \left( \frac{\partial T}{\partial \dot{q}_i} \right)
\end{aligned}$$

[2]This is easy to see. Suppose there is a point of the path $\mathbf{x}(t')$ where $m\ddot{\mathbf{x}} + \boldsymbol{\nabla} V(\mathbf{x})$ is nonzero. Then choose $\mathbf{h}$ parallel to this direction and nonvanishing only in an infinitesimal region about that point. Then the integral is approximately $|m\ddot{\mathbf{x}}(t') + \boldsymbol{\nabla} V(\mathbf{x}(t'))|\, h(t') > 0$, a contradiction.

$$= \quad -\sum_{i=1}^{N} \int_{t_1}^{t_2} dt\, h_i \frac{d}{dt}\left(\frac{\partial T}{\partial \dot{q}_i}\right)$$

The full expression is now,

$$\begin{aligned} 0 &= \sum_{i=1}^{N} \int_{t_1}^{t_2} dt\, h_i \left(\frac{\partial T}{\partial q_i} - \frac{d}{dt}\left(\frac{\partial T}{\partial \dot{q}_i}\right) - \frac{\partial V}{\partial q_i}\right) \\ &= \sum_{i=1}^{N} \int_{t_1}^{t_2} dt\, h_i \left(\frac{\partial L}{\partial q_i} - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_i}\right)\right) \end{aligned}$$

where $L = T - V$ and we use $\frac{\partial V}{\partial \dot{q}_i} = 0$ to replace $T$ with $L$ in the velocity derivative term. Now, since each $h_i$ is independent of the rest and arbitrary, each term in the sum must vanish separately. The result is the Euler-Lagrange equation,

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_i}\right) - \frac{\partial L}{\partial q_i} = 0$$