Physics 6010, Fall 2010

The Action, the Lagrangian and Hamilton's principle

Relevant Sections in Text: §1.4–1.6, §2.1–2.3

Variational Principles

A great deal of what we shall do in this course hinges upon the fact that one can describe a wide variety of dynamical systems using *variational principles*. You have probably already seen the Lagrangian and Hamiltonian ways of formulating mechanics; these stem from variational principles. We shall make a great deal of fuss about these variational formulations in this course. I will not try to completely motivate the use of these formalisms at this early point in the course since the motives are so many and so varied; you will see the utility of the formalism as we proceed. Still, it is worth commenting a little here on why these principles arise and are so useful.

The appearance of variational principles in classical mechanics can in fact be traced back to basic properties of *quantum* mechanics. This is most easily seen using Feynman's *path integral formalism*. For simplicity, let us just think about the dynamics of a point particle in 3-d. Very roughly speaking, in the path integral formalism one sees that the probability amplitude for a particle to move from one place, $\mathbf{r}_1$, to another, $\mathbf{r}_2$, is given by adding up the probability amplitudes for all possible paths connecting these two positions (not just the classically allowed trajectory). The amplitude for a given path $\mathbf{r}(t)$ is of the form $e^{\frac{i}{\hbar}S[\mathbf{r}]}$, where $S[\mathbf{r}]$ is the *action functional* for the trajectory. The action functional assigns a number to each path connecting $\mathbf{r}_1$ to $\mathbf{r}_2$. The specific way in which the action assigns numbers to paths depends upon the physics (degrees of freedom, masses, potentials, *etc.*) of the system being considered. In a classical limit (usually when various parameters characterizing the system are in some sense "macroscopic") it can be shown that the dominant paths in the sum over paths come from critical (or "stationary") points of the action functional. These are paths which have the property that "nearby" paths do not change $S[\mathbf{r}]$ appreciably, and this is the essence of a variational principle. The critical points of the action are the classically allowed paths; we see that the derivation of classical equations of motion from variational principles is preordained by quantum mechanics. This is a satisfying state of affairs given the fact that classical mechanics can be viewed as a macroscopic approximation to quantum mechanics.

Of course, the variational principles of mechanics (19th century) came much earlier than quantum mechanics (1920's), let alone Feynman's path integral approach (1940's). This is a testament to the great minds (Euler, Lagrange, Hamilton, Jacobi, ...) that found these variational principles! These principles came into favor because they provide a very powerful way to organize information about a dynamical system. In particular,

using a single quantity (the Lagrangian or the Hamiltonian) one can deduce (in principle) essentially all aspects of a dynamical system, *e.g.,* equations of motion, symmetries, conservation laws, ..., even the basic strategy for building the associated quantum system. In fact, modern approaches to modeling dynamical systems take the variational principle as fundamental: we begin by building the Lagrangian or Hamiltonian for the system. As mentioned before, one can think of the discovery of the variational principles of mechanics as really a discovery of a footprint left on the macroscopic world of the quantum world.

Hopefully this is enough motivation to get us started, we shall see the power of the variational principles of mechanics throughout this course.

A Simple Example: a Newtonian particle in one dimension

Before getting into the generalities, let us get a feel for what is going on with a simple example. Let us consider a particle (or some other system with a one-dimensional configuration space) moving in one dimension under the influence of a force. We parametrize the configuration space with $x \in R^1$, and the force is $\vec{f} = f(x,t)\hat{i}$. The equation of motion *a la* Newton is then

$$m\ddot{x}(t) = f(x(t),t).$$

We note that all one-dimensional forces admit a potential energy function $V(x,t)$ such that

$$f(x,t) = -\frac{\partial V(x,t)}{\partial x}.$$

(As an exercise you should prove this!) So the dynamical law can be written as

$$m\ddot{x} + \frac{\partial V}{\partial x} = 0.$$

We can derive this dynamical law from a variational principle as follows. We begin by considering paths $x(t)$, which range between fixed initial and final points, x_1 at $t = t_1$ and x_2 at $t = t_2$, that is,

$$x(t_1) = x_1, \quad x(t_2) = x_2.$$

For example,

$$x(t) = \frac{t - t_2}{t_1 - t_2}x_1 + \frac{t - t_1}{t_2 - t_1}x_2,$$

and

$$x(t) = x_1 \cos\left(\frac{\pi}{2}\frac{t - t_1}{t_2 - t_1}\right) + x_2 \sin\left(\frac{\pi}{2}\frac{t - t_1}{t_2 - t_1}\right),$$

are such paths. There are, of course, infinitely many paths connecting any given endpoints. Note that these paths will not, in general, satisfy Newton's second law for the given force.

Next we define a *functional* $S[x]$ on the set of paths described above. $S[x]$ is called the *action* functional; the action is a rule that associates a number to each path satisfying the given boundary conditions. For the Newtonian particle of mass m moving in potential V we define $S[x]$ by

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

You will recognize the integrand, called the *Lagrangian*,* as the difference of kinetic ($T(t)$) and potential ($V(t)$) energies along the curve $x(t)$,

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

$$L(t) = T(t) - V(t).$$

Given a curve, $x(t)$, it is easy to see how to compute the number assigned by the action functional to $x(t)$ from the formula above: just compute $L(t)$ for that curve and integrate.

As an example, suppose $V(x, t) = mgx$, *i.e.,* we have a particle moving in a uniform gravitational field. Let us evaluate the action for the path

$$x(t) = \frac{t - t_2}{t_1 - t_2} x_1 + \frac{t - t_1}{t_2 - t_1} x_2.$$

We get (exercise)

$$S = \frac{m}{t_2 - t_1} \left\{ \frac{1}{2}(x_2 - x_1)^2 - \frac{1}{2} g(t_2 - t_1)^2 (x_2 + x_1) \right\} .$$

We now consider the problem of finding *critical points* of the action functional $S[x]$. Recall from elementary calculus that the critical points x_0 of a function $f(x)$ are points where the derivative of f vanishes: $f'(x_0) = 0$. What this means is that a small displacement of x_0 does not change the value of the function to first-order in the displacement. To see this, just write out the Taylor series (exercise). So, if x_0 is a critical point for the function f we have

$$f(x_0 + \epsilon) = f(x_0) + \text{terms of order } \epsilon^2.$$

Likewise we say that *a curve $x(t)$ is a critical point of $S[x]$ if a small change in the function does not alter the value of $S[x]$ to first order in the change in the function.* So, if a curve $x(t)$ is a critical point of the action, then if we change the curve, say, to $x(t) + \delta x(t)$, where $\delta x(t)$ is an arbitrary function (except for boundary conditions – see below) the action should be unchanged to first order in δx.

* As we shall see, strictly speaking the Lagrangian should be viewed as a function on the extended velocity phase space. The integrand of the action integral is actually the Lagrangian evaluated on a curve.

Note: The function $\delta x(t)$ is called the *variation* of $x(t)$.

Recall we are considering paths that begin and end at some fixed points (x_1 and x_2). And since $x(t)$ is already assumed to have those endpoints, in order for the varied path $x(t) + \delta x(t)$ to have the correct boundary conditions the variation must satisfy

$$\delta x(t_1) = \delta x(t_2) = 0.$$

Let us now compute the change in action to first order in the variation. Computing to first order in the variation we get (exercise)

$$S[x + \delta x(t)] = S[x] + \int_{t_1}^{t_2} dt \, \left(m\dot{x}(t)\delta\dot{x}(t) - \frac{\partial V(x,t)}{\partial x}\Big|_{x=x(t)} \delta x(t) \right) + \mathcal{O}(\delta x^2).$$

The strategy is now to see what conditions the curve $x(t)$ must satisfy so that the $\mathcal{O}(\delta x)$ term vanishes for any choice of $\delta x(t)$.† To this end, we integrate by parts in the first term of that integral; the endpoint terms do not contribute because δx vanishes at the endpoints:

$$\int_{t_1}^{t_2} dt\, m\dot{x}(t)\delta\dot{x}(t) = m\dot{x}(t_2)\delta x(t_2) - m\dot{x}(t_1)\delta x(t_1) - \int_{t_1}^{t_2} dt\, m\ddot{x}(t)\delta x(t) = -\int_{t_1}^{t_2} dt\, m\ddot{x}(t)\delta x(t)$$

So we get for our critical point condition (good exercise)

$$-\int_{t_1}^{t_2} \left(m\ddot{x}(t) + \frac{\partial V(x,t)}{\partial x}\Big|_{x=x(t)} \right) \delta x(t)\, dt = 0$$

Since this must hold for any function $\delta x(t)$ in the interval $t_1 < t < t_2$ (subject to its vanishing at the endpoints), it follows that the critical point $x(t)$ must satisfy Newton's second law:

$$m\ddot{x} + \frac{\partial V(x,t)}{\partial x}\Big|_{x=x(t)} = 0, \quad t_1 < t < t_2.$$

This can be made quite rigorous given appropriate statements about the smoothness of the functions being used. The idea of the proof is that we can choose $\delta x(t)$ to be arbitrarily well localized about any point t in the interval $t_1 < t < t_2$, and this forces the rest of the integrand to vanish in an arbitrarily small neighborhood of that point. Continuity does the rest.

To summarize: Newton's second law (for a particle moving in 1-d) can be viewed as arising from a *variational principle:** Physical trajectories $x(t)$ (obeying the second law) are critical points of the functional $S[x] = \int dt\, L$, where $L = T - V$.

† The order δx term is called the *first variation* of the action and is usually denoted by δS. The critical point condition is thus expressed as $\delta S = 0$.

* The term "variational principle" arises because we consider the change in the functional as we *vary* the possible paths in the vicinity of a critical point.

It is often asserted that the action is minimized by a curve satisfying the equations of motion, but this is by no means necessary. As in ordinary calculus, the existence of a critical point signals the existence of either a local maximum/minimum or a saddle point. We can investigate this a bit further and show that if the time interval $T = t_2 - t_1$ is sufficiently short the action analyzed above is minimized, however. Let's briefly see how this goes.

For later simplicity, we set $t_1 = 0$ and $t_2 = T$. We can decide on the nature of the critical point by expanding the action to *second* order in the variations. Granted that $x(t)$ is a critical point, we have

$$S[x + \delta x] = S[x] + \delta^2 S + \mathcal{O}(\delta x^3),$$

where $\delta^2 S$ is called the *second variation* of the action about the critical point. A simple computation shows that, for the 1-d Newtonian system we have that (exercise)

$$\delta^2 S = \int_0^T dt\, \frac{1}{2}\left[m(\delta\dot{x})^2 - f(t)\delta x^2\right],$$

where

$$f(t) = \frac{\partial^2 V(x,t)}{\partial x^2}\Big|_{x=x(t)}.$$

Our goal is to see if the second variation is positive, negative, or zero — corresponding to $x(t)$ being a local minimum, maximum, saddle point, respectively. To this end we assume that $f(t)$ is a continuous function of t; we then have the simple estimate

$$-\int_0^T dt\, \frac{1}{2} f(t)\delta x^2 \geq -C\int_0^T dt\, \frac{1}{2}\delta x^2,$$

where the constant C is given by

$$C = \sup_t(f(t)).$$

Thus we have

$$\delta^2 S \geq \int_0^T dt\, \frac{1}{2}\left[m(\delta\dot{x})^2 - C\delta x^2\right],$$

I think you can see that the kinetic energy term provides a positive contribution, while the potential energy term provides a negative contribution. Thus, in general, one cannot assert that a minimum occurs. Still, we can say a bit more. Recall that δx is a function on the interval $[0, T]$ which vanishes at the end points. We can express it as a Fourier series:

$$\delta x = \sum_{n=1}^{\infty} a_n \sin(\frac{n\pi t}{T}).$$

This gives (exercise)

$$\delta^2 S \;\geq\; \frac{T}{4}\sum_{n=1}^{\infty}\Big[m(\frac{n\pi}{T})^2 - C\Big]\, a_n^2.$$

For a given potential energy function C is a fixed constant. You can see from the above expression that, given C, we can always pick T small enough such that the first term in square brackets dominates the second. Thus for T sufficiently small we have that the second variation is positive and $x(t)$ defines a local minimum of the action functional.

Hamilton's Principle

The variational principle used to obtain Newton's second law for a particle moving in one dimension is known as *Hamilton's principle*. We now give a general version.

To a physical system described by generalized coordinates q^i, $i = 1, 2, \ldots, n$ we associate a *Lagrangian*, which is a function of $2n + 1$ variables:

$$L = L(q^i, \dot{q}^i, t).$$

We have *not* evaluated this function on a curve yet! The t dependence indicated is only of the "explicit" type. So, the Lagrangian is, in general, a time dependent function, *i.e.,* a one parameter family of functions, on the velocity phase space. Alternatively, the Lagrangian can be viewed as a function on the extended velocity phase space.* How do we choose the Lagrangian? For a Newtonian system, as we shall see, we take the difference between kinetic and potential energies. In this case the challenge is to find a set of generalized coordinates and to decide what is the correct potential energy function. More generally, finding the correct Lagrangian is tantamount to finding the physically correct description of the system. So, determining the Lagrangian is one of the essential arts of being a physicist. We will, of course, explore a lot of standard Lagrangians so you can see how to go about building them.

The *action* integral assigns a number to each curve $q^i = q^i(t)$ joining the fixed endpoints,

$$q^i(t_1) = q_1^i, \quad q^i(t_2) = q_2^i,$$

via

$$S[q] = \int_{t_1}^{t_2} L(q^i(t), \dot{q}^i(t), t)\, dt.$$

* Many Lagrangians of physical interest do not in fact have any explicit t dependence, *i.e.,*

$$\frac{\partial L}{\partial t} = 0.$$

For example, consider the Lagrangian for any conservative Newtonian dynamical system.

Here we have evaluated the Lagrangian on the curve $q^i(t)$, so that the integrand has time dependence of the explicit and implicit sort. We have chosen the Lagrangian to depend only upon the curve and its tangent vector because, as we shall see, this leads to second-order equations of motion, which are physically most relevant. Higher order equations of motion can be accommodated by letting the Lagrangian depend upon higher order derivatives of $q^i(t)$. You will explore this in a homework problem

Hamilton's principle asserts that physical trajectories are those curves between q_1 and q_2 which are critical points of the action integral. As in our simple example, we can derive a differential equation that these trajectories must satisfy. We do this as follows.

We consider a *variation* in the putative critical curve:

$$q^i(t) \longrightarrow q^i(t) + \delta q^i(t),$$

where $\delta q^i(t)$ is arbitrary except for the endpoint conditions

$$\delta q^i(t_1) = 0 = \delta q^i(t_2).$$

We next compute the first order change, δS, in the action,

$$\delta S[q] := S[q + \delta q] - S[q] \quad \text{to first order in } \delta q.$$

Using multi-variable Taylor series, we get

$$L(q + \delta q, \dot{q} + \delta\dot{q}, t) = L(q, \dot{q}, t) + \frac{\partial L(q, \dot{q}, t)}{\partial q^i}\delta q^i + \frac{\partial L(q, \dot{q}, t)}{\partial \dot{q}^i}\delta\dot{q}^i + \mathcal{O}(\delta q^2).$$

Here we've introduced the *Einstein summation convention* wherein a subscript and superscript with the same index label are to be summed, *e.g.,*

$$\frac{\partial L}{\partial q^i}\delta q^i \equiv \sum_{i=1}^{n} \frac{\partial L}{\partial q^i}\delta q^i.$$

Thus the first order change in the action is given by

$$\delta S[q] = \int_{t_1}^{t_2} \left(\frac{\partial L}{\partial q^i}\delta q^i + \frac{\partial L}{\partial \dot{q}^i}\delta\dot{q}^i \right) dt.$$

There are a number of remarks we must make about this formula.

* In the formula for δS, the partial derivatives of the Lagrangian are taken while viewing L as a function on the (extended) velocity phase space. The partial derivatives are thus functions of the $2n + 1$ variables q^i, $\dot{q}^i$, and t. These functions are evaluated on the curve, *i.e.,* in the integral one substitutes

$$q^i = q^i(t), \quad \dot{q}^i = \frac{d}{dt}q^i(t).$$

For example (in the formula for the variation)

$$\frac{\partial L}{\partial q^i}\delta q^i \equiv \left.\frac{\partial L}{\partial q^i}\right|_{\substack{q=q(t)\\ \dot{q}=\frac{dq(t)}{dt}}} \delta q^i(t),$$

with the same meaning for the second term in the formula. This makes the integrand a function of t which can then be integrated. Henceforth we follow the standard practice of omitting the notation which makes explicit where the functions of time are, *i.e.,* we are dropping all the "(t)" which we kept in our earlier example.

* Along similar lines, the variation $\delta\dot{q}^i$ means either the change induced in the tangent vector to $q^i(t)$ by the variation in the curve, or it means the time derivative of $\delta q^i(t)$. These are the same thing (exercise):

$$\delta\dot{q}^i(t) = \delta\left(\frac{d}{dt}q^i(t)\right) = \frac{d}{dt}\delta q^i(t).$$

* We call $\delta q^i(t)$ the *variation* of the path (curve, trajectory, motion, *etc.*) of the system in configuration space. We call δS the *first variation in the action* induced by the variation in $q^i(t)$.

* One can view δ as an operation – called "the variation" – which can be applied to functionals $A[q]$. This operation can be defined by the procedure of expanding in the variations as we did above. An equivalent, more elegant approach is as follows. Consider a 1-parameter family of curves $q^i(\alpha, t)$ containing the curve of interest $q^i(t)$, where $q^i(0,t) = q^i(t)$. Define

$$\delta q^i(t) = \frac{\partial}{\partial\alpha}q^i(\alpha,t)\Big|_{\alpha=0}.$$

and

$$\delta A[q] = \frac{d}{d\alpha}A[q(\alpha)]\Big|_{\alpha=0}.$$

It is not hard to see that applying δ to A is the same as finding the change in A obtained by evaluating A on $q+\delta q$ and expanding to first-order in δq. For example, with an action integral

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

we have

$$\begin{aligned}
\delta S &= \frac{d}{d\alpha}\left[\int_{t_1}^{t_2} dt\, L(q(\alpha,t),\dot{q}(\alpha,t),t)\right]_{\alpha=0} \\
&= \left[\int_{t_1}^{t_2}\left(\frac{\partial L(q(\alpha,t),\dot{q}(\alpha,t),t)}{\partial q^i}\frac{\partial q^i(\alpha,t)}{\partial\alpha} + \frac{\partial L(q(\alpha,t),\dot{q}(\alpha,t),t)}{\partial\dot{q}^i}\frac{\partial\dot{q}^i(\alpha,t)}{\partial\alpha}\right) dt\right]_{\alpha=0} \\
&= \int_{t_1}^{t_2}\left(\frac{\partial L}{\partial q^i}\delta q^i + \frac{\partial L}{\partial\dot{q}^i}\delta\dot{q}^i\right) dt.
\end{aligned}$$

Let us return to our analysis of the variation of the action and the condition for a critical point. As in our initial example, we now integrate by parts in the second term in our formula for δS to get (exercise)

$$\delta S = \frac{\partial L}{\partial \dot{q}^i} \delta q^i \Big|_{t_1}^{t_2} + \int_{t_1}^{t_2} \left[\frac{\partial L}{\partial q^i} - \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}^i} \right) \right] \delta q^i \, dt.$$

The endpoint terms do not contribute because the variations in the curve vanish at the endpoint. So we get

$$\delta S = \int_{t_1}^{t_2} \left[\frac{\partial L}{\partial q^i} - \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}^i} \right) \right] \delta q^i \, dt.$$

Hamilton's principle can be expressed by the demand that the physical trajectory $q^i(t)$ is such that, for all $\delta q^i(t)$,

$$\delta S[q] = 0.$$

This implies, as in our example,

$$\frac{\partial L}{\partial q^i} - \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}^i} \right) = 0.$$

These are the famous *Euler-Lagrange equations* (EL equations) for the physical trajectories.

Remarks:

* It is understood in the EL equations that the partial derivatives of L are evaluated on the trajectory $q^i(t)$. Thus the EL equations are a system of n second-order differential equations for the curve $q^i(t)$, not, *e.g.*, equations for L as might appear from the way the equation is written. We take the point of view that L is given and the curve is to be found from the differential equations. The equations are second order and are of the form

$$\frac{\partial L}{\partial q^i} - \frac{\partial^2 L}{\partial \dot{q}^i \partial t} - \frac{\partial^2 L}{\partial \dot{q}^i \partial q^j} \dot{q}^j - \frac{\partial^2 L}{\partial \dot{q}^i \partial \dot{q}^j} \ddot{q}^j = 0.$$

  The first three terms in this equation involve up to first derivatives of the curve. The last term explicitly displays the second derivatives, with coefficients which depend on up to first derivatives.

* We can check that when $q^i \to x$ and

$$L = \frac{1}{2} m \dot{x}^2 - V(x,t)$$

  that the EL equations reproduce the equations of motion obtained from Newton's second law.

We have

$$\frac{\partial L}{\partial x} = -\frac{\partial V}{\partial x}, \quad \frac{\partial L}{\partial \dot{x}} = m\dot{x}.$$

Hence

$$\frac{\partial L}{\partial x} - \frac{d}{dt}\frac{\partial L}{\partial \dot{x}} = -\frac{\partial V}{\partial x} - m\ddot{x}$$

as desired.

In 3 dimensions, with $q^i \to \mathbf{r}$ and

$$L = \frac{1}{2}m\dot{\mathbf{r}}^2 - V(\mathbf{r}, t)$$

the EL equations give the usual equations of motion.

$$m\ddot{\mathbf{r}} + \nabla V = 0.$$

To see this, just note that

$$\frac{\partial L}{\partial x} = -\frac{\partial V}{\partial x}, \quad \frac{\partial L}{\partial y} = -\frac{\partial V}{\partial y}, \quad \frac{\partial L}{\partial z} = -\frac{\partial V}{\partial z},$$

and

$$\frac{\partial L}{\partial \dot{x}} = m\dot{x}, \quad \frac{\partial L}{\partial \dot{y}} = m\dot{y}, \quad \frac{\partial L}{\partial \dot{z}} = m\dot{z},$$

and you can easily see that the three EL equations (one for each degree of freedom) give Newton's second law.

The total time derivative

The time derivative $\frac{d}{dt}$ appearing in the EL equations is sometimes called the *total time derivative* since it sees the *implicit* and *explicit* time dependence. More precisely, we have (exercise)

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{q}^i} = \frac{\partial^2 L}{\partial q^j \partial \dot{q}^i}\dot{q}^j + \frac{\partial^2 L}{\partial \dot{q}^j \partial \dot{q}^i}\ddot{q}^j + \frac{\partial^2 L}{\partial \dot{q}^i \partial t}.$$

Thus the EL equations are a system of second-order ordinary differential equations. (It is not hard to handle Lagrangians which lead to higher-order differential equations. You will explore this in your homework.)

This total time derivative bears some additional emphasis, as it pops up fairly often. Given a function $G(q, \dot{q}, t)$ on the extended velocity phase space, we define the *total time derivative* of G to be the function of coordinates, velocities, *and accelerations*, $\ddot{q}$ given by

$$\frac{d}{dt}G(q, \dot{q}, t) = \frac{\partial G}{\partial q^i}\dot{q}^i + \frac{\partial G}{\partial \dot{q}^i}\ddot{q}^i + \frac{\partial G}{\partial t}.$$

For example, for a system with one degree of freedom q, suppose

$$G(q, \dot{q}, t) = \sin q\, \dot{q}^3 + te^q.$$

We have

$$\frac{\partial G}{\partial q} = \cos q\, \dot{q}^3 + te^q, \quad \frac{\partial G}{\partial \dot{q}} = 3 \sin q\, \dot{q}^2, \quad \frac{\partial G}{\partial t} = e^q,$$

and so

$$\frac{dG}{dt} = 3 \sin q\, \dot{q}^2 \ddot{q} + \cos q\, \dot{q}^4 + e^q + te^q \dot{q}.$$

On a given trajectory in configuration space, we have

$$q^i = q^i(t), \quad \dot{q}^i = \frac{d}{dt} q^i(t), \quad \ddot{q}^i = \frac{d^2}{dt^2} q^i(t),$$

etc. If we evaluate dG/dt, as *defined* above, on the curve* we get (exercise)

$$\left\{ \frac{d}{dt} G(q, \dot{q}, \ldots, t) \right\}_{q=q(t)} = \left(\frac{\partial G}{\partial q^i} \right)_{q=q(t)} \frac{d}{dt} q^i(t) + \left(\frac{\partial G}{\partial \dot{q}^i} \right)_{q=q(t)} \frac{d^2}{dt^2} q^i(t) + \ldots + \left(\frac{\partial G}{\partial t} \right)_{q=q(t)}.$$

This is, of course, just the ordinary time derivative of the function of t obtained by evaluating G on the curve, but computed using the chain rule. Thus the total time derivative, viewed as an operator on $G(q, \dot{q}, \ldots, t)$, is designed to satisfy

$$\left(\frac{dG}{dt} \right) \Big|_{q=q(t)} = \frac{d}{dt} \left(G \Big|_{q=q(t)} \right).$$

As a nice exercise you can check in our little example above that this is indeed what happens.

The idea behind the total time derivative is that $G(q, \dot{q}, \ldots, t)$ is some physical quantity describing the system (e.g., the force, or the total energy). This quantity may change in time by its definition – this is the explicit time dependence. The total time derivative gives a formula for the rate of change of G in terms of positions, velocities, etc. valid *as we move along a given curve.* As we mentioned before, the change in G then comes from two sources (i) the fact that G changes in time by its definition ("explicit time dependence") and the fact that G will *also* change as we move along a curve because the position, velocity, *etc.* are changing in time ("implicit time dependence").

* This means substituting

$$q^i = q^i(t), \quad \dot{q}^i = \frac{d}{dt} q^i(t), \quad \ddot{q}^i = \frac{d^2}{dt^2} q^i(t), \quad \text{etc.}$$

which we schmatically denote as "$q = q(t)$" in the following formulas.

Lack of uniqueness of the Lagrangian

Since we only use the Lagrangian to compute the equations of motion via the EL equations, the Lagrangian is not uniquely specified by the equations of motion.* In particular there are a number of changes we can make to a Lagrangian without altering the content of the equations of motion.

There are two ways to modify a Lagrangian which are pretty simple. They stem from the fact that the EL equations are obtained from a formula which is constructed by applying a linear differential operator (in jet space) to the Lagrangian. First, if we add a constant to the Lagrangian the EL equations do not change. This is because the EL equations only involve derivatives of the Lagrangian. Second, it is easy to see that if we multiply the Lagrangian by a constant, $L \to cL$, the Euler-Lagrange equations of motion get multiplied by an overall constant (the EL equations are built linearly from the Lagrangian), which does not disturb the meaning of the equations of motion (exercise).

A less trivial way to change a Lagrangian is as follows. Given a Lagrangian $L(q, \dot{q}, t)$ for each function $f(q, t)$, we can make a new Lagrangian $\tilde{L}(q, \dot{q}, t)$ via

$$\tilde{L}(q, \dot{q}, t) = L(q, \dot{q}, t) + \frac{d}{dt} f(q, t).$$

The EL equations coming from $\tilde{L}$ are identical to those coming from L. You can understand this result from the variational principle point of view – the two action functionals, S and $\tilde{S}$ defined by the Lagrangians L and $\tilde{L}$ have the same critical points. This is because (exercise)

$$\tilde{S} = S + f(q_2, t_2) - f(q_1, t_1),$$

and hence

$$\delta\tilde{S} = \delta S + \delta f(q_2, t_2) - \delta f(q_1, t_1).$$

But

$$\delta f = \frac{\partial f}{\partial q^i} \delta q^i,$$

and the endpoints are fixed,

$$\delta q_1 = \delta q_2 = 0,$$

so that

$$\delta\tilde{S} = \delta S$$

thanks to the endpoint conditions on $\delta q^i(t)$ (exercise). Thus, adding a total time derivative to the Lagrangian will not change the critical points and has no effect on the EL equations

* This is reasonable: the Lagrangian is ultimately a quantum mechanical object – the classical limit does not determine the quantum system uniquely.

obtained via Hamilton's principle. In your homework you will see this another way. You will show that the EL equations, computed directly from $L(q, \dot{q}, t)$ and from

$$L'(q, \dot{q}, t) = L(q, \dot{q}, t) + \frac{df(q)}{dt},$$

are the same. Put differently, the EL expression for the Lagrangian $L = \frac{df}{dt}$ vanishes identically.

Coordinate Invariance of EL Equations

There is an even bigger type of non-uniqueness of the Lagrangian. Any two Lagrangians which are related by a change of coordinates in configuration space will give equivalent EL equations. Another way to view this is to say that the EL equations can be computed in *any* set of generalized coordinates and they are guaranteed to be correct. This is a very useful result, as we shall see.

To see the non-triviality of the statement, let us compare the Lagrangian situation to the Newtonian one. In the Lagrangian description of a Newtonian system, you simply choose your favorite generalized coordinates, express $L = T - V$ in terms of them and compute the EL equations. The Lagrangian can, of course, have a variety of functional forms depending upon your choice of coordinates. But it is always $T - V$, and the equations you get are the correct ones (as you prove in your homework). In the Newtonian description, one determines the force and uses Newton's second law. But Newton's second law is only valid in an inertial reference frame. Mathematically, the transformation to a non-inertial reference frame can be viewed as a change of generalized coordinates. For example, for a reference frame (x^1, x^2, x^3) rotating uniformly about the z-axis we get (exercise).

$$x^1 = x \cos \omega t + y \sin \omega t,$$

$$x^2 = y \cos \omega t - x \sin \omega t,$$

$$x^3 = z.$$

With this change of coordinates you can't really use $\vec{F} = m\vec{a}$ as usual. For example, suppose that the particle is isolated (*i.e.,* does not interact with anything), so that $\vec{F} = 0$. In an inertial reference frame we get the equation of motion for the curve

$$\vec{r}(t) = (x(t), y(t), z(t))$$

given by

$$\frac{d^2\vec{r}(t)}{dt^2} = 0.$$

The equations of the curve in the non-inertial frame, $\vec{x} = (x^1(t), x^2(t), x^3(t))$ are not the same since (exercise)

$$\frac{d^2\vec{x}(t)}{dt^2} \neq 0$$

in this reference frame. So, in general, you can't just make arbitrary coordinate transformations and keep using $\vec{F} = m\vec{a}$.

On the other hand you *can* do it in the Lagrangian. For the free particle Lagrangian we have

$$L = \frac{1}{2}m(\dot{x}^2 + \dot{y}^2 + \dot{z}^2) = \frac{1}{2}m(\dot{x}^{12} + \dot{x}^{22} + \dot{x}^{32}) + \frac{1}{2}m\omega^2(x^{12} + x^{22}) + m\omega(\vec{x} \times \dot{\vec{x}})_3 = \tilde{L}(\vec{x}, \dot{\vec{x}})$$

As a nice exercise you should compute the EL equations for $\vec{x}(t)$ from $\tilde{L}$ and verify that they are equivalent to

$$\ddot{x}^1 - \omega^2 x^1 + 2\omega\dot{x}^2 = 0$$

$$\ddot{x}^2 - \omega^2 x^2 - 2\omega\dot{x}^1 = 0,$$

$$\ddot{x}^3 = 0$$

which is precisely what you'd get from combining (i) the definition of the rotating reference frame with (ii) $\ddot{\vec{r}} = 0$.

Let me be more precise about what it means to change coordinates in the Lagrangian. Suppose you compute the EL equations in one system of coordinates q^i, as we have described. Suppose somebody else computes using a different set of coordinates, say, $\tilde{q}^i$. For example, we could use the inertial and rotating coordinates defined above. The two coordinate systems will be related by equations of the form:

$$\tilde{q}^i = f^i(q, t), \quad q^i = g^i(\tilde{q}, t).$$

The velocities are then related by (exercise)

$$\dot{\tilde{q}}^i = \frac{\partial f^i}{\partial q^j}\dot{q}^j + \frac{\partial f^i}{\partial t},$$

and

$$\dot{q}^i = \frac{\partial g^i}{\partial \tilde{q}^j}\dot{\tilde{q}}^j + \frac{\partial g^i}{\partial t}.$$

The two coordinate systems will yield two different Lagrangians, $L(q, \dot{q}, t)$ and $\tilde{L}(\tilde{q}, \dot{\tilde{q}}, t)$. The Lagrangians will be related via

$$\tilde{L}(\tilde{q}, \dot{\tilde{q}}, t) := L(q(\tilde{q}, t), \dot{q}(\tilde{q}, \dot{\tilde{q}}, t), t).$$

The key point is that if $q^i = q^i(t)$ solves the EL equations

$$\frac{\partial L}{\partial q^i} - \frac{d}{dt}\frac{\partial L}{\partial \dot{q}^i} = 0,$$

then

$$\tilde{q}^i(t) = f^i(q(t), t)$$

will satisfy

$$\frac{\partial \tilde{L}}{\partial \tilde{q}^i} - \frac{d}{dt}\frac{\partial \tilde{L}}{\partial \dot{\tilde{q}}^i} = 0.$$

This coordinate invariance of the EL equations stems from the fact that the curves in configuration space which are critical points of the action are critical points irrespective of how they are mathematically represented. This is a distinct practical advantage of the EL equations; they facilitate the use of non-trivial coordinate systems to compute the equations of motion. We shall study several examples shortly. You will prove this invariance of the equations in the homework.