Physics 303/573

The calculus of variations and the Euler-Lagrange equations

October 7, 2015

1 Motivation

In this lecture, we will develop *variational* techniques that let us reformulate Newton's laws in a different language–as Lagrangian and Hamiltonian mechanics. These reformulations make it possible to find the equations that describe complicated systems with constraints in a very direct way, and ultimately allowed the development of quantum mechanics.

We consider two simple examples of problems that are naturally solved using variational techniques. We know that the shortest distance between two points is a straight line–how can we prove this? We consider an *arbitrary* path between the two points and minimize the length of the path. We know how to minimize functions of a finite number of variables–we take the derivative and set it to zero. However, the length of a path doesn't depend on a finite number of variables–it depends on a whole function!

Suppose that we describe the path by a graph $y(x)$. Then the length ℓ is given by the integral

$$\ell[y] = \int_{x_1}^{x_2} dx \sqrt{1 + \left(\frac{dy}{dx}\right)^2} \tag{1.1}$$

It is a *functional* of the path; given a function $y(x)$, we compute the length $\ell[y]$, that is, a number.

For simplicity, we have written the path as a function $y(x)$; more generally, we can consider more complicated paths that double back or even intersect themselves by writing the path parametrically: $\vec{r}(t) = x(t)\hat{x} + y(t)\hat{y}$. Then the length of the path would be written as

$$\ell[\vec{r}] = \int_{t_1}^{t_2} dt \sqrt{\left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2} \tag{1.2}$$

Before describing how to solve this problem, let us consider another example: Fermat's principle.

Recall that if light leaves a medium with an index of refraction n_1 at an angle θ_1 to the normal to the interface with a medium whose index of refraction is n_2, it will propagate in the second medium at an angle θ_2–see the figure below–where θ_2 is given by Snell's law:

$$n_1 \sin\theta_1 = n_2 \sin\theta_2 \tag{1.3}$$

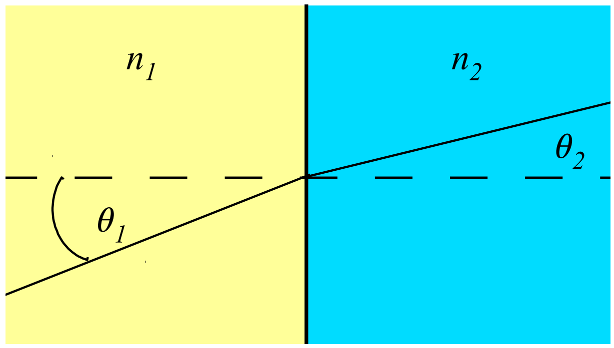

Figure 1: Snell's Law: $n_1 \sin\theta_1 = n_2 \sin\theta_2$

In 1021, the great arabic scientist Alhazen realized that light moved more slowly in dense materials, and interpreted the bending of light as a response to that–in fact, the index of refraction is just the ratio of the speed of light in vacuum to the speed of light in the medium: $n = c/v$. In 1662, Pierre Fermat formulated the general principle, namely that light moves from point 1 to point 2 in such a way as to minimize the time it travels. In Figure 1 we see that instead of moving in a straight line, it follows a path that spends more time in the medium with the larger velocity: $v_1 = c/n_1 > v_2 = c/n_2$. Suppose that the speed of light is a known function of the position $v(x, y)$; then the time it takes for the light is

$$t[y] = \int_{x_1}^{x_2} dx\, \frac{\sqrt{1 + \left(\frac{dy}{dx}\right)^2}}{v(x,y)} = \frac{1}{c}\int_{x_1}^{x_2} dx\, n(x,y)\sqrt{1 + \left(\frac{dy}{dx}\right)^2} \tag{1.4}$$

where we use the local index of refraction $n(x, y) = c/v(x, y)$. In this case, t, the time that the light takes, is a functional of the path $y(x)$.

2 Calculus of variations

In both the examples above, we have a a functional that depends on the path and we want to extremize the functional.

Consider a general functional of the form

$$S[y] = \int_{x_1}^{x_2} dx\, L\left(y(x), \frac{dy}{dx}, x\right) \tag{2.5}$$

To find the extremum, we need to do the equivalent of "setting the derivative to zero". For a function $f(x)$, at an extremum x_0, if we go to a nearby point $x_0 + \delta$, we have

$$\left.\frac{df}{dx}\right|_{x_0} = 0 \qquad \Rightarrow \qquad f(x_0 + \delta) = f(x_0) + O(\delta^2) \tag{2.6}$$

We want to copy this–so we consider two functions $y(x)$ and $y(x) + \delta y(x)$ that both go from point 1 to point 2 and differ by a small but arbitrary deviation $\delta y(x)$:

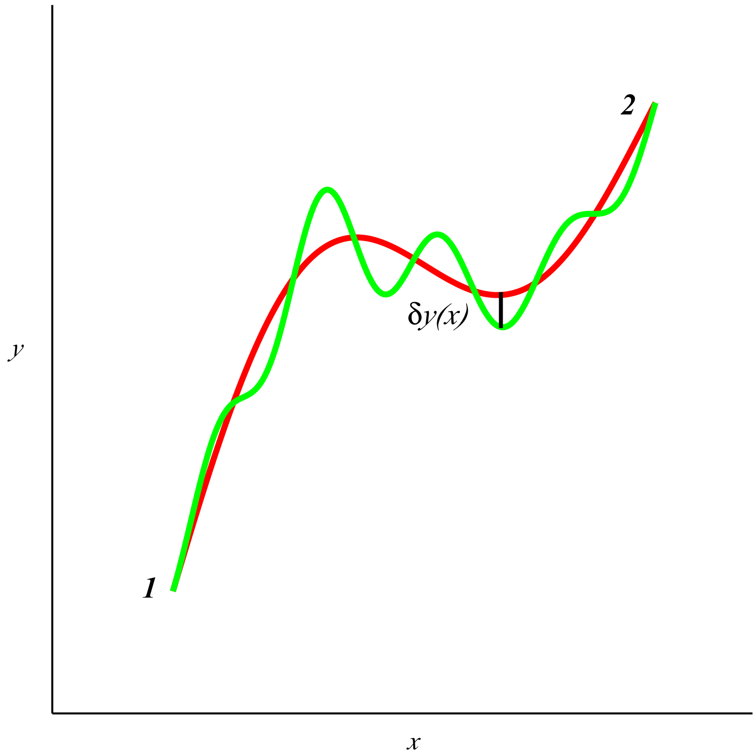

Figure 2: Two paths $y(x)$ and $y(x) + \delta y(x)$ from point 1 to point 2 that differ by a small variation $\delta y(x)$.

Notice that since both paths start at point 1 and end at point 2, the deviation must vanish at the endpoints: $\delta y(x_1) = \delta y(x_2) = 0$. We now evaluate

$$\begin{aligned}\delta S[y] = S[y+\delta y] - S[y] &= \int_{x_1}^{x_2} dx\, L\bigg(y(x)+\delta y(x), \frac{dy(x)}{dx} + \frac{d}{dx}\delta y(x), x\bigg) - L\bigg(y(x), \frac{dy}{dx}, x\bigg) \\ &= \int_{x_1}^{x_2} dx\, \left(\frac{\partial L}{\partial y}\delta y(x) + \frac{\partial L}{\partial\left(\frac{dy}{dx}\right)}\frac{d}{dx}\delta y(x)\right) + O((\delta y)^2) \end{aligned} \tag{2.7}$$

We now integrate the last term by parts; because $\delta y(x_1) = \delta y(x_2) = 0$, we get

$$\delta S[y] = \int_{x_1}^{x_2} dx\, \left(\frac{\partial L}{\partial y} - \frac{d}{dx}\frac{\partial L}{\partial\left(\frac{dy}{dx}\right)}\right)\delta y(x) + O((\delta y)^2) \tag{2.8}$$

We want the variation $\delta S[y]$ to vanish to leading order for any $\delta y(x)$; since we can choose $\delta y(x)$ to vanish almost everywhere with just a narrow peak at any x, the only way we can ensure this is if the whole expression multiplying $\delta y(x)$ vanishes point by point:[1]

$$\frac{\partial L}{\partial y} - \frac{d}{dx}\frac{\partial L}{\partial\left(\frac{dy}{dx}\right)} = 0 \tag{2.9}$$

[1]Strictly speaking, this is true if we assume that our functions are sufficiently well-behaved; for example, if they have a fourier series, then everything works, because the integral $\int fg$ of two functions f and g becomes a scalar product on an infinite dimensional vector space.

This is called the Euler-Lagrange equation. Let us apply it to our two examples. For (1.1), we have $\partial L/\partial y = 0$ and hence

$$\frac{d}{dx}\frac{\frac{dy}{dx}}{\sqrt{1+\left(\frac{dy}{dx}\right)^2}} = 0 \tag{2.10}$$

which we can integrate to find

$$\frac{\frac{dy}{dx}}{\sqrt{1+\left(\frac{dy}{dx}\right)^2}} = a \tag{2.11}$$

where a is an integration constant. This implies

$$\frac{dy}{dx} = \frac{a}{\sqrt{1-a^2}} = A \quad \Rightarrow \quad y = Ax + B \tag{2.12}$$

where A, B are constants. Thus we have shown that the shortest (strictly speaking, extremal) path between two points is a straight line.

For Fermat's principle, we have

$$\left(\frac{\partial}{\partial y}n(x,y)\right)\sqrt{1+\left(\frac{dy}{dx}\right)^2} = \frac{d}{dx}\frac{n(x,y)\frac{dy}{dx}}{\sqrt{1+\left(\frac{dy}{dx}\right)^2}} \tag{2.13}$$

This becomes particularly simple if the index of refraction varies only along the x axis; then $n(x,y) = n(x)$ and the left-hand side of (2.13) vanishes. We integrate as before and find

$$\frac{n(x)\frac{dy}{dx}}{\sqrt{1+\left(\frac{dy}{dx}\right)^2}} = a \quad \Rightarrow \quad \frac{dy}{dx} = \frac{a}{\sqrt{n^2-a^2}} \tag{2.14}$$

But dy/dx is the slope of the path, which is the tangent of the angle. It is easy to check that (2.14) implies that

$$n(x)\sin(\theta(x)) = a \tag{2.15}$$

which is the local form of Snell's law. In the more general case, when $n = n(x,y)$, we have to use the full equation (2.13).

Fermat's principle gives us a nice example of the fact that we are really only looking for extrema and not a maximum or a minimum–consider the example of a lens that focuses light emitted from point 1 onto point 2; then there are many paths that all take the same time and are extremal:

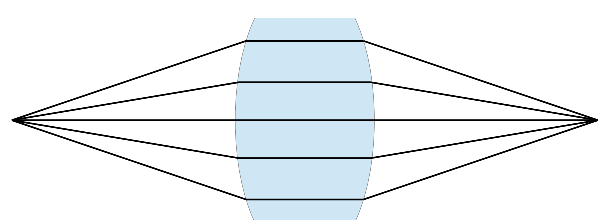

Figure 3: When a lens focuses, many paths extremize the time that the light travels.

3 General Euler-Lagrange Equations

I mentioned above (see 1.2) that in general we are interested in paths that are described parametrically. But we have seen many times already that once we allow more than one coordinate in this case x and y–we might as well allow an arbitrary finite number. So consider the variational problem for a function $L(\{q_i\}, \{\dot{q}_i\}, t)$ where t is a parameter, and the paths that we are considering are described by n coordinates $\{q_i\} = \{q_1(t), q_2(t), ..., q_{n-1}(t), q_n(t)\}$ and their derivatives with respect to t: $\{\dot{q}_i\} = \{\dot{q}_1(t), \dot{q}_2(t), ..., \dot{q}_{n-1}(t), \dot{q}_n(t)\}$. Then

$$S[\{q_i\}] = \int_{t_1}^{t_2} dt\, L(\{q_i\}, \{\dot{q}_i\}, t) \quad , \quad \dot{q}_i \equiv \frac{d}{dt} q_i \tag{3.16}$$

Then introducing independent variations $\{\delta q_i(t)\}$ that vanish at the endpoints t_1, t_2, we find the extremal conditions:

$$\frac{\partial}{\partial q_i} L - \frac{d}{dt} \frac{\partial}{\partial \dot{q}_i} L = 0 \tag{3.17}$$

This is a set of n differential equations for the n variables $q_i(t)$.