6 The Calculus of Variations

The techniques we are about to discuss are of central importance in the theory of dynamics, and more generally, for finding the equations of motion for many physical systems.

A classic example is given by the Brachistochrone problem of John Bernoulli. Consider a particle, displaced by a horizontal distance a, and a vertical distance, from a particular point, A. We assume the particle begins at the origin, O. The problem is then: along which path from O to A will a particle slide (from rest at O), under gravity, with no friction, *in the least time.*

Begin by writing the energy of the particle at any time, as satisfying

$$\frac{1}{2}v^2 = gy \ , \tag{1}$$

(where I've set the mass to be unity). We may then write the time taken for a particular motion as

$$\begin{aligned} T &= \int_0^T dt \ , \\ &= \int \frac{ds}{v} \ , \\ &= \int_0^a \frac{(1+y'^2)^{1/2}}{(2gy)^{1/2}}\,dx \ . \end{aligned} \tag{2}$$

T is a *functional* of the path $y(x)$ in question. We write $T[y(x)]$. Given $y(x)$, we can calculate $T[y]$. The question is, which y minimizes T?

6.1 The Euler-Lagrange Equations

To answer this, let us consider a more general problem. Given a functional $F(x, y, y')$, where $y = y(x)$, form the functional

$$I[y] = \int_a^b dx\, F(x, y, y') \ . \tag{3}$$

We seek to find stationary values of I between fixed end points, such that $y(a)$, and $y(b)$ are given. Our aim is to choose y so that $I[y]$ is stationary; i.e.

$$\delta y = 0 \ . \tag{4}$$

More explicitly, we will require $\delta y = \varepsilon\eta(x)$, for ε small and $\eta(x)$ arbitrary except that $\eta(a) = \eta(b) = 0$.

We study

$$\delta I = \varepsilon \left.\frac{\partial I}{\partial \varepsilon}\right|_{\varepsilon \to 0} = 0 \ . \tag{5}$$

Proceed in the following way. Write

$$I(\varepsilon) = \int_a^b dx\, F(x, y + \varepsilon\eta, y' + \varepsilon\eta') \ . \tag{6}$$

Then

$$\begin{aligned} \frac{\partial I}{\partial \varepsilon} &= \int_a^b dx \left(\frac{\partial F}{\partial y}\eta + \frac{\partial F}{\partial y'}\eta' \right) \ , \\ &= \int_a^b dx\, \eta \left(\frac{\partial F}{\partial y} - \frac{d}{dx}\frac{\partial F}{\partial y'} \right) + \eta \left.\frac{\partial F}{\partial y'}\right|_a^b \ , \end{aligned} \tag{7}$$

by an integration by parts. Now, the last term is zero, since $\eta(a) = \eta(b) = 0$. Therefore, $\partial I/\partial\varepsilon = 0$ (evaluated at $\varepsilon = 0$) will hold for (otherwise) arbitrary $\eta(x)$ if y obeys

$$\frac{\partial F}{\partial y} - \frac{d}{dx}\frac{\partial F}{\partial y'} = 0 \ , \tag{8}$$

which is the *Euler-Lagrange* equation.

Note that $\partial F/\partial x$, $\partial F/\partial y$, $\partial F/\partial y'$ have natural meanings for $F(x, y, y')$. However, dF/dx means the total derivative of F with respect to x, with F viewed as a function of x via $F(x, y(x), y'(x))$. i.e.

$$\frac{dF}{dx} = \frac{\partial F}{\partial x} + \frac{\partial F}{\partial y}\frac{dy}{dx} + \frac{\partial F}{\partial y'}\frac{d^2y}{dx^2} \ . \tag{9}$$

As a matter of notation, we define

$$\frac{\delta F}{\delta y} \equiv \frac{\partial F}{\partial y} - \frac{d}{dx}\frac{\partial F}{\partial y'} \ . \tag{10}$$

We refer to $\delta F/\delta y$ as the *functional derivative* of F with respect to y. Then, the Euler-Lagrange equation is

$$\frac{\delta F}{\delta y} = 0 \ . \tag{11}$$

An alternative form of the Euler-Lagrange equation is

$$\frac{\delta F}{\delta x} = \frac{d}{dx}\left(F - y\frac{\partial F}{\partial y'} \right) \ . \tag{12}$$

To prove this, look at the right hand side of the expression.

$$\begin{aligned} \text{RHS} &= \frac{\partial F}{\partial x} + y'\frac{\partial F}{\partial y} + y''\frac{\partial F}{\partial y'} - \left(y''\frac{\partial F}{\partial y'} + y'\frac{d}{dx}\frac{\partial F}{\partial y'}\right) \ , \\ &= \frac{\partial F}{\partial x} + y'\left(\frac{\partial F}{\partial y} - \frac{d}{dx}\frac{\partial F}{\partial y'}\right) \ , \\ &= \frac{\partial F}{\partial x} \ . \end{aligned} \tag{13}$$

In general, the Euler-Lagrange equations lead to a second order nonlinear ordinary differential equation. This is generally very hard to solve directly. However, we shall see that there can be a number of possible simplifications.

6.2 First Integrals

There are two particular simplifications of the Euler-Lagrange equations that are very easy to prove. I'll state them here, but you should prove them for yourselves.

1. Suppose $F = F(x, y')$. i.e., F does not depend on y. In this case, the Euler-Lagrange equations imply
$$\frac{\partial F}{\partial y'} = \text{const} \ . \tag{14}$$
This is a first order ordinary differential equation, and consequently much easier to deal with.

2. Suppose $F = F(y, y')$. i.e., F does not depend on x. In this case, the alternative form of the Euler-Lagrange equation implies
$$F - y'\frac{\partial F}{\partial y'} = \text{const} \ . \tag{15}$$

6.3 Some Examples

Let's see how this all works in some concrete physical examples.

6.3.1 The Brachistochrone

The relevant functional is

$$\sqrt{2g}T \equiv I = \int_0^a dx\, F(y, y') \ , \tag{16}$$

with

$$F(y,y') = \sqrt{\frac{1+y'^2}{y}} \ . \tag{17}$$

By our first integrals, the Euler-Lagrange equation implies

$$\sqrt{\frac{1+y'^2}{y}} - y'\sqrt{\frac{1}{y}}\frac{y'}{\sqrt{1+y'^2}} = K \ , \tag{18}$$

where K is a constant. This implies that

$$y' = \sqrt{\frac{2c}{y} - 1} \ , \tag{19}$$

where $2c = 1/K^2$. We solve this parametrically, by setting

$$\begin{aligned} y &= 2c\sin^2\theta \ , \\ &= c(1-\cos 2\theta) \ , \end{aligned} \tag{20}$$

so that $\theta = 0$ at the origin. Then

$$\begin{aligned} x &= \int \frac{\sqrt{2c}\sin\theta}{\sqrt{2c}\cos\theta} 2c\sin\left(\frac{\theta}{2}\right)\cos\left(\frac{\theta}{2}\right) d\theta \ , \\ &= c\int d\theta\,(1-\cos\theta) \ , \\ &= c\left(\theta - \frac{1}{2}\sin 2\theta\right) \ , \end{aligned} \tag{21}$$

These are the parametric equations of a cycloid.

6.3.2 Fermat's Principle of Least Time

This states that a ray of light between two fixed points in a medium travels along that path it can traverse in the least time. The refractive index, μ, in the medium is related to the speed of light, c, in the medium by

$$\mu = \frac{1}{c} \ . \tag{22}$$

To find the path, minimize

$$T = \int dt = \int \frac{ds}{c} = \int \mu\, ds \ . \tag{23}$$

Given $\mu(x,y)$ in two dimensions, we seek y to secure $\delta T = 0$, where

$$T = \int dx\,\sqrt{1+y'^2}\mu(x,y) \ . \tag{24}$$

For μ independent of y, our first integrals give

$$\frac{\mu y'}{\sqrt{1+y'^2}} = K \ , \tag{25}$$

with K a constant. If μ is also independent of x, then

$$y' = \text{constant} \ . \tag{26}$$

We can write $y' = \tan\theta$. Therefore,

$$y' = \frac{y'}{\sqrt{1+y'^2}} = \text{constant} \tag{27}$$

also. Looking at an interface, we then get

$$K = \mu_1 \sin\theta_1 = \mu_2 \sin\theta_2 \ , \tag{28}$$

which is *Snell's Law.*

6.3.3 Geodesics

A *geodesic* is the path of minimum length between two fixed points on some surface. As an example, consider the unit sphere.

$$S = \int ds \ , \tag{29}$$

where

$$ds^2 = d\theta^2 + \sin^2(\theta) d\phi^2 \ . \tag{30}$$

Let's choose θ as the independent variable, and therefore seek a solution of the form $\phi(\theta)$:

$$S = \int d\theta \, F(\theta, \phi, \phi') \ , \tag{31}$$

with

$$F = \sqrt{1 + \sin^2(\theta)\phi'^2} \ , \tag{32}$$

and a prime denotes differentiation with respect to θ. Now, F is independent of ϕ, and so this implies that $\partial F/\partial\phi'$ is a constant:

$$\frac{\sin^2(\theta)\phi'}{\sqrt{1+\sin^2(\theta)\phi'^2}} = K \ , \tag{33}$$

a constant. This equation yields the great circle paths.

6.4 More General Cases

Let's see how the ideas presented in this section can be extended in more general contexts.

6.5 $n > 1$ Dependent Variables, and 1 Independent Variable

We now consider the case $\mathbf{y} = (y_1(x), y_2(x), \ldots, y_n(x))$. The objective is to find $\mathbf{y}(x)$ such that $\delta I[\mathbf{y}] = 0$, where

$$I[\mathbf{y}] = \int_a^b F(x, \mathbf{y}, \mathbf{y}')\, dx \ , \tag{34}$$

for variations $\delta\mathbf{y}$ which leave the end points fixed, i.e. $\mathbf{y}(a)$ =fixed, $\mathbf{y}(b)$ =fixed. A trivial extension of our previous analysis yields a set of n Euler-Lagrange equations, given by

$$\frac{\delta F}{\delta y_i} \equiv \frac{\partial F}{\partial y_i} - \frac{d}{dx}\frac{\partial F}{\partial y_i'} = 0 \ . \tag{35}$$

Also, if F is independent of x, then we obtain a conserved quantity (first integral), given by

$$F - \sum_{i=1}^{n} y_i \frac{\partial F}{\partial y_i'} = \text{constant} \ . \tag{36}$$

6.5.1 Several Independent Variables, 1 Dependent Variable

Here, I is defined as an integral over a surface, volume, or higher-dimensional hypersurface. For example

$$I[u] = \int_V dV\, F(\mathbf{x}, u, \nabla u) \ . \tag{37}$$

We seek $\delta I = 0$ for variations of u that leave the values of u on the boundary, $S = \partial V$, fixed. Again, applying our variational machinery, we obtain n separate Euler-Lagrange equations:

$$\frac{\partial F}{\partial u} - \frac{d}{dx_i}\frac{\partial F}{\partial u_i} = 0 \ , \tag{38}$$

where $u_i \equiv \partial u/\partial x_i$.