Hamiltonian Systems and Variational Problems

We have seen that Hamiltonian systems arise naturally in Classical Mechanics. Now we will see that they also arise in general problems in the Calculus of Variations.

Consider a real-valued function $L(q, \dot{q}, t)$ of the variables $(q, \dot{q}, t) \in \mathbf{R}^{2n+1}$. Let $t_1 < t_2$ be real numbers, $\mathbf{a}, \mathbf{b}$ be two fixed elements in $\mathbf{R}^n$, and suppose we seek to find conditions on C^2 curves $\gamma : q = q(t)$ defined on the interval $[t_1, t_2]$ such that

$$q(t_1) = \mathbf{a}, \ \ q(t_2) = \mathbf{b} \tag{1}$$

and

$$I(\gamma) = \int_{t_1}^{t_2} L(q(t), \dot{q}(t), t) dt \tag{2}$$

is stationary for nearby curves η with the same boundary conditions (1) and $\dot{q}(t)$ is the derivative $\frac{dq}{dt}(t)$.

This means we consider one-parameter families $q(t, \alpha)$ of C^2 curves with $q(t, 0) = \gamma(t)$ such that

$$q(t_1, \alpha) = \mathbf{a}, \ \ q(t_2, \alpha) = \mathbf{b} \tag{3}$$

for all α and

$$\frac{dI}{d\alpha} \mid_{\alpha=0} = 0 \tag{4}$$

for

$$I(\alpha) = \int_{t_1}^{t_2} L(q(t, \alpha), \dot{q}(t, \alpha), t) dt. \tag{5}$$

One sometimes writes the condition (4) as

$$\delta \int_{t_1}^{t_2} L(q(t), \dot{q}(t), t) dt = 0$$

The operator δ is used to denote the fact that we are not considering an ordinary derivative, but rather, a stationary value of the integral as a family of curves changes.

Note that if γ were a curve for which the integral (2) assumed a minimum for all nearby curves with the given boundary conditions, then it would be stationary in the sense of condition (4).

To express the derivative $\frac{dI}{d\alpha}$ more conveniently, we introduce some notation. Write $q = (q_1, \ldots, q_n), \dot{q} = (\dot{q}_1, \ldots, \dot{q}_n)$, L_{q_k} for the partial derivative of L with respect to q_k, $L_{\dot{q}_k}$ for the partial derivative of L with respect to $\dot{q}_k$ with $1 \leq k \leq n$. Also, we denote differentiation with respect to t by 'dot' and that with respect to α by 'prime'.

Consider the condition $\frac{dI}{d\alpha} \mid_{\alpha=0} = 0$.

We have

$$0 = I'(\alpha) = \int_{t_1}^{t_2} L_q \cdot q' + L_{\dot{q}} \cdot \dot{q}' dt \tag{6}$$

where $L_q \cdot q', \; L_{\dot{q}} \cdot \dot{q}'$ respectively stand for

$$\sum_{k=1}^{n} L_{q_k} q_k'$$

and

$$\sum_{k=1}^{n} L_{\dot{q}_k} \dot{q}_k'$$

The constraints of our variation curves make $q(t, \alpha)$ constant in α at the boundary points t_1, t_2. So, we have the dot product

$$s(t, \alpha) = L_{\dot{q}} \cdot q' = \sum_{k=1}^{n} L_{\dot{q}_k} q_k'$$

vanishes at t_1, t_2.

Hence,

$$0 = \int_{t_1}^{t_2} \frac{ds}{dt} dt \tag{7}$$

which gives

$$0 = \int_{t_1}^{t_2} \frac{dL_{\dot{q}}}{dt} \cdot q' + L_{\dot{q}} \cdot \dot{q}' dt \tag{8}$$

Subtracting (8) from (6), we get

$$0 = \frac{dI}{d\alpha}\mid_{\alpha=0} = \int_{t_1}^{t_2} (L_q - \frac{d}{dt}L_{\dot{q}})q' dt \tag{9}$$

Now, the 'prime' derivatives q' can be made arbitrary, so, (9) implies

$$L_q - \frac{d}{dt}L_{\dot{q}} = 0 \tag{10}$$

or, written out completely,

$$\frac{d}{dt}L_{\dot{q}_k} = L_{q_k},\ k = 1, \ldots, n \tag{11}$$

The equations (11) are called the Euler-Lagrange equations.

At a curve $(q(t), \dot{q}(t), t)$ which makes the integral (2) stationary, we have that $(q(t), \dot{q}(t))$ satisfies

$$\frac{d}{dt}L_{\dot{q}_k}(q(t), \dot{q}(t), t) = L_{q_k}(q(t), \dot{q}(t), t)$$

for $k = 1, \ldots, n$.

Note that these are second order differential equations.

Examples.

1. Suppose that we consider the problem of finding the curve γ of shortest length shortest joining two points $\mathbf{a}, \mathbf{b} \in \mathbf{R}^2$

   Writing $\gamma(t) = (x(t), y(t)),\ t_1 \le t \le t_2$, we seek to minimize the function

   $$I(\gamma) = \int_{t_1}^{t_2} \sqrt{\dot{x}^2 + \dot{y}^2} dt$$

   over all such curves. Let $L(x, y, \dot{x}, \dot{y}, t) = \sqrt{\dot{x}^2 + \dot{y}^2}$.

   The Euler-Lagrange equations become

   $$\frac{d}{dt}L_{\dot{x}} = L_x,\ \frac{d}{dt}L_{\dot{y}} = L_y$$

   Since, L is independent of x, y, t, we get

   $$\frac{d}{dt}L_{\dot{x}} = 0,\ \frac{d}{dt}L_{\dot{y}} = 0,\ L_t = 0$$

These equations become

$$\begin{aligned} \frac{d}{dt} L_{\dot{x}} &= \frac{d}{dt} \frac{\dot{x}}{L} \\ &= \frac{L\ddot{x} - \dot{x} L_t}{L^2} \\ &= \frac{\ddot{x}}{L} = 0, \end{aligned}$$

and

$$\begin{aligned} \frac{d}{dt} L_{\dot{y}} &= \frac{d}{dt} \frac{\dot{y}}{L} \\ &= \frac{L\ddot{y} - \dot{y} L_t}{L^2} \\ &= \frac{\ddot{y}}{L} = 0. \end{aligned}$$

Using that L is never zero, we see the only solutions are those $(x(t), y(t))$ for which $\ddot{x} = 0, \ \ddot{y} = 0$. That is, the only solutions are straight lines. With the above boundary condition, we get a unique line segment joining **a** to **b**.

2. Consider the open unit disk D in the complex plane, $D = \{z \in \mathbf{C} : \mid z \mid < 1\}$. Define the function

$$I(\gamma) = \int_\gamma \frac{2 \mid dz \mid}{1 - \mid z \mid^2}$$

Then, the curves which make this integral a minimum are the straight lines through the origin and the circles orthogonal to the boundary of D. In elementary differential geometry, it is proved that the above functional I gives the arclength of the curve γ with the Riemannian metric on D with constant Gaussian curvature equal to -1. The simplest way to prove that the minimizing curves are as indicated is to use a

Linear Fractional Transformation T (e.g. $z \to \frac{-i(z-1)}{z+1}$) to map D to the upper half-plane H. The integrals are transformed to

$$I(T(\gamma)) = \int_{T(\gamma)} \frac{|\, dz \,|}{Im\ z} \tag{12}$$

The integrand (12) is preserved by transformations of the form $S(z) = \frac{az+b}{cz+d}$ where a, b, c, d are real and $ad - bc > 0$. Also, such transformations map circles and lines orthogonal to the real axis into other such curves. Given a pair of points, $\mathbf{a} \neq \mathbf{b}$ in $\mathbf{C}$, we can find an S which maps $\mathbf{a}, \mathbf{b}$ into the same vertical line. It is easy to see that a minimizing curve for (12) connecting points on the same vertical line must be a curve along that vertical line. Thus, minimizing curves in H must be pieces of lines and circles orthogonal to the real axis. Pulling back to D by T^{-1} gives the desired result for D.

We now show that, in an open set G in $\mathbf{R}^{2n+1}$ in which the matrix function

$$L_{\dot{q}_k, \dot{q}_\ell}(q, \dot{q}, t) \tag{13}$$

is non-singular, we can choose coordinates in which the Euler-Lagrange equations become a Hamiltonian system. This is one of the main reasons that Hamiltonian systems are important.

So, assume that we have the independent coordinates $(q, \dot{q}, t)$ in an open set G in $\mathbf{R}^{2n+1}$, that $L(q, \dot{q}, t)$ is a C^2 real-valued function in G, and that the matrix function (13) is non-singular in G.

Consider the set of equations

$$p_k = L_{\dot{q}_k}(q, \dot{q}, t), \ \ k = 1, \ldots, n \tag{14}$$

Because of the assumption that (13) is non-singular, the Implicit Function Theorem gives us a set of C^2 functions $S_k(q, p, t)$ for $k = 1, \ldots, n$, such that (14) holds if and only if

$$\dot{q}_k = S_k(q, p, t), \ \ k = 1, \ldots, n \tag{15}$$

Let

$$\begin{aligned} H(q, p, t) &= \sum_k p_k \dot{q}_k - L(q, \dot{q}, t) \\ &= \sum_k p_k S_k(q, p, t) - L(q, S(q, p, t), t) \end{aligned}$$

Then, for $\ell = 1, \ldots, n$,

$$\begin{aligned} \frac{\partial H}{\partial p_\ell} &= S_\ell(q,p,t) + \sum_k p_k \frac{\partial S_k(q,p,t)}{\partial p_\ell} - \sum_k L_{\dot{q}_k} \frac{\partial S_k}{\partial p_\ell} \\ &= S_\ell(q,p,t) = \dot{q}_\ell \end{aligned}$$

since $p_k = L_{\dot{q}_k}(q, \dot{q}, t)$.
Also,

$$\begin{aligned} -\frac{\partial H}{\partial q_\ell} &= -\sum_k p_k \frac{\partial S_k(q,p,t)}{\partial q_\ell} + L_{q_\ell} + \sum_k L_{\dot{q}_k} \frac{\partial S_k}{\partial q_\ell} \\ &= L_{q_\ell} \\ &= \frac{d}{dt} L_{\dot{q}_\ell} \text{ by Euler-Lagrange} \\ &= \frac{d}{dt} p_\ell \text{ by definition of } p_\ell \end{aligned}$$

In the (q,p,t), coordinates, we therefore have a Hamiltonian system with Hamiltonian function H. If $L(q,\dot{q},t) = L(q,\dot{q})$ is independent of time t, then so is H. However, in the general case, both L and H are time dependent. Note that if H is time-dependent, then the function H is *not* constant on solutions to the Hamiltonian system.

Let us now return to the Conservative Mechanical system with potential energy U we studied in the last section.

Using, position q and momentum p as coordinates, we saw that the equations of motion were a (time-independent) Hamiltonian system with Hamiltonian function

$$H(q,p) = \frac{1}{2} \sum_k \frac{p_k^2}{m_k} + U(q_1, \ldots, q_n)$$

and the velocity $\dot{q}$ satisfied $\dot{q}_k = \frac{p_k}{m_k}$.

If we assume that this Hamiltonian system comes from a variational problem as above, we are led to write

$$H = \sum_k p_k \dot{q}_k - L(q,p)$$

or

$$
\begin{aligned}
L &= \sum_k p_k \dot{q}_k - H \\
&= \sum_k p_k \dot{q}_k - \frac{1}{2}\sum_k \frac{p_k^2}{m_k} - U \\
&= \sum_k p_k \frac{p_k}{m_k} - \frac{1}{2}\sum_k \frac{p_k^2}{m_k} - U \\
&= \frac{1}{2}\sum_k \frac{p_k^2}{m_k} - U \\
&= K - U
\end{aligned}
$$

where K denotes the kinetic energy. The function $L = K - U$ is called the *Lagrangian function* or *action function*, as opposed to the function $T = K + U$ which was called the Energy function.

From the above, we are led to guess that conservative mechanical systems would satisfy the Euler-Lagrange equations for the function $L = K - U$. This is indeed the case as can be easily verified. In this case, it can be verified that the integral

$$\int L(q, \dot{q})dt$$

is actually minimized by the solution curves $(q, \dot{q})$, not just made stationary. This is known as Hamilton's *Principal of Least Action.*