5 Calculus of Variations

Given a function F(y, y', x) where y = y(x) is unspecified, then we consider

$$I(y) = \int_{x_1}^{x_2} F(y, y', x) dx,$$

called the functional. If y(x) is subject to some boundary conditions $y(x_1) = y_1$ and $y(x_2) = y_2$, then I(y) has a maximum or minimum value if F satisfies the Euler-Lagrange Equation:

$$\frac{\partial F}{\partial y} - \frac{d}{dx} \left(\frac{\partial F}{\partial y'} \right) = 0.$$

Proof: Suppose y satisfies the given boundary conditions and consider small variations in y:

$$y \to y + \delta y = y + \epsilon \eta(x)$$

where ϵ is a small parameter independent of x and $\eta(x)$ is an arbitrary function such that $\eta(x_1) = \eta(x_2) = 0$. (This is necessary so the boundary conditions are satisfied.) Then $\delta y = \epsilon \eta(x)$ is the variation in y.

We wish to find y(x) such that $I(y + \epsilon \eta(x))$ takes a maximum or minimum value at $\epsilon = 0$, for any choice of $\eta(x)$. A necessary condition for this is that

$$\left. \frac{d}{d\epsilon} I(y + \epsilon \eta(x)) \right|_{\epsilon=0} = 0$$

for all $\eta(x)$ satisfying $\eta(x_1) = \eta(x_2) = 0$. Then at $\epsilon = 0$ we must have

$$0 = \frac{d}{d\epsilon} I(y + \epsilon \eta(x)) \Big|_{\epsilon=0}$$
$$= \int_{x_1}^{x_2} \frac{d}{d\epsilon} F(y + \epsilon \eta(x), y' + \epsilon \eta'(x), x) \Big|_{\epsilon=0} dx.$$

Consider the substitutions

$$u = y + \epsilon \eta(x)$$
 \Rightarrow $\frac{du}{d\epsilon} = \eta$ $v = y' + \epsilon \eta'(x)$ \Rightarrow $\frac{dv}{d\epsilon} = \eta'$.

But by the chain rule

$$\frac{\partial F}{\partial u} = \frac{\partial F}{\partial y} \frac{\partial y}{\partial u} = \frac{\partial F}{\partial y},$$

$$\frac{\partial F}{\partial v} = \frac{\partial F}{\partial y'} \frac{\partial y'}{\partial v} = \frac{\partial F}{\partial y'}.$$

Hence

$$\frac{d}{d\epsilon}F(u,v,x)\Big|_{\epsilon=0} = \left(\frac{\partial F}{\partial u}\frac{du}{d\epsilon} + \frac{\partial F}{\partial v}\frac{dv}{d\epsilon}\right)\Big|_{\epsilon=0} \\
= \left(\frac{\partial F}{\partial u}\frac{du}{d\epsilon} + \frac{\partial F}{\partial v}\frac{dv}{d\epsilon}\right)\Big|_{\epsilon=0} + \frac{\partial F}{\partial v}\frac{dv}{d\epsilon} + \frac{\partial F}{\partial v}\frac$$

Substituting this into our integral, we find

$$0 = \int_{x_1}^{x_2} \left(\frac{\partial F}{\partial y} \eta + \frac{\partial F}{\partial y'} \eta' \right) dx.$$

Now by the product rule

$$\frac{d}{dx}\left(\frac{\partial F}{\partial y'}\eta\right) = \frac{d}{dx}\left(\frac{\partial F}{\partial y'}\right)\eta + \frac{\partial F}{\partial y'}\eta',$$

SO

$$0 = \int_{x_1}^{x_2} \left(\frac{\partial F}{\partial y} \eta - \frac{d}{dx} \left(\frac{\partial F}{\partial y'} \right) \eta \right) dx + \int_{x_1}^{x_2} \frac{d}{dx} \left(\frac{\partial F}{\partial y'} \eta \right) dx$$
$$= \int_{x_1}^{x_2} \eta \left(\frac{\partial F}{\partial y} - \frac{d}{dx} \left(\frac{\partial F}{\partial y'} \right) \right) dx + \underbrace{\left[\eta \frac{\partial F}{\partial y'} \right]_{x_1}^{x_2}}_{0}.$$

The last term equals zero because of the boundary conditions for η . Hence we have:

$$\int_{x_1}^{x_2} \eta \left(\frac{\partial F}{\partial y} - \frac{d}{dx} \left(\frac{\partial F}{\partial y'} \right) \right) dx = 0 \quad \forall \eta(x).$$

As this holds for all $\eta(x)$, we must have

$$\frac{\partial F}{\partial y} - \frac{d}{dx} \left(\frac{\partial F}{\partial y'} \right) = 0,$$

which is the Euler-Lagrange equation.

We can extend the result for systems with $F(y_1, \ldots, y_n, y'_1, \ldots, y'_n, x)$, where $y_j(x)$, $(1 \le j \le n)$ are n independent functions of x whose values at x_1 and x_2 are specified and we have the functional:

$$I(y_1, \dots, y_n) = \int_{x_1}^{x_2} F(y_1, \dots, y_n, y_1', \dots, y_n', x) dx$$

The functional will have a minimum or maximum only if:

$$\frac{d}{dx}\left(\frac{\partial F}{\partial y_i'}\right) - \frac{\partial F}{\partial y_i} = 0 \qquad \forall i = 1 \dots n$$

Examples:

1) Consider the special case where F depends only on y, y', and is independent of x, that is, F = F(y, y'). Then the Euler-Lagrange equations give:

$$0 = y' \left[\frac{\partial F}{\partial y} - \frac{d}{dx} \left(\frac{\partial F}{\partial y'} \right) \right]$$

$$= y' \frac{\partial F}{\partial y} - \left[\frac{d}{dx} \left(y' \frac{\partial F}{\partial y'} \right) - y'' \frac{\partial F}{\partial y'} \right]$$
(This trick is due to the product rule on $\frac{d}{dx} \left(y' \frac{\partial F}{\partial y'} \right)$)
$$= \frac{dF}{dx} - \frac{d}{dx} \left(y' \frac{\partial F}{\partial y'} \right)$$

$$0 = \frac{d}{dx} \left[F - y' \frac{\partial F}{\partial y'} \right].$$

5.1 The Brachistochrone Problem

Jacob Bernoulli solved this long-standing problem in 1696 by what turned out to be an application of the Calculus of Variations. A particle starting at rest moves along a frictionless wire under it's own weight from the origin to point P(a, b). What is the path y = y(x) for which the time taken is least?

The system is conserved, so we have conservation of energy:

$$T + V = \frac{1}{2}mv^2 - mgy = E.$$

Initially y = 0 and v = 0, and hence E = 0. Therefore

$$\frac{1}{2}mv^2 = mgy \qquad \Rightarrow v^2 = 2gy$$

But

$$v = \sqrt{\left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2}.$$

Now

$$\sqrt{\left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2}$$

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

$$= \sqrt{1 + y'^2} \frac{dx}{dt}.$$

Therefore

$$\sqrt{2gy} = \sqrt{1 + {y'}^2} \frac{dx}{dt} \qquad \Rightarrow \qquad \frac{dt}{dx} = \frac{\sqrt{1 + {y'}^2}}{\sqrt{2gy}}.$$

Thus the time taken is given by

$$t = \int_0^a \sqrt{\frac{1 + {y'}^2}{2gy}} \ dx,$$

or

$$t = \int_0^a F(y, y') dx$$
, where $F(y, y') = \sqrt{\frac{1 + {y'}^2}{2gy}}$.

We can find y so that t is a minimum if F(y, y') satisfies the specialised Euler-Lagrange equation from the above example,

$$F - y' \frac{\partial F}{\partial y'} = c.$$

Hence

$$c = \sqrt{\frac{1 + {y'}^2}{2gy}} - \frac{{y'}^2}{\sqrt{2gy(1 + {y'}^2)}} = \frac{1}{\sqrt{2gy(1 + {y'}^2)}}.$$

Therefore

$$y(1+{y'}^2) = 2c, \quad \text{const.}$$

SO

$$y'^2 = \frac{2c}{y} - 1.$$

On physical grounds, y' > 0, so that

$$y' = \sqrt{\frac{2c - y}{y}}.$$

We solve by substituting $y=2c\sin^2\theta$, so that $dy=4c\sin\theta\cos\theta d\theta$ to give

$$dx = \sqrt{\frac{2c\sin^2\theta}{2c\cos^2\theta}} 4c\sin\theta\cos\theta d\theta$$
$$= 4c\sin^2\theta d\theta = 2c(1-\cos2\theta)d\theta$$
$$x = 2c(\theta - \frac{1}{2}\sin2\theta)$$

provided x = 0 at $\theta = 0$. Thus

$$x = c(2\theta - \sin 2\theta), \quad y = 2c\sin^2 \theta = c(1 - \cos 2\theta).$$

Setting $\phi = 2\theta$ gives

$$x = c(\phi - \sin \phi), \quad y = c(1 - \cos \phi)$$

which are the parametric equations of a cycloid.

Note: The origin at (x, y) = (0, 0) corresponds to $\phi = 0$. We may determine c from the requirement that (x, y) = (a, b) at the other end point.

6 Principle of Least Action - Hamilton's principle

In investigating dynamics, we often strive to look for basic concepts, things which are at the heart of our theory and encompass it all. When you are first presented dynamics, we often consider energy to be a fundamental concept. However, there turns out to be a more fundamental concept than energy: action. The action of a system is related to how the system follows a particular trajectory.

Consider a conservative system with generalised coordinates $q_1, ..., q_m$ and Lagrangian $L(q_i, \dot{q}_i, t) = T(q_i, \dot{q}_i, t) - V(q_i, t)$. According to Hamilton's principle, the observed motion of the system from time t_1 to t_2 is given by that trajectory which minimises the **action** integral:

$$A = \int_{t_1}^{t_2} L(q_i, \dot{q}_i, t) dt.$$

Since the generalised coordinates q_i are independent, the Euler-Lagrange equations state that a necessary condition for a minimum is

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_i}\right) - \frac{\partial L}{\partial q_i} = 0, \quad i = 1, ..., n,$$

which are Lagrange's equations of motion derived previously.

Example: Projectile Motion.

Consider a particle launched from the origin with a particular velocity and angle. From this angle and velocity, we can work out our horizontal and vertical initial velocities. Now the question is, what path does it travel? By the Principle of Least Action, we have the functional:

$$A = \int_{t_1}^{t_2} L(q_i, \dot{q}_i, t) dt.$$

Assume there are no forces other than gravity acting on the particle (e.g. no air resistance). Then our Lagrangian is given by:

$$L = T - V = \frac{1}{2}m\dot{x}^2 + \frac{1}{2}m\dot{y}^2 + mgy.$$

Now the Principle of Least Action states that the system will evolve according to the trajectory that minimises the action functional A. Therefore the Euler-Lagrange equations must hold for both generalised coordinates. So

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

$$\frac{\dot{d}}{dt} (m\dot{x}) - 0 = 0$$

$$\ddot{x} = 0.$$

This is what we expect as there are no horizontal forces. Now let's look at the Euler-Lagrange equations for y:

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

$$\frac{d}{dt} (m\dot{y}) + mg = 0$$

$$\ddot{y} = -g.$$

This is also consistent with our system (and experience). So what about our trajectory? It's easy to see (by integrating twice), that we get:

$$\begin{array}{rcl}
x & = & v_x t, \\
y & = & y_0 + v_y t - \underline{gt^2}.
\end{array}$$

These two parametric equations will trace out a parabola, as per our experience with basic mechanics.

6.1 Generalised momenta and conservation

In the case of n free particles with cartesian coordinates (x_i, y_i, z_i) , $1 \le i \le n$,

$$L = T = \sum_{i} \frac{1}{2} m_i (\dot{x}_i^2 + \dot{y}_i^2 + \dot{z}_i^2).$$

We have $\frac{\partial L}{\partial \dot{x}_i} = m_i \dot{x}_i = x$ —component of the momentum \mathbf{p}_i of particle i. (Similarly for the y, z components).

In general for a system with constraints and generalised coordinates q_i , i = 1, ..., m, we define the **generalised momentum** associated with q_i by

$$p_i = \frac{\partial L}{\partial \dot{q}_i}, \quad i = 1, ..., m.$$

However p_i does not usually have dimensions of linear momentum.

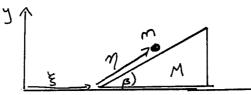
If L depends on \dot{q}_1 but not on q_1 then

$$\frac{\partial L}{\partial q_1} = 0 \Rightarrow \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}_1} \right) = 0,$$

which means

$$\frac{dp_1}{dt} = 0 \Rightarrow p_1 = \text{const.}$$

Several of the generalised momenta p_i may be conserved in a problem. The relevant q_i are called **cyclic** or **ignorable** coordinates.



18) A particle of mass m slides down the smooth face of a wedge of mass M which itself is free to slide on a smooth horizontal plane. Suppose the angle of inclination of the wedge is β and at t=0 the wedge is at rest and the mass is at rest a distance b along the face of the wedge. Determine the time taken for the mass to reach the bottom of the wedge.

Solution Introduce coordinates η, ξ as shown so that at t=0 we may assume $\eta=b, \ \xi=0, \ \dot{\eta}=\dot{\xi}=0$. Let v be the velocity of the mass so that

$$\mathbf{v} = (\dot{\xi} + \dot{\eta}\cos\beta)\,\hat{\mathbf{i}} + \dot{\eta}\sin\beta\,\hat{\mathbf{j}}.$$

Then

$$T = \frac{1}{2}M\dot{\xi}^2 + \frac{1}{2}m(\dot{\eta}^2 + \dot{\xi}^2 + 2\dot{\xi}\dot{\eta}\cos\beta), \quad V = mg\eta\sin\beta$$

and

$$L = \frac{1}{2}M\dot{\xi}^{2} + \frac{1}{2}m(\dot{\eta}^{2} + \dot{\xi}^{2} + 2\dot{\xi}\dot{\eta}\cos\beta) - mg\eta\sin\beta.$$

Note that the P.E. of the wedge is a constant so may be dropped. Therefore ξ is ignorable with generalised momentum

$$p_{\xi} = \frac{\partial L}{\partial \dot{\xi}} = (m+M)\dot{\xi} + m\dot{\eta}\cos\beta = \text{const.}$$

Since $\dot{\eta} = \dot{\xi} = 0$ at t = 0 we find that const. = 0. Then

$$\dot{\xi} = -\frac{m\cos\beta}{m+M}\dot{\eta}\tag{6}$$

The generalised momentum corresponding to η is

$$p_{\eta} = \frac{\partial L}{\partial \dot{\eta}} = m\dot{\eta} + m\dot{\xi}\cos\beta$$
$$= \frac{m(M + m\sin^2\beta)\dot{\eta}}{m + M}$$

Therefore using Lagrange's equation we get

$$\frac{m(M+m\sin^2\beta)\ddot{\eta}}{m+M} = \frac{dp_{\eta}}{dt} = \frac{\partial L}{\partial \eta} = -mg\sin\beta$$

so

$$\ddot{\eta} = -\frac{g(M+m)\sin^2\beta}{m\sin^2\beta + M} = \text{const.}$$

Given at t = 0, $\dot{\eta} = 0$, $\eta = b$ we find

$$\eta = b - \frac{g(M+m)t^2 \sin \beta}{2(m \sin^2 \beta + M)}.$$

Therefore mass reaches the bottom of the wedge $(\eta = 0)$ when

$$t = \sqrt{\frac{2b(M + m\sin^2\beta)}{g(M + m)\sin\beta}}.$$

Note: Since $\xi = 0$, $\eta = b$ at t = 0, equation (6) gives us

$$\xi = \frac{m\cos\beta}{m+M}(b-\eta)$$

which determines a straight line in (ξ, η) space.

7 Hamiltonian Mechanics

Consider a conservative system of n particles with generalised coordinates q_1, \ldots, q_m and corresponding generalised momenta

$$p_j = \frac{\partial L}{\partial \dot{q}_j}$$
 $j = 1 \dots m$, $L = L(q_i, \dot{q}_i, t)$.

Then from Lagrange's equations

$$dL = \sum_{i=1}^{m} \frac{\partial L}{\partial q_i} dq_i + \sum_{i=1}^{m} \frac{\partial L}{\partial \dot{q}_i} d\dot{q}_i + \frac{\partial L}{\partial t} dt$$

$$= \sum_{i=1}^{m} \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}_i} \right) dq_i + \sum_{i=1}^{m} \frac{\partial L}{\partial \dot{q}_i} d\dot{q}_i + \frac{\partial L}{\partial t} dt$$

$$= \sum_{i=1}^{m} \dot{p}_i dq_i + \sum_{i=1}^{m} p_i d\dot{q}_i + \frac{\partial L}{\partial t} dt.$$

Let us set

$$H = \sum_{i=1}^{m} p_i \dot{q}_i - L,$$

called the **Hamiltonian** of the system. Then

$$dH = \sum_{i=1}^{m} p_i d\dot{q}_i + \sum_{i=1}^{m} \dot{q}_i dp_i - \sum_{i=1}^{m} \frac{\partial L}{\partial q_i} dq_i - \sum_{i=1}^{m} \frac{\partial L}{\partial \dot{q}_i} d\dot{q}_i - \frac{\partial L}{\partial t} dt$$
$$= -\sum_{i=1}^{m} \dot{p}_i dq_i + \sum_{i=1}^{m} \dot{q}_i dp_i - \frac{\partial L}{\partial t} dt.$$

Thus changes in H depend only on changes in p_i , q_i , t which suggests that H is a function of the generalised coordinates and their corresponding generalised momenta, and possibly t.

$$H = H(p_i, q_i, t).$$

The equation for dH then implies

$$\dot{q}_i = \frac{\partial H}{\partial p_i}, \qquad \dot{p}_i = -\frac{\partial H}{\partial q_i}, \qquad \frac{dH}{dt} = -\frac{\partial L}{\partial t} = \frac{\partial H}{\partial t}.$$

These are known as Hamilton's equations, the last one arising from

$$\frac{dH}{dt} = -\sum_{i=1}^{m} \dot{p}_{i} \frac{dq_{i}}{dt} + \sum_{i=1}^{m} \dot{q}_{i} \frac{dp_{i}}{dt} - \frac{\partial L}{\partial t} = -\frac{\partial L}{\partial t}.$$

Note: The 2m first order Hamilton equations above replace the m second order Lagrange equations!

7.1 Physical significance of H

In the case L and thus H has no explicit dependence on t, we have from Hamilton's equations

$$\frac{dH}{dt} = \frac{\partial H}{\partial t} = -\frac{\partial L}{\partial t} = 0$$

so H = constant.

Assume \mathbf{r}_j is only a function of the generalised coordinates q_i , so

$$\frac{\partial \mathbf{r}_j}{\partial t} = 0.$$

Then

$$\dot{\mathbf{r}}_i = \sum_{j=1}^m \frac{\partial \mathbf{r}_i}{\partial q_j} \dot{q}_j$$

and

$$T = \frac{1}{2} \sum_{k} m_k \, \dot{\mathbf{r}}_k \cdot \dot{\mathbf{r}}_k.$$

Assuming $V = V(q_i, t)$, then

$$p_{i} = \frac{\partial L}{\partial \dot{q}_{i}}$$

$$= \frac{\partial T}{\partial \dot{q}_{i}}$$

$$= \sum_{k} m_{k} \dot{\mathbf{r}}_{k} \cdot \frac{\partial \dot{\mathbf{r}}_{k}}{\partial \dot{q}_{i}}$$

$$= \sum_{k} m_{k} \dot{\mathbf{r}}_{k} \cdot \frac{\partial \mathbf{r}_{k}}{\partial q_{i}}$$

Now

$$T = \frac{1}{2} \sum_{k} m_{k} \dot{\mathbf{r}}_{k} \cdot \dot{\mathbf{r}}_{k}$$

$$= \frac{1}{2} \sum_{i,k} m_{k} \dot{\mathbf{r}}_{k} \cdot \frac{\partial \mathbf{r}_{k}}{\partial q_{i}} \dot{q}_{i}$$

$$= \frac{1}{2} \sum_{i} p_{i} \dot{q}_{i}$$

$$(7)$$

or

$$\sum_{i} p_i \dot{q}_i = 2T.$$

Hence

$$H = \sum_{i} p_{i}\dot{q}_{i} - L$$

$$= 2T - (T - V)$$

$$= T + V$$

$$= E$$

which is constant since the system is conservative. Thus H denotes the total energy of the system when the \mathbf{r}_i have no explicit time dependence.

Note It is possible to have $\mathbf{r}_i = \mathbf{r}_i(q_j, t)$ but $H = H(p_i, q_i)$. In that case H is still conserved, but does not denote the total energy (see tutorial sheet).

Example 19. Simple pendulum.

$$T = \frac{1}{2}ml^2\dot{\theta}^2, \quad V = -mgl\cos\theta$$

Next

$$L = T - V = \frac{1}{2}ml^2\dot{\theta}^2 + mgl\cos\theta$$

So

$$p_{\theta} = \frac{\partial L}{\partial \dot{\theta}} = ml^2 \dot{\theta}$$

and

$$H = \dot{\theta}p_{\theta} - L = ml^{2}\dot{\theta}^{2} - \frac{1}{2}ml^{2}\dot{\theta}^{2} - mgl\cos\theta = \frac{1}{2}ml^{2}\dot{\theta}^{2} - mgl\cos\theta = T + V.$$

In terms of generalised momentum,

$$H = \frac{p_{\theta}^2}{2ml^2} - mgl\cos\theta.$$

Hamilton's equations are

$$\dot{\theta} = \frac{\partial H}{\partial p_{\theta}} = \frac{p_{\theta}}{ml^2}, \quad \dot{p}_{\theta} = -\frac{dH}{d\theta} = -mgl\sin\theta$$

or

$$ml^2\ddot{\theta} = -mgl\sin\theta.$$

In this example neither H nor f have explicit time dependence, so H=T+V (constant).

Example 20. Bead on a wire rotating with angular velocity ω .

$$V = mga\cos\theta, \quad T = \frac{1}{2}ma^2[\dot{\theta}^2 + \omega^2\sin^2\theta]$$

and $L = T - V = \frac{1}{2}ma^2[\dot{\theta}^2 + \omega^2\sin^2\theta] - mga\cos\theta$. Now $p_{\theta} = \frac{\partial T}{\partial \dot{\theta}} = ma^2\dot{\theta}$ and

$$H = \dot{\theta}p_{\theta} - L$$

$$= ma^{2}\dot{\theta}^{2} - \frac{1}{2}ma^{2}[\dot{\theta}^{2} + \omega^{2}\sin^{2}\theta] + mga\cos\theta$$

$$= \frac{1}{2}ma^{2}[\dot{\theta}^{2} - \omega^{2}\sin^{2}\theta] + mga\cos\theta$$

$$= \frac{p_{\theta}^{2}}{2ma^{2}} - \frac{1}{2}ma^{2}o^{2}\sin^{2}\theta + mga\cos\theta.$$
(8)

Then Hamilton's equations become

$$\dot{\theta} = \frac{\partial H}{\partial p_{\theta}} = \frac{p_{\theta}}{ma^{2}},$$

$$\dot{p}_{\theta} = -\frac{\partial H}{\partial \theta} = ma^{2}\omega^{2}\sin\theta\cos\theta + mga\sin\theta$$

$$= ma^{2}\sin\theta[\omega^{2}\cos\theta + g/a]$$

or

$$\ddot{\theta} = \sin \theta (\omega^2 \cos \theta + g/a). \tag{9}$$

Note: In this example $\frac{\partial H}{\partial t}=0$ so H is constant. However the position depends explicitly on time, so $H\neq T+V$. In fact,

$$H - (T + V) = -ma^2\omega^2\sin^2\theta$$

which is not constant, so energy is not conserved. This is as we expect, as work is being done on the system to drive it with constant angular velocity ω .