5. CALCULUS OF VARIATIONS

Consider the problem of finding the shortest distance between two points on a surface.

If the surface is flat then the shortest distance is a straight line.

If the surface is a sphere then the shortest distance is a along a great circle.

We can find the extremum of a function $y(x)$ by finding the points where $\frac{df}{dx} = 0$.

In the calculus of variations we are concerned with finding a function or a curve, rather than a value of some variable, that makes a given quantity stationary.

Some examples:

1. calculating the shortest distance between two points
2. calculating the shape of a wire such that a bead slides down it in the minimum time
3. calculating the shape taken up by a soap film between two circular wires such that the surface area is a minimum
4. calculating the path of a light ray in a medium of a given refractive index.

5.1 EULER EQUATION

The calculus of variations typically involves problems in which a quantity to be minimized (or maximized) appears as a functional, i.e. a quantity whose arguments are themselves functions as well as variables.

Consider a functional J of y defined by

$$J[y] = \int_{x_1}^{x_2} f\left(y(x), \frac{dy}{dx}(x), x\right) dx$$

The value of J depends on the behaviour of $y(x)$ throughout the range of x, $x_1 \leq x \leq x_2$.

We might be required to find a function $y(x)$ that makes J stationary relative to small changes in y.

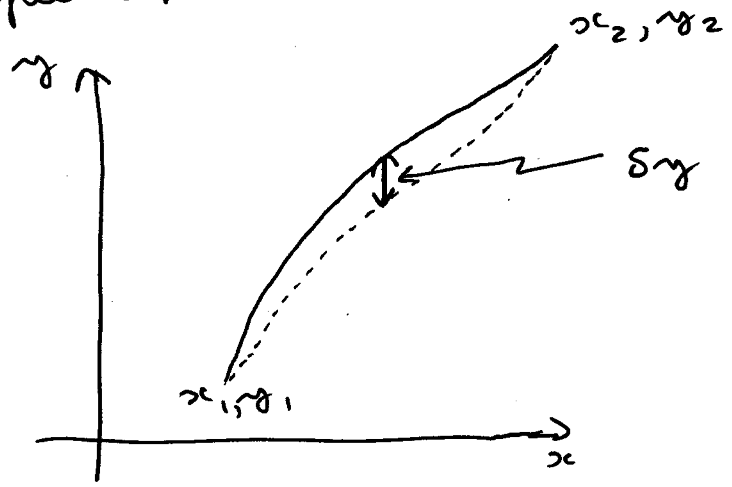
We indicate the variation in J by

$$\delta J = \delta \int_{x_1}^{x_2} f(y, y_x, x)\, dx$$

where $y_x = \dfrac{dy}{dx}$.

It is helpful to think of $y(x)$ as a path or curve connecting the values $y(x_1)$ and $y(x_2)$.

We assume the existence of an optimum path, i.e. a function $y(x)$ for which J is stationary, and then compare J for the optimum path with that for the infinite number of neighbouring paths.



Let δy denote the variation of y.

Describe δy by a new function, $\eta(x)$, and a scale factor α that controls the magnitude of the variation.

We will have

$$\eta(x_1) = \eta(x_2) = 0$$

and
$$y(x, \alpha) = y(x, 0) + \alpha \eta(x)$$

where $y(x,0)$ is the as yet unknown path that will minimize J. Relative to $y(x,0)$ the variation δy is then

$$\delta y = \alpha \eta(x)$$

and we can now write

$$J(\alpha) = \int_{x_1}^{x_2} f\left(y(x,\alpha),\, y_x(x,\alpha),\, x\right) dx .$$

J is now a function of the scale factor α rather that a functional of y. This means we now know how to optimize it.

The condition is

$$\left[\frac{\partial J(\alpha)}{\partial \alpha}\right]_{\alpha=0} = 0$$

(analogous to finding $dy/dx = 0$ in differential calculus)

$$\frac{\partial J(\alpha)}{\partial \alpha} = \int_{x_1}^{x_2} \left[\frac{\partial f}{\partial y}\frac{\partial y}{\partial \alpha} + \frac{\partial f}{\partial y_x}\frac{\partial y_x}{\partial \alpha}\right] dx = 0$$

But $$\frac{\partial y}{\partial \alpha}(x,\alpha) = \eta(x)$$

and $$\frac{\partial y_x}{\partial \alpha}(x,\alpha) = \frac{d\eta(x)}{dx}$$

$$\Rightarrow \frac{\partial J(\alpha)}{\partial \alpha} = \int_{x_1}^{x_2} \left(\frac{\partial f}{\partial y}\eta(x) + \frac{\partial f}{\partial y_x} \cdot \frac{d\eta(x)}{dx}\right) dx = 0$$

But $$\int_{x_1}^{x_2} \frac{d\eta(x)}{dx} \cdot \frac{\partial f}{\partial y_x}\, dx = \left[\eta(x)\frac{\partial f}{\partial y_x}\right]_{x_1}^{x_2} - \int_{x_1}^{x_2} \eta(x)\frac{d}{dx}\frac{\partial f}{\partial y_x} \cdot dx$$

and $\eta(x_2) = \eta(x_1) = 0$.

Hence

$$\frac{\partial J(\alpha)}{\partial \alpha} = \int_{x_1}^{x_2} \left[\frac{\partial f}{\partial y} - \frac{d}{dx}\frac{\partial f}{\partial y_x}\right]\eta(x)\, dx = 0$$

This equation must be satisfied for arbitrary $\eta(x)$ and gives a condition on $y(x)$.

We can rewrite this as

$$\delta J = \int_{x_1}^{x_2} \left(\frac{\partial f}{\partial y} - \frac{d}{dx}\frac{\partial f}{\partial y_x} \right) \eta(x)\, \delta\alpha \, dx = 0$$

or $$\delta J = \int_{x_1}^{x_2} \left(\frac{\partial f}{\partial y} - \frac{d}{dx}\frac{\partial f}{\partial y_x} \right) \delta y \, dx = 0$$

$$\Rightarrow \qquad \frac{\partial f}{\partial y} - \frac{d}{dx}\frac{\partial f}{\partial y_x} = 0 \qquad (1)$$

This is the <u>Euler Equation</u>.

It is important to note the meaning of d/dx in the Euler equation. If $f = f(y(x), y_x, x)$ then

$$\frac{df}{dx} = \underbrace{\frac{\partial f}{\partial x}}_{\substack{\text{EXPLICIT} \\ x \text{ DEPENDENCE}}} + \underbrace{\frac{\partial f}{\partial y}\cdot\frac{dy}{dx} + \frac{\partial f}{\partial y_x}\frac{d^2 y}{dx^2}}_{\text{IMPLICIT } x \text{ DEPENDENCE (VIA } y \text{ AND } y_x\text{)}} \qquad (2)$$

where we have used $\frac{dy_x}{dx} = \frac{d^2 y}{dx^2}$.

There is an equivalent form of the Euler equation:

$$\frac{\partial f}{\partial x} - \frac{d}{dx}\left(f - y_x \frac{\partial f}{\partial y_x} \right) = 0 \qquad (3)$$

This can be derived as follows:

From (2) we have

$$\frac{\partial f}{\partial x} - \frac{df}{dx} = -\frac{\partial f}{\partial y}\frac{dy}{dx} - \frac{\partial f}{\partial y_x}\frac{d^2y}{dx^2}$$

But $$\frac{d}{dx}\left(y_x \frac{\partial f}{\partial y_x}\right) = \frac{d^2y}{dx^2}\cdot\frac{\partial f}{\partial y_x} + y_x \frac{d}{dx}\left(\frac{\partial f}{\partial y_x}\right)$$

Hence

$$\frac{\partial f}{\partial x} - \frac{d}{dx}\left(f - y_x\frac{\partial f}{\partial y_x}\right)$$

$$= \frac{\partial f}{\partial x} - \frac{df}{dx} + \frac{d}{dx}\left(y_x \frac{\partial f}{\partial y_x}\right)$$

$$= -\frac{\partial f}{\partial y}\frac{dy}{dx} - \cancel{\frac{\partial f}{\partial y_x}\frac{d^2y}{dx^2}} + \cancel{\frac{\partial f}{\partial y_x}\frac{d^2y}{dx^2}} + y_x\frac{d}{dx}\left(\frac{\partial f}{\partial y_x}\right)$$

$$= -\frac{\partial f}{\partial y}y_x + y_x\frac{d}{dx}\left(\frac{\partial f}{\partial y_x}\right)$$

$$= -y_x\left(\frac{\partial f}{\partial y} - \frac{d}{dx}\left(\frac{\partial f}{\partial y_x}\right)\right) = 0 \quad \text{by (1)}$$

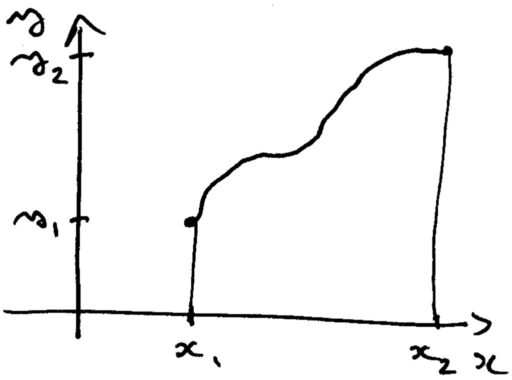
(assuming $y_x \neq 0$)

From the alternative form, (3), if f does not depend on x explicitly, then

$$f - y_x\frac{\partial f}{\partial y_x} = \text{const.}$$

5.2 EXAMPLES OF EULER EQUATION'S APPLICATION

5.2.1 STRAIGHT LINE



To show that the shortest distance between two points in a plane is a straight line.

The length of the path is given by

$$L = \int_{x_1}^{x_2} \sqrt{1 + \left(\frac{dy}{dx}\right)^2}\, dx$$

Let $f = \sqrt{1 + y_x^2}$

$$\frac{\partial f}{\partial y} = 0$$

$$\frac{\partial f}{\partial y_x} = \frac{y_x}{\sqrt{1 + y_x^2}}$$

But
$$\frac{\partial f}{\partial y} - \frac{d}{dx}\left(\frac{\partial f}{\partial y_x}\right) = 0$$

$$\Rightarrow \quad -\frac{d}{dx}\left[\frac{y_x}{\sqrt{1 + y_x^2}}\right] = 0$$

$$\Rightarrow \quad \frac{y_x}{\sqrt{1 + y_x^2}} = \sqrt{C} \quad \text{(constant)}$$

$$\Rightarrow \quad y_x^2 = c\left(1 + y_x^2\right)$$

$$\Rightarrow \quad y_x^2 (1 - c) = c$$

$$y_x = \sqrt{\frac{c}{1-c}}$$

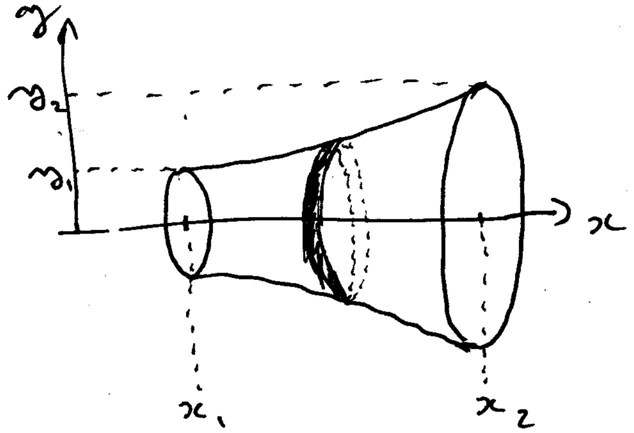
$$\Rightarrow \quad y = \sqrt{\frac{c}{1-c}}\, x + c' \qquad \text{(by integration)}$$

This is the equation of a straight line.

Explicitly,

$$y = \left(\frac{y_2 - y_1}{x_2 - x_1}\right) x + \left(\frac{x_2 y_1 - x_1 y_2}{x_2 - x_1}\right)$$

5.2.2 SOAP FILM



Consider the surface composed of a soap film connecting two parallel, coaxial, circular wires. We are required to minimize the surface area.

$$I = \int_{x_1}^{x_2} 2\pi y(x) \sqrt{1 + \left(\frac{dy}{dx}\right)^2}\, dx \qquad \Rightarrow \quad f = 2\pi y \sqrt{1 + y_x^2}$$

$$\frac{\partial f}{\partial y} = 2\pi \sqrt{1 + y_x^2} \quad ; \quad \frac{\partial f}{\partial y_x} = \frac{2\pi y\, y_x}{\sqrt{1 + y_x^2}}$$

Now use the second form of the Euler equation

$$f - y_x \frac{\partial f}{\partial y_x} = 2\pi y\sqrt{1+y_x^2} - \frac{2\pi y\, y_x^2}{\sqrt{1+y_x^2}} = 2\pi c_1$$

(const.)

$$\Rightarrow \quad y(1+y_x^2) - y\,y_x^2 = c_1\sqrt{1+y_x^2}$$

$$\Rightarrow \quad y = c_1\sqrt{1+y_x^2}$$

$$\text{or} \quad y_x = \sqrt{\frac{y^2}{c_1^2} - 1}$$

$$\Rightarrow \quad \int dx = \int \frac{dy}{\sqrt{(y^2/c_1^2)-1}}$$

$$\Rightarrow \quad x = c_1 \cosh^{-1}\left(\frac{y}{c_1}\right) + c_2$$

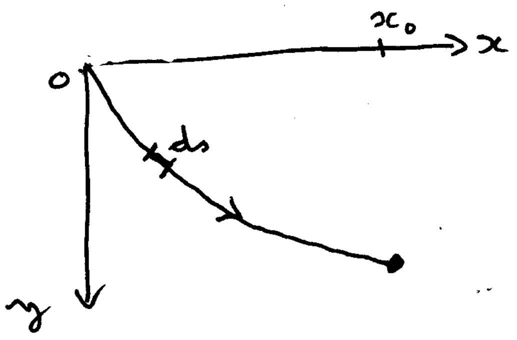
$$\Rightarrow \quad y = c_1 \cosh\left(\frac{x - c_2}{c_1}\right)$$

c_1 and c_2 are obtained from (x_1, y_1) and (x_2, y_2)

For $x_1 = -x_0$, $x_2 = x_0$, $y_1 = y_2 = 1$ we have $c_2 = 0$ by symmetry and

$$c_1 \cosh\left(\frac{x_0}{c_1}\right) = 1$$

5.2.3 THE BRACHISTOCHRONE PROBLEM



What shape wire minimizes the time it takes for a bead to slide down it?

From the energy equation : $\frac{1}{2}mv^2 - mgy = 0$

$$\Rightarrow \quad v = \sqrt{2gy}$$

The integral we want to minimize is

$$\int dt = \int \frac{ds}{v} = \int \frac{ds}{\sqrt{2gy}} = \frac{1}{\sqrt{2g}} \int_0 \frac{\sqrt{1+y_x^2}}{\sqrt{y}}\, dx$$

We have $\quad f(y, y_x, x) = f(y, y_x) = \dfrac{\sqrt{1+y_x^2}}{\sqrt{y}}$

$$\Rightarrow \frac{\partial f}{\partial x} = 0 \quad \Rightarrow \quad f - y_x \frac{\partial f}{\partial y_x} = \text{const.}$$

$$\left(\text{from} \quad \frac{\partial f}{\partial x} - \frac{d}{dx}\left(f - y_x \frac{\partial f}{\partial y_x} \right) = 0 \right)$$

$$\frac{\partial f}{\partial y_x} = \frac{y_x}{\sqrt{y}\sqrt{1+y_x^2}}$$

$$f - y_x \frac{\partial f}{\partial y_x} = \frac{\sqrt{1+y_x^2}}{\sqrt{y}} - \frac{y_x^2}{\sqrt{y}\sqrt{1+y_x^2}} = \frac{1}{\sqrt{y}}\left(\frac{1+y_x^2 - y_x^2}{\sqrt{1+y_x^2}} \right)$$

$$\Rightarrow \quad f - y_x \frac{\partial f}{\partial y_x} = \frac{1}{\sqrt{y}\sqrt{1+y_x^2}} = \text{const.}$$

$$\Rightarrow \quad y\left(1+y_x^2\right) = a \quad (\text{const.})$$

Make the substitution $y = a\sin^2\theta$

$$\Rightarrow \quad y_x^2 = \frac{a}{y} - 1 = \frac{1}{\sin^2\theta} - 1 = \frac{1-\sin^2\theta}{\sin^2\theta}$$

$$= \frac{\cos^2\theta}{\sin^2\theta}$$

$$\Rightarrow \quad y_x = \cot\theta$$

$$\Rightarrow \quad \frac{d\theta}{dx} = \frac{dy}{dx} \Big/ \frac{dy}{d\theta} = \frac{\cot\theta}{2a\sin\theta\cos\theta} = \frac{1}{2a}\operatorname{cosec}^2\theta$$

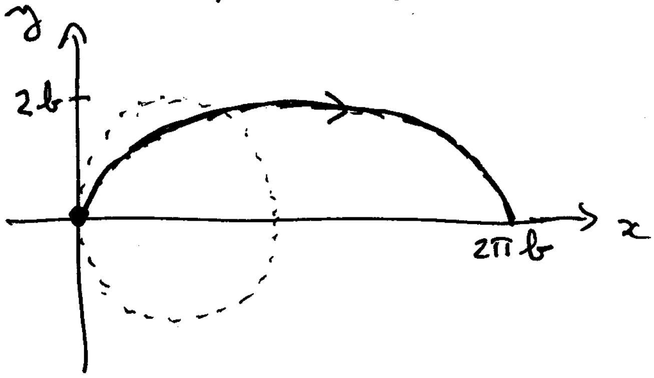
$$\Rightarrow \quad dx = 2a\sin^2\theta \, d\theta$$

$$\Rightarrow \quad x = \frac{a}{2}\left(2\theta - \sin 2\theta\right)$$ (using $y=0$ when $x=0$)

We can write this as a parametric solution

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

where $\phi = 2\theta$, $b = a/2$.



This is the parametric equation of a cycloid – the path followed by a point on the rim of a moving circle.

5.3 EULER EQUATION FOR SEVERAL DEPENDENT VARIABLES

We assume that f is a function of several dependent variables $y_1, y_2, \ldots, y_n$ which all depend on the independent variable x.

Then $$J = \int f(y_1(x), \ldots, y_n(x), y_1'(x), \ldots, y_n'(x), x)\,dx$$

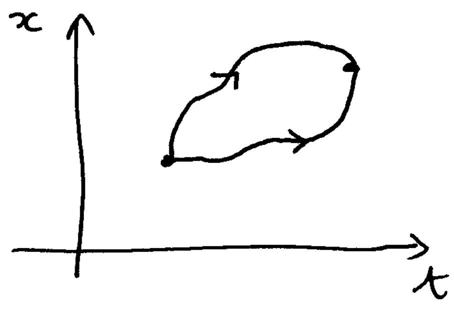
where $'$ denotes d/dx.

We have
$$\delta J = \int \left[\sum_i \frac{\partial f}{\partial y_i} \delta y_i + \sum_i \frac{\partial f}{\partial y_i'} \delta y_i' \right] dx = 0$$

$$\Rightarrow \quad \frac{\partial f}{\partial y_i} - \frac{d}{dx}\left(\frac{\partial f}{\partial y_i'}\right) = 0 \quad \forall i$$

(compare with first form of the Euler equation,
$$\frac{\partial f}{\partial y} - \frac{d}{dx}\frac{\partial f}{\partial y'} = 0 \quad)$$

5.4 THE LAGRANGIAN



Consider different paths between fixed start and end points in spacetime.

The Lagrangian, L, is given by

$$L = K.E. - P.E. = \frac{1}{2} m \dot{x}^2 - V(x)$$

The action, S, is then given by

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

The action principle requires that $\delta S = 0$ and so a variational principle applies.

Since

$$\frac{\partial L}{\partial x} = -\frac{\partial V}{\partial x} \quad \text{and} \quad \frac{\partial L}{\partial \dot{x}} = m\dot{x}$$

Euler equation gives

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

$$\Rightarrow \qquad m\ddot{x} = -\frac{dV}{dx}$$

which is the equation of motion when a force $-dV/dx$ is applied.

In the 3D case the Lagrangian is

$$L = \frac{1}{2} m \underset{\sim}{v}^2 - V(\underset{\sim}{x})$$

where $\underset{\sim}{x} = (x, y, z)$

$$\Rightarrow \quad m\ddot{\underset{\sim}{x}} = -\nabla V(\underset{\sim}{x})$$

Hamilton's principle of classical mechanics asserts that the motion of a system from time t_1 to t_2 is such that the time integral of the Lagrangian (i.e. the action) has a stationary value.

i.e. $$\delta \int_{t_1}^{t_2} L(x_1, x_2, \ldots, x_n, \dot{x}_1, \dot{x}_2, \ldots, \dot{x}_n; t)\, dt = 0$$

The resulting Euler equations are usually called the Lagrangian equations of motion:

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{x}_i} - \frac{\partial L}{\partial x_i} = 0 \qquad \forall i$$

The Lagrangian equations can be derived from Newton's equations and vice versa. However, there are many advantages to using Lagrange's equations (scalar quantities, coordinates selected to match physical problem, equations invariant w.r.t. choice of coordinate system, etc.). The Lagrangian formulation can be extended to other fields and leads to their unification (e.g. quantization of Lagrangian particle mechanics provided model for quantization of EM fields and led to gauge theory of quantum electrodynamics.)

There is a link between symmetry and conservation laws. Noether's Theorem states that every differentiable symmetry of the action of a system has a corresponding conservation law.

e.g. $$\frac{\partial L}{\partial t} = 0 \quad \Rightarrow \quad H \equiv L - \dot{x}\frac{\partial L}{\partial \dot{x}} = \text{const.}$$

$\Rightarrow$ the energy of the system (i.e. the Hamiltonian) is conserved.

5.5 EULER EQUATION FOR SEVERAL INDEPENDENT VARIABLES

Consider
$$J = \iiint f(u, u_x, u_y, u_z, x, y, z)\,dx\,dy\,dz$$

where f is a function of u which itself depends on many independent variables. Euler equation is now

$$\frac{\partial f}{\partial u} - \frac{\partial}{\partial x}\frac{\partial f}{\partial u_x} - \frac{\partial}{\partial y}\frac{\partial f}{\partial u_y} - \frac{\partial}{\partial z}\frac{\partial f}{\partial u_z} = 0.$$

For example, in electrostatics the energy density is $\frac{1}{2}\varepsilon_0 E^2 = \frac{1}{2}\varepsilon_0 (\nabla\varphi)^2$

and we want to minimize
$$J = \iiint (\varphi_x^2 + \varphi_y^2 + \varphi_z^2)\,dx\,dy\,dz$$

$$\Rightarrow \quad \varphi_{xx} + \varphi_{yy} + \varphi_{zz} = 0 \quad \text{(Laplace's equation)}$$

For a scalar field

$$S = \int dt \int d^3x \, \frac{1}{2} \left\{ (\partial_0 \varphi)^2 - (\partial_i \varphi)^2 - (m^2 \varphi^2) \right\}$$

POTENTIAL

where $L = (\partial_0 \varphi)^2 - (\partial_i \varphi)^2 - (m^2 \varphi^2)$

$$\partial_0 \frac{\partial L}{\partial(\partial_0 \varphi)} + \partial_i \frac{\partial L}{\partial(\partial_i \varphi)} = \frac{\partial L}{\partial \varphi}$$

$$\Rightarrow \quad \Box \varphi = m^2 \varphi \quad \text{(Klein-Gordon equation)}$$

where $\Box \equiv \dfrac{\partial^2}{\partial t^2} - \dfrac{\partial^2}{\partial x^2} - \dfrac{\partial^2}{\partial y^2} - \dfrac{\partial^2}{\partial z^2}$

Einstein's field equations can be derived from the action,

$$S = \int R \sqrt{-g} \, d^4x$$

giving $\quad R_{\mu\nu} - \frac{1}{2} R \, g_{\mu\nu} = T_{\mu\nu}$