An example of a variational problem

An illustrative example of a variational problem is the one where one is asked to find the curve that corresponds to the shortest path between two points on a surface. The problem has a particularly simple solution on flat surface (plane).

In rectangular coordinates a straight line is uniquely determined by fixing two points and it can also be described by the differential equation $\frac{d^2y}{dx^2}=0$ combined with the condition that the values of $y(x)$ at $x=x_1$ (point 1) and $x=x_2$ (point 2) are provided.

Another and more general way to look at the problem is to minimize the path between the two points, i.e.

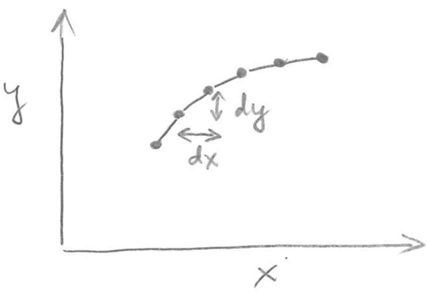
$$\int dl = \min .$$

One can imagine that the two (fixed) points are being connected by all possible curves and among these curves we select the one which yields the smallest possible value of the integral. An important property of such a formulation of the problem is that it does not depend on the choice of particular coordinates.

After introducing rectangular coordinates x and y (as an example), the problem boils down to finding a function $y(x)$ for which $y(x_1)$ and $y(x_2)$ have given fixed values and the integral

$$I = \int_{x_1}^{x_2} \sqrt{1+[y'(x)]^2}\, dx$$

takes the minimum value



$$dy = y'dx$$
$$dl = \sqrt{dx^2+dy^2}$$
$$= \sqrt{1+\left(\frac{dy}{dx}\right)^2}\, dx$$

when $y(x) = y_0(x)$

In the vicinity of $y_0(x)$, $y(x)$ can be represented as

$$y_0(x) + \delta y(x) = y_0(x) + \varepsilon \eta(x)$$

where $\eta(x)$ is an arbitrary function that satisfies

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

and ε is an infinitesimal parameter. The derivative of y in the same vicinity could be represented as

$$y_0'(x) + \varepsilon \eta'(x)$$

For the minimum of functional $I[y(x)]$ we must require that $\delta I = 0$, or, alternatiely,

$$I[y(x) + \varepsilon \eta(x)] - I[y(x)] = 0$$

up to the first order in ε. Taking into account that

$$\sqrt{1+(y_0' + \varepsilon \eta')^2} = \sqrt{1+y_0'^2} + \varepsilon \frac{y_0' \eta'}{\sqrt{1+y_0'^2}} + O(\varepsilon^2)$$

we obtain:

$$I[y_0 + \varepsilon \eta] - I[y_0] = \int_{x_1}^{x_2} \left(\sqrt{1+y_0'^2} + \varepsilon \frac{y_0' \eta'}{\sqrt{1+y_0'^2}} + \dots \right) dx - \int_{x_1}^{x_2} \sqrt{1+y_0'^2}\, dx$$

$$= \varepsilon \int_{x_1}^{x_2} \frac{y_0' \eta'}{\sqrt{1+y_0'^2}}\, dx = 0$$

The above equality should take place for any η (not η') that satisfies the condition $\eta(x_1) = \eta(x_2)$. Now let us integrate by parts. This will give us an integral that contains η (not η'):

$$\int_{x_1}^{x_2} \frac{y_0'}{\sqrt{1+y_0'^2}} \frac{d\eta}{dx}\, dx = \underbrace{\left\{ \frac{y_0' \eta}{\sqrt{1+y_0'^2}} \right\} \Bigg|_{x_1}^{x_2}}_{\substack{\| \\ 0,\ \text{because } \eta(x_1) = \eta(x_2) = 0}} - \int_{x_1}^{x_2} \eta \frac{d}{dx} \left(\frac{y_0'}{\sqrt{1+y_0'^2}} \right) dx = 0$$

The last integral vanishes only if

$$\frac{d}{dx}\left(\frac{y_0'}{\sqrt{1+y_0'^2}}\right) = 0$$

for if it were not satisfied everywhere we could choose $\eta(x)$ in such a way that it is always positive where $\frac{d}{dx}\left(\frac{y_0'}{\sqrt{1+y_0'^2}}\right)$ is positive and choose it negative where this expression is negative. Therefore it follows that

$$y_0' = \text{const} \qquad \text{and} \qquad y_0'' = 0$$

The latter differential equation yields a straight line.

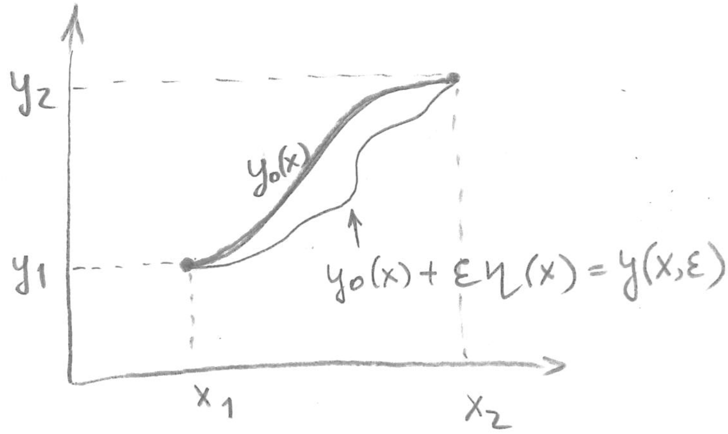
General discussion of variational principles

Let us consider the following one-dimensional problem. Given an integrable function $F(y, y', x)$ defined on a path $y = y(x)$ between two values x_1 and x_2, where $y' = \frac{dy}{dx}$, we wish to find a particular path such as the line integral

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

takes an extremum value (e.g. I has a stationary value relative to path differing infinitesimally from the $y(x)$ which realizes the extremum)

Our problem is transformed into an elementary extremum value problem by covering the ensemble of all physically meaningful paths by a parametric representation:



$$y(x, \varepsilon) = y_0(x) + \varepsilon \eta(x)$$

where ε is a parameter constant for every path, $\eta(x)$ is an arbitrary differentiable funtion that vanishes at x_1 and x_2: $\eta(x_1) = \eta(x_2) = 0$.

The desired curve is given by $y_0(x) = y(x, 0)$

For any such parametric family of curves I is also a funtion of ε

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

The condition for an extremum value of integral I is then

$$\left.\frac{dI}{d\varepsilon}\right|_{\varepsilon=0} = 0$$

The differentiation under the integral symbol (allowed if F is continuously differentiable with respect to ε) yields

$$\frac{dI}{d\varepsilon} = \int_{x_1}^{x_2} \left(\frac{\partial F}{\partial y}\frac{\partial y}{\partial \varepsilon} + \frac{\partial F}{\partial y'}\frac{\partial y'}{\partial \varepsilon} + \frac{\partial F}{\partial x}\underbrace{\frac{\partial x}{\partial \varepsilon}}_{0} \right) dx = \int_{x_1}^{x_2} \left(\frac{\partial F}{\partial y}\eta + \frac{\partial F}{\partial y'}\eta' \right) dx$$

The second term can be integrated by parts

$$\int_{x_1}^{x_2} \frac{\partial F}{\partial y'}\frac{\partial \eta}{\partial x}\, dx = \underbrace{\left[\frac{\partial F}{\partial y'}\eta\right]\Bigg|_{x_1}^{x_2}}_{0,\ \text{because } \eta(x_1)=\eta(x_2)=0} - \int_{x_1}^{x_2} \left(\frac{d}{dx}\frac{\partial F}{\partial y'} \right) \eta\, dx$$

Then the extremum condition becomes

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

The "fundamental lemma" of the calculus of variations says that if

$$\int_{x_1}^{x_2} G(x)\, \eta(x)\, dx = 0$$

for any arbitrary function $\eta(x)$ continuous through the second derivative then $G(x)$ must vanish. Therefore we can write

$$\frac{\partial F}{\partial y} - \frac{d}{dx}\left(\frac{\partial F}{\partial y'}\right) = 0 \quad \leftarrow \text{Euler-Lagrange equation}$$

The solution to the Euler-Lagrange equation, a differential equation of second order, together with boundary conditions, yields the path sought.

Let us define the variation of a function $y(x,\varepsilon)$ as

$$\delta y = y(x,\varepsilon) - y(x,0) = \left.\frac{\partial y}{\partial \varepsilon}\right|_{\varepsilon=0} \cdot \varepsilon \qquad \varepsilon \to 0$$

Then the variational problem can be formulated as

$$\delta \int_{x_1}^{x_2} F(y(x), y'(x), x)\, dx = 0$$

Function F can also include constraints by means of Lagrange multipliers.

<u>Generalization to many independent variables and derivation of Lagrange's equations</u>

A variation of the integral I in case of multiple independent variables $y_i(x)$

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

is obtained by considering I as a function of a parameter that labels sets of curves

$$y_1(x,\varepsilon) = y_1(x,0) + \varepsilon \eta_1(x)$$

$$\vdots$$

$$y_n(x,\varepsilon) = y_n(x,0) + \varepsilon \eta_n(x)$$

where $\eta_1 \ldots \eta_n$ are independent arbitrary functions that vanish at x_1 and x_2

$$\delta I = \frac{\partial I}{\partial \varepsilon}\varepsilon = \int_{x_1}^{x_2} \sum_{i=1}^{n} \left(\frac{\partial F}{\partial y_i}\frac{\partial y_i}{\partial \varepsilon}\varepsilon + \frac{\partial F}{\partial y_i'}\frac{\partial y_i'}{\partial \varepsilon}\varepsilon \right) dx$$

Again we integrate by parts the second term

$$\int_{x_1}^{x_2} \frac{\partial F}{\partial y_i'}\frac{\partial^2 y_i}{\partial \varepsilon \partial x}\, dx = \underbrace{\frac{\partial F}{\partial y_i'}\frac{\partial y_i}{\partial \varepsilon}\Big|_{x_1}^{x_2}}_{0} - \int_{x_1}^{x_2} \frac{\partial y_i}{\partial \varepsilon}\frac{d}{dx}\left(\frac{\partial F}{\partial y_i'}\right) dx$$

Then

$$\delta I = \int_{x_1}^{x_2} \sum_{i=1}^{n} \left(\frac{\partial F}{\partial y_i} - \frac{d}{dx}\frac{\partial F}{\partial y_i'} \right) \delta y_i \, dx \qquad \text{where } \delta y_i = \frac{\partial y_i}{\partial \varepsilon}\Big|_{\varepsilon=0} \cdot \varepsilon$$

By the extension of the "fundamental lemma" the condition of extrema requires that all coefficients by δy_i vanish:

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

If we now change notations and replace

$$F \to L \qquad y_i \to q_i \qquad y_i' \to \dot{q}_i \qquad x \to t$$

then the principle of least action $\delta S = 0$

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

yields

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{q}_i} - \frac{\partial L}{\partial q_i} = 0 \qquad i = 1, \dots, n$$

under assumption that q_i are all independent, which requires that the constraints be holonomic