4. Calculus of Variations

Introduction - Typical Problems

The calculus of variations generalises the theory of maxima and minima.

Example (a): *Shortest distance between two points.* On a given surface (e.g. a plane), find the shortest curve between two points. The length L of a curve $y = y(x)$ between the values $x = a$ and $x = b$ is given by the integral

$$L = \int_a^b \sqrt{1 + \left(\frac{dy}{dx}\right)^2} dx = \int_a^b \sqrt{1 + (y')^2} dx. \tag{4.1}$$

The value of L depends upon the function $y(x)$ which appears as an argument in the integrand, an arbitrary continuous function with piecewise continuous drivative.

Example (b) *Minimal surface of revolution. Let the curve* $y = y(x) \geq 0$ which passes through the points $y(a) = y_1$, $y(b) = y_2$ be rotated about the x-axis. The resulting surface between $x = a$ and $x = b$ has surface area A given by

$$A = 2\pi \int_a^b y\sqrt{1 + (y')^2} dx.$$

The curve $y = y(x)$ which gives the smallest surface of revolution is found by minimising the integral.

Example (c): *Isoperimetric problem.* Find a closed plane curve of given perimeter which encloses the greatest area. The area may be written as

$$A = \int_*^{**} y dx$$

between some limits (and taking care about regions above and below the x-axis) and is subject to a *constraint* of the form equ. (4.1) where now L (the perimeter) is fixed in length.

Common ingredients:

(1) An integral with an integrand containing an arbitrary function

(2) A problem which asks for a minimum or maximum.

Note the geometric language and that the concepts of "curve" and "function" do not coincide.

Functionals

Let S be a (vector) space (i.e. set of functions + algebraic structure - closed under addition of functions $f + g$ and multiplication by a scalar, λf).

Example: Let $C^1(a, b)$ denote the set of functions continuous on the closed interval $[a, b]$ with piecewise continuous first order derivatives.

4.2 Definition: A *functional* is a mapping (function) from a space of functions into the underlying field (usually the real or complex numbers)

$$\begin{aligned} \phi &: S \to \mathbb{R}(\text{or } \mathbb{C}) \\ \phi &: y \mapsto c \\ y &\in S, \qquad c \in \mathbb{R}(\ \mathbb{C}) \\ \phi(y) &= c, \qquad \text{or} \qquad \phi y = c. \end{aligned}$$

S is called the *domain* of the functional and the *space of admissible functions.*
Examples (a): Evaluating a function is a functional, e.g. for $d \in [a, b]$

$$\phi(y(x)) = y(c)$$

(b) $\phi(y(x)) = y''(7) + y(3)$.
(c) The Dirac delta "function" is a functional:

$$\int_a^b \delta(x-d) f(x)\, dx = f(d)$$

(d) Let $y(x) \in C^1(a, b)$,

$$\phi(y(x)) = \int_a^b \left[(y(x))^2 - (y'(x))^2 \right] dx.$$

(e) The area A of a surface $z = z(x, y)$ lying above the region G in the xy-plane is given by

$$A = \iint\limits_G \sqrt{1 + z_x^2 + z_y^2}\, dx dy$$

where $z_x = \partial z / \partial x$, $z_y = \partial z / \partial y$ and is a functional of the argument function $z(x, y)$.

The calculus of variations is concerned with finding extrema or stationary values of functionals.

Consider functionals (defined by integrals) of the form

$$I(y) = \int_a^b F(x, y, y')\, dx. \tag{4.3}$$

The integrand F depends on the function $y(x)$, its derivative $y'(x)$ and the independent variable x.
In order to discuss maxima and minima we need to define what is meant by two functions being "close together", i.e. we need a notion of distance.
4.4 Definition:
Given $h \in \mathbb{R}$, $h > 0$, a function $y_1(x)$ lies in the *neighbourhood* $N_h(y)$ of the function $y(x)$ if

$$|y(x) - y_1(x)| < h$$

$\forall x \in [a, b]$.
Sometimes it is necessary to use a more refined definition:
4.4a Definition:
Given $h \in \mathbb{R}$, $h > 0$, a function $y_1(x)$ lies in the *first order neighbourhood* $N_h(y)$ of the function $y(x)$ if

$$|y(x) - y_1(x)| < h$$

and

$$|y'(x) - y_1'(x)| < h$$

$\forall x \in [a, b]$.
4.5 Fundamental problem of the Calculus of Variations: Find a function $y = y_0(x) \in S(a, b)$ for which the functional $I(y)$ takes an extremal value (i.e. maximum or minimum) value with respect to all $y(x) \in S(a, b)$ in $N_h(y)$ for sufficiently small h.
$y = y_0(x)$ is called an *extremal function.*
Note: There is no guarantee a solution exists for this problem (unlike maxima and minima of functions continuous on a closed interval where existence is guaranteed).
Example: The shortest distance between two points A, B is a straight line but there is no curve of shortest length which departs from A and arrives at B at right angles to the line segment AB.

Euler-Lagrange Equations

4.6 Fundamental lemma in the Calculus of Variations

Let $f(x)$ be continuous in $[a, b]$ and let $\eta(x)$ be an arbitrary function on $[a, b]$ such that η, η', η'' are continuous and $\eta(a) = \eta(b) = 0$. If

$$\int_a^b f(x)\, \eta(x)\, dx = 0$$

for **all** such $\eta(x)$ then $f(x) \equiv 0$ on $[a, b]$.
Proof: Suppose to the contrary w.l.o.g. that $f(x) > 0$ at, say, $x = \xi$.
Then there is a neighbourhood N, $\xi_0 < x < \xi_1$ in which $f(x) > 0$. Let

$$\eta(x) = \begin{cases} (x - \xi_0)^4 (x - \xi_1)^4 & \text{for } x \in N \\ 0 & \text{elsewhere} \end{cases}.$$

Then

$$\int_a^b f(x)\, \eta(x)\, dx > 0$$

contradicting the hypothesis**.**
4.7 Theorem: Euler-Lagrange Equations
The extremal function $y = y_0(x)$ for the functional (4.3)

$$I(y) = \int_a^b F(x, y, y')\, dx$$

where a, b, $y(a)$, $y(b)$ are given, F is twice continuously differentiable w.r.t. its arguments and $y''(x)$ is continuous, satisfies the equation

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

This is a necessary, but not sufficient, condition for an extremal function.
(Every extremal function $y_0(x)$ satisfies the Euler-Lagrange equation,
Not every function $f(x)$ which satisfies the Euler-Lagrange equation is an extremal function.)
Proof: Let $y = y(x)$ be a variable function and let $y = y_0(x)$ be an extremal function for the functional $I(y)$, i.e. $I(y_0)$ takes an extreme value (maximum or minimum).
Let $\eta(x)$ be as in lemma (4.6) and define the function

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

where $\varepsilon > 0$ is a parameter, and write

$$y = y_0 + \delta y.$$

Then $\delta y = \varepsilon\eta(x)$ is called the *variation* of $y = y_0(x)$.
For ε sufficiently small, y lies in an arbitrarily small neighbourhood $N_h(y_0)$ of $y_0(x)$.
Now, the integral $I(y) = I(y_0 + \varepsilon\eta)$ is a function of $\Phi(\varepsilon)$ of ε.
Let

$$\delta I = I(y_0 + \varepsilon\eta) - I(y_0) = \Phi(\varepsilon) - \Phi(0)$$

then

$$\delta I = \int_a^b \left[F(x, y_0 + \varepsilon\eta, y_0' + \varepsilon\eta') - F(x, y_0, y_0')\right] dx.$$

Expanding $\Phi(\varepsilon)$ in a Maclaurin series with respect to ε to first order gives

$$\Phi(\varepsilon) = \Phi(0) + \Phi'(0)\,\varepsilon$$

and so to first order in ε

$$\delta I = \Phi'(0)\,\varepsilon = \left(\int_a^b \left[\frac{\partial F}{\partial y}\frac{dy}{d\varepsilon} + \frac{\partial F}{\partial y'}\frac{dy'}{d\varepsilon}\right] dx\bigg|_{y=y_0}\right)\varepsilon$$

so

$$\Phi'(0) = \int_a^b \left(\frac{\partial F}{\partial y}\eta + \frac{\partial F}{\partial y'}\eta'\right) dx.$$

Now, since $y = y_0(x) + \varepsilon\eta(x)$ where $y_0(x)$ is the extremal function, it follows that $\delta I = 0$ for all $\eta(x)$ with $y \in N_h(y_0)$.
For suppose not, then replace $\eta(x)$ by $-\eta(x)$ and the sign of δI changes, contradicting the fact that $y_0(x)$ is the extremal function!

Hence $I(y)$ takes a *stationary value* for $y = y_0(x)$.
Thus

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

Integrating the second integral by parts with

$$u = \frac{\partial F}{\partial y'}, \qquad v' = \eta'$$

then

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

gives

$$\left[\frac{\partial F}{\partial y'}\eta \right]_a^b + \int_a^b \left(\frac{\partial F}{\partial y}\eta - \frac{d}{dx}\left(\frac{\partial F}{\partial y'} \right)\eta \right) dx = 0.$$

Now, $\eta(a) = \eta(b) = 0$ so

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

Hence, since $\eta(x)$ is arbitrary, using the Fundamental Lemma of the Calculus of Variations, we obtain

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

Note: For any such $f(x, y, y')$

$$\frac{df}{dx} = \frac{\partial f}{\partial x} + \frac{\partial f}{\partial y}\frac{dy}{dx} + \frac{\partial f}{\partial y'}\frac{dy'}{dx}$$

so putting $f = \partial F/\partial y'$ gives

$$\frac{\partial^2 F}{\partial y'^2} y''(x) + \frac{\partial^2 F}{\partial y' \partial y} y'(x) + \frac{\partial^2 F}{\partial x \partial y'} - \frac{\partial F}{\partial y} = 0.$$

This is a 2nd. order ode!

Example: Find the extremal function for

$$I(y) = \int_0^{\pi/2} \left[(y')^2 - y^2 \right] dx$$

with $y(0) = 0$, $y\left(\frac{1}{2}\pi\right) = 1$.
Here,

$$F(x, y, y') = (y')^2 - y^2$$

so

$$\frac{\partial F}{\partial y} = -2y, \qquad \frac{\partial F}{\partial y'} = 2y'.$$

The Euler-Lagrange equation

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

becomes

$$\frac{d}{dx}(2y')-(-2y)=0$$

i.e.

$$y''+y=0.$$

The general solution is

$$y(x)=A\cos x+B\sin x$$

$y(0)=0\Rightarrow A=0$ and $y\left(\frac{1}{2}\pi\right)=1\Rightarrow B=1$ hence

$$y=\sin x$$

and $I(\sin x)$ is stationary and

$$I(\sin x)=\int_0^{\pi/2}\left(\cos^2 x-\sin^2 x\right)dx=0.$$

Special cases: First integrals of the Euler-Lagrange equations

For special forms of $F(x,y,y')$ the 2nd. order order ode

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

may be integrated to give a "first integral".

4.8 Theorem: Let $F\equiv F(x,y')$ (no y dependence) then

$$\frac{\partial F}{\partial y'}=\text{constant}.$$

Proof:

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

so

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

which integrates w.r.t. x to give the result.

4.9 Theorem: Let $F\equiv F(y,y')$ (no x dependence) then

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

Proof: In general $F \equiv F(x, y, y')$

$$\begin{aligned}\frac{d}{dx}\left(F - y'\frac{\partial F}{\partial y'}\right) &= \frac{\partial F}{\partial x} + \frac{\partial F}{\partial y}y' + \frac{\partial F}{\partial y'}y'' - y''\frac{\partial F}{\partial y'} - y'\frac{d}{dx}\left(\frac{\partial F}{\partial y'}\right) \\ &= \frac{\partial F}{\partial x} + y'\left[\frac{\partial F}{\partial y} - \frac{d}{dx}\left(\frac{\partial F}{\partial y'}\right)\right] \\ &= \frac{\partial F}{\partial x} \text{ by the Euler-Lagrange equation} \\ &= 0 \text{ since, in fact, } F \text{ is independent of } x.\end{aligned}$$

Extensions of the Euler-Lagrange equations

(a) Several unknown functions $y_1(x), ..., y_m(x)$.

Let

$$I(y_1, ..., y_m) = \int_a^b F(x, y_1, ..., y_m, y_1', ..., y_m')\, dx \qquad (4.10)$$

then

$$\frac{d}{dx}\left(\frac{\partial F}{\partial y_j'}\right) - \frac{\partial F}{\partial y_j} = 0, \qquad (4.11)$$

where $j = 1, ..., m$, i.e. m simultaneous ode's for the m unknowns y_j.

(b) Several independent variables $x_1, ..., x_n$. Let $y = y(x_1, ..., x_n)$ and

$$I(y) = \int ... \int_V F\left(x_1, ..., x_n, y, \frac{\partial y}{\partial x_1}, ..., \frac{\partial y}{\partial x_n}\right) dx_1...dx_n \qquad (4.12)$$

(a multiple integral) where V is a region in n-dimensional $(x_1, ..., x_n)$-space then, writing

$$\frac{\partial y}{\partial x_i} \equiv y_{,x_i} \equiv y_{,i} \equiv y_i$$

we have

$$\sum_{i=1}^{n} \frac{\partial}{\partial x_i}\left(\frac{\partial F}{\partial y_{,i}}\right) - \frac{\partial F}{\partial y} = 0, \qquad (4.13)$$

a partial differential equation.

(c) Additional integral constraint - Lagrange multipliers

Suppose that, in addition to Eq. (4.3), $y(x)$ also satisfies an integral constraint of the form

$$J(y) = \int_a^b G(x, y, y')\, dx = C \qquad (4.14)$$

where C is a constant. We wish to find stationary values of the functional

$$I(y) = \int_a^b F(x, y, y')\, dx$$

where $y(x)$ is now not free, but is subject to the constraint Equ.(4.13). In this case, form the functional

$$\begin{aligned} K(y) &= \int_a^b (F - \lambda G)\, dx \\ &= \int_a^b H(x, y, y', \lambda)\, dx, \text{ say} \end{aligned}$$

where λ is a *Lagrange multiplier.*
The corresponding Euler-Lagrange equation is

$$\frac{d}{dx}\left(\frac{\partial H}{\partial y'}\right) - \frac{\partial H}{\partial y} = 0. \qquad (4.15)$$

The general solution of equ. (4.15) contains λ and two integration constants. These are determined by the boundary conditions and the constraint equ. (4.14).
Given several constraints

$$J_k(y) = \int_a^b G_k(x, y, y')\, dx = C_k$$

where $k = 1, ..., p$, define

$$H = F - \sum_{k=1}^{p} \lambda_k G_k.$$

Imposing the constraints gives the values of λ_k (usually of little physical significance).

(d) Several functions of several variables.

Now consider m functions $y_1, ...y_m$ each of n variables $x_1, ...x_n$,

$$y_j = y_j(x_1, ...x_n) = y_j(x_i),$$

where $j = 1, ..., m$ and $i = 1, ..., n$. Let

$$y_{ji} = \frac{\partial y_j}{\partial x_i}$$

and let V denote a region in the n-dimensional $(x_1, ...x_n)$-space, $dV = dx_1...dx_n$.
Consider the functional

$$I(y_1, ..., y_m) = \int_V F\left(x_1, ..., x_n, y_1, ..., y_m, \frac{\partial y_1}{\partial x_1}, ..., \frac{\partial y_m}{\partial x_n}\right) dx_1...dx_n \qquad (4.16)$$

which may be abbreviated by

$$I(y_j) = \int_V F(x_i, y_j, y_{ji})\, dx_1...dx_n.$$

The Euler-Lagrange equations are the m equations

$$\sum_{i=1}^{n} \frac{\partial}{\partial x_i} \left(\frac{\partial F}{\partial y_{ji}} \right) - \frac{\partial F}{\partial y_j} = 0, \qquad (4.17)$$

where $j = 1, ..., m$. Note that, using the double suffix summation convention (dssc) we may abbreviate this as

$$\frac{\partial}{\partial x_i} \left(\frac{\partial F}{\partial y_{ji}} \right) - \frac{\partial F}{\partial y_j} = 0,$$

where, any term containing a repeated subscript (here i) is summed over all possible values (here i takes the values $1, ..., n$). A single subscript occurring in every term of an equation (here j) is callled a free subscript and indicates that there is an equation for each value (here m).

Example: Dirichlet's integral and Poisson's equation.
The above notation was developed as a generalisation of of the original notation for the simplest case, the functional (4.3) and the original Euler-Lagrange equation in theorem (4.7). It is not particularly convenient for considering a single partial differential equation. So, we now change notation. Let $n = 3$ and let the independent variables x_1, x_2, x_3 be replaced by x, y, z (more suitable for problems in three space dimensions). Also replace the dependent variable y (now being used as a coordinate) by u. Thus, let $u = u(x, y, z)$ be a function of three variables. Recall the notation for the gradient function:

$$\nabla u = \left(\frac{\partial u}{\partial x}, \frac{\partial u}{\partial y}, \frac{\partial u}{\partial z} \right)$$

and the notation

$$(\nabla u)^2 = \nabla u \cdot \nabla u = \left(\frac{\partial u}{\partial x} \right)^2 + \left(\frac{\partial u}{\partial y} \right)^2 + \left(\frac{\partial u}{\partial z} \right)^2 = u_x^2 + u_y^2 + u_x^2.$$

Now, Dirichlet's integral is defined by

$$I(u) = \int_V \left[\frac{1}{2} (\nabla u)^2 + fu \right] dV$$

where V is a volume in $\mathbb{R}^3$, $f = f(x, y, z)$ is a known function of x, y, z. Here,

$$F = \frac{1}{2} \left[\left(\frac{\partial u}{\partial x} \right)^2 + \left(\frac{\partial u}{\partial y} \right)^2 + \left(\frac{\partial u}{\partial z} \right)^2 \right] + fu.$$

Hence

$$\frac{\partial F}{\partial u} = f, \qquad \frac{\partial F}{\partial u_x} = u_x, \qquad \frac{\partial F}{\partial u_y} = u_y, \qquad \frac{\partial F}{\partial u_z} = u_z.$$

The Euler Lagrange equations is

$$\frac{\partial}{\partial x} \left(\frac{\partial F}{\partial u_x} \right) + \frac{\partial}{\partial y} \left(\frac{\partial F}{\partial u_y} \right) + \frac{\partial}{\partial z} \left(\frac{\partial F}{\partial u_z} \right) - \frac{\partial F}{\partial u} = 0$$

i.e.

$$\frac{\partial}{\partial x}\left(u_x\right)+\frac{\partial}{\partial y}\left(u_y\right)+\frac{\partial}{\partial z}\left(u_z\right)-f=0$$

or

$$\nabla^2 u=f$$

where ∇^2 is the Laplacian operator

$$\nabla^2=\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}+\frac{\partial^2}{\partial z^2}.$$

Thus, the Dirichlet integral is stationary when the function $u\left(x,y,z\right)$ satisfies Poisson's equation. This is known as Dirichlet's Principle. This is a simple example of a *variational principle*, which turns out to be a very important application of the calculus of variations.