Calculus of Variations with Constraints

We begin with some examples.

**Example 1**
What curve through the points \((x_1, y_1)\) and \((x_2, y_2)\) of given length \(L\) has the maximum area between the curve and the \(x\)–axis?

If \(y(x)\) is a **single-valued function** of \(x\), then

\[
A[y(x)] = \int_{x_1}^{x_2} y(x) \, dx
\]

whereas

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

Therefore the problem is to maximize (1) subject to the constraint (2) and the conditions \(y(x_1) = y_1, y(x_2) = y_2\).

**Example 2**
If in the previous example we do not assume that \(y(x)\) is a single-valued function of \(x\), then it is convenient to suppose that \(x\) and \(y\) are given parametrically, i.e.,

\[
x = x(t), \quad y = y(t) \quad t_1 \leq t \leq t_2
\]

where \(x(t_1) = x_1, x(t_2) = x_2, y(t_1) = y_1, \) and \(y(t_2) = y_2\). Then we have the constraint

\[
\int_{t_1}^{t_2} \sqrt{\left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2} \, dt = L
\]

Also

\[
dA = \frac{1}{2} (xdy - ydx)
\]

since by Green’s Theorem

\[
\oint_C P \, dx + Q \, dy = \iint_R (Q_x - P_y) \, dA
\]

so that

\[
\oint_C \frac{1}{2} (xdy - ydx) = \frac{1}{2} \iint_R [1 - (-1)] \, dA = \iint_R dA = \text{area of } R
\]

Thus
\[ A[x(t), y(t)] = \frac{1}{2} \int_0^{t_2} \left[ x \frac{dy}{dt} - y \frac{dx}{dt} \right] dt \] (4)

The problem is to minimize (4) subject to the constraint (3) and the conditions \( x(t_1) = x_1, x(t_2) = x_2, \) \( y(t_1) = y_1, \) and \( y(t_2) = y_2. \)

The above two examples illustrate problems in which one desires to minimize (or maximize) a given integral subject to a constraint. Several examples of such problems are

\[ \delta \int_{x_1}^{x_2} F(x, y, y') dx = 0 \quad \text{subject to} \quad \int_{x_1}^{x_2} G(x, y, y') dx = \text{constant} \] (5)

\[ \delta \int_{x_1}^{x_2} F(t, x, y, \dot{x}, \dot{y}) dx = 0 \quad \text{subject to} \quad \int_{x_1}^{x_2} G(t, x, y, \dot{x}, \dot{y}) dx = \text{constant} \] (6)

\[ \delta \int_{x_1}^{x_2} F(x, u, v, u', v') dx = 0 \quad \text{subject to} \quad \phi(u, v) = 0 \] (7)

where \( u = u(x), v = v(x). \)

We deal with (5) first. Thus our problem is to make

\[ J[y(x)] = \int_{x_1}^{x_2} F(x, y, y') dx = \text{minimum or maximum} \]

where \( y \) is prescribed as \( y(x_1) = y_1 \) and \( y(x_2) = y_2 \) subject to the condition

\[ J[y(x)] = \int_{x_1}^{x_2} G(x, y, y') dx = k \]

where \( k \) is a given constant.

This problem cannot be attacked by the earlier method of forming \( y + \epsilon \eta \) where \( \eta \) vanishes on the boundary only, for in general such functions do not satisfy the subsidiary condition in a neighborhood of \( \epsilon = 0 \) except at \( \epsilon = 0. \) Since we have two requirements, we therefore consider the function

\[ y(x) + \epsilon_1 \eta_1(x) + \epsilon_2 \eta_2(x) \]

where \( \eta_1 \) and \( \eta_2 \) have continuous derivatives and

\[ \eta_1(x_1) = \eta_1(x_2) = \eta_2(x_1) = \eta_2(x_2) = 0. \]

Then we have

\[ \Phi(\epsilon_1, \epsilon_2) = J[y(x) + \epsilon_1 \eta_1(x) + \epsilon_2 \eta_2(x)] = \int_{x_1}^{x_2} F(x, y + \epsilon_1 \eta_1 + \epsilon_2 \eta_2, y' + \epsilon_1 \eta_1' + \epsilon_2 \eta_2') dx \]

subject to

\[ \Psi(\epsilon_1, \epsilon_2) = J[y(x) + \epsilon_1 \eta_1(x) + \epsilon_2 \eta_2(x)] - k = \int_{x_1}^{x_2} G(x, y + \epsilon_1 \eta_1 + \epsilon_2 \eta_2, y' + \epsilon_1 \eta_1' + \epsilon_2 \eta_2') dx - k = 0 \]

Thus we desire that the function \( \Phi(\epsilon_1, \epsilon_2) \) take on a relative minimum or maximum at all \( \epsilon_1 = \epsilon_2 = 0 \) with respect to all sufficiently small values of \( \epsilon_1 \) and \( \epsilon_2 \) for which \( \Psi(\epsilon_1, \epsilon_2) = 0. \)

This problem is of the form we treated earlier by means of Lagrange multipliers. In particular, if \( \lambda \) is the Lagrange multiplier, then we require
\[
\frac{\partial (\Phi + \lambda \Psi)}{\partial \epsilon_1} \bigg|_{\epsilon_1=\epsilon_2=0} = 0 \tag{1}
\]
\[
\frac{\partial (\Phi + \lambda \Psi)}{\partial \epsilon_2} \bigg|_{\epsilon_1=\epsilon_2=0} = 0 \tag{2}
\]

and
\[
\Psi(\epsilon_1, \epsilon_2) = 0
\]

Now (1) ⇒
\[
\int_{x_1}^{x_2} [F_{y_1} + F_{y_1}'] dx + \lambda \int_{x_1}^{x_2} [G_{y_1} + G_{y_1}'] dx = 0
\]

and (2) ⇒
\[
\int_{x_1}^{x_2} [F_{y_2} + F_{y_2}'] dx + \lambda \int_{x_1}^{x_2} [G_{y_2} + G_{y_2}'] dx = 0
\]

Let
\[
[H]_y = H_y - \frac{d}{dx} H_y
\]

Then as before an integration by parts yields
\[
\int_{x_1}^{x_2} \{[F]_y + \lambda[G]_y\} \eta_1 dx = 0 \tag{3}
\]

and
\[
\int_{x_1}^{x_2} \{[F]_y + \lambda[G]_y\} \eta_2 dx = 0 \tag{4}
\]

If \([G]_y \neq 0\), then we can, say, choose \(\eta_2\) such that \(\int_{x_1}^{x_2} [G]_y \eta_2 dx \neq 0\), and thus \(\lambda\) may be chosen so that (4) holds. However, since \(\eta_1\) is arbitrary, \(\lambda\) will not be such that (3) holds. Therefore it follows from (3) that
\[
[F]_y + \lambda[G]_y = 0
\]
or
\[
\frac{\partial (F + \lambda G)}{\partial y'} - \frac{d}{dx} \left( \frac{\partial (F + \lambda G)}{\partial y'} \right) = 0
\]

is the necessary condition. The general solution of this equation will involve two constants of integration and the constant parameter \(\lambda\). Thus we have 3 constants to satisfy the 3 conditions \(y(x_1) = y_1, y(x_2) = y_2\), and \(\int_{x_1}^{x_2} G(x, y, y') dx = k\).

The above results may be summarized as follows:

**Theorem.** In order to minimize (or maximize) \(\int_{x_1}^{x_2} F dx\) subject to a constraint \(\int_{x_1}^{x_2} G dx = k\) we first write \(H = F + \lambda G\), where \(\lambda\) is a constant, and minimize (or maximize) \(\int_{x_1}^{x_2} H dx\) subject to no constraints. Carry the Lagrange multiplier \(\lambda\) through the calculations, and determine it, together with the constants of integration arising in the Euler equation, so that the constraint \(\int_{x_1}^{x_2} G dx = k\) holds, and the end conditions are satisfied.

**Example** Maximize
\[
A[y(x)] = \int_{x_1}^{x_2} y(x) dx
\]
subject to
\[ L = \int_{x_1}^{x_2} \sqrt{1 + (y')^2} \, dx \]  
and \( y(x_1) = y_1, y(x_2) = y_2 \).

Solution: Without loss of generality we may move the axis and change the scale so that the curve is to pass through \((0, 0)\) and \((1, 0)\). Thus we must maximize
\[ \int_0^1 y(x) \, dx \]
subject to \( y(0) = y(1) = 0 \) and the constraint
\[ \int_0^1 \sqrt{1 + (y')^2} \, dx = L \quad \text{where} \quad L > 1 \]
We form
\[ H = y + \lambda \left( 1 + (y')^2 \right)^{\frac{1}{2}} \]
The Euler equation
\[ -H_y + \frac{d}{dx} H_{y'} = 0 \]
implies
\[ -1 + \lambda \frac{d}{dx} \left\{ \frac{y'}{\left( 1 + (y')^2 \right)^{\frac{1}{2}}} \right\} = 0 \]
Then integrating we have
\[ \lambda \frac{y'}{\left( 1 + (y')^2 \right)^{\frac{1}{2}}} = x - c_1 \]
or
\[ \left\{ \lambda^2 - (x - c_1)^2 \right\} (y')^2 = (x - c_1)^2 \]
Therefore
\[ y' = \pm \frac{(x - c_1)}{\sqrt{\lambda^2 - (x - c_1)^2}} \]
so
\[ y = \pm \left( \lambda^2 - (x - c_1)^2 \right)^{\frac{1}{2}} + c_2 \]
or finally
\[ (y - c_2)^2 + (x - c_1)^2 = \lambda^2 \]
The required curves are arcs of circles. We have three constants to determine. They are determined so that the arc passes through \((0, 0), (0, 1)\) and has length \(L\).

Remark: When \( L = \frac{\pi}{2} \), we have a semicircle since the circumference of the circle is \(2\pi \lambda\). For \( L > \pi/2 \) \( y \) is no longer a single-valued function of \( x \). For such a case it is convenient to employ a parametric representation expressing \( x \) and \( y \) as functions of \( t \), i.e., \( x = x(t) \) and \( y = y(t) \). We are led to the problem
Maximize

\[ I = \frac{1}{2} \int_{t_1}^{t_2} (x \dot{y} - y \dot{x}) \, dt \]

where \( x(t_1) = 0, x(t_2) = 1, y(t_1) = y(t_2) = 0 \) and

\[ J = \int_{t_1}^{t_2} \left[ (\dot{x})^2 + (\dot{y})^2 \right]^{\frac{1}{2}} \, dt = L \]

Remark: It is easily seen that the problem

\[ \delta \int_{x_1}^{x_2} F(t,x,y,\dot{x},\dot{y}) \, dx = 0 \quad \text{subject to} \quad \int_{x_1}^{x_2} G(t,x,y,\dot{x},\dot{y}) \, dx = \text{constant} \]

leads to the Euler equations

\[ \frac{d}{dt} H_x - H_{\dot{x}} = 0 \]
\[ \frac{d}{dt} H_y - H_{\dot{y}} = 0 \]

where \( H = F + \lambda G \).

For our problem

\[ H = \frac{1}{2} \left( x \dot{y} - y \dot{x} \right) + \lambda \left( \dot{x}^2 + \dot{y}^2 \right)^{\frac{1}{2}} \]

These again leads to arcs of circles.

Remark: For the problem

\[ \delta \int_{x_1}^{x_2} F(x,u,v,u',v') \, dx = 0 \quad \text{subject to} \quad \phi(u,v) = 0 \]

the Euler equations are

\[ \frac{d}{dx} \left( \frac{\partial F}{\partial u'} \right) - \frac{\partial F}{\partial u} - \lambda \phi = 0 \]
\[ \frac{d}{dx} \left( \frac{\partial F}{\partial v'} \right) - \frac{\partial F}{\partial v} - \lambda \phi = 0 \]