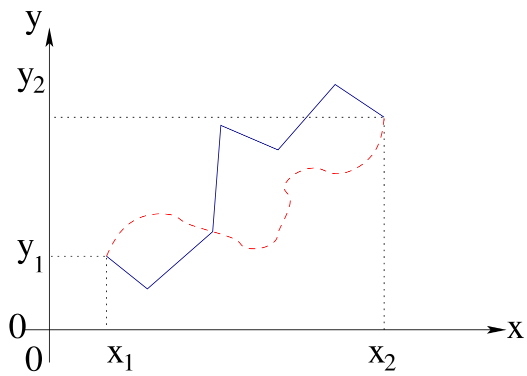

Figure 7:

4 Calculus of Variations

4.1 Introduction

The 'calculus of variations' deals with extremal problems in which the 'thing' being maximized or minimized depends upon a function; the problem is solved by finding which function, from a given set F of available functions, that gives the extreme value. Usually the 'thing' being maximized or minimized is given as an integral involving the function.

Example 4.1. *Let* (x_1, y_1) *and* (x_2, y_2) *be points in the plane. We seek the curve of shortest length joining these points (see Figure 9).*
We restrict the class of curves by insisting that each curve can be written as a graph $y = f(x)$*. Suppose* $x_1 < x_2$*. Then, we could take for our set* F *all graphs of functions that have continuous derivatives and which satisfy* $y_i = f(x_i),\ i = 1, 2,$ *i.e.,*

$$F := \{f \in C^1[x_1, x_2] \ : \ f(x_1) = y_1 \text{ and } f(x_2) = y_2\}.$$

The length of the curve f *or any function* $f \in F$ *is*

$$J(f) := \int_{x_1}^{x_2} (1 + (f'(x))^2)^{1/2} dx.$$

Find $f \in F$ *such that* $J(f)$ *is minimized.*

Definition 4.1. *A functional is a mapping that assigns a real number to each function* f *in some set* F *of functions. (That is, a functional is a mapping* $J : F \to \mathbb{R}$ *where* F *is a set of functions.)*

Ordinary Calculus
Suppose we have $g : S \to \mathbb{R}$ where $S \subset \mathbb{R}^n$ and wish to find a maximum or a minimum. ($g = g(x_1, x_2, \cdots, x_n)$, $\underline{x} \in S$.)
If g is differentiable, this leads to a necessary condition in terms of the partial derivatives of g. Indeed, if $\underline{x}^*$ is a maximizer or a minimizer of g then necessarily

$$\nabla g(\underline{x}^*) = 0.$$

We can also use the concept of directional derivatives.
Recall that the *directional derivatives* of g at $\underline{x_0} \in S$ in the direction of a unit vector $\underline{n}$ is given by

$$D_{\underline{n}} g(\underline{x_0}) := \lim_{\varepsilon \to 0} \frac{g(\underline{x_0} + \varepsilon \underline{n}) - g(\underline{x_0})}{\varepsilon}. \tag{4.1}$$

Then the necessary condition for g to have an extreme value at $\underline{x_0}$ is that the directional derivatives (4.1) should be zero for all directions $\underline{n}$. For a sufficient condition more is needed: we need to know whether the extreme point is a hill, valley or saddle. These are usually reflected in terms of second derivatives.

Remark 4.1.

- (4.1) *also can be written as*
$$\frac{d}{d\varepsilon} g(\underline{x_0} + \varepsilon \underline{n}) =: D_n g(\underline{x_0}).$$
- *Fix* $\underline{n} = e_i$.
$$\frac{d}{d\varepsilon} g(\underline{x_0} + \varepsilon e_i) = \frac{\partial g}{\partial x_i}(\underline{x_0}).$$

Variational Calculus
For an extremal problem involving a functional $J : F \to \mathbb{R}$, we can proceed in an analogous manner by using an analogue of the directional derivative.

Definition 4.2. *Let* $J : F \to \mathbb{R}$ *be a functional defined on a set by functions of* F*. Then the first variation of* J *at* $y_0 \in F$ *in the direction* h *is*

$$\delta J(y_0, h) := \lim_{\varepsilon \to 0} \frac{J(y_0 + \varepsilon h) - J(y_0)}{\varepsilon}. \tag{4.2}$$

Here 'h' is a function such that $y_0 + \varepsilon h \in F$ *for sufficiently small* ε*. Such a direction* h *is called admissible.*

Before we look at an example we note the following useful theorem.

Lemma 4.1 (Fundamental Lemma of Calculus of Variations)**.**
If $f \in C[a, b]$ *and*
$$\int_a^b f(x)h(x)dx = 0 \qquad \forall h \in C[a, b]$$
satisfying $h(a) = h(b) = 0$*, then*
$$f(x) = 0 \qquad \forall x \in [a, b].$$

Proof. Suppose $\exists\, x_0 \in (a, b)$ for which $f(x_0) > 0$ (without loss of generality; the proof when we assume $f(x_0) < 0$ is identical). Then $\exists\, (c, d) \subset (a, b)$ with $x_0 \in (c, d)$ and $f(x) > 0$ in (c, d) by continuity of f. Set h to the special function

$$h(x) = \begin{cases} (x - c)^3 (d - x)^3, & x \in [c, d], \\ 0, & \text{otherwise,} \end{cases}$$

$h \in C^2[a, b],\ h(a) = h(b) = 0.$
Now we have
$$0 = \int_a^b f(x)h(x)dx = \int_c^d f(x)h(x)dx = f(\bar{x}) \int_c^d h(x)dx > 0.$$
(Since $f(x) > 0$ in (c, d) and $h(x) > 0$ in (c, d), $\bar{x} \in (c, d)$ by integral mean value theorem.)
Thus we have a contradiction and hence the lemma is proved since $f(x) = 0, x \in (a, b)$ and $f \in C[a, b] \Rightarrow f(a) = f(b) = 0$. □

Remark 4.2. *This result is similar to the lemma which says that:*

$$f \in C[a, b] \text{ and } \int_c^d f(x)dx = 0\ \forall (c, d) \subset [a, b],$$
$$\Rightarrow\ f(x) = 0\ \forall x \in [a, b].$$

Remark 4.3. *This fundamental lemma is true without imposing boundary conditions on* h*. The same argument holds.*

Remark 4.4. *In forming the variation* (4.2) *we add functions and multiply by scalars (i.e. we form* $y_0 + \varepsilon h$*). So* F *should be a subset of a space of functions (i.e. a subset of a (linear) 'vector' space) in which these operations are defined. For* f_1, f_2 *and* $x \in [a, b]$,

$$(f_1 + f_2)(x) := f_1(x) + f_2(x).$$

In Example 4.1 we have the subset of $C^1[x_1, x_2]$ *defined* $f(x_1) = y_1$ *and* $f(x_2) = y_2$.
$F = \{f \in C^1[x_1, x_2] \ : \ f(x_1) = y_1 \text{ and } f(x_2) = y_2\}$ *is not a linear space.*

$$f_1, f_2 \in F \ \Rightarrow \ f_1 + f_2 \notin F.$$

The admissible functions h *that give the first variation must belong to the following subspace of* $C^1[x_1, x_2]$ *:-*

$$H := \{h \in C^1[x_1, x_2] \ : \ h(x_1) = 0 \text{ and } h(x_2) = 0\}$$

so that $y_0 \in F$ *and* $h \in H \ \Rightarrow \ y_0 + \varepsilon h \in F$.
In general we would also need a concept of 'distance' (or a concept of 'smallness'); to assert that J *has a local minimum at* y_0. *We need to consider functions which are close to* y_0. *Formally this is carried out in a 'normed' linear space;* $\|y\|$ *gives a measure of 'smallness' of* y *and* $\|y - \overline{y}\|$ *gives a measure of 'closeness' of* y *to* $\overline{y}$. *The use of norms and the concept of convergence for normed vector spaces involving 'spaces of functions' is a hard subject in analysis. We avoid this.*
Less formally, we take it as clear that $y_0 + \varepsilon h$ *is closed to* y_0 *when* ε *is closed to* 0.

Definition 4.3. *An* **extremum** *is the extreme value (max or min) of* $J(y)$ *over* F.
An **extremal** *is the function or element of* F *which achieves the extremum.*

4.2 Necessary Conditions for an Extremum

We regard y_0 as fixed and then fix an admissible direction h and define

$$V(\varepsilon) := J(y_0 + \varepsilon h).$$

V is a function of the real variable ε and we note that $V(0) = J(y_0)$ and

$$V'(0) = \lim_{\varepsilon \to 0} \frac{V(\varepsilon) - V(0)}{\varepsilon} = \lim_{\varepsilon \to 0} \frac{J(y_0 + \varepsilon h) - J(y_0)}{\varepsilon} = \delta J(y_0, h).$$

We see in the remark below that a necessary condition for J to have an extremal at y_0 must be that for any admissible direction h, the function $V(\varepsilon)$ has an extremum at $\varepsilon = 0$, i.e. from Calculus that $V'(0) = 0$.

Remark:
If J has an extremal at y_0 then

$$\begin{aligned} & J(y_0) \le J(y) \quad \forall\, y \\ \Rightarrow \ & J(y_0) \le J(y_0 + \varepsilon h) \quad \forall\, \varepsilon \\ \Rightarrow \ & V(0) \le V(\varepsilon) \quad \forall\, \varepsilon \\ \Rightarrow \ & 0 \text{ is a minimum of } V \\ \Rightarrow \ & V'(0) = 0. \end{aligned}$$

Theorem 4.1. *Let* $J : F \to \mathbb{R}$ *be a functional on a set of functions* F. *If* $y_0 \in F$ *gives an extremum of* J *(max or min) then*

$$\delta J(y_0, h) = 0$$

for all admissible directions.

Remark 4.5. *We have ignored the 'difficulty' or 'notion' that* $\delta J(y_0, h)$ *might not be defined. The limiting process in* (4.2) *requires some properties on* J *in order that the limit is defined.*

Example 4.2. *Minimize*

$$J(y) = \int_0^1 \{1 + (y'(x))^2\} dx$$

for $y \in C^1[0, 1]$ *with* $y(0) = 0$, $y(1) = 1$.

Solution:
We need to consider $y_0(x) + \varepsilon h(x) \in F$ where $y_0 \in \{y \in C^1[0,1] \ : \ y(0) = 0, y(1) = 1\}$, so $h \in \{y \in C^1[0,1] \ : \ y(0) = 0, y(1) = 0\}$. We require $\delta J(y_0, h) = V'(0)$=0. Since

$$\begin{aligned} V(\varepsilon) &= \int_0^1 [1 + \{(y_0(x) + \varepsilon h(x))'\}^2] dx \\ &= \int_0^1 \{(1 + y_0'(x)^2) + 2\varepsilon y_0'(x) h'(x) + \varepsilon^2 h'(x)^2\} dx \\ &= \int_0^1 (1 + y_0'(x)^2) dx + 2\varepsilon \int_0^1 y_0'(x) h'(x) dx + \varepsilon^2 \int_0^1 h'(x)^2 dx \end{aligned}$$

and

$$\frac{dV}{d\varepsilon}(\varepsilon) = \lim_{\varepsilon \to 0} \frac{V(\varepsilon) - V(0)}{\varepsilon} \ = \ 2 \int_0^1 y_0'(x) h'(x) dx + 2\varepsilon \int_0^1 h'(x)^2 dx,$$

we have

$$\delta J(y_0, h) \equiv V'(0) = 2 \int_0^1 y_0'(x) h'(x) dx.$$

Thus a necessary condition for y_0 to be an extremal is that

$$\delta J(y_0, h) = 0,$$

or

$$\int_0^1 y_0'(x) h'(x) dx = 0 \tag{4.3}$$

for all $h \in C^1[0,1]$ and $h(0) = h(1) = 0$.

Equation (4.3) is a 'variational' equation. Such equations imply that y_0 solves a differential equation.

If $y_0 \in C^2[0,1]$ then we can integrate (4.3) by parts to obtain

$$0 = [y_0'(x) h(x)]_0^1 - \int_0^1 y_0''(x) h(x) dx.$$

Since $h(0) = h(1) = 0$ we have

$$0 = \int_0^1 y_0''(x) h(x) dx \quad \forall h \in C^1[0,1] \quad \text{with } h(0) = h(1) = 0.$$

Using the Fundamental theorem of calculus gives

$$0 = y_0''(x) \qquad \forall x \in (0,1).$$

Since $y_0 \in F$ we have $y_0(0) = 0$ and $y_0(1) = 1$

$$\Rightarrow \ y_0(x) = x.$$

Remark 4.6. *In Example 4.2 we had $y_0, h \in C^1[0,1]$ initially but to complete the problem we integrated by parts and introduced y_0'', for which we need $y_0 \in C^2[0,1]$.*
For simplicity from now on we will impose this 'regularity' on F from the beginning. That is we will look for solutions of the extremal problem in a set F which is contained in $C^2[a,b]$ for some interval $[a,b]$.

4.3 'Simple Problems' :- Euler Equations

Our main result is for functionals of the kind

$$J(y) := \int_a^b L(x, y, y')dx$$

and $F = \{y \in C^2[a,b] \ : \ y(a) = \alpha, y(b) = \beta\}$. The integral $L(x, y, y')$ is called a *Lagrangian.*

Theorem 4.2. *If a function y is an extremal of $J(y)$ for $y \in F$ (defined as above) then y satisfies*

$$L_y(x, y, y') - \frac{d}{dx} L_{y'}(x, y, y') = 0, \tag{4.4}$$

for $x \in (a, b)$. Here

$$L_y := \frac{\partial L}{\partial y} \text{ and } L_{y'} := \frac{\partial L}{\partial y'},$$

where we regard L as a function of three variables labelled x, y and y'.

Note that (4.4) is an ordinary differential equation involving the functions $y(x)$ and $dy(x)/dx$.

Proof. As in Example 4.2, we consider

$$\begin{aligned} V(\varepsilon) &:= \int_a^b L(x, y + \varepsilon h, y' + \varepsilon h')dx \\ &= J(y + \varepsilon h). \end{aligned}$$

Here h is chosen so that $y + \varepsilon h \in F$, i.e., $h \in C^2[a, b]$ and $h(a) = h(b) = 0$, in other words h is admissible.
A necessary condition for y to be an extremal is that $0 = \delta J(y, h) = V'(0)$.
Since by the chain rule we have

$$\begin{aligned} V'(\varepsilon) &= \int_a^b \frac{d}{d\varepsilon} L(x, y + \varepsilon h, y' + \varepsilon h')dx \\ &= \int_a^b \left\{ \frac{\partial L}{\partial y}(x, y + \varepsilon h, y' + \varepsilon h')h(x) + \frac{\partial L}{\partial y'}(x, y + \varepsilon h, y' + \varepsilon h')h'(x) \right\} dx \end{aligned}$$

it follows that

$$\begin{aligned} V'(0) = 0 \quad &\Rightarrow \lim_{\varepsilon \to 0} \int_a^b \left\{ \frac{\partial L}{\partial y}(x, y + \varepsilon h, y' + \varepsilon h')h(x) + \frac{\partial L}{\partial y'}(x, y + \varepsilon h, y' + \varepsilon h')h'(x) \right\} dx = 0 \\ &\Rightarrow \int_a^b \{L_y(x, y, y')h + L_{y'}(x, y, y')h'\}dx = 0. \end{aligned} \tag{4.5}$$

Performing integration by parts gives

$$0 = \int_a^b L_y(x, y, y')hdx - \int_a^b \frac{d}{dx}(L_{y'}(x, y, y'))hdx - [L_{y'}(x, y, y')h]_a^b.$$

Since h is admissible we have $h(a) = h(b) = 0$ and hence

$$\int_a^b \left\{ L_y(x, y, y') - \frac{d}{dx}(L_{y'}(x, y, y')) \right\} h(x)dx = 0$$

for all admissible h.
By Fundamental Lemma of Calculus of Variations we conclude that

$$L_y(x, y, y') - \frac{d}{dx} L_{y'}(x, y, y') = 0, \ \forall x \in [a, b].$$

□

Equation (4.4) is known as the *Euler equation* (or *Euler-Lagrange equation*) for the Lagrangian $L(x, y, y')$.

Example 4.3. *Find an extremal for*

$$J(y) = \int_0^1 ({y'}^2 + 3y + 2x)dx$$

with y *satisfying* $y(0) = 0$, $y(1) = 1$.

Solution:
The Lagrangian is $L(x, y, y') = {y'}^2 + 3y + 2x$ which gives

$$L_y = \frac{\partial L}{\partial y} = 3, \quad L_{y'} = \frac{\partial L}{\partial y'} = 2y'.$$

Hence the Euler equation is

$$\begin{aligned} & L_y - \frac{d}{dx} L_{y'} = 0. \\ \Rightarrow\ & 3 - \frac{d}{dx}(2y') = 0, \text{ i.e., } 3 - 2y'' = 0. \\ \Rightarrow\ & y'' = \frac{3}{2} \ \Rightarrow\ y(x) = \frac{3}{4}x^2 + \frac{1}{4}x. \end{aligned}$$

Thus the function $y(x) = (3x^2 + x)/4$ is an extremal for $J(y)$ with the boundary condition $y(0) = 0$, $y(1) = 1$.

Example 4.4 (Example 4.1 continued)**.** *This is the shortest curve problem:*
Find an extremal for

$$J(y) = \int_{x_1}^{x_2} \sqrt{1 + y'(x)^2}dx.$$

Solution:
The Lagrangian is

$$L(x, y, y') = (1 + {y'}^2)^{1/2}.$$

Since $L_y = 0$ and $L_{y'} = y'(1 + {y'}^2)^{-1/2}$, the Euler equation is

$$\begin{aligned} & 0 - \frac{d}{dx}\left(\frac{y'}{\sqrt{1 + {y'}^2}}\right) = 0 \\ \Rightarrow\ & \frac{y'(x)}{\sqrt{1 + {y'(x)}^2}} = A \\ \Rightarrow\ & {y'(x)}^2 = A^2(1 + {y'(x)}^2) \\ \Rightarrow\ & y'(x) = B \\ \Rightarrow\ & y(x) = Bx + C. \end{aligned}$$

Boundary conditions $y(x_1) = y_1$, $y(x_2) = y_2$.
Hence, the shortest curve written as a graph $y = f(x)$ connecting two points (x_1, y_1) and (x_2, y_2) in the plane is the straight line connecting them (as we would hopefully expect).
Clearly this is a minimizer rather than a maximizer.

4.4 Special Lagrangians

1) y' absent from L, i.e., $L = L(x, y)$.

$$\frac{\partial L}{\partial y'} \equiv 0.$$

$\Rightarrow$ Euler equation is $L_y(x, y) = 0$.
This is not a differential equation.

2) <u>y absent from L</u>, i.e., $L = L(x, y')$.

$$\Rightarrow \frac{\partial L}{\partial y} \equiv 0.$$

Euler equation is $-dL_{y'}/dx = 0$.

$$\Rightarrow L_{y'} = \text{ constant}.$$

3) <u>x absent from L</u>, i.e., $L = (y, y')$.
In this case we obtain the first integral

$$L - y'L_{y'} = \text{constant}.$$

Proof. In general consider

$$\begin{aligned}
\frac{d}{dx}(L - y'L_{y'}) &= \frac{d}{dx}L(x, y(x), y'(x)) - \frac{d}{dx}(y'L_{y'}) \\
&= (L_x + y'L_y + y''L_{y'}) - \left(y''L_{y'} + y'\frac{d}{dx}L_{y'}\right) \\
&= L_x + y'\left(L_y - \frac{d}{dx}L_{y'}\right) \\
&= L_x + 0y' \text{ using Euler equation.}
\end{aligned}$$

Now if $L_x \equiv 0$, we obtain the desired result that

$$\frac{d}{dx}(L - y'L_{y'}) = 0, \text{ or } L - y'L_{y'} = \text{constant}.$$

□

Example 4.5 (Brachistochrone Problem)**.**
A bead slides down a smooth wire (no friction) in the (x, y) plane (which is vertical) under gravity. Suppose the wire connects (x_1, y_1) to (x_2, y_2) where $x_1 < x_2$ and $y_1 > y_2$ and suppose the shape of the wire is a graph $y = y(x)$. In the bead is released from rest, which shape of wire gives the smallest time of descent?

Solution:
The curve which solves this problem is called the *brachistochrone.* The particle's mechanics are best written down using conservation of energy:

<u>Kinetic Energy + Potential Energy = Constant</u>

at any moment in time (and location of bead on wire), the speed is v so $KE = mv^2/2$.
The potential energy is mgy. At time $t = 0$, $v = 0$ (the bead is at rest)

$$\Rightarrow \frac{1}{2}mv^2 + mgy = 0 + mgy.$$

Hence the problem does not depend on mass of the bead. We have the speed is given by

$$v^2 = 2g(y_1 - y) \Rightarrow v = \sqrt{2g(y_1 - y)}.$$

Recall that arc length is

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

Note that $ds/dt = v$, hence the time for descent is $t = \int_0^t dt = \int_{s_1}^{s_2} \frac{ds}{dt} dt = \int_{s_1}^{s_2} ds/v = \text{time}$ ($s_1 = 0$, s_2 is the length of the curve).

$$\text{Time} = \int_{x_1}^{x_2} \frac{\sqrt{1 + y'^2}}{\sqrt{2g(y_1 - y)}}dx.$$

The functional is

$$\int_{x_1}^{x_2} \sqrt{\frac{1+{y'}^2}{2g(y_1-y)}}dx,\ \ y=y(x).$$

This is a problem in the calculus of variations. We look for a critical point of this functional. The Lagrangian is :-

$$L(y,y') = \sqrt{\frac{1+{y'}^2}{2g(y_1-y)}}. \tag{4.6}$$

Since x is absent from the Lagrangian, we use the remark given earlier to observe that a critical point satisfies:

$$\begin{aligned}
& L - y'L_{y'} = \text{constant.} \\
\Rightarrow & \sqrt{\frac{1+{y'}^2}{2g(y_1-y)}} - y'\frac{2y'(1+{y'}^2)^{-1/2}/2}{\sqrt{2g(y_1-y)}} = C \\
\Rightarrow & \sqrt{\frac{1+{y'}^2}{y_1-y}} - \frac{{y'}^2(1+{y'}^2)^{-1/2}}{\sqrt{y_1-y}} = A \qquad \text{where } A=\sqrt{2g}C \\
\Rightarrow & 1 = A\sqrt{y_1-y}\sqrt{1+{y'}^2}.
\end{aligned}$$

We want a differential equation, i.e. solve for y'.

$$\begin{aligned}
& \frac{1}{A^2(y_1-y)} = 1+{y'}^2, \\
& B=\frac{1}{A^2}, \quad {y'}^2 = \frac{B}{y_1-y}-1, \\
& \frac{dy}{dx} = -\left(\frac{B}{y_1-y}-1\right)^{1/2} = -\left(\frac{B-(y_1-y)}{y_1-y}\right)^{1/2},
\end{aligned}$$

observing that we want the slope of the graph to be negative.
This is a first order equation of the standard type $dy/dx = f(y)$, and hence it can be solved by using separation of variables:

$$\int \frac{dy}{f(y)} = \int dx \quad \Rightarrow \quad \int dx = -\int \frac{\sqrt{y_1-y}}{\sqrt{B-(y_1-y)}}dy.$$

In this case try the substitution

$$y_1 - y = B\sin^2\theta \tag{4.7}$$

$$\Rightarrow \frac{dy}{d\theta} = -B\sin\theta\cos\theta$$

Integral becomes

$$\begin{aligned}
x &= \int \frac{\sin\theta}{\cos\theta}\frac{dy}{d\theta}d\theta \\
&= \int \frac{\sin\theta}{\cos\theta}\left(-B\sin\theta\cos\theta\right)d\theta \\
&= -B\int \sin^2\theta d\theta \qquad (4.8) \\
&= B\int(\cos 2\theta - 1)d\theta \\
&= \frac{B}{2}(\sin 2\theta - 2\theta) + C.
\end{aligned}$$

Instead of solving for y in terms of x (which is horrible), (4.7) and (4.8) give a parametric representation of the curve.

$$\theta = 0 \ \Rightarrow \ \begin{array}{ll} y = y_1 & \text{by (4.7),} \\ x_1 = C & \text{by (4.8).} \end{array}$$

The parametric equations are

$$\left\{\begin{array}{l} x = x_1 + \frac{B}{2}(\sin 2\theta - 2\theta), \\ y = y_1 - B\sin^2\theta, \end{array}\right\} \text{ CYCLOID.}$$

B is defined by $y(x_2) = y_2$, i.e.

$$\begin{array}{l} y_2 = y_1 - B\sin^2\bar{\theta}, \\ x_2 = x_1 + \frac{B}{2}(\sin 2\bar{\theta} - 2\bar{\theta}) \end{array}$$

are solved for B and $\bar{\theta}$

4.5 Free End Point Problems

We now look at problems where the value of y is not prescribed at an end point. Indeed it is left free to vary.

Theorem 4.3 (Necessary Condition)**.** *If a function y gives an extreme value of the functional*

$$J(y) = \int_a^b L(x, y, y')dx$$

for $y \in F = \{y \in C^2[a, b] \ : \ y(a) = \alpha\}$, then y satisfies the Euler equation

$$L_y(x, y, y') - \frac{d}{dx}L_{y'}(x, y, y') = 0$$

and the boundary condition

$$L_{y'}(b, y(b), y'(b)) = 0$$

holds at $x = b$.

Proof. As for the case in which both end points are specified we consider

$$V'(0) = 0$$

where

$$V(\varepsilon) = \int_a^b L(x, y + \varepsilon h, y' + \varepsilon h')dx$$

with $h \in C^2[a, b]$ and $h(a) = 0$.
As in the earlier calculation in the proof of Theorem 4.2 we have (4.5), i.e.

$$\int_a^b \{L_y(x, y, y')h(x) + L_{y'}(x, y, y')h'(x)\}\, dx = 0$$

and integrating by parts yields

$$\int_a^b \left\{L_y(x, y, y') - \frac{d}{dx}L_{y'}(x, y, y')\right\} h(x)dx + [L_{y'}(x, y, y')h(x)]_a^b = 0. \tag{4.9}$$

We know that $h(a) = 0$ for h to be admissible.

1) Since an h with $h(b) = 0$ is admissible we have

$$\int_a^b \left[L_y - \frac{d}{dx}L_{y'}\right] dx = 0,$$

$$\Rightarrow\ L_y - \frac{d}{dx}L_{y'} = 0, \quad \text{(the Euler equation)},$$

for $x \in [a, b]$.

2) We now have

$$[L_{y'}(x,y,y')h(x)]_a^b = 0$$

for any admissible h.
Since $h(a) = 0$ we have

$$L_{y'}(b, y(b), y'(b))h(b) = 0 \quad \text{for all } h(b)$$

$$\Rightarrow \; L_{y'}(b, y(b), y'(b)) = 0.$$

□

Example 4.6 (Boat crossing river)**.** *The river has banks at $x = 0$ and $x = b$. Consider a boat crossing the river of width b whose sides are straight and parallel. Suppose $v(x)$ is the velocity of water at position x. Suppose that the speed of the boat relative to still water is c where $c > v(x)$ $\forall x \in [0, b]$. Starting from a fixed position on one bank what is the minimum time to reach the other bank? Here the point on the other bank is not specified.*

Solution:
Assuming that the boat is capable of extremely rapid acceleration in order to maintain a constant given speed c relative to the water, then for a path $y(x)$ the time taken is

$$J(y) = \int_0^b \frac{\left\{c^2(1 + y'^2) - v^2\right\}^{1/2} - vy'}{c^2 - v^2} dx,$$

unless the water velocity is a constant the Euler equation is difficult to solve. We content ourselves here to find the boundary condition at $x = b$. By the theory for free end point problem we have

$$L_{y'}(b, y(b), y'(b)) = 0$$

where

$$L(x, y, y') = \frac{\left\{c^2(1 + y'^2) - v^2\right\}^{1/2} - vy'}{c^2 - v^2},$$

$$L_{y'} = \frac{\frac{1}{2}\left\{c^2(1 + y'^2) - v^2\right\}^{-1/2}(2c^2y') - v}{c^2 - v^2}.$$

$$\Rightarrow \frac{c^2 y'(b)}{\{c^2(1 + y'(b)) - v^2(b)\}^{1/2}} = v(b).$$

After simplifying we obtain

$$y'(b) = \frac{v(b)}{c}.$$

We learn from this boundary condition in the context of this problem, that the angle of approach to the bank is given by the ratio of the water to speed of boat.

Example 4.7. *Find possible extremals of*

$$J(y) = \int_{1/27}^1 (9x^2(y')^2 + 6y' - 2y^2)dx \qquad y\left(\frac{1}{27}\right) = 0, \;\; y(1) \textit{ free}$$

Solution: Since $L(x, y, y') = 9x^2(y')^2 + 6y' - 2y^2$ we have

$$L_y = -4y, \;\; L_{y'} = 18x^2y' + 6 \quad \text{and} \quad \frac{d}{dx}L_{y'} = 36xy' + 18x^2y''.$$

The Euler-Lagrange equation is

$$L_y - \frac{d}{dx}L_{y'} = 0$$

$$\Rightarrow \ -4y - 36xy' - 18x^2y'' = 0$$

Looking for a solution of the form $y = x^r$, $\Rightarrow\ y' = rx^{r-1},\ y'' = r(r-1)x^{r-2}$ yields

$$18r(r-1) + 36r + 4 = 0 \ \Rightarrow\ 18r^2 + 18r + 4 = 0 \ \Rightarrow\ (3r+1)(6r+4) = 0$$

$$\Rightarrow\ r = -\frac{1}{3} \ \text{ and } \ r = -\frac{2}{3}.$$

Hence

$$y = Ax^{-1/3} + Bx^{-2/3}.$$

Using the boundary condition $y(1/27) = 0$ yields $3A + 9B = 0$ and hence

$$y = B(x^{-2/3} - 3x^{-1/3}) \ \Rightarrow\ y' = B\left(-\frac{2}{3}x^{-5/3} + x^{-4/3}\right).$$

The free end point condition is

$$L_{y'}(1, y(1), y'(1)) = 0$$

i.e.

$$18y'(1) + 6 = 0 \ \Rightarrow\ 18B\left(-\frac{2}{3} + 1\right) + 6 = 0 \ \Rightarrow\ B = -1.$$

Thus

$$y = (3x^{-1/3} - x^{-2/3}).$$

4.6 Higher Dimensional Problems

Suppose we have $J(y_1, y_2, \cdots, y_n)$ where each $y_i \in C^2[a,b]$, we can think of

$$\underline{y}(x) = (y_1(x), y_2(x), \cdots, y_n(x)),\ x \in [a,b]$$

as giving a curve in $\mathbb{R}^n$ using $x \in [a,b]$ as a parameter. If we specify

$$\begin{array}{ll} y_1(a) = \alpha_1, \cdots, y_n(a) = \alpha_n, & (\underline{y}(a) = \underline{\alpha}), \\ y_1(b) = \beta_1, \cdots, y_n(b) = \beta_n, & (\underline{y}(b) = \underline{\beta}), \end{array}$$

then we have a fixed end point problem: the ends of the curve are fixed at $\underline{\alpha}$ and $\underline{\beta} \in \mathbb{R}^n$.
Using vector notation we consider fixed point problems given by

$$J(\underline{y}) = \int_a^b L(x, \underline{y}, \underline{y}')dx$$

with $\underline{y}(a) = \underline{\alpha}$ and $\underline{y}(b) = \underline{\beta}$.

Theorem 4.4 (Necessary condition)**.** *If a vector function $\underline{y}$ is an extremal of*

$$J(\underline{y}) = \int_a^b L(x, \underline{y}, \underline{y}')dx$$

for $\underline{y}$ with $y_i \in C^2[a,b]$ and $\underline{y}(a) = \underline{\alpha}$ and $\underline{y}(b) = \underline{\beta}$, then $y_1, y_2, \cdots, y_n$ satisfy the system of Euler equation

$$L_{y_i}(x, \underline{y}, \underline{y}') - \frac{d}{dx}L_{y_i'}(x, \underline{y}, \underline{y}') = 0,$$

$i = 1, 2, \cdots, n$.

Proof. As for scalar case consider $V'(0)$ where

$$V(\varepsilon) = \int_a^b L(x, \underline{y} + \varepsilon\underline{h}, \underline{y}' + \varepsilon\underline{h}')dx$$

where $\underline{h} = (h_1, \cdots, h_n)$ with $h_i \in C^2[a,b]$ and $h_i(a) = h_i(b) = 0,\ i = 1, 2, \cdots, n$. Performing the differentiation with respect to ε, setting $\varepsilon = 0$ and integrating by parts gives

$$V'(0) = \int_a^b \sum_{i=1}^n \left(L_{y_i} - \frac{d}{dx} L_{y_i'} \right) h_i(x) dx$$

for all admissible $\underline{h}$.
$V'(0) = 0$ seems to be one equation.

$$\int_a^b \sum_{i=1}^n f_i(x) h_i(x) dx = 0, \qquad \forall \text{ admissible } \underline{h}.$$

Choosing $\underline{h} = h_i(x)\underline{e}_i$, where $\underline{e}_i$ is the unit coordinate vector, gives

$$\int_a^b \left(L_{y_i} - \frac{d}{dx} L_{y_i'} \right) h_i(x) dx = 0,\ \forall h_i(x),$$

$$\Rightarrow\ L_{y_i} - \frac{d}{dx} L_{y_i'} = 0.$$

□

Example 4.8. *Find possible extremals of*

$$J(y, x) = \int_0^{\pi/4} (4y^2 + z^2 + y'z') dx$$

satisfying

$$y(0) = 0, \qquad z(0) = 0,$$
$$y(\pi/4) = 1, \qquad z(\pi/4) = 1.$$

Solution:
Here $L = 4y^2 + z^2 + y'z'$ and the Euler equations are

$$L_y - \tfrac{d}{dx} L_{y'} = 0,$$
$$L_z - \tfrac{d}{dx} L_{z'} = 0.$$

Since we have $L_y = 8y,\ L_{y'} = z',\ L_z = 2z$ and $L_{z'} = y'$ we have

$$8y - \frac{d^2 z}{dx^2} = 0, \quad 8y = z'',$$
$$2z - \frac{d^2 y}{dx^2} = 0, \quad 2z = y''.$$

Eliminating $z \ \Rightarrow\ 16y = y''''$.
This is a constant coefficient fourth order ODE, so we look for a solution of the form $y = e^{mx}$ where m satisfies $m^4 - 16 = 0$

$$\Rightarrow\ (m^2 - 4)(m^2 + 4) = 0$$

$$\Rightarrow\ (m - 2)(m + 2)(m + 2i)(m - 2i) = 0.$$

Hence we have

$$y(x) = A\cosh 2x + B\sinh 2x + C\cos 2x + D\sin 2x$$

$$\Rightarrow\ z(x) = \frac{1}{2} y'' = 2A\cosh 2x + 2B\sinh 2x - 2C\cos 2x - 2D\sin 2x.$$

Using $y(0) = 0$ and $z(0) = 0$ gives

$$A + C = 0 \quad \text{and} \quad 2(A - C) = 0$$

$$\Rightarrow\ A = C = 0$$

Using $y(\pi/4) = 1$ and $z(\pi/4) = 1$ gives

$$B\sinh(\pi/2) + D = 1 \quad \text{and} \quad 2B\sinh(\pi/2) - 2D = 1$$

$$\Rightarrow\ B = 3/4\,\text{cosech}(\pi/2) \quad \text{and} \quad D = 1/4.$$

4.7 Problems with Higher Order Derivatives

For

$$J(y) = \int_a^b L(x, y, y', y'')dx$$

with

$$y(a) = \alpha, \quad y(b) = \beta,$$
$$y'(a) = \gamma, \quad y'(b) = \delta,$$

($\alpha, \beta, \gamma, \delta$ are all given), the following theorem gives a necessary condition for an extremal.

Theorem 4.5 (Necessary condition)**.** *If y gives an extremal value of*

$$J(y) = \int_a^b L(x, y, y', y'')dx$$

for $y \in C^4[a,b]$ with $y(a) = \alpha$, $y(b) = \beta$, $y'(a) = \gamma$, $y'(b) = \delta$ all specified, then y satisfies the Euler equation

$$L_y - \frac{d}{dx}L_{y'} + \frac{d^2}{dx^2}L_{y''} = 0. \tag{4.10}$$

Example 4.9. *Find possible extremals of*

$$J(y) = \int_0^1 (yy' + (y'')^2)dx \qquad y(0) = 0,\ y'(0) = 1,\ y(1) = 2,\ y'(1) = 4.$$

Solution: Since $L(y) = yy' + (y'')^2$, the Euler-Lagrange equation is

$$L_y - \frac{d}{dx}L_{y'} + \frac{d^2}{dx^2}L_{y''} = 0$$

and we have

$$y' - \frac{d}{dx}y + \frac{d^2}{dx^2}(2y'') = 0 \quad \Rightarrow 2y'''' = 0.$$

Hence

$$y = Ax^3 + Bx^2 + Cx + D$$

and

$$y' = 3Ax^2 + 2Bx + C.$$

Using the boundary data $y(0) = 0$, $y'(0) = 1$, $y(1) = 2$, $y'(1) = 4$ yields

$$D = 0,\ C = 1,\ 2 = A + B + C,\ 4 = 3A + 2B + C \quad \Rightarrow\ B = 0 \text{ and } A = 1$$

Thus

$$y = x^3 + x.$$

Remark 4.7.

1) *Note that the Euler equation (4.10) is a 4th order equation, hence the requirement that $y \in C^4[a,b]$. Its solution needs 4 boundary conditions which for fixed end point problems are the given conditions $y(a) = \alpha$, $y(b) = \beta$, $y'(a) = \gamma$, $y'(b) = \delta$. Naturally one could also consider free end point problems in which fewer than 4 conditions on y and y' at a and b are specified. In that case, we would also have extra boundary conditions from consideration of the Lagrangian.*

2) *The result generalizes to $L(x, y, y', y'', \cdots, y^{(n)})$ where the Euler equation can be shown to be*

$$L_y - \frac{d}{dx}L_{y'} + \frac{d^2}{dx^2}L_{y''} - \cdots + (-1)^n \frac{d^n}{dx^n}L_{y^{(n)}},$$

which has order $2n$. Typically one requires the following $2n$ boundary conditions to define the class of admissible functions y :-

$$y^{(j)}(a); y^{(j)}(b) \quad j = 0, 1, \cdots, n-1$$

to be specified.

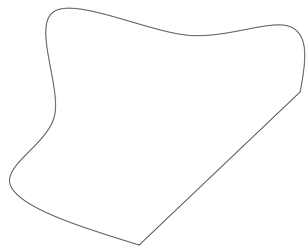

Figure 8: Maximise area subject to length of the boundary being fixed.

3) Problems involving y'' are natural in elastically theory for example the elastic energy of a 'beam' is given by

$$J(y) = \int_a^b \left\{ \frac{1}{2} E(y'')^2 - f(y) \right\} dx$$

where $y(x)$ is the displacement of the beam in the vertical direction and f is the load on the beam.

4.8 Isoperimetric Problems

An *isoperimetric problem is a constrained extremal problem such as finding the shape of the region which has maximum area amongst all regions with the same perimeter.*

$$\text{'Isoperimetric'} = \textit{having the same perimeter.}$$

For example, what shape of region defined by the rope (i.e. a curve) of fixed length L maximizes the area enclosed by the rope (the curve)? In this particular case (Figure 8) one side of the region is part of a given straight line. The basic theorem that can be used for such problems is an analogue of the following theorem in $n-$dimensional calculus.

Theorem 4.6. *Suppose $f, g : S \to \mathbb{R}$ are $C^1(S)$ where $S \subset \mathbb{R}^n$. The equation*

$$g(\underline{x}) = C, \tag{4.11}$$

(C a given constant) constrains $\overline{x}$ to a subset of S. (Indeed (4.11) is the equation of a hyper surface.) If $x_0 \in S$ and satisfies (4.11) and it is an extremum of f subject to the constraint (4.11), then provided $\underline{x_0}$ does not give an extreme value of g (i.e. $\nabla g(\underline{x_0}) \neq 0$) there exists a constant multiplier λ such that

$$\frac{\partial f}{\partial x_i}(\underline{x_0}) = \lambda \frac{\partial g}{\partial x_i}(\underline{x_0}), \quad (i = 1, 2, \cdots, n). \tag{4.12}$$

This theorem is the basis of the method of Lagrange multipliers for finding extreme values subject to constraints. Note that (4.12) can be written as

$$\nabla f(\underline{x_0}) = \lambda \nabla g(\underline{x_0}).$$

We could extend this to a constraint.

We consider analogues in the calculus of variations where we seek to find an extremal $y \in C^2[a, b]$ of

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

where y is subject to the constraint

$$K(y) = \int_a^b G(x, y, y')dx = C$$

(C given constant) and subject to fixed end point condition.

Now commonly used for general constrained problems is the calculus of variations (side conditions are expressed as integrals).

Theorem 4.7 (Necessary condition). *If $y \in C^2[a, b]$ gives an extreme value of*

$$J(y) = \int_a^b F(x, y, y')dx, \tag{4.13}$$

where y is subject to the integral constraint

$$K(y) = \int_a^b G(x, y, y')dx = C \tag{4.14}$$

(C given constant) and satisfies fixed end point conditions, then provided y is not an extremal of K, there exists a constant multiplier λ such that y satisfies

$$F_y - \frac{d}{dx}F_{y'} = \lambda\left(G_y - \frac{d}{dx}G_{y'}\right).$$

Remark of proof

1) Proof uses Lagrange multiplier theorem in $\mathbb{R}^2$.

2) If we start as usual and vary y in the direction h we would consider

$$V(\varepsilon) = \int_a^b F(x, y + \varepsilon h, y' + \varepsilon h')dx$$

and impose fixed end point conditions $h(a) = h(b) = C$. However such a variation would give

$$K(y + \varepsilon h) = \int_a^b G(x, y + \varepsilon h, y' + \varepsilon h')dx$$

which in general is not equal to C. To get around this we use a 2-parameter family of variations

$$y + \varepsilon_1 h_1 + \varepsilon_2 h_2.$$

Proof. The set of admissible directions is

$$\tilde{H} = \{h \in C^2[a, b] : h(a) = h(b) = 0\}.$$

Suppose y is a constrained extremal of $J(y)$ subject to the constraint $K(y) = C$. Set

$$V(\varepsilon_1, \varepsilon_2) = J(y + \varepsilon_1 h_1 + \varepsilon_2 h_2)$$

and

$$W(\varepsilon_1, \varepsilon_2) = K(y + \varepsilon_1 h_1 + \varepsilon_2 h_2)$$

where $h_1, h_2 \in \tilde{H}$.
By hypothesis, y is not an extremal of $K(y)$ so there are values of $\varepsilon_1, \varepsilon_2$ (near to zero) such that

$$K(\varepsilon_1, \varepsilon_2) = C, \tag{4.15}$$

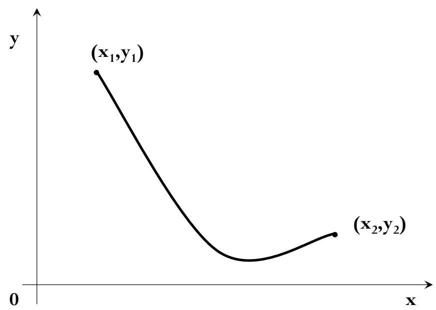

Figure 9: The hanging chain problem.

i.e. $W(\varepsilon_1, \varepsilon_2) = C$ satisfies the constraint (4.14). Furthermore, $V(\varepsilon_1, \varepsilon_2)$ has an extremum at $\varepsilon_1 = \varepsilon_2 = 0$. So Lagrange multiplier theorem in $\mathbb{R}^2$ gives a constant multiplier λ such that

$$\frac{\partial}{\partial \varepsilon}(V - \lambda W)(0, 0) = 0. \tag{4.16}$$

Writing $H = F - \lambda G$, then

$$V(\varepsilon_1, \varepsilon_2) - \lambda W(\varepsilon_1, \varepsilon_2) = J(y + \varepsilon_1 h_1 + \varepsilon_2 h_2) - \lambda K(y + \varepsilon_1 h_1 + \varepsilon_2 h_2)$$

$$= \int_a^b F(x, y + \varepsilon_1 h_1 + \varepsilon_2 h_2, y' + \varepsilon_1 h_1' + \varepsilon_2 h_2')dx - \lambda \int_a^b G(x, y + \varepsilon_1 h_1 + \varepsilon_2 h_2, y' + \varepsilon_1 h_1' + \varepsilon_2 h_2')dx$$

$$= \int_a^b [F(x, y + \varepsilon_1 h_1 + \varepsilon_2 h_2, y' + \varepsilon_1 h_1' + \varepsilon_2 h_2') - \lambda G(x, y + \varepsilon_1 h_1 + \varepsilon_2 h_2, y' + \varepsilon_1 h_1' + \varepsilon_2 h_2')]dx$$

$$= \int_a^b H(x, y + \varepsilon_1 h_1 + \varepsilon_2 h_2, y' + \varepsilon_1 h_1' + \varepsilon_2 h_2')$$

we then calculate $\partial(V(\varepsilon_1, \varepsilon_2) - \lambda W(\varepsilon_1, \varepsilon_2)/\partial\varepsilon$ and set the derivative equal to zero when $\varepsilon_1 = \varepsilon_2 = 0$. This yields an integral similar to the unconstrained case.

$$\int_a^b \{H_y(x, y, y')h_i + H_{y'}(x, y, y')h_i'\}dx = 0.$$

$$\Rightarrow \text{ using the unconstrained argument,}$$

$$\Rightarrow H_y - \frac{d}{dx}H_{y'} = 0,$$

as required. □

Example 4.10 (Hanging Chain)**.** *A classical isoperimetric problem is that of finding the shape of a hanging chain (under gravity) of fixed length and fixed to 2 points. The chain hangs in the plane of the two points.The chain will hang so that the energy is minimized. Suppose that the shape of the chain is described by the graph of* $y(x)$ *with* $y \in C^2[x_1, x_2]$.
Density ρ *fixed constant. The potential energy of the chain is*

$$\rho E = \int_0^\beta \rho g y ds$$
$$= \int_{x_1}^{x_2} (\rho g) y \sqrt{1 + y'^2} dx$$

: gravitational energy.
Constraint is

$$K(y) = \int_{x_1}^{x_2} \sqrt{1+y'^2}dx = length,$$
$$K(y) = B.$$

$$\min J(y) = \int_{x_1}^{x_2} y\sqrt{1+y'^2}dx,$$
$$subject\ to\ K(y) = B.$$

It follows that

$$F(y,y') = y\sqrt{1+(y')^2}$$

and

$$F(y,y') = \sqrt{1+(y')^2}.$$

Hence

$$H(y,y') = F(y,y') - \lambda G(y,y') = (y - \lambda\sqrt{1+(y')^2}).$$

Since x is absent from H, the Euler equation

$$H_y(y,y') - \frac{d}{dx}H_{y'}(y,y') = 0$$

implies that

$$H(y,y') - y'H_{y'}(y,y') = A,$$

where A is a constant. It follows that

$$(y-\lambda)\sqrt{1+(y')^2} - y'(y-\lambda)\frac{y'}{\sqrt{1+(y')^2}} = A$$
$$(y-\lambda)[(1+(y')^2) - (y')^2] = A\sqrt{1+(y')^2}$$
$$(y-\lambda) = A\sqrt{1+(y')^2}$$
$$(y-\lambda)^2 = A^2[1+(y')^2]$$
$$y' = \pm\frac{\sqrt{(y-\lambda)^2 - A^2}}{A}$$
$$\frac{1}{\sqrt{(y-\lambda)^2 - A^2}}\frac{dy}{dx} = \pm\frac{1}{A}$$
$$\int \frac{1}{\sqrt{(y-\lambda)^2 - A^2}}dy = \pm\frac{x}{A} + B$$
$$arccosh(\frac{y-\lambda}{A}) = \pm\frac{x}{A} + B$$
$$y = \lambda + Acosh(\pm\frac{x}{A}) + B.$$

The three constants λ, A and B are determined by the endpoints (x_1, y_1) and (x_2, y_2) of the chain and the length L of the chain.