CALCULUS OF VARIATIONS

PROF. ARNOLD ARTHURS

1. Introduction

Example 1.1 (Shortest Path Problem). Let A and B be two fixed points in a space. Then we want to find the shortest distance between these two points. We can construct the problem diagrammatically as below.

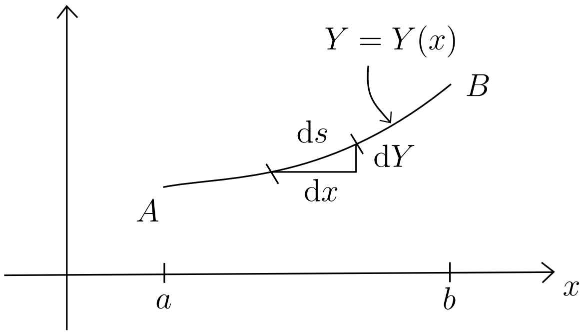

Figure 1. A simple curve.

From basic geometry (i.e. Pythagoras' Theorem) we know that

$$\begin{aligned} \mathrm{d}s^2 &= \mathrm{d}x^2 + \mathrm{d}Y^2 \\ &= \big\{1 + (Y')^2\big\}\mathrm{d}x^2. \end{aligned} \tag{1.1}$$

The second line of this is achieved by noting $Y' = \frac{\mathrm{d}Y}{\mathrm{d}x}$. Now to find the path between the points A and B we integrate $\mathrm{d}s$ between A and B, i.e. $\int_A^B \mathrm{d}s$. We however replace $\mathrm{d}s$ using equation (1.1) above and hence get the expression of the length of our curve

$$J(Y) = \int_a^b \sqrt{1 + (Y')^2}\ \mathrm{d}x.$$

To find the shortest path, i.e. to minimise J, we need to find the extremal function.

Example 1.2 (Minimal Surface of Revolution - Euler). This problem is very similar to the above but instead of trying to find the shortest distance, we are trying to find the smallest surface to cover an area. We again display the problem diagrammatically.

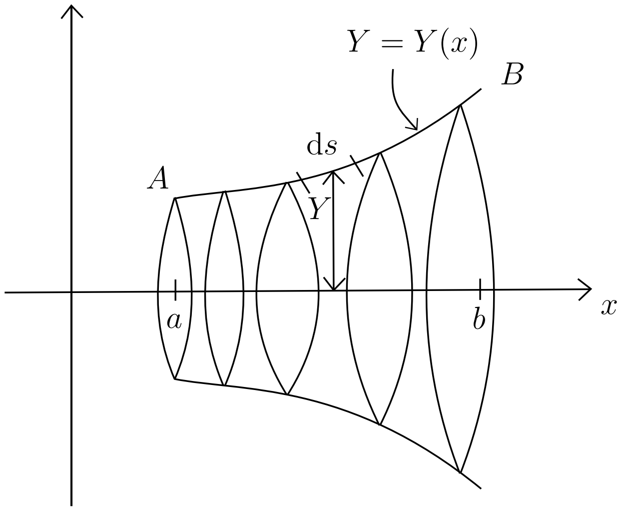

FIGURE 2. A volume of revolution, of the curve $Y(x)$, around the line $y = 0$.

To find the surface of our shape we need to integrate $2\pi Y \mathrm{d}s$ between the points A and B, i.e. $\int_A^B 2\pi Y \mathrm{d}s$. Substituting $\mathrm{d}s$ as above with equation (1.1) we obtain our expression of the size of the surface area

$$J(Y) = \int_a^b 2\pi Y \sqrt{1 + (Y')^2}\, \mathrm{d}x.$$

To find the minimal surface we need to find the extremal function.

EXAMPLE 1.3 (Brachistochrone). This problem is derived from Physics. If I release a bead from O and let it slip down a frictionless curve to point B, accelerated only by gravity, what shape of curve will allow the bead to complete the journey in the shortest possible time. We can construct this diagrammatically below.

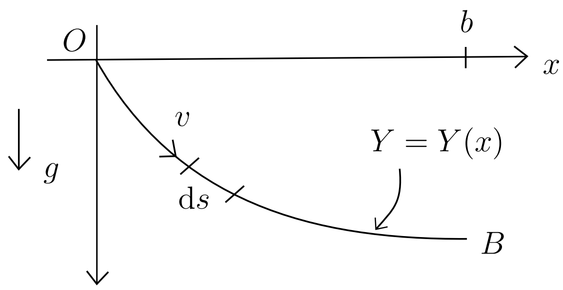

FIGURE 3. A bead slipping down a frictionless curve from point O to B.

In the above problem, we want to minimise the variable of time. So, we construct our integral accordingly and consider the total time taken T as a function of the curve Y.

$$T(Y) = \int_{x=0}^{b} \mathrm{d}t$$

now using $v = \frac{\mathrm{d}s}{\mathrm{d}t}$ and rearranging we achieve

$$= \int_{x=0}^{b} \frac{\mathrm{d}s}{v}.$$

Finally using the formula $v^2 = 2gY$ we obtain

$$= \int_{0}^{b} \sqrt{\frac{1 + (Y')^2}{2gY}}\, \mathrm{d}x.$$

Thus to find the smallest possible time taken we need to find the extremal function.

Example 1.4 (Isoperimetric problems). These are problems with constraints. A simple example of this is trying to find the shape that maximises the area enclosed by a rectangle of fixed perimeter p.

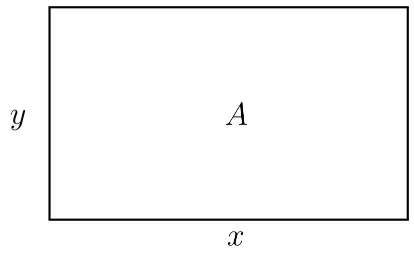

Figure 4. A rectangle with sides of length x and y.

We can see clearly that the constraint equations are

$$A = xy \tag{1.2}$$
$$p = 2x + 2y. \tag{1.3}$$

By rearranging equation (1.3) in terms of y and substituting into (1.2) we obtain that

$$A = \frac{x}{2}p - 2x$$
$$\Rightarrow \frac{\mathrm{d}A}{\mathrm{d}x} = \frac{p}{2} - 2x$$

we require that $\frac{\mathrm{d}A}{\mathrm{d}x} = 0$ and thus

$$\frac{p}{2} - 2x = 0 \Rightarrow x = \frac{p}{4}$$

and finally substituting back into equation (1.3) gives us

$$y = \frac{1}{2}\left(p - \frac{1}{2}p\right) = \frac{p}{4}.$$

Thus a square is the shape that maximises the area.

EXAMPLE 1.5 (Chord and Arc Problem). Here we are seeking the curve of a given length that encloses the largest area above the x-axis. So, we seek the curve $Y = y$ (y is reserved for the solution and Y is used for the general case). We describe this diagrammatically below.

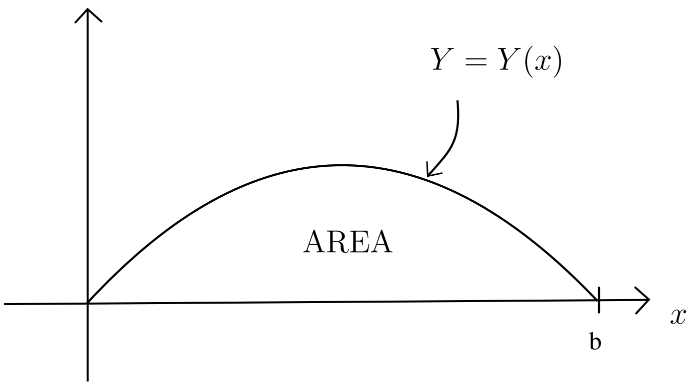

FIGURE 5. A curve $Y(x)$ above the x-axis.

We have the area of the curve $J(Y)$ to be

$$J(Y) = \int_0^b Y \, \mathrm{d}x$$

where $J(Y)$ is maximised subject to the length of the curve

$$K(Y) = \int_0^b \sqrt{1 + (Y')^2} \, \mathrm{d}x = c,$$

where c is a given constant.

2. The Simplest / Fundamental Problem

Examples 1.1, 1.2 and 1.3 are all special cases of the simplest/fundamental problem.

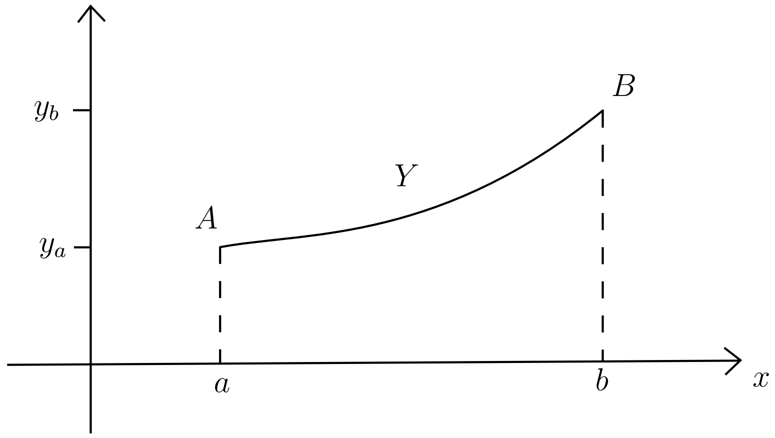

FIGURE 6. The Simplest/Fundamental Problem.

Suppose $A(a, y_a)$ and $B(b, y_b)$ are two fixed points and consider a set of curves

$$Y = Y(x) \tag{2.1}$$

joining A and B. Then we seek a member $Y = y(x)$ of this set which minimises the integral

$$J(Y) = \int_a^b F(x, Y, Y')\ \mathrm{d}x \tag{2.2}$$

where $Y(a) = y_a$, $Y(b) = y_b$. We note that examples 1.1 to 1.3 correspond to a specification of the above general case with the integrand

$$F = \sqrt{1 + (Y')^2},$$
$$F = 2\pi Y \sqrt{1 + (Y')^2},$$
$$F = \sqrt{\frac{1 + (Y')^2}{2gY}}.$$

An extra example is

$$J(Y) = \int_a^b \left\{ \frac{1}{2} p(x)(Y')^2 + \frac{1}{2} q(x) Y^2 + f(x) Y \right\}\ \mathrm{d}x.$$

Now, the curves Y in (2.2) may be continuous, differentiable or neither and this affects the problem for $J(Y)$. We shall suppose that the functions $Y = Y(x)$ are continuous and have continuous derivatives a suitable number of times. Thus the functions (2.1) belong to a set Ω of admissible functions. We define Ω precisely as

$$\Omega = \left\{ Y \,\middle|\, Y \text{ continuous and } \frac{\mathrm{d}^k Y}{\mathrm{d}x^k} \text{ continuous, } k = 1, 2, \dots, n \right\}. \tag{2.3}$$

So the problem is to minimise $J(Y)$ in (2.2) over the functions Y in Ω where

$$Y(a) = y_a \quad \text{and} \quad Y(b) = y_b.$$

This basic problem can be extended to much more complicated problems.

Example 2.1. We can extend the problem by considering more derivatives of Y. So the integrand becomes

$$F = F(x, Y, Y', Y''),$$

i.e. F depends on Y'' as well as x, Y, Y'.

Example 2.2. We can consider more than one function of x. Then the integrand becomes

$$F = F(x, Y_1, Y_2, Y_1', Y_2'),$$

so F depends on two (or more) functions Y_k of x.

Example 2.3. Finally we can consider functions of more than one independent variable. So, the integrand would become

$$F = F(x, y, \Phi, \Phi_x, \Phi_y),$$

where subscripts denote partial derivatives. So, F depends on functions $\Phi(x, y)$ of two independent variables x, y. This would mean that to calculate $J(Y)$ we would have to integrate more than once, for example

$$J(Y) = \iint F \,\mathrm{d}x\mathrm{d}y.$$

Note. The integral $J(Y)$ is a *numerical-valued* function of Y, which is an example of a *functional*.

Definition: Let $\mathbb{R}$ be the real numbers and Ω a set of functions. Then the function $J : \Omega \to \mathbb{R}$ is called a *functional*.

Then we can say that the calculus of variations is concerned with *maxima* and *minima* (*extremum*) of functionals.

3. Maxima and Minima

3.1. The First Necessary Condition

(i) We use ideas from elementary calculus of functions $f(u)$.

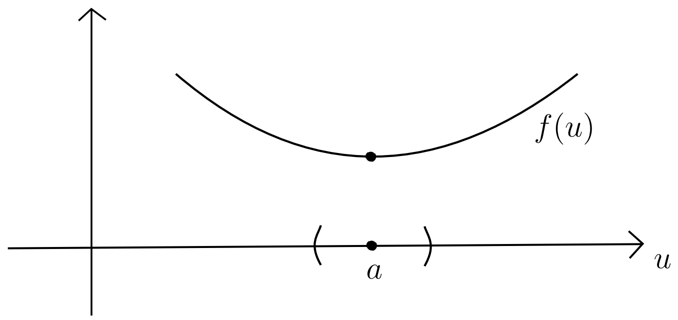

FIGURE 7. Plot of a function $f(u)$ with a minimum at $u = a$.

If $f(u) \geqslant f(a)$ for all u near a on both sides of $u = a$ this means that there is a minimum at $u = a$. The consequences of this are often seen in an expansion. Let us assume that there is a minimum at $f(a)$ and a *Taylor expansion* exists about $u = a$ such that

$$f(a+h) = f(a) + hf'(a) + \frac{h^2}{2}f''(a) + \mathrm{O}_3 \qquad (h \neq 0). \tag{3.1}$$

Note that we define $\mathrm{d}f(a,h) := hf'(a)$ to be the first differential. As there exists a minimum at $u = a$ we must have

$$f(a+h) \geqslant f(a) \quad \text{for} \quad h \in (-\delta, \delta) \tag{3.2}$$

by the above comment. Now, if $f'(a) \neq 0$, say it is positive and h is sufficiently small, then

$$\operatorname{sign}\{\triangle f = f(a+h) - f(a)\} = \operatorname{sign}\{\mathrm{d}f = hf'(a)\} \quad (\neq 0). \tag{3.3}$$

In equation (3.3) the L.H.S. $\geqslant 0$ because f has a minimum at a and hence equation (3.2) holds and also the R.H.S. > 0 if $h > 0$. However this is a contradiction, hence $\mathrm{d}f = 0$ which $\Rightarrow f'(a) = 0$.

(ii) For functions $f(u,v)$ of *two* variables; similar ideas hold. Thus if (a,b) is a minimum then $f(u,v) \geqslant f(a,b)$ for all u near a and v near b. Then for some intervals $(-\delta_1, \delta_1)$ and $(-\delta_2, \delta_2)$ we have that

$$\begin{cases} a - \delta_1 \leqslant u \leqslant a + \delta_1 \\ b - \delta_2 \leqslant v \leqslant b + \delta_2 \end{cases} \tag{3.4}$$

gives a minimum / maximum at $(a,b) \leqslant f(a,b)$. The corresponding Taylor expansion is

$$f(a+h,b+k) = f(a,b) + hf_u(a,b) + kf_v(a,b) + \mathrm{O}_2. \tag{3.5}$$

We note that in this case the first derivative is $\mathrm{d}f(a,b,h,k) := hf_u(a,b) + kf_v(a,b)$. For a minimum (or a maximum) at (a,b) it follows, as in the previous case that a necessary condition is

$$\mathrm{d}f = 0 \Rightarrow \frac{\partial f}{\partial u} = \frac{\partial f}{\partial v} = 0 \text{ at } (a,b). \tag{3.6}$$

(iii) Now considering functions of multiple (say n) variables, i.e. $f = f(u_1, u_2, \ldots, u_n) = f(\mathbf{u})$ we have the Taylor expansion to be

$$f(\mathbf{a}+\mathbf{h}) = f(\mathbf{a}) + \mathbf{h} \cdot \nabla f(\mathbf{a}) + \mathrm{O}_2. \tag{3.7}$$

Thus the statement from the previous case (3.6) becomes

$$\mathrm{d}f = 0 \Rightarrow \nabla f(\mathbf{a}) = 0. \tag{3.8}$$

3.2. **Calculus of Variations**

Now we consider the integral

$$J(Y) = \int_a^b F(x, Y, Y')\ \mathrm{d}x. \tag{3.9}$$

Suppose $J(Y)$ has a minimum for the curve $Y = y$. Then

$$J(Y) \geqslant J(y) \tag{3.10}$$

for $Y \in \Omega = \{Y \mid Y \in C_2,\, Y(a) = y_a, Y(b) = y_b\}$. To obtain information from (3.10) we *expand* $J(Y)$ about the curve $Y = y$ by taking the so-called *varied curves*

$$Y = y + \varepsilon\xi \tag{3.11}$$

like $u = a + h$ in (3.2). We can represent the consequences of this expansion diagrammatically.

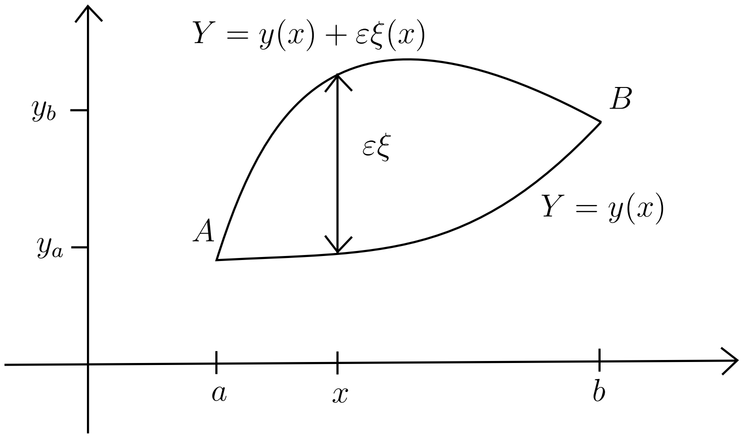

FIGURE 8. Plot of $y(x)$ and the expansion $y(x) + \varepsilon\xi(x)$.

Since all curves Y, including y, go through A and B, it follows that

$$\xi(a) = 0 \quad \text{and} \quad \xi(b) = 0. \tag{3.12}$$

Now substituting equation (3.11) into (3.10) gives us

$$J(Y) = J(y + \varepsilon\xi) \geqslant J(y) \tag{3.13}$$

for all $y + \varepsilon\xi \in \Omega$ and substituting (3.11) into (3.9) gives us

$$J(y + \varepsilon\xi) = \int_a^b F(x, y + \varepsilon\xi, y' + \varepsilon\xi') \, \mathrm{d}x. \tag{3.14}$$

Now to deal with this expansion we take a *fixed* x in (a, b) and treat y and y' as independent variables. Recall Y and Y' are independent and the Taylor expansion of two variables (equation (3.5)) from above. Then we have

$$f(u + h, v + k) = f(u, v) + h\frac{\partial f}{\partial u} + k\frac{\partial f}{\partial v} + \mathrm{O}_2. \tag{3.15}$$

Now we take $u = y$, $h = \varepsilon\xi$, $v = y'$, $k = \varepsilon\xi'$, $f = F$. Then (3.14) implies

$$\begin{aligned} J(Y) = J(y + \varepsilon\xi) &= \int_a^b \left\{ F(x, y, y') + \varepsilon\xi\frac{\partial F}{\partial y} + \varepsilon\xi'\frac{\partial F}{\partial y'} + \mathrm{O}(\varepsilon^2) \right\} \, \mathrm{d}x \\ &= J(y) + \delta J + \mathrm{O}_2. \end{aligned}$$

We note that δJ is calculus of variations notation for $\mathrm{d}J$ where we have

(3.16)
$$\begin{aligned}\delta J &= \varepsilon \int_a^b \left\{ \xi \frac{\partial F}{\partial y} + \xi' \frac{\partial F}{\partial y'} \right\} \, \mathrm{d}x \\ &= \text{linear terms in } \varepsilon = \text{ first } \textit{variation} \text{ of } J\end{aligned}$$

and we also have that

(3.17)
$$\frac{\partial F}{\partial y} = \left\{ \frac{\partial F(x, Y, Y')}{\partial Y} \right\}_{Y=y, Y'=y'}.$$

Now δJ is analogous to the linear terms in (3.1), (3.5), (3.7). We now prove that if $J(Y)$ has a minimum at $Y = y$, then

(3.18)
$$\boxed{\delta J = 0}$$

the first necessary condition for a minimum.

Proof. Suppose $\delta J \neq 0$ then $J(Y) - J(y) = \delta J + \mathrm{O}_2$. For small enough $\varepsilon\xi$ then

$$\operatorname{sign}\{J(Y) - J(y)\} = \operatorname{sign}\{\delta J\}.$$

We have that $\delta J > 0$ or $\delta J < 0$ for some varied curves corresponding to $\varepsilon\xi$. However there is a minimum of J at $Y = y \Rightarrow$ L.H.S. $\geqslant 0$. This is a contradiction. $\square$

For our J, we have by (3.17)

(3.19)
$$\delta J = \varepsilon \int_a^b \left(\xi \frac{\partial F}{\partial y} + \xi' \frac{\partial F}{\partial y'} \right) \, \mathrm{d}x.$$

If we integrate 2nd term by parts we obtain

(3.20)
$$\delta J = \varepsilon \int_a^b \xi \left\{ \frac{\partial F}{\partial y} - \frac{\mathrm{d}}{\mathrm{d}x} \frac{\partial F}{\partial y'} \right\} \, \mathrm{d}x + \varepsilon \underbrace{\left[\xi \frac{\partial F}{\partial y'} \right]_a^b}_{=0}$$

and so we have

(3.21)
$$\boxed{\delta J = \varepsilon \int_a^b \xi \left\{ \frac{\partial F}{\partial y} - \frac{\mathrm{d}}{\mathrm{d}x} \frac{\partial F}{\partial y'} \right\} \, \mathrm{d}x.}$$

Note. For notational purposes we write

(3.22)
$$\langle f, g \rangle = \int_a^b f(x) g(x) \, \mathrm{d}x$$

which is an inner (or scalar) product. Also, for our J, write

$$J'(y) = \frac{\partial F}{\partial y} - \frac{\mathrm{d}}{\mathrm{d}x}\frac{\partial F}{\partial y'} \tag{3.23}$$

as the derivative of J. Then we can express δJ as

$$\delta J = \big\langle \varepsilon\xi, J'(y) \big\rangle. \tag{3.24}$$

This gives us the Taylor expansion

$$J(y + \varepsilon\xi) = J(y) + \underbrace{\big\langle \varepsilon\xi, J'(y) \big\rangle}_{\delta J} + \mathrm{O}_2. \tag{3.25}$$

We compare this with the previous cases of Taylor expansion (3.1), (3.5) and (3.7). Now collecting our results together we get

Theorem 3.1. *A necessary condition for $J(Y)$ to have an extremum (maximum or minimum) at $Y = y$ is*

$$\boxed{\delta J = \big\langle \varepsilon\xi, J'(y) \big\rangle = 0} \tag{3.26}$$

for all admissible ξ. i.e.

$$\text{“}J(Y) \textit{ has an extremum at } Y = y\text{”} \Rightarrow \text{“}\delta J(y, \varepsilon\xi) = 0\text{”}.$$

Definition: y is a *critical curve* (or an *extremal*), i.e. y is a solution of $\delta J = 0$. $J(y)$ is a *stationary value* of J and (3.26) is a *stationary condition*.

To establish an extremum, we need to examine the sign of

$$\triangle J = J(y + \varepsilon\xi) - J(y) = \text{total variation of } J.$$

We consider this later in part two of this course.

4. Stationary Condition ($\delta J = 0$)

Our next step is to see what can be deduced from the condition

$$\delta J = 0. \tag{4.1}$$

In detail this is

$$\big\langle \varepsilon\xi, J'(y) \big\rangle = \int_a^b \varepsilon\xi J'(y)\ \mathrm{d}x = 0. \tag{4.2}$$

For our case $J'(Y) = \int_a^b F(x, Y, Y')\ \mathrm{d}x$, we have

$$J'(y) = \frac{\partial F}{\partial y} - \frac{\mathrm{d}}{\mathrm{d}x}\frac{\partial F}{\partial y'}. \tag{4.3}$$

To deal with equations (4.2) and (4.3) we use the Euler-Legrange Lemma. Using this Lemma we can state

Theorem 4.1. *A necessary condition for*

$$J(Y) = \int_a^b F(x, Y, Y')\ \mathrm{d}x, \tag{4.4}$$

with $Y(a) = y_a$ and $Y(b) = y_b$, to have an extremum at $Y = y$ is that y is a solution of

$$J'(y) = \frac{\partial F}{\partial y} - \frac{\mathrm{d}}{\mathrm{d}x}\frac{\partial F}{\partial y'} = 0 \tag{4.5}$$

with $a < x < b$ and $F = F(x, y, y')$. This is known as the Euler-Lagrange equation.

The above theorem is the Euler-Lagrange variational principle and it is a *stationary* principle.

Example 4.1 (Shortest Path Problem). We now revisit Example 1.1 with the mechanisms that we have just set up. So we had that $F(x, Y, Y') = \sqrt{1 + (Y')^2}$ previously but instead we write $F(x, y, y') = \sqrt{1 + (y')^2}$. Now

$$\frac{\partial F}{\partial y} = 0, \qquad \frac{\partial F}{\partial y'} = \frac{2y'}{2\sqrt{1 + (y')^2}} = \frac{y'}{\sqrt{1 + (y')^2}}.$$

The Euler-Lagrange equation for this example is

$$\begin{aligned} 0 - \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{y'}{\sqrt{1 + (y')^2}}\right) &= 0 \qquad a < x < b \\ \Rightarrow \frac{y'}{\sqrt{1 + (y')^2}} &= \text{const.} \\ \Rightarrow y' &= \alpha \\ \Rightarrow y &= \alpha x + \beta \end{aligned}$$

where α, β are arbitrary constants. So, $y = \alpha x + \beta$ defines the critical curves. We require more information to establish a minimum.

Example 4.2 (Minimum Surface Problem). Now revisiting Example 1.2, we had $F(x, Y, Y') = 2\pi Y\sqrt{1 + (Y')^2}$ but again we write $F(x, y, y') = 2\pi y\sqrt{1 + (y')^2}$. We drop the 2π to give us

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

So, we are left with the Euler-Lagrange equation

$$\sqrt{1+(y')^2} - \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{yy'}{\sqrt{1+(y')^2}}\right) = 0.$$

Now solving the differential equation leaves us with

$$\left(1+(y')^2\right)^{-\frac{3}{2}}\left\{1+(y')^2 - yy''\right\} = 0,$$

for finite y' in (a, b). We therefore have

$$\boxed{1+(y')^2 = yy''.} \tag{4.6}$$

Start by rewriting y'' in terms of y and y'. Then substituting gives us

$$\begin{aligned} y'' &= \frac{\mathrm{d}y'}{\mathrm{d}x} \\ &= \frac{\mathrm{d}y'}{\mathrm{d}y} \cdot \frac{\mathrm{d}y}{\mathrm{d}x} \\ &= y'\frac{\mathrm{d}y'}{\mathrm{d}y} = \frac{1}{2}\frac{\mathrm{d}(y')^2}{\mathrm{d}y}. \end{aligned} \tag{4.7}$$

So, substituting (4.7) into (4.6), the Euler-Lagrange equation implies that

$$\begin{aligned} 1+(y')^2 &= \frac{1}{2}y\frac{\mathrm{d}(y')^2}{\mathrm{d}y} \\ \Rightarrow \frac{\mathrm{d}y}{y} &= \frac{1}{2}\frac{\mathrm{d}(y')^2}{1+(y')^2} = \frac{1}{2}\frac{\mathrm{d}z}{1+z} \\ \Rightarrow \ln y &= \frac{1}{2}\ln\left(1+(y')^2\right) + C \\ \Rightarrow y &= C\sqrt{1+(y')^2}, \end{aligned}$$

which is a first integral. Now integrating again gives us

$$y = \cosh\left(\frac{x+C_1}{C}\right)$$

where C_1 and C are arbitrary constants. We note that this is a *catenary* curve.

5. Special Cases of the Euler-Lagrange Equation for the Simplest Problem

For $F(x, y, y')$ the Euler-Lagrange equation is a 2nd order differential equation in general. There are special cases worth noting.

(1) $\frac{\partial F}{\partial y} = 0 \Rightarrow F = F(x, y')$ and hence y is missing. The Euler-Lagrange equation is

$$\begin{aligned} \underbrace{\frac{\partial F}{\partial y}}_{=0} - \frac{\mathrm{d}}{\mathrm{d}x}\frac{\partial F}{\partial y'} &= 0 \\ \Rightarrow \frac{\mathrm{d}}{\mathrm{d}x}\frac{\partial F}{\partial y'} &= 0 \\ \Rightarrow \frac{\partial F}{\partial y'} &= c \end{aligned}$$

on orbits (extremals or critical curves), first integral. If this can be for y' thus

$$y' = f(x, c)$$

then $y(x) = \int f(x, c)\ \mathrm{d}x + c_1$.

Example 5.1. $F = x^2 + x^2(y')^2$. Hence $\frac{\partial F}{\partial y} = 0$. So Euler-Lagrange equation has a first integral $\frac{\partial F}{\partial y'} = c \Rightarrow 2x^2 y' = c \Rightarrow y' = \frac{c}{2x^2} \Rightarrow y' = -\frac{c}{2x} + c_1$, which is the solution.

(2) $\frac{\partial F}{\partial x} = 0$ and so $F = F(y, y')$ and hence x is missing. For any differentiable $F(x, y, y')$ we have by the chain rule

$$\begin{aligned} \frac{\mathrm{d}F}{\mathrm{d}x} &= \frac{\partial F}{\partial x} + \frac{\partial F}{\partial y} + \frac{\partial F}{\partial y'}y'' \\ &= \frac{\partial F}{\partial x} + \left(\frac{\mathrm{d}}{\mathrm{d}x}\frac{\partial F}{\partial y'}\right)y' + \frac{\partial F}{\partial y'}y'' \\ &= \frac{\partial F}{\partial x} + \frac{\mathrm{d}}{\mathrm{d}x}\left(y'\frac{\partial F}{\partial y'}\right) \end{aligned}$$

on orbits. Now, we have

$$\Rightarrow \frac{\mathrm{d}}{\mathrm{d}x}\left(F - y'\frac{\partial F}{\partial y'}\right) = \frac{\partial F}{\partial x}$$

on orbits. When $\frac{\partial F}{\partial x} = 0$, this gives

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(F - y'\frac{\partial F}{\partial y'}\right) = 0 \text{ on orbits}$$
$$\Rightarrow \boxed{F - y'F_{y'} = c}$$

on orbits. First integral.

Note. $G := F - y'F_{y'}$ is the Jacobi function and $G = G(x, y, y')$.

EXAMPLE 5.2. $F = 2\pi y\sqrt{1 + (y')^2}$. An example of case 2. So we have

$$F_{y'} = \frac{2\pi y y'}{\sqrt{1 + (y')^2}}.$$

then the first integral is

$$\begin{aligned} F &= y'F_{y'} \\ &= 2\pi y\{\sqrt{1 + (y')^2} - \frac{(y')^2}{\sqrt{1 + (y')^2}}\} \\ &= \frac{2\pi y}{\sqrt{1 + (y')^2}} \\ &= \text{const.} = 2\pi c \end{aligned}$$

$\Rightarrow y = c\sqrt{1 + (y')^2}$ - First Integral.

This arose in Example 4.2 as a result of integrating the Euler Lagrange equation once.

(3) $\frac{\partial F}{\partial y'} = 0$ and hence $F = F(x, y)$, i.e. y' is missing. The Euler-Lagrange equation is

$$0 = \frac{\partial F}{\partial y} - \frac{\mathrm{d}}{\mathrm{d}x}\frac{\partial F}{\partial y'} = \frac{\partial F}{\partial y} = 0$$

but this isn't a differential equation in y.

EXAMPLE 5.3. $F(x, y, y') = -y\ln y + xy$ and then the Euler-Lagrange equation gives

$$\frac{\partial F}{\partial y} - \frac{\mathrm{d}}{\mathrm{d}x}\frac{\partial F}{\partial y'} = 0 \Rightarrow -\ln y - 1 + x = 0$$
$$\ln y = -1 + x \Rightarrow y = e^{-1+x}$$

6. Change of Variables

In the shortest path problem above, we note that $J(Y)$ does not depend on the co-ordinate system chosen. Suppose we require the extremal for

$$J(r) = \int_{\theta_0}^{\theta_1} \sqrt{r^2 + (r')^2}\ \mathrm{d}\theta \tag{6.1}$$

where $r = r(\theta)$ and $r' = \frac{\mathrm{d}r}{\mathrm{d}\theta}$. The Euler-Lagrange equation is

$$\frac{\partial F}{\partial r} - \frac{\mathrm{d}}{\mathrm{d}\theta}\frac{\partial F}{\partial r'} = 0 \Rightarrow \frac{r}{\sqrt{r^2 + (r')^2}} - \frac{\mathrm{d}}{\mathrm{d}\theta}\left(\frac{r'}{\sqrt{r^2 + (r')^2}}\right) = 0. \tag{6.2}$$

To simplify this we can

(a) change variables in (6.2), or
(b) change variables in (6.1).

For example, in (6.1) take

$$\begin{aligned} x &= r\cos\theta & \qquad \mathrm{d}x &= \mathrm{d}r\cos\theta - r\sin\theta\,\mathrm{d}\theta \\ y &= r\sin\theta & \qquad \mathrm{d}y &= \mathrm{d}r\sin\theta + r\cos\theta\,\mathrm{d}\theta. \end{aligned}$$

So, we obtain

$$\begin{aligned} \mathrm{d}x^2 + \mathrm{d}y^2 &= \mathrm{d}r^2 + r^2\mathrm{d}\theta^2 \\ \big(1 + (y')^2\big)\mathrm{d}x^2 &= \left\{\left(\frac{\mathrm{d}r}{\mathrm{d}\theta}\right)^2 + r^2\right\}\ \mathrm{d}\theta^2 \\ \sqrt{1 + (y')^2}\mathrm{d}x &= \sqrt{r^2 + (r')^2}\mathrm{d}\theta. \end{aligned}$$

This is just the smallest path problem, hidden in polar coordinates. We know the solution to this is

$$J(r) \to \bar{J}(y) = \int \sqrt{1 + (y')^2}\ \mathrm{d}x.$$

Now the Euler-Lagrange equation is $y'' = 0 \Rightarrow y = \alpha x + \beta$. Now,

$$r\sin\theta = \alpha r\cos\theta + \beta$$

$$\boxed{r = \frac{\beta}{\sin\theta - \alpha\cos\theta}}$$

7. Several Independent functions of One Variable

Consider a general solution of the simplest problem.

$$J(Y_1, Y_2, \ldots, Y_n) = \int_a^b F(x, Y_1, \ldots, Y_n, Y_1', \ldots, Y_n')\ \mathrm{d}x. \tag{7.1}$$

Here each curve Y_k goes through given end points

$$Y_k(a) = y_{ka},\ Y_k(b) = y_{kb} \qquad \text{for } k = 1, \ldots, n. \tag{7.2}$$

Find the critical curves y_k, $k = 1, \ldots, n$. Take

$$Y_k = y_k(x) + \varepsilon\xi_k(x) \tag{7.3}$$

where $\xi_k(a) = 0 = \xi_k(b)$. By taking a Taylor expansion around the point $(x, y_1, \ldots, y_n, y_1', \ldots, y_n')$ we find that the first variation of J is

$$\delta J = \sum_{k=1}^{n} \big\langle \varepsilon\xi_k, J'(y_1, \ldots, y_n)\big\rangle. \tag{7.4}$$

with

$$J_k'(y_1, \ldots, y_n) = \frac{\partial F}{\partial y_k} - \frac{\mathrm{d}}{\mathrm{d}x}\frac{\partial F}{\partial y_k'} \tag{7.5}$$

for $k = 1, \ldots, n$. The stationary condition $\delta J = 0 \Rightarrow$

$$\sum_{k=1}^{n} \langle \varepsilon\xi_k, J_k'\rangle = 0 \tag{7.6}$$

$$\Rightarrow \langle \varepsilon\xi_k, J_k'\rangle = 0 \quad \text{by linear independence.} \tag{7.7}$$

for all $k = 1, \ldots, n$. Thus, by the Euler-Lagrange Lemma, this implies

$$J_k' = 0 \tag{7.8}$$

for $k = 1, \ldots, n$. i.e.

$$\frac{\partial F}{\partial y_k} - \frac{\mathrm{d}}{\mathrm{d}x}\frac{\partial F}{\partial y_k'} = 0 \tag{7.9}$$

for $k = 1, \ldots, n$. (7.9) is a system of n Euler-Lagrange equations. These are solved subject to (7.3).

EXAMPLE 7.1. $F = y_1 y_2^2 + y_1^2 y_2 + y_1' y_2'$ and so we obtain equations

$$\left.\begin{aligned} J_1' = 0 &\Rightarrow y_2^2 + 2y_1y_2 - \frac{\mathrm{d}}{\mathrm{d}x}(y_2') = 0 \\ J_2' = 0 &\Rightarrow 2y_1y_2 + y_1^2 - \frac{\mathrm{d}}{\mathrm{d}x}(y_1') = 0 \end{aligned}\right\}$$

Consider

$$\frac{\mathrm{d}F}{\mathrm{d}x} = \frac{\partial F}{\partial x} + \sum_{k=1}^{n} \left(\frac{\partial F}{\partial y_k} y_k' + \frac{\partial F}{\partial y_k'} y_k'' \right)$$

the general chain rule.

$$= \frac{\partial F}{\partial x} + \sum_{k=1}^{n} \left(y_k' \frac{\mathrm{d}}{\mathrm{d}x} \frac{\partial F}{\partial y_k'} + \frac{\partial F}{\partial y_k'} y_k'' \right)$$

on orbits.

$$= \frac{\partial F}{\partial x} + \sum_{k=1}^{n} \frac{\mathrm{d}}{\mathrm{d}x} \left(y_k' \frac{\partial F}{\partial y_k'} \right)$$

and this is again on orbits. Now we obtain

$$\Rightarrow \frac{\mathrm{d}}{\mathrm{d}x} \left\{ F - \sum_{k=1}^{n} y_k' \frac{\partial F}{\partial y_k'} \right\} = \frac{\partial F}{\partial x}$$

on orbits. We define $G = -\{\ \}$. Then

$$-\frac{\mathrm{d}G}{\mathrm{d}x} = \frac{\partial F}{\partial x}.$$

If $\frac{\partial F}{\partial x} = 0$, i.e. F does not depend *implicitly* on x, we have

$$-\frac{\mathrm{d}G}{\mathrm{d}x} = \frac{\mathrm{d}}{\mathrm{d}x} \left\{ \quad \right\} = 0$$

on orbits. i.e.

$$\boxed{F - \sum_{k=1}^{n} y_k' \frac{\partial F}{\partial y_k'} = c}$$

on orbits, where c is a constant. This is a first integral. i.e. G is constant on orbits. This is the Jacobi integral.

8. Double Integrals

Here we look at functions of 2 variables, i.e. surfaces,

$$\Phi = \Phi(x, y). \tag{8.1}$$

We then take the integral

$$J(\Phi) = \iint_R F(x, y, \Phi, \Phi_x, \Phi_y)\ \mathrm{d}x\mathrm{d}y \tag{8.2}$$

with $\Phi - \varphi_B =$ given on the boundary ∂R of R. Suppose $J(\Phi)$ has a minimum for $\Phi = \varphi$. Hence R is some closed regioon in the xy-plane and $\Phi_x = \frac{\partial \Phi}{\partial x}$, $\Phi_y = \frac{\partial \Phi}{\partial y}$. Assume F has continuous first and second derivatives with respect to x, y, Φ, Φ_x, Φ_y. Consider

$$J(\varphi + \varepsilon\xi) = \iint_R F(x, y, \varphi + \varepsilon\xi, \varphi_x + \varepsilon\xi_x, \varphi_y + \varepsilon\xi_y)\ \mathrm{d}x\mathrm{d}y$$

and expand in Taylor series

$$\begin{aligned} &= \iint_R \left\{ F(x, y, \varphi, \varphi_x, \varphi_y) + \varepsilon\xi\frac{\partial F}{\partial \varphi} + \varepsilon\xi_x\frac{\partial F}{\partial \varphi_x} + \varepsilon\xi_y\frac{\partial F}{\partial \varphi_y} + \mathrm{O}_2 \right\}\ \mathrm{d}x\mathrm{d}y \\ &= J(\varphi) + \delta J + \mathrm{O}_2 \end{aligned} \tag{8.3}$$

where we have

$$\delta J = \varepsilon \iint_R \left\{ \xi\frac{\partial F}{\partial \varphi} + \xi_x\frac{\partial F}{\partial \varphi_x} + \xi_y\frac{\partial F}{\partial \varphi_y} \right\}\ \mathrm{d}x\mathrm{d}y \tag{8.4}$$

is the first variation. For an extremum at φ, it is necessary that

$$\delta J = 0. \tag{8.5}$$

To use this we rewrite (8.4)

(8.6)
$$\delta J = \varepsilon \iint_R \left\{ \xi\frac{\partial F}{\partial \varphi} + \frac{\partial}{\partial x}\left(\xi\frac{\partial F}{\partial \varphi_x}\right) + \frac{\partial}{\partial y}\left(\xi\frac{\partial F}{\partial \varphi_y}\right) - \xi\frac{\partial}{\partial x}\left(\frac{\partial F}{\partial \varphi_x}\right) - \xi\frac{\partial}{\partial y}\left(\frac{\partial F}{\partial \varphi_y}\right) \right\}\ \mathrm{d}x\mathrm{d}y.$$

Now by Green's Theorem we have

$$\begin{aligned} &\iint_R \frac{\partial P}{\partial x}\ \mathrm{d}x\mathrm{d}y = \int_{\partial R} P\ \mathrm{d}y && P = \xi\frac{\partial F}{\partial \varphi_x} \\ &\iint_R \frac{\partial Q}{\partial y}\ \mathrm{d}x\mathrm{d}y = -\int_{\partial R} Q\ \mathrm{d}x && Q = \xi\frac{\partial F}{\partial \varphi_y} \end{aligned} \tag{8.7}$$

So,

$$\delta J = \varepsilon \iint_R \xi \left\{ \frac{\partial F}{\partial \varphi} - \frac{\partial}{\partial x}\frac{\partial F}{\partial \varphi_x} - \frac{\partial}{\partial y}\frac{\partial F}{\partial \varphi_y} \right\} \, \mathrm{d}x\mathrm{d}y + \varepsilon \int_{\partial R} \left(\xi \frac{\partial F}{\partial \varphi_x}\mathrm{d}y - \xi \frac{\partial F}{\partial \varphi_y}\mathrm{d}x \right). \tag{8.8}$$

If we choose all functions Φ including φ to statisfy

$$\begin{aligned} \varphi &= \varphi_B && \text{on } \partial R \\ \text{then } \varphi + \varepsilon\xi &= \varphi_B && \text{on } \partial R \\ \Rightarrow \xi &= 0 && \text{on } \partial R. \end{aligned}$$

Hence (8.8) simplifies to

$$\begin{aligned} \delta J &= \varepsilon \iint_R \xi(x,y) \left\{ \frac{\partial F}{\partial \varphi} - \frac{\partial}{\partial x}\frac{\partial F}{\partial \varphi_x} - \frac{\partial}{\partial y}\frac{\partial F}{\partial \varphi_y} \right\} \, \mathrm{d}x\mathrm{d}y \\ &\equiv \big\langle \varepsilon\xi, J'(\varphi) \big\rangle. \end{aligned}$$

By a simple extension of the Euler-Lagrange lemma, since ξ is arbitrary in R, we have $\delta J = 0 \to \varphi(x,y)$ is a solution of

$$\boxed{J'(\varphi) = \frac{\partial F}{\partial \varphi} - \frac{\partial}{\partial x}\frac{\partial F}{\partial \varphi_x} - \frac{\partial}{\partial y}\frac{\partial F}{\partial \varphi_y} = 0.} \tag{8.9}$$

This is the Euler-Lagrange Equation - a partial differential equation. We seek the solution φ which takes the given values φ_B on ∂R.

Example 8.1. $F = F(x, y, \varphi, \varphi_x, \varphi_y) = \frac{1}{2}\varphi_x^2 + \frac{1}{2}\varphi_y^2 + f\varphi$, where $f = f(x,y)$ is given. The Euler-Lagrange equation is

$$\frac{\partial F}{\partial \varphi} = f \qquad \frac{\partial F}{\partial \varphi_x} = \varphi_x \qquad \frac{\partial F}{\partial \varphi_y} = \varphi_y.$$

So,

$$\frac{\partial F}{\partial \varphi} - \frac{\partial}{\partial x}\frac{\partial F}{\partial \varphi_x} - \frac{\partial}{\partial y}\frac{\partial F}{\partial \varphi_y} = 0 \Rightarrow f - \frac{\partial}{\partial x}\varphi_x - \frac{\partial}{\partial y}\varphi_y = 0.$$

So, i.e. we are left with

$$\boxed{\frac{\partial^2 \varphi}{\partial x^2} + \frac{\partial^2 \varphi}{\partial y^2} = f,}$$

which is Poisson's Equation.

EXAMPLE 8.2. $F = F(x, t, \varphi, \varphi_x, \varphi_t) = \frac{1}{2}\varphi_x^2 - \frac{1}{2c^2}\varphi_t^2$. The Euler-Lagrange equation is

$$\begin{aligned} \frac{\partial F}{\partial x} - \frac{\partial}{\partial x}\frac{\partial F}{\partial \varphi_x} - \frac{\partial}{\partial t}\frac{\partial F}{\partial \varphi_t} &= 0 \\ \Rightarrow 0 - \frac{\partial}{\partial x}\varphi_x - \frac{\partial}{\partial t}\left(\frac{1}{c^2}\varphi_t\right) &= 0 \\ \Rightarrow \boxed{\frac{\partial^2 \varphi}{\partial x^2} = \frac{1}{c^2}\frac{\partial^2 \varphi}{\partial t^2}} \end{aligned}$$

This is the classical wave equation.

9. Canonical Euler-Equations (Euler-Hamilton)

9.1. The Hamiltonian

We have the Euler-Lagrange equations

$$\frac{\partial F}{\partial y_k} - \frac{\mathrm{d}}{\mathrm{d}x}\frac{\partial F}{\partial y_{k'}} = 0 \tag{9.1}$$

which give the critical curves $y_1, \ldots, y_n$ of

$$J(Y_1, \ldots, Y_n) = \int F(x, Y_1, \ldots, Y_n, Y_1', \ldots, Y_n')\ \mathrm{d}x \tag{9.2}$$

Equations (9.1) form a system of n 2nd order differential equations. We shall now rewrite this as a system of $2n$ first order differential equations. First we introduce a new variable

$$p_i = \frac{\partial F}{\partial y_i} \qquad i = 1, \ldots, n. \tag{9.3}$$

p_i is said to be the variable *conjugate* to y_i. We suppose that equations (9.3) can be solved to give y' as a function ψ_i of x, y_j, p_j $(j = 1, \ldots, n)$. Then it is possible to define a new function H by the equation

$$H(x, y_1, \ldots, y_n, p_1, \ldots, p_n) = \sum_{i=1}^{n} p_i y_i' - F(x, y_1, \ldots, y_n, y_1', \ldots, y_n') \tag{9.4}$$

where $y_i' = \psi_i(x, y_j, p_j)$. The function H is called the *Hamiltonian* corresponding to (9.2). Now look at the differential of H which by (9.4) is

$$dH = \sum_{i=1}^{n}(\cancel{p_i \mathrm{d}y_i'} + y_i'\mathrm{d}p_i) - \frac{\partial F}{\partial x}\mathrm{d}x - \sum_{i=1}^{n}\left(\frac{\partial F}{\partial y_i}\mathrm{d}y_i + \cancel{\frac{\partial F}{\partial y_i'}\mathrm{d}y_i'}\right)$$

$$= -\frac{\partial F}{\partial x}\mathrm{d}x + \sum_{i=1}^{n}\left(y_i'\mathrm{d}p_i - \frac{\partial F}{\partial y_i}\right) \tag{9.5}$$

using (9.3). For $H = (x, y_1, \ldots, y_n, p_1, \ldots, p_n)$ then we have

$$\mathrm{d}H = \frac{\partial H}{\partial x} + \sum_{i=1}^{n}\left(\frac{\partial H}{\partial y_i}\mathrm{d}y_i + \frac{\partial H}{\partial p_i}\mathrm{d}p_i\right).$$

Comparison with (9.5) gives

$$\begin{aligned} y_i' &= \frac{\partial H}{\partial p_i} & \qquad -\frac{\partial F}{\partial y_i} &= \frac{\partial H}{\partial y_i} \\ \Rightarrow \frac{\mathrm{d}y_i}{\mathrm{d}x} &= \frac{\partial H}{\partial p_i} & \qquad -\frac{\mathrm{d}p_i}{\mathrm{d}x} &= \frac{\partial H}{\partial y_i} \end{aligned} \tag{9.6}$$

for $i = 1, \ldots, n$. Equations (9.6) are the *canonical Euler-Lagrange equations* associated with the integral (9.2).

Example 9.1. Take

$$J(Y) = \int_a^b \left(\alpha(Y')^2 + \beta Y^2\right)\,\mathrm{d}x \tag{9.7}$$

where α, β are given functions of x. For this $F(x, y, y') = \alpha(y')^2 + \beta y^2$ if so

$$\phi = \frac{\partial F}{\partial y'} = 2\alpha y' \Rightarrow y' = \frac{1}{2\alpha}\phi.$$

The Hamiltonian H is, by (9.4),

$$\begin{aligned} H &= py' - F \\ &= py' - \alpha(y')^2 - \beta y^2 \qquad \text{with } y' = \frac{1}{2\alpha}\phi \\ &= p\frac{1}{2\alpha}p - \alpha\frac{1}{4\alpha^2}p^2 - \beta y^2 \\ &= \frac{1}{4\alpha}p^2 - \beta y^2 \end{aligned}$$

in correct variables x, y, p. Canonical equations are then

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{\partial H}{\partial p} = \frac{1}{2\alpha}p \qquad -\frac{\mathrm{d}p}{\mathrm{d}x} = \frac{\partial H}{\partial y} = -2\beta y. \tag{9.8}$$

The ordinary Euler-Lagrange equation for $J(Y)$ is

$$\frac{\partial F}{\partial y} - \frac{\mathrm{d}}{\mathrm{d}x}\frac{\partial F}{\partial y'} = 0 \Rightarrow 2\beta y - \frac{\mathrm{d}}{\mathrm{d}x}(2\alpha y') = 0$$

which is equivalent to (9.8)

9.2. **The Euler-Hamilton (Canonical) variational principle.**

Let

$$I(P,Y) = \int_a^b \left\{ P\frac{\mathrm{d}Y}{\mathrm{d}x} - H(x,Y,P) \right\} \mathrm{d}x$$

be defined for any admissible independent functions P and Y with $Y(a) = y_a$, $Y(b) = y_b$, i.e.

$$Y = y_B = \begin{cases} y_a & \text{at } x = a \\ y_b & \text{at } x = b. \end{cases}$$

Suppose $I(P,Y)$ is stationary at $Y = y$, $P = p$. Take varied curves

$$Y = y + \varepsilon\xi \qquad P = p + \varepsilon\eta.$$

Then we obtain

$$\begin{aligned} I(p+\varepsilon\eta, y+\varepsilon\xi) &= \int_a^b \left\{ (p+\varepsilon\eta)\frac{\mathrm{d}}{\mathrm{d}x}(y+\varepsilon\xi) - H(x, y+\varepsilon\xi, p+\varepsilon\eta) \right\} \mathrm{d}x \\ &= \int_a^b \left\{ p\frac{\mathrm{d}y}{\mathrm{d}x} + \varepsilon\eta\frac{\mathrm{d}y}{\mathrm{d}x} + \varepsilon p\frac{\mathrm{d}\xi}{\mathrm{d}x} + \varepsilon^2\eta\frac{\mathrm{d}\xi}{\mathrm{d}x} - H(x,y,p) - \varepsilon\xi\frac{\partial H}{\partial y} - \varepsilon\eta\frac{\partial H}{\partial p} - \mathrm{O}_2 \right\} \mathrm{d}x \\ &= I(p,y) + \delta I + \mathrm{O}_2. \end{aligned}$$

where we have

$$\begin{aligned} \delta I &= \varepsilon \int_a^b \left\{ \eta\frac{\mathrm{d}y}{\mathrm{d}x} + p\frac{\mathrm{d}\xi}{\mathrm{d}x} - \xi\frac{\partial H}{\partial y} - \eta\frac{\partial H}{\partial p} \right\} \mathrm{d}x \\ &= \varepsilon \int_a^b \left\{ \eta\left(\frac{\mathrm{d}y}{\mathrm{d}x} - \frac{\partial H}{\partial p}\right) - \xi\left(\frac{\mathrm{d}p}{\mathrm{d}x} + \frac{\partial H}{\partial y}\right) \right\} \mathrm{d}x + \underbrace{\big[\varepsilon p\xi\big]_a^b}_{=0} \end{aligned}$$

If all curves Y, including y, go through (a, y_a) and (b, y_b) then $\xi = 0$ at $x = a$ and $x = b$. Then $\delta I = 0 \Rightarrow (p, y)$ are solutions of

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{\partial H}{\partial p} \qquad -\frac{\mathrm{d}p}{\mathrm{d}x} = \frac{\partial H}{\partial y} \qquad (a < x < b)$$

with

$$y = y_B = \begin{cases} y_a & \text{at } x = a \\ y_b & \text{at } x = b \end{cases}$$

Note. If $Y = y_B$ no boundary conditions are required on p.

9.3. **An extension.**

Modify $I(P, Y)$ to be

$$I_{\text{mod}}(P, Y) = \int_a^b \left\{ P\frac{\mathrm{d}Y}{\mathrm{d}x} - H(x, P, Y) \right\} \, \mathrm{d}x - \Big[P(Y - y_B) \Big]_a^b$$

Here P and Y are any admissible functions. I_{mod} is stationary at $P = p$, $Y = y$. $Y = y + \varepsilon\xi$, $P = p + \varepsilon y$ and then make $\delta I_{\text{mod}} = 0$. You should find that (y, p) solve

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{\partial H}{\partial p} \qquad -\frac{\mathrm{d}p}{\mathrm{d}x} = \frac{\partial H}{\partial y}$$

with $y = y_B$ on $[a, b]$.

10. First Integrals of the Canonical Equations

A first integral of a system of differential equations is a function, which has constant value along each solution of the differential equations. We now look for first integrals of the canonical system

$$\frac{\mathrm{d}y_i}{\mathrm{d}x} = \frac{\partial H}{\partial p_i} \qquad -\frac{\mathrm{d}p_i}{\mathrm{d}x} = \frac{\partial H}{\partial y_i} \qquad (i = 1, \ldots, n) \tag{10.1}$$

and hence of the system

$$\frac{\mathrm{d}y_i}{\mathrm{d}x} = y_i' \qquad \frac{\partial F}{\partial y_i} - \frac{\mathrm{d}}{\mathrm{d}x}\frac{\partial F}{\partial y_i'} = 0 \qquad (i = 1, \ldots, n) \tag{10.2}$$

which is equivalent to (10.1). Take the case where

$$\frac{\partial F}{\partial x} = 0.$$

i.e. $F = F(y_1, \ldots, y_n, y_1', \ldots, y_n')$. Then

$$H = \sum_{i=1}^{n} p_i y_i' - F$$

is such that $\frac{\partial H}{\partial x} = 0$ and hence

$$\frac{\mathrm{d}H}{\mathrm{d}x} = \cancelto{0}{\frac{\partial H}{\partial x}} + \sum_{i=1}^{n} \left(\frac{\partial H}{\partial y_i} \frac{\mathrm{d}y_i}{\mathrm{d}x} + \frac{\partial H}{\partial p_i} \frac{\mathrm{d}p_i}{\mathrm{d}x} \right).$$

On critical curves (orbits) this gives

$$\begin{aligned} &= \sum_{i=1}^{n} \left(\frac{\partial H}{\partial y_i} \frac{\partial H}{\partial p_i} + \frac{\partial H}{\partial p_i} (-1) \frac{\partial H}{\partial y_i} \right) \\ &= 0 \end{aligned}$$

which implies H is constant on orbits. Consider now an arbitrary differentiable function $W = W()x, y_1, \ldots, y_n, p_1, \ldots, p_n)$. Then

$$\frac{\mathrm{d}W}{\mathrm{d}x} = \frac{\partial W}{\partial x} + \sum_{i=1}^{n} \left(\frac{\partial W}{\partial y_i} \frac{\mathrm{d}y_i}{\mathrm{d}x} + \frac{\partial W}{\partial p_i} \frac{\mathrm{d}p_i}{\mathrm{d}x} \right) = \frac{\partial W}{\partial x} + \sum_{i=1}^{n} \left(\frac{\partial W}{\partial y_i} \frac{\partial H}{\partial p_i} - \frac{\partial W}{\partial p_i} \frac{\partial H}{\partial y_i} \right)$$

on orbits.

Definition: We define the *Poisson bracket* of X and Y to be

$$[X, Y] = \sum_{i=1}^{n} \left(\frac{\partial X}{\partial y_i} \frac{\partial Y}{\partial p_i} - \frac{\partial X}{\partial p_i} \frac{\partial Y}{\partial y_i} \right)$$

with $X = X(y_i, p_i)$ and $Y = Y(y_i, p_i)$.

Then we have

$$\frac{\mathrm{d}W}{\mathrm{d}x} = \frac{\partial W}{\partial x} + [W, H]$$

on orbits. In the case when $\frac{\partial W}{\partial x} = 0$, we have

$$\frac{\mathrm{d}W}{\mathrm{d}x} = [W, H].$$

So,

$$\frac{\mathrm{d}W}{\mathrm{d}x} = 0 \Leftrightarrow [W, H] = 0.$$

11. Variable End Points (In the y Direction)

We let the end points be variable in the y direction

$$J(Y) = \int_a^b F(x, Y, Y') \, \mathrm{d}x.$$

So $y(x)$ is NOT prescribed at the end points. Suppose J is stationary for $Y = y$. Take $Y = y + \varepsilon\xi$. Then

$$\begin{aligned} J(y + \varepsilon\xi) &= \int_a^b F(x, y + \varepsilon\xi, y' + \varepsilon'\xi') \, \mathrm{d}x \\ &= \int_a^b \left\{ F(x, y, y') + \varepsilon\xi \frac{\partial F}{\partial y} + \varepsilon\xi' \frac{\partial F}{\partial y'} + \mathrm{O}_2 \right\} \, \mathrm{d}x. \end{aligned}$$

Thus

$$\begin{aligned} \delta J &= \int_a^b \left\{ \varepsilon\xi \frac{\partial F}{\partial y} + \varepsilon\xi' \frac{\partial F}{\partial y'} \right\} \, \mathrm{d}x \\ &= \int_a^b \varepsilon\xi \left\{ \frac{\partial F}{\partial y} - \frac{\mathrm{d}}{\mathrm{d}x} \frac{\partial F}{\partial y'} \right\} \, \mathrm{d}x + \left[\varepsilon\xi \frac{\partial F}{\partial y'} \right]_a^b. \end{aligned}$$

If $J(Y)$ has a minimum for $Y = y$, we need $\delta J = 0$. There are four possible cases in which we require this to happen. We realise these diagrammatically as

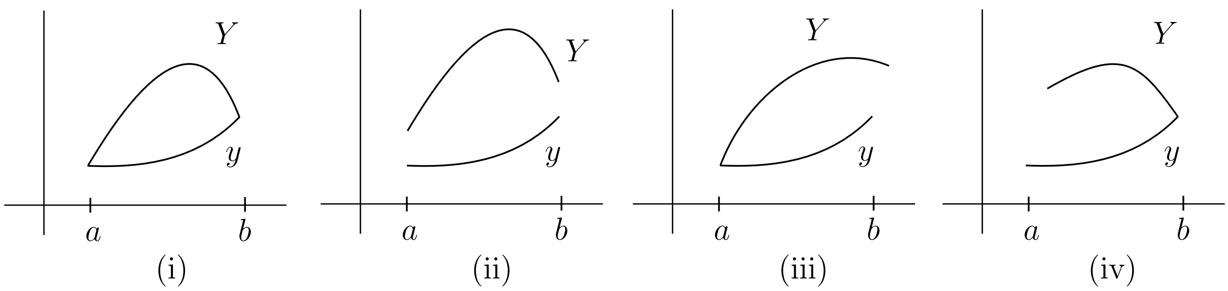

Figure 9. The four possible cases of varying end points in the direction of y.

i) In this case we have that $\xi = 0$ at $x = a$ and $x = b$. This gives us that

$$\begin{aligned} \delta J = \int_a^b \varepsilon\xi \left\{ \frac{\partial F}{\partial y} - \frac{\mathrm{d}}{\mathrm{d}x} \frac{\partial F}{\partial y'} \right\} \, \mathrm{d}x = 0 \\ \Rightarrow \frac{\partial F}{\partial y} - \frac{\mathrm{d}}{\mathrm{d}x} \frac{\partial F}{\partial y'} = 0 \qquad (a < x < b). \end{aligned}$$

which is just our standard Euler-Lagrange equation.

ii) In this case we have the complete opposite situation where $\xi \neq 0$ for either $x = a$ or $x = b$. This gives us that

$$\delta J = \left[\varepsilon\xi\frac{\partial F}{\partial y'}\right]_a^b = 0.$$

iii) In this case we are given that only one end point gives $\xi = 0$, namely $\xi(a) = 0$ but $\xi \neq 0$ for $x = b$. This gives us the criteria that

$$\varepsilon\xi\frac{\partial F}{\partial y'} = 0 \qquad \text{at } x = b.$$

iv) Now our final case is where $\xi(b) = 0$ but $\xi(a) \neq 0$. This gives us the condition

$$\varepsilon\xi\frac{\partial F}{\partial y'} = 0 \qquad \text{at } x = a.$$

Hence $\delta J = 0 \Rightarrow y$ satisfies the Euler-Lagrange equation

$$\frac{\partial F}{\partial y} - \frac{\mathrm{d}}{\mathrm{d}x}\frac{\partial F}{\partial y'} = 0 \qquad\qquad (a < x < b)$$

with

$$\begin{aligned} \xi\frac{\partial F}{\partial y'} &= 0 \qquad\qquad \text{at } x = a \\ \xi\frac{\partial F}{\partial y'} &= 0 \qquad\qquad \text{at } x = b. \end{aligned}$$

If Y is not prescribed at an end point, e.g. $x = b$, then we require $\frac{\partial F}{\partial y'} = 0$ at $x = b$. Such a condition is called a *natural boundary condition.*

Example 11.1 (The simplest problem, revisited). We go back to the Simplest Problem from chapter 1 where we have $F = (y')^2$ and thus

$$J(Y) = \int_0^1 (y')^2 \, \mathrm{d}x$$

with $Y(0) = 0$ and $Y(1) = 1$. The Euler-Lagrange equation is $y'' = 0$ which implies $y = ax + b$. Using the boundary conditions we get

$$\left.\begin{aligned} y(0) = 0 \Rightarrow b = 0 \\ y(1) = 1 \Rightarrow a = 1 \end{aligned}\right\} \qquad \Rightarrow \qquad y = x$$

which is our critical curve.

Example 11.2. Let us re-examine this problem in the light of our new theory. We have again that

$$J(Y) = \int_0^1 (Y')^2 \, \mathrm{d}x$$

but it is NOT prescribed at $x = 0, 1$. The Euler-Lagrange equation is $y'' = 0$ which implies that $y = \alpha x + \beta$. Now we have from the boundary conditions that at $x = 0$, 1 we have

$$\frac{\partial F}{\partial y'} = 0 \quad \Rightarrow \quad y'(1) = \alpha = 0 \quad \Rightarrow \quad \alpha = 0$$

and hence $y = \beta$.

12. Variable End Points: Variable in x and y Directions

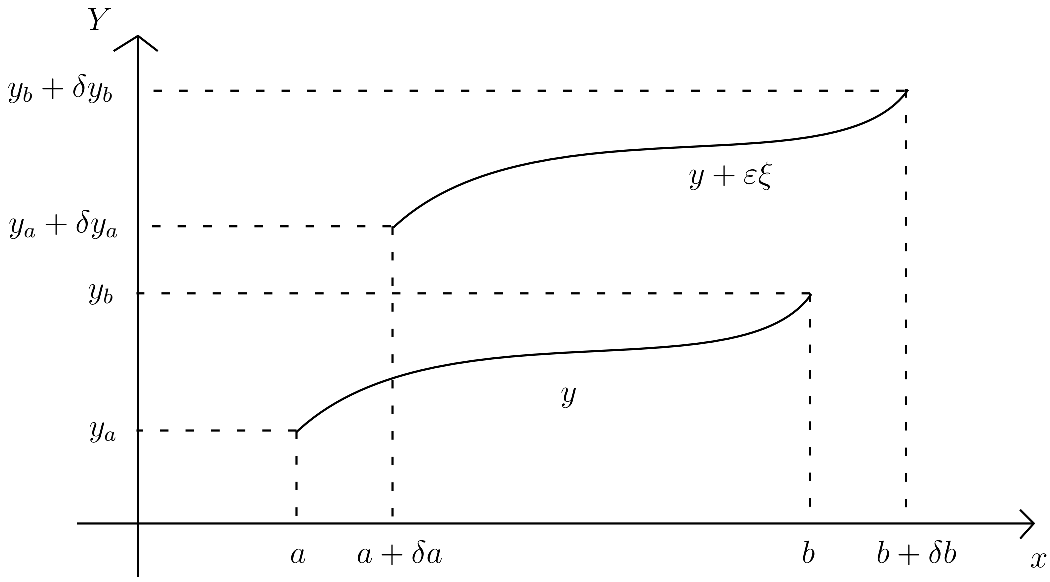

FIGURE 10. Variable end points in two directions.

The Euler-Lagrange equations ($a < x < b$) are

$$\xi \frac{\partial F}{\partial y'} = 0 \qquad \text{at } x = a$$
$$\xi \frac{\partial F}{\partial y'} = 0 \qquad \text{at } x = b.$$

Let our J to be minimised be

$$J(y) = \int_a^b F(x, y, y') \, \mathrm{d}x. \tag{12.1}$$

Here y is an extremal and write

$$y(a) = y_a \qquad y(b) = y_b. \tag{12.2}$$

Suppose that the varied curve $Y = y + \varepsilon\xi$ is defined over $(a + \delta a, b + \delta b)$ so that

$$J(y + \varepsilon\xi) = \int_{a+\delta a}^{b+\delta b} F(x, y + \varepsilon\xi, y' + \varepsilon\xi')\ \mathrm{d}x \tag{12.3}$$

and at the end points of this curve we write

$$\begin{cases} y + \varepsilon\xi = y_a + \delta y_a & \text{at} \quad x = a + \delta a \\ \qquad\quad\ = y_b + \delta y_b & \text{at} \quad x = b + \delta b. \end{cases} \tag{12.4}$$

To find the first variation δJ we must obtain the terms in $J(y+\varepsilon\xi)$ which are linear in the first order quantities $\varepsilon\xi$, δa, δb, δy_a, δy_b. The total variation (difference of)

$$\begin{aligned} \Delta J &= J(y + \varepsilon\xi) - J(y) \\ &= \int_{a+\delta a}^{b+\delta b} F(x, y + \varepsilon\xi, y' + \varepsilon\xi')\ \mathrm{d}x - \int_a^b F(x, y, y')\ \mathrm{d}x \end{aligned}$$

which we rewrite as

$$\begin{aligned} \Delta J = \int_a^b &\{F(x, y + \varepsilon\xi, y' + \varepsilon\xi') - F(x, y, y')\}\ \mathrm{d}x \\ &+ \int_b^{b+\delta b} F(x, y + \varepsilon\xi, y' + \varepsilon\xi')\ \mathrm{d}x \\ &- \int_a^{a+\delta a} F(x, y + \varepsilon\xi, y' + \varepsilon\xi')\ \mathrm{d}x. \end{aligned} \tag{12.5}$$

We have to assume here that y and $\varepsilon\xi$ are defined on the interval $(a, b+\delta b)$, i.e. extensions of y and $y + \varepsilon\xi$ are needed and we suppose that this is done (by Taylor expansion - see later). Then, expanding $F(x, y + \varepsilon\xi, y' + \varepsilon\xi')$ about (x, y, y').

$$\begin{aligned} \Delta J = \int_a^b &\left\{\varepsilon\xi\frac{\partial F}{\partial y} + \varepsilon\xi\frac{\partial F}{\partial y'} + \mathrm{O}(\varepsilon^2)\right\}\ \mathrm{d}x \\ &+ \int_b^{b+\delta b} \{F(x, y, y') + \mathrm{O}(\varepsilon)\}\ \mathrm{d}x \\ &- \int_a^{a+\delta a} \{F(x, y, y') + \mathrm{O}(\varepsilon)\}\ \mathrm{d}x \end{aligned} \tag{12.6}$$

and the linear terms in this case are

$$\begin{aligned}
\delta J &= \int_a^b \left\{ \varepsilon\xi \frac{\partial F}{\partial y} + \varepsilon\xi' \frac{\partial F}{\partial y'} \right\} \, \mathrm{d}x + F(x,y,y') \Big|_{x=b} \delta b - F(x,y,y') \Big|_{x=a} \delta a \\
&= \int_a^b \left\{ \varepsilon\xi \frac{\partial F}{\partial y} + \varepsilon\xi' \frac{\partial F}{\partial y'} \right\} + \Big[F(x,y,y')\delta x \Big]_{x=a}^{x=b} \\
&= \int_a^b \varepsilon\xi \left\{ \frac{\partial F}{\partial y} - \frac{\mathrm{d}}{\mathrm{d}x} \frac{\partial F}{\partial y'} \right\} \, \mathrm{d}x + \left[\varepsilon\xi \frac{\partial F}{\partial y'} + F(x,y,y')\delta x \right]_a^b . \qquad (12.7)
\end{aligned}$$

The final step is to find ξ in terms of δy, δa, δb at $x = a$ and $x = b$. Use (12.4) to obtain

$$\begin{aligned}
y_a + \delta y_b &= y(a + \delta a) + \varepsilon\xi(a + \delta a) \\
&= y(a) + \delta a y'(a) + \delta + \varepsilon\xi(a) + \delta a \varepsilon \xi'(a) + \cdots
\end{aligned}$$

First order terms are then

$$\begin{aligned}
\delta y_a &= \delta a y'(a) + \varepsilon\xi(a) \\
\Rightarrow \varepsilon\xi(a) &= \delta y_a - \delta a y'(a) \qquad (12.8) \\
\varepsilon\xi(b) &= \delta y_b - \delta b y'(b). \qquad (12.9)
\end{aligned}$$

We can now write (12.7) as

$$\delta J = \int_a^b \varepsilon\xi \left\{ \frac{\partial F}{\partial y} - \frac{\mathrm{d}}{\mathrm{d}x} \frac{\partial F}{\partial y'} \right\} \, \mathrm{d}x + \left[\frac{\partial F}{\partial y'} \delta y - \left(y' \frac{\partial F}{\partial y'} - F \right) \delta x \right]_a^b . \qquad (12.10)$$

This is the General First Variation of the integral J. If we introduce the canonical variable p conjugate to y defined by

$$p = \frac{\partial F}{\partial y'} \qquad (12.11)$$

and the Hamiltonian H given by

$$H = py' - F \qquad (12.12)$$

we can rewrite (12.10) as

$$\delta J = \int_a^b \varepsilon\xi \left(\frac{\partial F}{\partial y} - \frac{\mathrm{d}}{\mathrm{d}x} \frac{\partial F}{\partial y'} \right) \, \mathrm{d}x + \Big[p\delta y - H\delta x \Big]_a^b . \qquad (12.13)$$

If J has an extremum for $Y = y$, then $\delta J = 0 \Rightarrow y$ is a solution of

$$\frac{\partial F}{\partial y} - \frac{\mathrm{d}}{\mathrm{d}x} \frac{\partial F}{\partial y'} = 0 \qquad (a < x < b) \qquad (12.14)$$

with

$$\Big[p\delta y - H\delta x\Big]_a^b = 0. \tag{12.15}$$

13. Hamilton-Jacobi Theory

13.1. Hamilton-Jacobi Equations

For our standard integral

$$J(y) = \int_a^b F(x, y, y')\ \mathrm{d}x. \tag{13.1}$$

The general First Variation is

$$\delta J = \int_a^b \varepsilon\xi \left(F_y - \frac{\mathrm{d}}{\mathrm{d}x}F_{y'}\right)\ \mathrm{d}x + \Big[p\delta y - H\delta x\Big]_a^b \tag{13.2}$$

where $p = \frac{\partial F}{\partial y'}$ and $H = py' - F$.

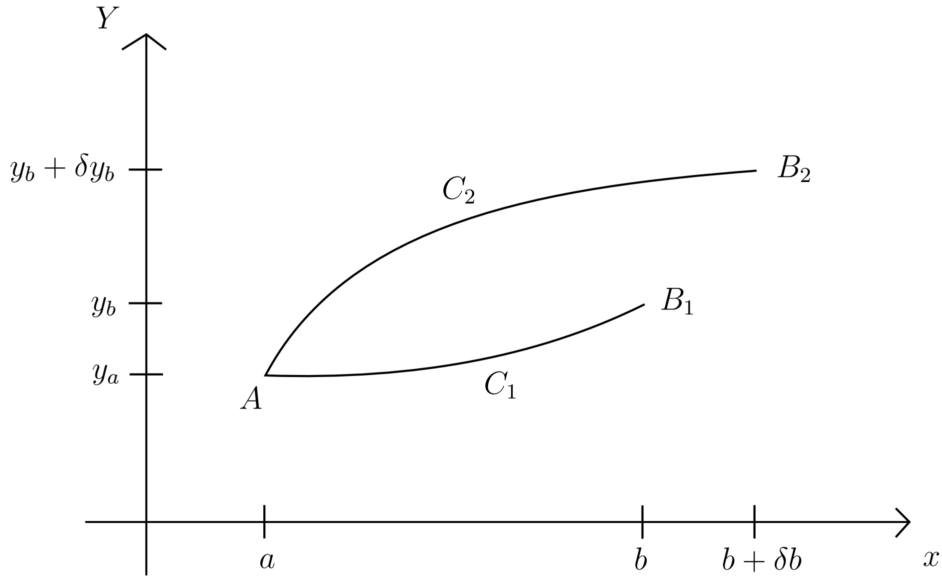

Figure 11. Variable end points in the y direction.

We apply (13.2) to the case in the above Figure, where A is fixed at (a, y_a) and C_1, C_2 are both extremal curves to $B_1(b, y_b)$ and $B_2(b + \delta b, y_b + \delta y_b)$ respectively. Then we have

$$\begin{aligned} J(C_1) &= \text{function of } B_1 = S(b, y_b) \\ J(C_2) &= \text{function of } B_2 = S(b + \delta b, y_b + \delta y_b) \end{aligned}$$

then $\Delta S \Rightarrow \delta S = p\delta y_b - H\delta b$. Consider the total variation

$$\Delta S = S(b + \delta b, y_b + \delta y_b) - S(b, y_b).$$

By (2), the first order terms of (3) are

$$\delta S = p\delta y_b - H\delta b. \tag{13.3}$$

Now, $B_1(b, y_b)$ may be *any* end point $B(x, y)$ say and by (4)

$$\delta S(x, y) = p\delta y - H\delta x. \tag{13.4}$$

Compare with, given $S(x, y)$,

$$\delta S = S_y \delta y + S_x \delta x$$

which gives

$$S_y = p \qquad S_x = -H. \tag{13.5}$$

Hence,

$$\boxed{S_x + H(x, y, S_y) = 0} \tag{13.6}$$

which is the Hamilton-Jacobi equation, where $p(= \frac{\partial F}{\partial y'}) = S_y$, y' denoting the derivative $\frac{\mathrm{d}y}{\mathrm{d}x}$ calculated at $B(x, y)$ for the extremal going from A to B.

Example 13.1 (Hamilton-Jacobi). Let our integral be

$$J(y) = \int_a^b (y')^2 \, \mathrm{d}x$$

and thus we have $F(x, y, y') = (y')^2$. Thus we have

$$\begin{aligned} p &= \frac{\partial F}{\partial y'} = 2y' \Rightarrow y' = \frac{1}{2}p \\ H &= py' - F \\ &= p\frac{1}{2}p - \left(\frac{1}{2}p\right)^2 \\ &= \frac{1}{4}p^2 \end{aligned}$$

Then the Hamilton-Jacobi equation is

$$\begin{aligned} &S_x + H(x, y, p = S_y) = 0 \\ \Rightarrow &S_x + \frac{1}{4}(S_y)^2 = 0 \end{aligned}$$

(1) How do we solve this for $S = S(x, y)$?
(2) Then, how do we find the extremals?

13.2. Hamilton-Jacobi Theorem

See hand out.

13.3. Structure of S

For $S = S(x, y, \alpha)$ we have

$$\frac{\mathrm{d}S}{\mathrm{d}x} = \frac{\partial S}{\partial x} + \frac{\partial S}{\partial y}\frac{\mathrm{d}y}{\mathrm{d}x}$$

by the chain rule. Now, in a critical curve C_0 we have

$$\frac{\partial S}{\partial x} = -H \qquad \frac{\partial S}{\partial y} = p.$$

So,

$$\begin{aligned} \frac{\mathrm{d}S}{\mathrm{d}x} &= -H + p\frac{\mathrm{d}y}{\mathrm{d}x} && \text{on } C_0 \\ \mathrm{d}S &= -H\mathrm{d}x + p\mathrm{d}y && \text{on } C_0. \end{aligned}$$

Integrating we obtain

$$\begin{aligned} S(x, y, \alpha) &= \int_{C_0} (p\mathrm{d}y - H\mathrm{d}x) \\ &= \int_{C_0} F \ \mathrm{d}x \end{aligned}$$

up to additive constant. So, Hamilton's principal function S is equal to $J(y)$, where $J(y)$ corresponds to $J(y)$ evaluated along a critical curve.

(1) Case $\frac{\partial H}{\partial x} = 0$, i.e. $H = H(y, p)$. In this case, by the result in 10, $H = \text{const.} = \alpha$ say in C_0. This $p = f(\alpha, y)$ say. Then

$$\begin{aligned} S &= \int_{C_0} (p\mathrm{d}y - H\mathrm{d}x) \\ &= \int_{C_0} \{f(\alpha, y)\mathrm{d}y - \alpha\mathrm{d}x\} \\ &= W(y) - \alpha x. \end{aligned}$$

Put this in the Hamilton-Jacobi equation to find $W(y)$ and hence S. Note: it is *additive* rather than $S = X(x)Y(y)$.

(2) Case $\frac{\partial H}{\partial y} = 0$, hence $H = H(x, p)$. On C_0

$$-\frac{\mathrm{d}p}{\mathrm{d}x} = \frac{\partial H}{\partial y} = 0 \Rightarrow p = \text{const.}$$

on C_0, say c. Then

$$\begin{aligned} S &= \int_{C_0} (p\mathrm{d}y - H\mathrm{d}x) \\ &= \int_{C_0} (c\mathrm{d}y - H(x, p = c)\mathrm{d}x) \\ &= cy - V(x). \end{aligned}$$

Again, this *separates* the x and y variables for the Hamilton-Jacobi equation. In this case, find $V(x)$ and hence S, by evaluating

$$V(x) = \int H(x, p = c)\ \mathrm{d}x.$$

13.4. **Examples**

Example 13.2. Let our integral be

$$J(y) = \int_a^b (y')^2\ \mathrm{d}x.$$

Then we have that $F = (y')^2$. Thus we have

$$\begin{aligned} p = F_{y'} = 2y' &\Rightarrow y' = \frac{1}{2}p \\ H = py' - F &\Rightarrow H = \frac{1}{4}p^2 \end{aligned}$$

The Hamilton-Jacobi equation is

$$\frac{\partial S}{\partial x} + H\left(x, y, p = \frac{\partial S}{\partial y}\right) = 0$$

$$(\star) \qquad \frac{\partial S}{\partial x} + \frac{1}{4}\left(\frac{\partial S}{\partial y}\right)^2 = 0$$

(1) Now, $\frac{\partial H}{\partial x} = 0$ here. So,

$$H = \text{const.} = \alpha$$

say n C_0. Hence, the the result of 13.3 we can write

$$S(x, y) = W(y) - \alpha x$$

Put this in $(\star)$ then

$$-\alpha + \frac{1}{4}\left(\frac{\mathrm{d}W}{\mathrm{d}y}\right)^2 = 0$$

$$\Rightarrow \frac{\mathrm{d}W}{\mathrm{d}y} = 2\sqrt{\alpha} \Rightarrow W(y) = 2\sqrt{\alpha}y$$

So, we have

$$\boxed{S = 2\sqrt{\alpha}y - \alpha x}.$$

The extemals are

$$\frac{\partial S}{\partial y} = p \Rightarrow p = 2\sqrt{\alpha}$$

$$\frac{\partial S}{\partial \alpha} = \beta \Rightarrow \frac{1}{\sqrt{\alpha}}y - x = \beta$$

(2) This is also an example of $\frac{\partial H}{\partial y} = 0$ on C_0

$$-\frac{\mathrm{d}p}{\mathrm{d}x} = \frac{\partial H}{\partial y} \to \frac{\mathrm{d}p}{\mathrm{d}x} = 0$$

$$= p = \text{const.} = c$$

say. Then

$$
\begin{aligned}
S(x,y) &= \int_{C_0} (p\mathrm{d}y - H\mathrm{d}x) \\
&= \int_{C_0} \left(c\mathrm{d}y - \frac{1}{2}c^2\mathrm{d}x \right) \\
&= cy - \frac{1}{2}c^2 x.
\end{aligned}
$$

Extremals are

$$
p = \frac{\partial S}{\partial y} = c \qquad\qquad \beta = \frac{\partial S}{\partial c} \qquad\qquad \beta = \text{const.}
$$

So, $\beta = \frac{\partial S}{\partial c} = y - cx \Rightarrow y = cx + \beta$, which are straight lines (c, β constants).

Example 13.3. We have that $F(x,y,y') = y\sqrt{1+(y')^2}$ and we note $\frac{\partial F}{\partial x} = 0$. Now,

$$
p = \frac{\partial F}{\partial y'} = \frac{yy'}{\sqrt{1+(y')^2}} \Rightarrow y' = \frac{p}{\sqrt{y^2 - p^2}}.
$$

Also,

$$
\begin{aligned}
H &= py' - F(x,y,y') \qquad\qquad \text{with } y' = \frac{p}{\sqrt{y^2-p^2}} \\
&= -\sqrt{y^2 - p^2}.
\end{aligned}
$$

The Hamilton-Jacobi equation $H + \frac{\partial S}{\partial x} = 0$ is in this case

$$
-\left\{ y^2 - \left(\frac{\partial S}{\partial y} \right)^2 \right\}^{\frac{1}{2}} + \frac{\partial S}{\partial x} = 0
$$

i.e. we have

$$
\boxed{\left(\frac{\partial S}{\partial x} \right)^2 + \left(\frac{\partial S}{\partial y} \right)^2 = y^2}
$$

Hence $\frac{\partial H}{\partial x} = 0$ and so $H = \text{const.} = \alpha$ say on C_0. So, we can write $S(x,y) = W(y) + \alpha x$

$$
\begin{aligned}
&\left(\frac{\mathrm{d}W}{\mathrm{d}y} \right)^2 = y^2 - \alpha^2 \\
&\Rightarrow W(y) = \int^y (v^2 - \alpha^2)^{\frac{1}{2}} \,\mathrm{d}v.
\end{aligned}
$$

So we have that

$$S(x, y) = \int^y (v^2 - \alpha^2)^{\frac{1}{2}} \, \mathrm{d}v + \alpha x.$$

To find the extremals we use the Hamilton-Jacobi theorem

$$\frac{\partial S}{\partial y} = p \qquad\qquad \frac{\partial S}{\partial \alpha} = \text{const.} = \beta$$

say. Then

$$\begin{aligned} \frac{\partial S}{\partial \alpha} = \beta \Rightarrow \beta &= -\int^y \frac{\alpha}{\sqrt{v^2 - \alpha^2}} \, \mathrm{d}v + x \\ &\Rightarrow y = \alpha \cosh\left(\frac{x - \beta}{\alpha}\right) \end{aligned}$$

Part II - Extremum Princples

14. The Second Variation - Introduction to Minimum Problems

So far we have considered the *stationary* aspect of variational problems (i.e. $\delta J = 0$). Now we shall look at the max, or minimum aspect. That is, we look for the function y (if there is one) which makes $J(Y)$ take a maximum or a minimum value in a certain class Ω of functions. So

$$\begin{aligned} J(y) &\leqslant J(Y), \quad Y \in \Omega \quad \text{a MINIMUM} \\ J(Y) &\leqslant J(y), \quad Y \in \Omega \quad \text{a MAXIMUM.} \end{aligned}$$

In what follows we shall discuss the MIN case, since the MAX case is equivalent to a MIN case for $-J(Y)$. For the integral

$$J(Y) = \int_a^b F(x, Y, Y') \, \mathrm{d}x$$

which is stationary for $Y = y$ ($\delta J = 0$), we have

$$\begin{aligned} J(y + \varepsilon\xi) &= J(y) + \frac{1}{2}\varepsilon^2 \int_a^b \left\{\xi^2 F_{yy} + 2\xi\xi' F_{yy'} + (\xi')^2 F_{yy'}\right\} \mathrm{d}x + \mathrm{O}_3 \\ &= J(y) + \delta^2 J + \mathrm{O}_3. \end{aligned}$$

So, providing ξ and ξ' are small enough, the second variation $\delta^2 J$ determines the sign of

$$\Delta J = J(Y) - J(y).$$

For quadratic problems $\big($e.g. $F = ay^2 + byy' + c(y')^2\big)$ the O_3 terms are zero and

$$\Delta J = J(Y) - J(y) = \delta^2 J.$$

15. Quadratic Problems

Consider

$$J(Y) = \int_a^b \left\{ \frac{1}{2} v (Y')^2 + \frac{1}{2} w Y^2 - fY \right\} \, \mathrm{d}x \tag{15.1}$$

with

$$Y(a) = y_a \qquad Y(b) = y_b \tag{15.2}$$

where v, w and f are given functions of x in general. The associated Euler-Lagrange equation is

$$-\frac{\mathrm{d}}{\mathrm{d}x}\left\{ v \frac{\mathrm{d}y}{\mathrm{d}x} \right\} + wy = f \qquad a < x < b \tag{15.3}$$

with

$$y(a) = y_a \qquad y(b) = y_b. \tag{15.4}$$

If y is a critical curve (solution of (15.3) and (15.4)) then (15.1)

$$\Delta J = J(y + \varepsilon\xi) - J(y) = \frac{1}{2} \int_a^b \left\{ v(\varepsilon\xi')^2 + w(\varepsilon\xi)^2) \right\} \, \mathrm{d}x \tag{15.5}$$

with $\xi = 0$ at $x = a$, b. Has this a definite sign?

$$\begin{aligned}
J(y + \varepsilon\xi) &= \int_a^b \left\{ \frac{1}{2} v (y' + \varepsilon\xi')^2 + \frac{1}{2} w (y + \varepsilon\xi)^2 - f(y + \varepsilon\xi) \right\} \, \mathrm{d}x \\
&= \int_a^b \left\{ \frac{1}{2} v (y')^2 + \frac{1}{2} w y^2 - fy \right\} \, \mathrm{d}x + \varepsilon \int_a^b \left\{ v y' \xi' + w y \xi - f \xi \right\} \, \mathrm{d}x \\
&\quad + \varepsilon^2 \int_a^b \left\{ \frac{1}{2} v (\xi')^2 + \frac{1}{2} w \xi^2 \right\} \, \mathrm{d}x \\
&= J(y) + \delta J + \delta^2 J
\end{aligned}$$

and we have

$$\begin{aligned}\delta J &= \varepsilon \int_a^b \{vy'\xi' + wy\xi - f\xi\} \; \mathrm{d}x \\ &= \varepsilon \int_a^b \left\{-\xi \frac{\mathrm{d}}{\mathrm{d}x}(vy') + wy\xi - f\xi\right\} \; \mathrm{d}x + \underbrace{[vy'\xi]_a^b}_{=0} \\ &= \varepsilon \int_a^b \xi \left\{-\frac{\mathrm{d}}{\mathrm{d}x}(vy') + wy - f\right\} \; \mathrm{d}x\end{aligned}$$

We consider

$$\begin{aligned}\Delta J &= J(y+\varepsilon\xi) - J(y) \\ &= \delta^2 J \\ &= \frac{1}{2}\int_a^b \left\{v(\varepsilon\xi')^2 + w(\varepsilon\xi)^2\right\} \; \mathrm{d}x. \end{aligned} \tag{15.6}$$

Has this a definite sign?

Special Case

Take v and w to be constants and let $v > 0$ (if $v < 0$ change the sign of J). Then

$$\Delta J = \frac{1}{2} v\varepsilon^2 \int_a^b \left\{(\xi')^2 + \frac{w}{v}\xi^2\right\} \; \mathrm{d}x \tag{15.7}$$

The sign of ΔJ depends only on sign of

$$K(\xi) = \int_a^b (\xi') + \frac{w}{v}\xi^2 \; \mathrm{d}x$$

where $\xi(a) = 0$ and $\xi(b) = 0$. Clearly $K \geqslant 0$ if $\frac{w}{v} \geqslant 0$. To go further, integrate ξ' therm by parts. Then

$$K(\xi) = \int_a^b \xi \left\{-\frac{\mathrm{d}^2}{\mathrm{d}x^2} + \frac{w}{v}\right\} \; \xi \mathrm{d}x + \underbrace{[\xi\xi']_a^b}_{=0}. \tag{15.8}$$

Try to imagine ξ is expressed as eigenfunctions of operator $\{\ \}$. Then,

$$-\frac{\mathrm{d}^2}{\mathrm{d}x^2}\varphi_n - \lambda_n\varphi_n \qquad \varphi_n = \sin\left\{\frac{\pi n(x-a)}{b-a}\right\} \qquad \lambda_n = \frac{n^2\pi^2}{(b-a)^2}.$$

Expand

$$\xi = \sum_{n=1}^{\infty} a_n \varphi_n(x)$$

where $\varphi_n = \sin\left\{\frac{\pi n(x-a)}{b-a}\right\}$. So,

$$K(\xi) = \int_a^b \sum_{m=1}^{\infty} a_m \varphi_m(x) \left\{-\frac{\mathrm{d}^2}{\mathrm{d}x^2} + \frac{w}{v}\right\} \sum_{n=1}^{\infty} a_n \varphi_n(x)\ \mathrm{d}x \tag{15.9}$$

where $\int_a^b \varphi_n \varphi_m\ \mathrm{d}x = 0$ for $n \neq m$. Then

$$= \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} a_m a_n \int_a^b \underbrace{\varphi_m \left(-\frac{\mathrm{d}^2}{\mathrm{d}x^2} + \frac{w}{v}\right) \varphi_n}_{=\varphi_m\left(\frac{n^2\pi^2}{(b-a)^2} + \frac{w}{v}\right)\varphi_n}\ \mathrm{d}x \tag{15.10}$$

$$= \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} a_m a_n \left(\frac{n^2\pi^2}{(b-a)^2} + \frac{w}{v}\right) \underbrace{\int_a^b \varphi_n \varphi_m\ \mathrm{d}x}_{=0 \text{ for } m\neq n} \tag{15.11}$$

$$= \sum_{n=1}^{\infty} a_n^2 \left(\frac{n^2\pi^2}{(b-a)^2} + \frac{w}{v}\right) \int_a^b \varphi_n^2\ \mathrm{d}x. \tag{15.12}$$

Hence, in (12), if

$$\frac{\pi^2}{(b-a)^2} + \frac{w}{v} \geqslant 0 \tag{15.13}$$

($n = 1$ term) then

$$K(\xi) \geqslant 0 \tag{15.14}$$
$$\Rightarrow \Delta J \geqslant 0, \tag{15.15}$$

which is the MINIMUM PRINCIPLE. Note (15.13) is equivalent to

$$w \geqslant -\frac{v\pi^2}{(b-a)^2}. \tag{15.16}$$

So, for some negative w we shall get a MIN principle of $J(Y)$. (15.13) gives us that if

$$\frac{w}{v} \geqslant -\frac{\pi^2}{(b-a)^2}$$

then we have (7)

$$\int_a^b \left\{(\xi')^2 + \frac{w}{v}\xi^2\right\} \geqslant 0.$$

i.e.

$$\int_a^b \left\{ (\xi')^2 - \frac{\pi^2}{(b-a)^2}\xi^2 \right\} \, \mathrm{d}x \geqslant 0$$

with $\xi(a) = 0 = \xi(b)$

Theorem 15.1 (Wirtinger's Inequality)**.**

$$\int_a^b (u')^2 \, \mathrm{d}x \geqslant \frac{\pi^2}{(b-a)^2} \int_a^b u^2 \, \mathrm{d}x \tag{15.17}$$

for all u such that $u(a) = 0 = u(b)$.

16. Isoperimetric (constraint) problems

16.1. Lagrange Multipliers in Elementary Calculus

Find the extremum of $z = f(x, y)$ subject to $g(x, y) = 0$. Suppose $g(x, y) = 0 \Rightarrow y = y(x)$. Then $z = f(x, y(x)) = z(x)$. Make $\frac{\mathrm{d}z}{\mathrm{d}x} = 0$, i.e.

$$\frac{\partial f}{\partial x} + \frac{\partial f}{\partial y}\frac{\mathrm{d}y}{\mathrm{d}x} = 0.$$

Now, from $g(x, y) = 0$

$$\frac{\partial g}{\partial x} + \frac{\partial g}{\partial y}\frac{\mathrm{d}y}{\mathrm{d}x} = 0.$$

We now eliminate $\frac{\mathrm{d}y}{\mathrm{d}x}$ in the above equations to obtain

$$\frac{\partial f}{\partial x} + \frac{\partial f}{\partial y}(-1)\frac{\frac{\partial g}{\partial x}}{\frac{\partial x}{\partial y}} = 0 \tag{16.1}$$

Example 16.1. Let $z = f(x, y) = xy$ with constraint $g(x, y) = x+y-1 = 0 \Rightarrow y = 1-x$. We have discussed this problem in Example 1.4. In this case we have

$$z = f(x, y(x)) = x(x-1)$$

which gives us that

$$\frac{\mathrm{d}z}{\mathrm{d}x} = 1 - 2x = 0.$$

Going back to our theory we obtain

$$y + x\frac{\mathrm{d}y}{\mathrm{d}x} = 0$$
$$1 - x + x\frac{\mathrm{d}y}{\mathrm{d}x} = 0.$$

We also obtain

$$1 + 1\frac{\mathrm{d}y}{\mathrm{d}x} = 0.$$

Thus we have $y + x(-1) = 0$, $1 - x - x = 0$, $1 - 2x = 0$. This implies

$$\begin{aligned} x &= \frac{1}{2} \\ y &= 1 - x = \frac{1}{2} \end{aligned}$$

Thus $z'' = -2 \Rightarrow$ MAX at $(\frac{1}{2}, \frac{1}{2})$.

16.2. **Lagrange Multipliers**

We form

$$V(x, y, \lambda) = f(x, y) + \lambda g(x, y)$$

with no constraints on V. Make V stationary:

$$\text{(a)} \qquad \frac{\partial V}{\partial x} = \frac{\partial f}{\partial x} + \lambda \frac{\partial g}{\partial x} = 0$$

$$\text{(b)} \qquad \frac{\partial V}{\partial y} = \frac{\partial f}{\partial y} + \lambda \frac{\partial g}{\partial y} = 0$$

$$\text{(c)} \qquad \frac{\partial V}{\partial \lambda} = g(x, y) = 0$$

Combining (a) and (b) we obtain

$$\frac{\partial f}{\partial x} + \frac{\partial g}{\partial x}(-1)\frac{\frac{\partial f}{\partial y}}{\frac{\partial g}{\partial y}} = 0.$$

As before in (16.1) we have that (c) is the constraint. We solve (a), (b) and (c) for x, y and λ.

Example 16.2. Returning to our example from before we want to establish a maximum at $(\frac{1}{2}, \frac{1}{2})$ for f. Take the points

$$x = \frac{1}{2} + h \qquad y = \frac{1}{2} + k.$$

Then we obtain that

$$\begin{aligned} f\left(\frac{1}{2} + h, \frac{1}{2} + k\right) &= \left(\frac{1}{2} + h\right)\left(\frac{1}{2} + k\right) \\ &= \frac{1}{2}\frac{1}{2} + \frac{1}{2}(h + k) + hk \end{aligned}$$

but this has linear terms in our stationary points. Don't forget the constraint! Now, the point $(\frac{1}{2}+h, \frac{1}{2}+k)$ must satisfy

$$\begin{aligned} g &= x+y-1=0 \\ \frac{1}{2}+h+\frac{1}{2}+k-1 &= 0 \Rightarrow h+k=0. \end{aligned}$$

Hence,

$$\begin{aligned} f\left(\frac{1}{2}+h, \frac{1}{2}+k\right) &= \frac{1}{2}\frac{1}{2}-h^2 \\ &\leqslant \frac{1}{2}\cdot\frac{1}{2} \Rightarrow \text{MAX at } \frac{1}{2}, \frac{1}{2}. \end{aligned}$$

Problem with 2 Constraints

Find extremum of $f(x,y,z)$ subject to $g_1(x,y,z)=0$ and $g_2(x,y,z)=0$. This could be quite complicated. But Lagrange multipliers gives a simple extension. Form

$$V(x,y,z,\lambda_1,\lambda_2) = f_1 + \lambda_1 g_1 + \lambda_2 g_2.$$

We make V stationary, so

$$V_x = 0 \qquad V_y = 0 \qquad V_z = 0 \qquad V_{\lambda_1} = 0 \qquad V_{\lambda_2} = 0.$$

16.3. Isoperimetric Problems: Constraint Problems in Calculus of Variations

Find an extremum of

$$J(Y) = \int_a^b F(x,Y,Y')\ \mathrm{d}x$$

with boundary conditions $Y(a)=y_a$, $Y(b)=y_b$. This is subject to the constraint

$$K(Y) = \int_a^b G(x,Y,Y')\ \mathrm{d}x = k$$

where k is a constant. Note that $g(Y)=K(y)-k=0$. If we take varied curves $Y=y+\varepsilon\xi$ then $J=J(\varepsilon)$ and $K=K(\varepsilon)$ but we note that J and K are functions of one variable. We need at least two variables! So, take the infinite set of varied curves $Y=y+\varepsilon_1\xi_1+\varepsilon_2\xi_2$ (where y is a solution curve). Using this formula for Y we have that $J(Y)$ becomes some function $f(\varepsilon_1,\varepsilon_2)$. Now,

$$K(Y)-k=0 \qquad g(\varepsilon_1,\varepsilon_2)=0 \Rightarrow \varepsilon_1,\varepsilon_2$$

not independent. Now use the method of Lagrange multipliers to form

$$E(\varepsilon_1, \varepsilon_2, \lambda) = f + \lambda g.$$

Make E stationary at $\varepsilon_1 = 0$, $\varepsilon_2 = 0$ (i.e. $Y = y$). Then,

$$\left.\begin{aligned} \frac{\partial f}{\partial \varepsilon_1} + \lambda \frac{\partial g}{\partial \varepsilon_1} = 0 \\ \frac{\partial f}{\partial \varepsilon_2} + \lambda \frac{\partial g}{\partial \varepsilon_2} = 0 \end{aligned}\right\} \text{ at } \varepsilon_1 = \varepsilon_2 = 0.$$

Now

$$\begin{aligned} f(\varepsilon_1, \varepsilon_2) &= J(y + \varepsilon_1 \xi_1 + \varepsilon_2 \xi_2) \\ &= J(y) + \langle \varepsilon_1 \xi_1 + \varepsilon_2 \xi_2, J'(y) \rangle + \mathrm{O}_2 \end{aligned}$$

and we get

$$\begin{aligned} g(\varepsilon_1, \varepsilon_2) &= K(y + \varepsilon_1 \xi_1 + \varepsilon_2 \xi_2) - k \\ &= K(y) - k \\ &= \langle \varepsilon_1 \xi_1 + \varepsilon_2 \xi_2, K'(y) \rangle + \mathrm{O}_2. \end{aligned}$$

Taking partial derivatives we obtain

$$\begin{aligned} \frac{\partial f}{\partial \varepsilon_1} &= \langle \xi_1, J'(y) \rangle \text{ at } \varepsilon_1 = \varepsilon_2 = 0 & \qquad \frac{\partial g}{\partial \varepsilon_1} &= \langle \xi_1, K'(y) \rangle \text{ at } \varepsilon_1 = \varepsilon_2 = 0 \\ \frac{\partial f}{\partial \varepsilon_2} &= \langle \xi_2, J'(y) \rangle \text{ at } \varepsilon_1 = \varepsilon_2 = 0 & \qquad \frac{\partial g}{\partial \varepsilon_2} &= \langle \xi_2, K'(y) \rangle \text{ at } \varepsilon_1 = \varepsilon_2 = 0. \end{aligned}$$

Hence the stationary conditions imply that

$$\begin{aligned} \langle \xi_1, J'(y) + \lambda K'(y) \rangle &= 0 \\ \langle \xi_2, J'(y) + \lambda K'(y) \rangle &= 0. \end{aligned}$$

Since ξ_1 and ξ_2 are arbitrary and independent functions in (a), (b) and y is a solution of

$$\begin{aligned} J'(y) + \lambda K'(y) = 0 \\ \Rightarrow \left(\frac{\partial F}{\partial y} - \frac{\mathrm{d}}{\mathrm{d}x} \frac{\partial F}{\partial y} \right) + \lambda \left(\frac{\partial g}{\partial y} - \frac{\mathrm{d}}{\mathrm{d}x} \frac{\partial g}{\partial y'} \right) = 0. \end{aligned}$$

Let $L = F + \lambda G$ then we have

$$\frac{\partial L}{\partial y} - \frac{\mathrm{d}}{\mathrm{d}x} \frac{\partial L}{\partial y'} = 0.$$

This is the first necessary condition.

Euler's Rule

Make $J + \lambda K$ stationary and satisfy $K(Y) = k$. Check for

$$\begin{cases} \min & J(y) \leqslant J(Y) \\ \max & J(y) \geqslant J(Y) \end{cases}$$

Here we assume $K'(y) \neq 0$, i.e. y is not an extremal of $K(y)$.

Example 16.3. We want to maximize

$$J(y) = \int_{-a}^{a} y \, \mathrm{d}x$$

subject to the constraints $y(-a) = y(a) = 0$ and

$$K(y) = \int_{-a}^{a} \sqrt{1 + (y')^2} \, \mathrm{d}x = L.$$

Now, use Euler's Rule. Let $F = y$, $G = \sqrt{1 + (y')^2}$ and form

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

This gives us

$$I(y) = \int_{-a}^{a} L \, \mathrm{d}x.$$

Make $I(y)$ stationary and hence solve

$$\begin{aligned}
\frac{\partial L}{\partial y} - \frac{\mathrm{d}}{\mathrm{d}x}\frac{\partial L}{\partial y'} &= 0 \\
\Rightarrow 1 - \frac{\mathrm{d}}{\mathrm{d}x}\frac{\lambda y'}{\sqrt{1 + (y')^2}} &= 0 \\
\Rightarrow \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\lambda y'}{\sqrt{1 + (y')^2}}\right) &= 1 \\
\Rightarrow \frac{\lambda y'}{\sqrt{1 + (y')^2}} &= x + c \\
\Rightarrow \frac{\lambda^2 (y')^2}{1 + (y')^2} &= (x + c)^2 \\
\Rightarrow y' &= \frac{x + c}{\sqrt{\lambda^2 - (x + c)^2}} \\
\Rightarrow y &= -\big(\lambda^2 - (x + c)^2\big)^{\frac{1}{2}} + c_1 \\
\Rightarrow (y - c_1)^2 &= \lambda^2 - (x + c)^2
\end{aligned}$$

which is (part of) a cricle, centred at $(-c, c_1)$ and radius $r = |\lambda|$. Or $L = y + \lambda\sqrt{1 + (y')^2}$. However, note that $\frac{\partial L}{\partial x} = 0$ and hence there is a first integral

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

The equation above can be written

$$(x - c)^2 + (y - c_1)^2 = \lambda^2.$$

This must satisfy the conditions that $y(-a) = 0 = y(a)$ and $K(y) = L$. This can be realised diagramatically as:

Example 16.4. Find the extremum of

$$J(y) = \int_0^1 y^2 \, \mathrm{d}x$$

subject to the constraints

$$K(y) = \int_0^1 (y')^2 \, \mathrm{d}x = \text{const.} = k^2 \qquad y(0) = 0 = y(1)$$

We use Euler's Rule and form

$$I(y) = J(y) + \lambda K(y) = \int_0^1 L \, \mathrm{d}x$$

where $L = y^2 + \lambda(y')^2$. Make $I(y)$ stationary solve

$$\frac{\partial L}{\partial y} - \frac{\mathrm{d}}{\mathrm{d}x}\frac{\partial L}{\partial y'} = 0$$

$$\Rightarrow 2y - \frac{\mathrm{d}}{\mathrm{d}x}(2\lambda y') = 0$$

$$\Rightarrow \boxed{y'' - \frac{1}{\lambda}y = 0}$$

There are three cases for this

(1) $\frac{1}{\lambda} > 0$ where $\sin\frac{1}{x} = \alpha^2$ $(\alpha \neq 0)$ then

$$y = A\cosh\alpha x + B\sinh\alpha x.$$

(2) $\frac{1}{\lambda} = 0$ then

$$y = ax + b.$$

(3) $\frac{1}{\lambda} < 0$ where $\sinh^{-1}\frac{1}{\lambda} = -\beta^2$ $(\beta \neq 0)$ then

$$y = C\cos\beta x + D\sin\beta x.$$

Now, from the constraints we have

$$\begin{aligned} y(0) = 0 &\Rightarrow A = 0, b = 0, C = 0 \\ y(1) = 0 &\Rightarrow B = 0, a = 0, D \sin\beta = 0 \\ &\Rightarrow \sin\beta = 0 \\ &\Rightarrow \beta = n\pi \end{aligned}$$

with $n = 1, 2, \ldots$. So, $y = D \sin n\pi x$ with $n = 1, 2, 3, \ldots$. Take $y_n = \sin n\pi x$. To include all these functions we take

$$y = \sum_{n=1}^{\infty} c_n y_n \qquad y_n = \sin n\pi x.$$

Now calculate $J(y)$ and $K(y)$ for this y function.

$$\begin{aligned} J(y) &= \int_0^1 y^2 \, \mathrm{d}x \\ &= \int_0^1 \left(\sum_{n=1}^{\infty} c_n \sin n\pi x \right) \left(\sum_{m=1}^{\infty} c_m \sin n\pi x \right) \, \mathrm{d}x \\ &= \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} c_n c_m \underbrace{\int_0^1 \sin(n\pi x) \sin(m\pi x) \, \mathrm{d}x}_{=\frac{1}{2}\delta_{nm}} \\ &= \sum_{n=1}^{\infty} c_n^2 \frac{1}{2}. \end{aligned}$$

Also,

$$\begin{aligned} k^2 = K(y) &= \int_0^1 (y')^2 \, \mathrm{d}x \\ &= \int_0^1 \left(\sum_{n=1}^{\infty} c_n n\pi \cos(n\pi x) \right) \left(\sum_{m=1}^{\infty} c_m \cos(m\pi x) \right) \, \mathrm{d}x \\ &= \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} c_n c_m n m \pi^2 \underbrace{\int_0^1 \cos(n\pi x) \cos(m\pi x) \, \mathrm{d}x}_{\frac{1}{2}\delta_{nm}} \\ &= \sum_{n=1}^{\infty} c_n^2 n^2 \pi^2 \frac{1}{2}. \end{aligned}$$

Now,

$$\begin{aligned}
J(y) &= \sum_{n=1}^{\infty} c_n^2 \frac{1}{2} \\
K(y) &= \sum_{n=1} c_n^2 n^2 \pi^2 \frac{1}{2} = k^2 \\
&\Rightarrow c_1^2 \frac{1}{2} = \frac{k^2}{\pi^2} - \sum_{n=2}^{\infty} c_n^2 n^2 \frac{1}{2}.
\end{aligned}$$

Write

$$\begin{aligned}
J(y) &= c_1^2 \frac{1}{2} + \sum_{n=2}^{\infty} c_n^2 \frac{1}{2} \\
&= \frac{k^2}{\pi^2} - \sum_{n=2}^{\infty} c_n^2 n^2 \frac{1}{2} + \sum_{n=2}^{\infty} c_n^2 \frac{1}{2} \\
&= \frac{k^2}{\pi^2} - \sum_{n=2}^{\infty} c_n^2 \frac{1}{2}(n^2 - 1) \\
&\leqslant \frac{k^2}{\pi^2}
\end{aligned}$$

with equality when $c_n = 0$ $(n = 2, 3, 4, \dots)$ leaving $c_1^2 = \frac{2k^2}{\pi^2}$. Then

$$y = c_1 \sin \pi x,$$

with $c_1 = \pm \frac{k\sqrt{2}}{\pi}$.

Note. We had the Wertinger Inequality (Theorem 15.1)

$$\int_a^b (y')^2 \, \mathrm{d}x \geqslant \frac{\pi^2}{(b-a)^2} \int_a^b y^2 \, \mathrm{d}x$$

where $y(a) = 0$ and $y(b) = 0$. Relating this to our example above we have $a = 0$ and $b = 1$. Thus,

$$\begin{gathered}
\int_0^1 (y')^2 \, \mathrm{d}x \geqslant \pi^2 \int_0^1 y^2 \, \mathrm{d}x \\
k^2 \geqslant \pi^2 J(y)
\end{gathered}$$

or $J(y) \leqslant \frac{k^2}{\pi^2}$ as found above.

17. Direct Methods

Let $J(Y)$ have a minimum for $Y = y$ thus $J(y) \leqslant J(Y)$ for all $Y \in \Omega$. Our approach to finding $J(y)$ so far has been via the Euler-Lagrange equation

$$J'(y) = 0.$$

Now, suppose we are unable to solve this equation for y. The idea is to approach the problem of finding $J(y)$ directly. In this case we probably have to make do with an approximate estimate of $J(y)$. To illustrate, consider the case

$$J(Y) = \int_0^1 \underbrace{\left\{ \frac{1}{2}(Y')^2 + \frac{1}{2}Y^2 - qY \right\}}_{=F} \, \mathrm{d}x$$

with $Y(0) = 0$ and $Y(1) = 0$. The Euler-Lagrange equation for y is

$$\begin{aligned} \frac{\partial F}{\partial y} - \frac{\mathrm{d}}{\mathrm{d}x}\frac{\partial F}{\partial y'} = 0 &\Rightarrow y - q - y'' = 0 \\ &\Rightarrow \boxed{-y'' + y = q} \quad 0 < x < 1 \end{aligned}$$

with $y(0) = 0 = y(1)$. For $Y = y + \varepsilon\xi$,

$$\begin{aligned} J(Y) &= J(y) + \frac{\varepsilon^2}{2}\int_0^1 \left((\xi')^2 + \xi^2\right) \mathrm{d}x \\ &\geqslant J(y) \end{aligned}$$

a MIN principle. Suppose we cannot solve the Euler-Lagrange equation. Then we try to minimize $J(Y)$ by any route. To take a very simple approach we choose a *trial* function

$$Y_1 = \alpha x(1 - x)$$

such that $Y_1(0) = 0 = Y_1(1)$. We then calculate

$$\min_{\alpha} J(Y_1) \qquad \text{where } Y_1 = \alpha\Phi(x), \qquad \Phi = x(1 - x).$$

Now

$$\begin{aligned} J(Y_1) &= J(\alpha\Phi) \\ &= \frac{1}{2}\alpha^2 \int_0^1 (\Phi'^2 + \Phi^2) \, \mathrm{d}x - \alpha \int_0^1 q\Phi \, \mathrm{d}x \\ &= \frac{1}{2}\alpha^2 A - \alpha B \end{aligned}$$

with, say

$$A = \int_0^1 (\Phi'^2 + \Phi^2)\ \mathrm{d}x \qquad B = \int_0^1 q\Phi\ \mathrm{d}x.$$

We have that

$$\frac{\mathrm{d}J}{\mathrm{d}\alpha} = \alpha A - B = 0 \Rightarrow \alpha_{\mathrm{opt}} = \frac{B}{A}$$

Then

$$J(\alpha_{\mathrm{opt}}\Phi) = \frac{1}{2}\frac{B^2}{A^2}A - \frac{B}{A}B = -\frac{1}{2}\frac{B^2}{A}.$$

Now for $q = 1$ we have $A = \frac{11}{30}$, $B = \frac{1}{6}$, $\alpha_{\mathrm{opt}} = \frac{B}{A} = \frac{5}{11}$ and thus $J(\alpha_{\mathrm{opt}}\Phi) = -\frac{5}{132}$. We note for later that this has decimal expansion

$$J(\alpha_{\mathrm{opt}}\Phi) = -\frac{5}{132} = -0.03\dot{7}\dot{8}.$$

We can continue by introducing more than one variable. For example take

$$Y_2 = \alpha_1\Phi_1 + \alpha_2\Phi_2$$

with, say $\Phi_1 = x(1-x)$ and $\Phi_2 = x^2(1-x)$. We find

$$\min_{\alpha_1,\alpha_2} J(Y_2) \leqslant \min_{\alpha_1} J(Y_1).$$

We now try and solve the Euler-Lagrange equation for $q = 1$, which is

$$-y'' + y = 1\ (0 < x < 1) \qquad y(0) = 0 = y(1).$$

We have the Particular Integral and Complementary Function to be

$$y_1 = 1 \qquad y_2 = ae^x + be^{-x}.$$

This gives our general solution to the differential equation to be

$$y = y_1 + y_2 = 1 + ae^x + be^{-x}.$$

Now we use the boundary conditions to solve for a and b. These boundary conditions give us

$$\begin{aligned} y(0) = 0 &\Rightarrow 1 + a + b = 0 \Rightarrow b = -1 - a \\ y(1) = 0 &\Rightarrow 1 + ae + be^{-1} = 0 \\ &\Rightarrow 1 + ae - (1+a)e^{-1} = 0 \\ &\Rightarrow a = -\frac{1}{1+e}. \end{aligned}$$

Then $b = -1 - a = -\frac{e}{1+e}$. Hence our general solution becomes

$$\begin{aligned} y &= 1 - \frac{e^x}{1+e} - \frac{e}{1+e}e^{-x} \\ &= 1 - \frac{1}{1+e}\{e^x + e^{1-x}\}. \end{aligned}$$

Now calculate $J(y)$.

$$\begin{aligned} J(y) &= \int_0^1 \left\{ \frac{1}{2}(y')^2 + \frac{1}{2}y^2 - y \right\} \, \mathrm{d}x \quad \text{since } q = 1 \\ &= \int_0^1 \left\{ -\frac{1}{2}yy'' + \frac{1}{2}y^2 - y \right\} \, \mathrm{d}x + \frac{1}{2}\underbrace{[yy']_0^1}_{=0} \\ &= \int_0^1 \left\{ \frac{1}{2}y(-y'' + y) - y \right\} \, \mathrm{d}x \end{aligned}$$

but $-y'' + y = 1$ by the Euler Lagrange equation and hence

$$= -\frac{1}{2}\int_0^1 y \, \mathrm{d}x.$$

So putting in our solution for y we obtain

$$\begin{aligned} J(y) &= -\frac{1}{2}\int_0^1 \left\{ 1 - \frac{1}{1+e}(e^x + e^{1-x}) \right\} \, \mathrm{d}x \\ &= -\frac{1}{2} + \frac{1}{2}\frac{1}{1+e}\left\{ e - 1 - (1-e) \right\} \, \mathrm{d}x \\ &= -\frac{1}{2} + \frac{e-1}{e+1} \\ &= -0.03788246\ldots \end{aligned}$$

which is very close to our approximation. In fact the difference between the two values is 0.00000368.

18. Complementary Variational Principles (CVP's) / Dual Extremum Principles (DEP's)

Now seek upper *and* lower bounds for $J(y)$ so

$$\text{lower bound} \leqslant J(y) \leqslant \text{upper bound}.$$

To do this we turn to the *canonical formalism*. Take the particular case

$$J(Y) = \int_a^b \left\{ \frac{1}{2}(Y')^2 + \frac{1}{2}wY^2 - qY \right\} \, \mathrm{d}x \qquad Y(a) = 0 = Y(b). \tag{18.1}$$

Hence we have

$$F = \frac{1}{2}(Y')^2 + \frac{1}{2}wY^2 - qY \tag{18.2}$$

$$P = \frac{\partial F}{\partial Y'} = Y' \tag{18.3}$$

$$\begin{aligned} H(x, P, Y) &= PY' - F \\ &= \frac{1}{2}P^2 - \frac{1}{2}wY^2 + qY \end{aligned} \tag{18.4}$$

Now replace $J(Y)$ by

$$I(P, Y) = \int_b^a \left\{ p\frac{\mathrm{d}Y}{\mathrm{d}x} - H(x, P, Y) \right\} \,\mathrm{d}x - [PY]_a^b \tag{18.5.i}$$

$$= \int_a^b \{-P'Y - H\} \,\mathrm{d}x. \tag{18.5.ii}$$

Then we know from section 9 that $I(P, Y)$ is stationary at (p, y) where

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{\partial H}{\partial p} \quad (a < x < b) \qquad \text{and } y = 0 \text{ on } \partial[a, b] \tag{18.6}$$

$$-\frac{\mathrm{d}p}{\mathrm{d}x} = \frac{\partial H}{\partial y} \quad (a < x < b). \tag{18.7}$$

Now we define dual integrals $J(y_1)$ and $G(p_2)$. Let

$$\Omega_1 = \left\{ (p, y) \,\middle|\, \frac{\mathrm{d}y}{\mathrm{d}x} = \frac{\partial H}{\partial p} \ (a < x < b) \text{ and } y = 0 \text{ on } \partial[a, b] \right\} \tag{18.8}$$

$$\Omega_2 = \left\{ (p, y) \,\middle|\, -\frac{\mathrm{d}p}{\mathrm{d}x} = \frac{\partial H}{\partial y} \ (a < x < b) \right\}. \tag{18.9}$$

Then we define

$$J(y_1) = I(p_1, y_1) \text{ with } (p_1, y_1) \in \Omega_1 \tag{18.10}$$

$$G(p_2) = I(p_2, y_2) \text{ with } (p_2, y_2) \in \Omega_2 \tag{18.11}$$

An extremal pair $(p, y) \in \Omega_1 \cap \Omega_2$ and

$$G(p) = I(p, y) = J(y). \tag{18.12}$$

Case

$$H(x, P, Y') = \frac{1}{2}P^2 - \frac{1}{2}wY^2 + qY \tag{18.13}$$

where $w(x)$ and $q(x)$ are prescribed. Then

$$\Omega_1 = \left\{ (p,y) \,\middle|\, \frac{\mathrm{d}y}{\mathrm{d}x} = \frac{\partial H}{\partial p} = p \ (a < x < b) \text{ and } y = 0 \text{ on } \partial[a,b] \right\} \tag{18.14}$$

$$\Omega_2 = \left\{ (p,y) \,\middle|\, -\frac{\mathrm{d}p}{\mathrm{d}x} = \frac{\partial H}{\partial y} = -wy + q, \text{ i.e. } y = \frac{p' + q}{w} \ (a < x < b) \right\}. \tag{18.15}$$

So $J(y_1) = I(p_1, y_1)$ with $p_1 = y_1'$ if $y_1 = 0$ on $\partial[a,b]$. Now

$$\begin{aligned} J(y_1) &= \int_a^b \left\{ y_1'^2 - \left(\frac{1}{2}(y')^2 - \frac{1}{2} w y_1^2 + q y_1 \right) \right\} \,\mathrm{d}x \\ &= \int_a^b \left\{ \frac{1}{2} y_1'^2 + \frac{1}{2} w y_1^2 - q y_1 \right\} \,\mathrm{d}x \end{aligned} \tag{18.16}$$

the original Euler-Lagrange integral J. Next $G(p_2) = I(p_2, y_2)$ with $y_2 = \frac{1}{w}(p_2' + q)$

$$\begin{aligned} G(p_2) &= I(p_2, y_2) \\ &= \int_a^b \left\{ -p_2' y_2 - \left(\frac{1}{2} p_2^2 - \frac{1}{2} w y_2^2 + q y_2 \right) \right\} \,\mathrm{d}x \end{aligned}$$

this uses (18.5.ii)

$$\begin{aligned} &= \int_a^b \left\{ -\frac{1}{2} p_2^2 + \frac{1}{2} w y_2^2 - y_2 \underbrace{(p_2' + q)}_{= w y_2} \right\} \,\mathrm{d}x \\ &= \int_a^b \left\{ -\frac{1}{2} p_2^2 - \frac{1}{2} w y_2^2 \right\} \,\mathrm{d}x \end{aligned}$$

with $y_2 = \frac{1}{w}(p_2' + q)$

$$= -\frac{1}{2} \int_a^b \left\{ p_2^2 + \frac{1}{w}(p_2' + q)^2 \right\} \,\mathrm{d}x \tag{18.17}$$

If we write

$$y_1 = y + \varepsilon \xi \qquad\qquad p_2 = p + \varepsilon \eta \tag{18.18}$$

in (18.16) and (18.17), expand about y and p and use the stationary property

$$\delta J(y, \varepsilon \xi) = 0 \qquad \delta G(p, \varepsilon \eta) = 0.$$

We have

$$\Delta J = J(y_1) - J(y) = \frac{1}{2}\int_a^b \left\{(\varepsilon\xi')^2 + w(\varepsilon\xi)^2\right\}\,\mathrm{d}x \tag{18.19}$$

$$\Delta G = G(p_2) - G(p) = -\frac{1}{2}\int_a^b \left\{(\varepsilon\eta)^2 + \frac{1}{w}(\varepsilon\eta')^2\right\}\,\mathrm{d}x \tag{18.20}$$

Hence, if

$$w(x) > 0 \tag{18.21}$$

we have $\Delta J \geqslant 0$ and $\Delta G \leqslant 0$, i.e.

$$G(p_2) \leqslant G(p) = I(p, y) = J(y) \leqslant J(y_1), \tag{18.22}$$

the complementary (dual) principles.

Example 18.1. Calculation of G in the case $w = 1$, $q = 1$ on $(0, 1)$. For this

$$G(p_2) = -\frac{1}{2}\int_0^1 \left\{p_2^2 + (p_2' + 1)^2\right\}\,\mathrm{d}x.$$

Here p_2 is *any* admissible function. To make a good choice for p_2 we note that the exact p and y are linked through

$$p = \frac{\mathrm{d}y}{\mathrm{d}x} \qquad y(0) = 0 = y(1).$$

So, we follow this and choose

$$p_2 = \frac{\mathrm{d}}{\mathrm{d}x}(\beta\Psi) \qquad \Psi = x(1-x)$$

Thus we obtain

$$\begin{aligned} G(p_2) &= G(\beta) \\ &= -\frac{1}{2}\int_0^1 \left\{\beta^2\Psi'^2 + (\beta\Psi'' + 1)^2\right\}\,\mathrm{d}x \\ &= -\frac{1}{2}A\beta^2 - B\beta - C \end{aligned}$$

where we define

$$A = \int_0^1 \left\{ (\Psi')^2 + (\Psi'')^2 \right\} \, \mathrm{d}x = \frac{13}{3}$$
$$B = \int_0^1 \Psi'' \, \mathrm{d}x = -2$$
$$C = \frac{1}{2}.$$

Now compute $\frac{\mathrm{d}G}{\mathrm{d}\beta} = -A\beta - B = 0 \Rightarrow \beta_{\mathrm{opt}} = -\frac{B}{A} = \frac{6}{13}$. So,

$$\underset{\beta}{\mathrm{MAX}}\ G(\beta) = G(\beta_{\mathrm{opt}}) = \frac{B^2}{2A} - C = -\frac{1}{26} = -0.038461\ldots$$

From the previous calculation of $J(y_1)$ we had in section 17 that

$$\underset{\alpha}{\mathrm{MIN}} = -0.03787878\ldots$$

Hence we have bounded $I(p, y)$ such that

$$-0.038461 \leqslant I(p, y) \leqslant -0.03787878$$

In this case we know that $I(p, y) = -0.03788246$

19. Eigenvalues

$$Ly = \lambda y. \tag{19.1}$$

For example

$$-\frac{\mathrm{d}^2 y}{\mathrm{d}x^2} = \lambda y \qquad 0 < x < 1 \tag{19.2}$$
$$y(0) = 0 \qquad y(1) = 0. \tag{19.3}$$

We have that (19.2) has general solution

$$y = A \sin \sqrt{\lambda} x + B \cos \sqrt{\lambda} x \tag{19.4}$$
$$y(0) = 0 \Rightarrow B = 0 \tag{19.5}$$
$$y(1) = 0 \Rightarrow A \sin \sqrt{\lambda} = 0$$

Take $\sin \sqrt{\lambda} = 0$ then we have

$$\sqrt{\lambda} = n\pi, n = 1, 2, 3, \ldots \qquad \text{or} \qquad \lambda = n^2\pi^2, n = 1, 2, 3 \ldots \tag{19.6}$$

Then we have that

$$\begin{cases} y_n = \sin n\pi x & n = 1, 2, 3, \dots \\ \lambda_n = n^2\pi^2 \end{cases}$$

In general, Eignevalue problems (19.1) cannot be solved exactly, so we need methods for estimating Eigenvalues. One method is due to Rayleigh: it provides an upper bound for λ_1, the lowest Eigenvalue of the problem. Consider

$$A(Y) = \int_0^1 Y\left(-\frac{\mathrm{d}^2 y}{\mathrm{d}x^2} - \lambda_1\right) Y \,\mathrm{d}x \tag{19.7}$$

for all $Y \in \Omega = \{C_2 \text{ functions such that } Y(0) = 0 = Y(1)\}$. We write

$$\begin{aligned} Y &= \sum_{n=1}^{\infty} a_n y_n \\ \{y_n\} &= \text{complete set of eigenfunctions of } -\frac{\mathrm{d}^2 y}{\mathrm{d}x^2} \\ &= \{\sin n\pi x\}. \end{aligned}$$

Then (19.7) becomes

$$\begin{aligned} A(Y) &= \int_0^1 \sum_{n=1}^{\infty} a_n y_n \left(-\frac{\mathrm{d}^2 y}{\mathrm{d}x^2} - \lambda_1\right) \sum_{m=1}^{\infty} a_m y_m \mathrm{d}x \\ &= \sum_{n=1}^{\infty}\sum_{m=1}^{\infty} a_n a_m \left(\lambda_m - \lambda_1\right) \underbrace{\int_0^1 y_n y_m}_{=0 (n \neq m)} \\ &= \sum_{n=1}^{\infty} a_n^2 (\lambda_n - \lambda_1) \int_0^1 y_n^2 \,\mathrm{d}x \\ &\geqslant 0 \end{aligned}$$

where $\lambda_n = n^2\pi^2$ and hence $0 < \lambda_1 < \lambda_2 < \lambda_3 < \dots$. So we have $A(Y) \geqslant 0$. Rearranging (19.7) we obtain

$$\int_0^1 Y\left(-\frac{\mathrm{d}^2 y}{\mathrm{d}x^2}\right) Y \,\mathrm{d}x \geqslant \lambda_1 \int_0^1 Y^2 \,\mathrm{d}x$$

$$\lambda_1 \leqslant \frac{\int_0^1 Y\left(-\frac{\mathrm{d}^2 y}{\mathrm{d}x^2}\right) \,\mathrm{d}x}{\int_0^1 Y^2 \,\mathrm{d}x} = \Lambda(Y), \tag{19.8}$$

which is the Rayleigh Bound. To illustrate: take $Y = x(x-1)$. Then

$$\Lambda(Y) = \frac{\frac{1}{3}}{\frac{1}{30}} = 10.$$

The exact $\Lambda_1 = \pi^2 \approx 9.87$.

EXAMPLE 19.1. From example 2 of section 16 we have

$$(u', u') \geqslant \lambda_0 (u, u) \tag{$\dagger$}$$

where $u(0) = 0$, $u'(1) = 0$ and

$$(u, u) = \int_0^1 u^2 \, \mathrm{d}x/$$

We know that $\lambda_0 = \frac{\pi^2}{4}$. We can use (†) to obtain an upper bound for λ_0:

$$\lambda_0 \leqslant \frac{(u', u')}{(u, u)}.$$

Take $u = 2x - x^2$ such that $u(0) = 0$, $u'(1) = 0$. So

$$\lambda_0 \leqslant \frac{\frac{4}{3}}{\frac{7}{15}} = \frac{20}{7} \approx 2.85. \tag{$\dagger\dagger$}$$

Note $\lambda_0 = \frac{\pi^2}{4} \approx \frac{9.87}{4} \approx 2.47$.

20. Appendix: Structure

Integration by parts has been used in the work on Euler-Lagrange and Canonical Equations.

$$\frac{\partial F}{\partial y} - \frac{\mathrm{d}}{\mathrm{d}x}\frac{\partial F}{\partial y'} = 0$$

$$\begin{cases} \dfrac{\mathrm{d}y}{\mathrm{d}x} = \dfrac{\partial H}{\partial p} \\ -\dfrac{\mathrm{d}p}{\mathrm{d}x} = \dfrac{\partial H}{\partial y}. \end{cases}$$

Simplest pairs of derivatives $\frac{\mathrm{d}}{\mathrm{d}x}$ and $-\frac{\mathrm{d}}{\mathrm{d}x}$ appear. We now attempt to generalise. We take an inner produce space S of functions $u(x)$ and suppose $\langle u, v\rangle$ denotes inner product. A very useful case is

$$\langle v, u\rangle = \int uvw(x) \, \mathrm{d}x$$

with $w(x) \geqslant 0$, where w is a weighting function. Properties

$$\begin{aligned}\langle u, v\rangle &= \langle v, u\rangle \\ \langle u, u\rangle &\geqslant 0.\end{aligned}$$

Write T as a linear differetial operator, e.g. $T = \frac{\mathrm{d}}{\mathrm{d}x}$ and define its *adjoint* T^* by

$$\langle u, Tv\rangle = \langle T^* u, v\rangle,$$

where we suppose boundary terms vanish.

EXAMPLE 20.1. $T = \frac{\mathrm{d}}{\mathrm{d}x}$ and $\langle u, v\rangle = \int uvw(x)\ \mathrm{d}x$. Then

$$\begin{aligned}\langle u, Tv\rangle &= \int u\frac{\mathrm{d}v}{\mathrm{d}x}w(x)\ \mathrm{d}x \\ &= \int (-1)\frac{\mathrm{d}}{\mathrm{d}x}(w(x)u)v\ \mathrm{d}x + \underbrace{[\quad]}_{=0} \\ &= \int (-1)\frac{1}{w}\frac{\mathrm{d}}{\mathrm{d}x}(wu)vw(x)\ \mathrm{d}x \\ = \langle T^* u, v\rangle.\end{aligned}$$

So,

$$T^* u = -\frac{1}{w}\frac{\mathrm{d}}{\mathrm{d}x}\big(w(x)u\big) \qquad T^* = -\frac{1}{w}\frac{\mathrm{d}}{\mathrm{d}x}(w)$$

From T and T^* we can construct $L = T^*T$ and hence

$$L = T^*T = -\frac{1}{w}\frac{\mathrm{d}}{\mathrm{d}x}\left(w\frac{\mathrm{d}}{\mathrm{d}x}\right).$$

Now look at

$$\begin{aligned}\langle u, Lu\rangle &= \int u(T^*Tu)w(x)\ \mathrm{d}x \\ &= \langle u, T^*Tu\rangle \\ &= \langle Tu, u\rangle \\ &= \int (Tu)(Tu)w(x)\ \mathrm{d}x \\ &\geqslant 0.\end{aligned}$$

So, L is a positive operator (e.g. $w = 1$, $T = \frac{\mathrm{d}}{\mathrm{d}x}$, $T^* = -\frac{\mathrm{d}}{\mathrm{d}x} \Rightarrow L = -\frac{\mathrm{d}^2}{\mathrm{d}x^2}$). Also

$$\begin{aligned}\langle u, Lv\rangle &= \langle u, T^*Tv\rangle \\ &= \langle Tu, TV\rangle \\ &= \langle T^*Tu, v\rangle \\ &= \langle Lu, v\rangle.\end{aligned}$$

But $\langle u, Lv\rangle = \langle L^*u, v\rangle$ by definition of adjoint. So, $L^* = L$. So, $L = T^*T$ is a *self-adjoint* operator.

The Canonical equations generalise to

$$Ty = \frac{\partial H}{\partial p} \qquad T^*p = \frac{\partial H}{\partial y}$$

with Euler-Lagrange equation

$$\frac{\partial F}{\partial y} + T^*\left(\frac{\partial F}{\partial (Ty)}\right) = 0.$$

Eigenvalue problems.

$$Lu = \lambda u$$

(i.e. $T^*Tu = \lambda u$). λ is an eigenvalue. Suppose there is a complete set of eigenfunctions $\{u_n\}$, corresponding to eigenvalues $\{\lambda_n\}$. Now

$$\langle u, Lu\rangle = \langle u, T * Tu\rangle = \langle Tu, Tu\rangle \geqslant 0$$

Let $u = u_n$ then

$$\langle u_n, Lu_n\rangle = \langle u_n, \lambda_n u_n\rangle = \lambda_n\langle u_n, u_n\rangle \geqslant 0$$

which implies that $\lambda_n > 0$. Suppose the λ_n are discrete, so if we order them

$$0 < \lambda_1 < \lambda_2 < \lambda_3 < \dots$$

Orthogonality of $u_n's$. Let

$$Lu_n = \lambda_n u_n \qquad \text{and} \qquad Lu_m = \lambda_m u_m$$

Then, taking innner products we have

$$\begin{aligned}\langle u_m, Lu_n\rangle &= \langle u_m, \lambda_n u_n\rangle = \lambda_n\langle u_m, u_n\rangle \\ \langle u_n, Lu_m\rangle &= \langle u_n, \lambda_m u_m\rangle = \lambda_m\langle u_n, u_m\rangle = \lambda_m\langle u_m, u_n\rangle\end{aligned}$$

But L self adjoint, so

$$\langle u_m, Lu_n\rangle = \langle Lu_m, u_n\rangle = \langle u_n, Lu_m\rangle$$

using $\langle u, v\rangle = \langle v, u\rangle$. So,

$$0 = (\lambda_n - \lambda_m)\langle u_m, u_n\rangle$$

which implies $\langle u_m, u_n\rangle = 0$ if $\lambda_n \neq \lambda_m$. Orthogonality. An example of this is

$$w = 1 \quad t = \frac{\mathrm{d}}{\mathrm{d}x} \quad T^* = -\frac{\mathrm{d}}{\mathrm{d}x} \quad L = -\frac{\mathrm{d}^2}{\mathrm{d}x^2} \quad u_n = \sin(n\pi x) \quad \lambda_n = n^2\pi^2.$$

Another example is

$$w = x \quad T = \frac{\mathrm{d}}{\mathrm{d}x} \quad T^* = -\frac{1}{x}\frac{\mathrm{d}}{\mathrm{d}x}(x) \quad L = -\frac{1}{x}\frac{\mathrm{d}}{\mathrm{d}x}\left(x\frac{\mathrm{d}}{\mathrm{d}x}\right) \quad u_n = J_0(a_n x) \quad \lambda_n = a_n^2.$$

Orthogonality allows us to expand functions in terms of eigenfunctions u_n. Thus

$$u(x) = \sum_{n=1}^{\infty} a_n u_n(x) \qquad Lu_n = \lambda_n u_n.$$

Then

$$\langle u_k, u\rangle = \sum_{n=1}^{\infty}\langle u_k, u_n\rangle = c_k\langle u_k, u_k\rangle.$$

Consider $A(u) = \langle Tu, Tu\rangle$ for example $\langle u', u'\rangle = \int u'^2 w(x)\ \mathrm{d}x$. We have that

$$A(u) = \langle u, T * Tu\rangle = \langle u, Lu\rangle$$

Expand u in eigenfunctions u_n of L. Then

$$\begin{aligned}
A(u) &= \left\langle \sum_{n=1}^{\infty} c_n u_n, L\sum_{m=1}^{\infty} c_m u_m \right\rangle \\
&= \sum_{n}^{\infty}\sum_{m}^{\infty} c_n c_m \langle u_n, Lu_m\rangle
\end{aligned}$$

using that $Lu_m = \lambda_m u_m$

$$
\begin{aligned}
&= \sum_n^\infty \sum_m^\infty c_n c_m \langle u_n, \lambda_m u_m \rangle \\
&= \sum_n^\infty \sum_m^\infty c_n c_m \lambda_m \langle u_n, u_m \rangle \\
&= \sum_{n=1}^\infty c_n^2 \lambda_n \langle u_n, u_n \rangle \\
&\geqslant \lambda_1 \sum_{n=1}^\infty c_n^2 \lambda_n \langle u_n, u_n \rangle \\
&\geqslant \lambda_1 \langle u, u \rangle
\end{aligned}
$$

i.e. $A(u) \geqslant \lambda_1 \langle u, u \rangle$ and so $\frac{\langle u, Lu \rangle}{\langle u, u \rangle} \geqslant \lambda_1$.