

Calculus of variations:

Calculus of variations is an area of mathematics that provides us with techniques to find curves / surfaces that minimise certain quantities.

For instance, calculus of variations can be used to show that the shortest path between two points is the a curve that minimises the distance

line segment that connects them. It can also be used to find the shortest path on a surface (such as a sphere) between two points of the surface.



The quantity we are aiming to minimise has to be an integral. For instance, when we are looking for the curve $(x, y(x))$ that minimises the distance between the points (x_1, y_1) , (x_2, y_2) , we want to minimise

$$\int_{(x_1, y_1)}^{(x_2, y_2)} ds = \int_{x_1}^{x_2} \sqrt{1 + (y'(x))^2} dx$$

\downarrow the arc-length measure on a given curve \downarrow known \downarrow unknown, as y is the unknown function.

(2)

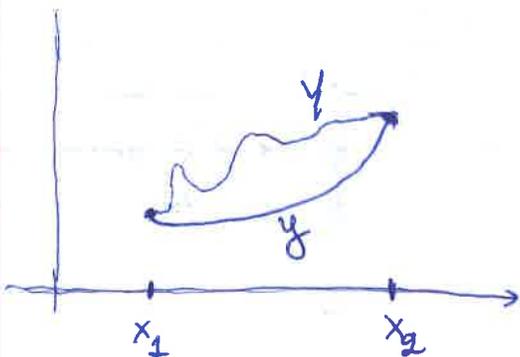
We first illustrate the technique in this example, and we will then find a technique that can deal with more general problems. So:

→ Show that the shortest path in \mathbb{R}^2 between two points $(x_1, y_1), (x_2, y_2) \in \mathbb{R}^2$ is the line segment that connects them:

We want to find the curve $\gamma(x, y(x)) : x \in [x_1, x_2]$ connecting (x_1, y_1) with (x_2, y_2) (i.e., with $y(x_1) = y_1, y(x_2) = y_2$) with the property that $\int_{(x_1, y_1)}^{(x_2, y_2)} ds$ is minimal, amongst all curves that connect $(x_1, y_1), (x_2, y_2)$.
the arc-length measure on the curve

Note that, on any curve γ , $ds = \sqrt{1 + (y'(x))^2} dx$.

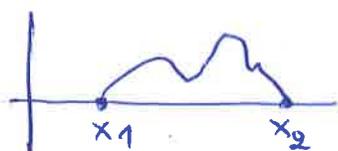
So, we are looking for $y(x), \forall x \in [x_1, x_2]$, such that $\int_{x_1}^{x_2} \sqrt{1 + (y'(x))^2} dx$ is minimal.



Suppose we have found this function y . Then, if we perturb it a little, to get a new function \tilde{y} , then \tilde{y} is such that $\int_{x_1}^{x_2} \sqrt{1 + (\tilde{y}'(x))^2} dx \geq \int_{x_1}^{x_2} \sqrt{1 + (y'(x))^2} dx$.

We now consider a particular kind of perturbation of our minimiser y :

fix an arbitrary function $\eta : [x_1, x_2] \rightarrow \mathbb{R}$,



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

with the extra property that η is twice continuously differentiable (so that we can apply integration by parts later care-free)

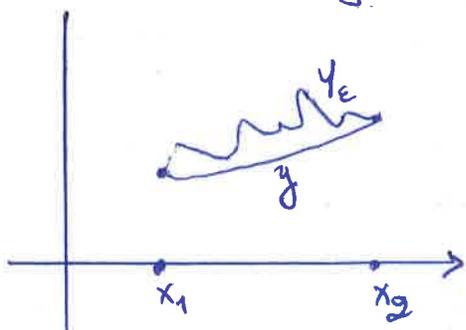
We assume the same for y .

Now, consider the perturbation Y_ϵ of y ,

with $Y_\epsilon : [x_1, x_2] \rightarrow \mathbb{R}$

$$Y_\epsilon(x) = y(x) + \epsilon \eta(x)$$

Notice that the graph of Y_ϵ is also a curve that connects $(x_1, y_1), (x_2, y_2)$, no matter what ϵ is :



And $y = Y_\epsilon$ for $\epsilon = 0$.
(i.e., $y = Y_0$).

So,

$$\underbrace{\int_{x_1}^{x_2} \sqrt{1 + (Y'_\epsilon(x))^2} dx}_{I'_y(\epsilon)} \geq \underbrace{\int_{x_1}^{x_2} \sqrt{1 + (y'(x))^2} dx}_{I'_y(0)}, \forall \epsilon \in \mathbb{R}.$$

We thus have that, for our fixed y and n ,

$$I_y^n(\epsilon) \geq I_y^n(0) \quad , \quad \forall \epsilon \in \mathbb{R}.$$

This means that the function

$$I_y^n: \mathbb{R} \rightarrow \mathbb{R}$$

← the ϵ 's

has a global minimum at $\epsilon = 0$;

therefore $\frac{dI_y^n(\epsilon)}{d\epsilon} \Big|_{\epsilon=0} = 0$ $\textcircled{*}$, no matter what n is.

And: For any $\epsilon \in \mathbb{R}$,

$$\frac{dI_y^n(\epsilon)}{d\epsilon} = \frac{d}{d\epsilon} \left(\int_{x_1}^{x_2} \sqrt{1 + (Y'_\epsilon(x))^2} \, dx \right) =$$

$$= \int_{x_1}^{x_2} \frac{d}{d\epsilon} \left(\sqrt{1 + (Y'_\epsilon(x))^2} \right) dx \quad \underline{\text{chain rule}}$$

$$= \int_{x_1}^{x_2} \frac{1}{2\sqrt{1 + (Y'_\epsilon(x))^2}} \cdot \frac{d}{d\epsilon} \left(1 + (Y'_\epsilon(x))^2 \right) dx$$

Now: • $Y'_\epsilon(x) = (y(x) + \epsilon n(x))' = y'(x) + \epsilon n'(x)$, $\forall x \in [x_1, x_2]$

• $\frac{d}{d\epsilon} \left[1 + (Y'_\epsilon(x))^2 \right] = \frac{d}{d\epsilon} \left((Y'_\epsilon(x))^2 \right) =$

(5)

$$= 2 y'_\epsilon(x) \cdot \frac{d}{d\epsilon} (y'_\epsilon(x)) = 2 y'_\epsilon(x) \cdot \frac{d}{d\epsilon} (y'(x) + \epsilon n'(x)) =$$

$$= 2 y'_\epsilon(x) \cdot n'(x).$$

$$\text{So, } y'_\epsilon(x) \Big|_{\epsilon=0} = y'(x), \quad \forall x \in [x_1, x_2],$$

$$\text{and } \frac{d}{d\epsilon} [1 + (y'_\epsilon(x))^2] \Big|_{\epsilon=0} = 2 y'_\epsilon(x) \Big|_{\epsilon=0} \cdot n'(x) = 2 y'(x) n'(x).$$

$$\text{Thus, } \frac{dI_f^n}{d\epsilon} \Big|_{\epsilon=0} = \int_{x_1}^{x_2} \frac{1}{2\sqrt{1+(y'_\epsilon(x))^2}} \cdot \frac{d}{d\epsilon} (1+(y'_\epsilon(x))^2) \Big|_{\epsilon=0} dx =$$

$$= \int_{x_1}^{x_2} \frac{1}{2\sqrt{1+(y'(x))^2}} \cdot 2 y'(x) n'(x) dx =$$

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

$$\text{So, by } \textcircled{*}_1, \int_{x_1}^{x_2} \frac{y'(x)}{\sqrt{1+(y'(x))^2}} n'(x) dx = 0 \text{ for this } n.$$

However, remember that n was an arbitrary twice continuously differentiable function with $n(x_1) = n(x_2) = 0$.

Thus, $\textcircled{*}_2$ $\int_{x_1}^{x_2} \frac{y'(x)}{\sqrt{1+(y'(x))^2}} n'(x) dx = 0, \quad \forall n: [x_1, x_2] \rightarrow \mathbb{R}$ twice cont. diff., with $n(x_1) = n(x_2) = 0$.

(6)

Thus, for any such n , integration by parts gives:

$$0 = \int_{x_1}^{x_2} \frac{y'(x)}{\sqrt{1+(y'(x))^2}} n'(x) dx =$$

$$= \underbrace{\left[\frac{y'(x)}{\sqrt{1+(y'(x))^2}} \cdot n(x) \right]_{x_1}^{x_2}}_{=0} - \int_{x_1}^{x_2} \frac{d}{dx} \left[\frac{y'(x)}{\sqrt{1+(y'(x))^2}} \right] \cdot n(x) dx$$

as $n(x_1) = n(x_2) = 0$

$$\Rightarrow \int_{x_1}^{x_2} \frac{d}{dx} \left[\frac{y'(x)}{\sqrt{1+(y'(x))^2}} \right] \cdot n(x) dx = 0,$$

for all $n: [x_1, x_2] \rightarrow \mathbb{R}$ twice continuously differentiable,

with $n(x_1) = n(x_2) = 0$.

This means that $\frac{d}{dx} \left[\frac{y'(x)}{\sqrt{1+(y'(x))^2}} \right] = 0, \forall x \in [x_1, x_2]!$

Indeed, the following theorem holds:

→ Thm:

Let $f: [a, b] \rightarrow \mathbb{R}$ be continuous.

If $\int_a^b f(x)g(x)dx = 0$ for all $g: [a, b] \rightarrow \mathbb{R}$ smooth, with $g(a)=g(b)=0$,

then $f \equiv 0$ on $[a, b]$.



Note that the above says that if $\langle f, g \rangle = 0$, i.e. if f is perpendicular to g for all smooth g with $g(a)=g(b)=0$, then f is the 0 vector (function).
(or, twice cont. diff...)

We can think about this this way:

If $\langle f, g \rangle = 0$ for all such g , then f is perpendicular to all functions, so f is the 0 vector function.

Eventually, we have:

↙ We assumed y is twice cont. diff., so this derivative is continuous.

$$\frac{d}{dx} \left[\frac{y'(x)}{\sqrt{1+(y'(x))^2}} \right] = 0, \quad \forall x \in [x_1, x_2]$$

$$\Rightarrow \frac{y'}{\sqrt{1+y'^2}} = \text{constant} \implies y' = \text{constant}$$

$$\forall x \in [x_1, x_2], \frac{y'(x)}{\sqrt{1+(y'(x))^2}} = c \implies c^2(1+(y'(x))^2) = (y'(x))^2 \quad \forall x \in [x_1, x_2]$$

$\implies \dots \implies y' = \text{constant}$
(Darboux property for derivatives)

8

So, the graph of y is a line segment - in fact, the unique line segment that connects (x_1, y_1) with (x_2, y_2) .



We will generalise this way of thinking to minimise more general integrals.

So, the graph of y is a line segment - in fact, the unique line segment that connects (x_1, y_1) with (x_2, y_2) .

Lecture 34:

We will generalise this way of thinking to minimise more general integrals.

First, we remember some basic facts:

① Let $f: \mathbb{R} \rightarrow \mathbb{R}$
 $x \rightarrow F(x)$. Let now x be a variable of t ,
i.e., $x = x(t)$, varying as t varies.

Then, $\left[\frac{\partial F}{\partial t} = \frac{\partial F}{\partial x} \cdot \frac{\partial x}{\partial t} \right]$ (*)

This is just the chain rule:

$$(F \circ x)'(t) = F'(x(t)) \cdot x'(t).$$

So, (*) really means that

$$\frac{\partial F}{\partial t} \Big|_{t=t_0} = \left(\frac{\partial F}{\partial x} \Big|_{x=x(t_0)} \right) \cdot \frac{\partial x}{\partial t} \Big|_{t=t_0}, \forall t_0.$$

the derivative of F w.r.t. its input variable.

In higher dimensions:

(2)

(2) Let $f: \mathbb{R}^n \rightarrow \mathbb{R}$
 $(x_1, \dots, x_n) \rightarrow f(x_1, \dots, x_n)$.

Suppose now that all the coordinates x_i are functions of t , $x_i(t)$.

Then:
$$\frac{\partial f}{\partial t} = \frac{\partial f}{\partial x_1} \cdot \frac{\partial x_1}{\partial t} + \frac{\partial f}{\partial x_2} \cdot \frac{\partial x_2}{\partial t} + \dots + \frac{\partial f}{\partial x_n} \cdot \frac{\partial x_n}{\partial t}$$

Just like in (1), this really means that

$$\frac{\partial f}{\partial t} \Big|_{t=t_0} = \left(\frac{\partial f}{\partial x_1} \Big|_{x_1=x_1(t_0)} \cdot \frac{\partial x_1}{\partial t} \Big|_{t=t_0} + \dots + \frac{\partial f}{\partial x_n} \Big|_{x_n=x_n(t_0)} \cdot \frac{\partial x_n}{\partial t} \Big|_{t=t_0} \right)$$

$\downarrow t_0$

the derivatives of f w.r.t. the coordinates of f , forgetting that these coordinates depend on t , and ^{these} thus all depend on each other.

→ The general situation:

Remember that in our previous example, we were looking

for the function $y: [x_1, x_2] \rightarrow \mathbb{R}$

(with $y(x_1) = y_1$, $y(x_2) = y_2$ given),

that minimised

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

We found the appropriate y by

observing that, for any perturbation $V_\epsilon = y + \epsilon \eta$
of the minimiser y , $I_y^\eta(\epsilon) \geq I$; $\int_{x_1}^{x_2} \sqrt{1 + (V_\epsilon'(x))^2} dx$
 $= I_y^\eta(0)$

thus, $\frac{dI_y^\eta(\epsilon)}{d\epsilon} \Big|_{\epsilon=0} = 0.$

In the general case, we will do something similar.

The question is: How can we find $y: [x_1, x_2] \rightarrow \mathbb{R}$,

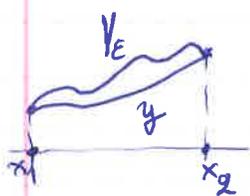
that minimises $I = \int_{x_1}^{x_2} \underbrace{F(x, y(x), y'(x))}_{\text{Some quantity}} dx$?

(4)

Suppose that $y: [x_1, x_2] \rightarrow \mathbb{R}$ is a minimiser for I . We perturb y as follows:

Let $n: [x_1, x_2] \rightarrow \mathbb{R}$ be a twice continuously differentiable function, with $n(x_1) = n(x_2) = 0$.

For each $\epsilon > 0$, let



$$y_\epsilon(x) := y(x) + \epsilon \cdot n(x), \quad \forall x \in [x_1, x_2]:$$

a perturbation of y . $y_\epsilon = y$ for $\epsilon = 0$.

$$\text{Let } I_y^n(\epsilon) := \int_{x_1}^{x_2} F(x, y_\epsilon(x), y_\epsilon'(x)) dx, \quad \forall \epsilon \in \mathbb{R}.$$

Since $y = y_0$ is a minimiser of the above, we have that $I_y^n(\epsilon)$ has a minimum at $\epsilon = 0$.

So,

$$\left. \frac{dI_y^n(\epsilon)}{d\epsilon} \right|_{\epsilon=0} = 0.$$

Everything we do from now on will give the y 's that satisfy this $\frac{d}{d\epsilon} I$ as above. **Not** just the minimisers y .

Now, we calculate $\frac{dI_y^n(\epsilon)}{d\epsilon}$:

• $y_\epsilon'(x) = y'(x) + \epsilon n'(x), \quad \forall \epsilon \in \mathbb{R}, \forall x \in [x_1, x_2].$

• So, $\left. \frac{dI_y^n(\epsilon)}{d\epsilon} \right|_{\epsilon=0} = \int_{x_1}^{x_2} \left. \frac{d}{d\epsilon} \left(F(x, y_\epsilon(x), y_\epsilon'(x)) \right) \right|_{\epsilon=0} dx.$

$$\frac{d}{d\epsilon} F(x, y_\epsilon(x), \underbrace{y'_\epsilon(x)}_{y'(x) + \epsilon n'(x)}) \Big|_{\epsilon=0} \quad \text{chain rule}$$

$$= \frac{\partial F}{\partial x} \cdot \frac{dx}{d\epsilon} \Big|_{\epsilon=0} \rightarrow 0, \text{ as } x \text{ indep. of } \epsilon$$

$$+ \frac{\partial F}{\partial y} \Big|_{y=y_0=y} \cdot \frac{dy_\epsilon}{d\epsilon} \Big|_{\epsilon=0}$$

just the derivative of F w.r.t. its second variable, evaluated at y: $= \frac{d}{d\epsilon} (y(x) + \epsilon n(x)) \Big|_{\epsilon=0} = n(x)$

so, $\frac{\partial F}{\partial y}$ (for $f=f(x,y,y')$)

$$+ \frac{\partial F}{\partial y'} \Big|_{y'_\epsilon=y'_0=y'}$$

$$\cdot \frac{dy'_\epsilon}{d\epsilon} \Big|_{\epsilon=0}$$

$\frac{\partial F}{\partial y'}$ (for $f=f(x,y,y')$) $\frac{d}{d\epsilon} (y'(x) + \epsilon n'(x)) = n'(x)$

$$= \frac{\partial F}{\partial y} \Big|_{(x)} \cdot n(x) + \frac{\partial F}{\partial y'} \Big|_{(x)} \cdot n'(x)$$

Here, y, y' are thought of as names for the second and third variables of $f(x,y,y')$



Note that here y and y' are seen as independent variables: the 2nd and 3rd coordinates of f 's input. We are forgetting that the one is the derivative of the other. So, if $F(x, y(x), y'(x)) = x + y(x) + y'(x)$, and that they depend on x ...

we see it as $F(x, y, y') = x + y + y'$, and

$$\frac{\partial F}{\partial y} = \frac{\partial y}{\partial y} = 1, \quad \frac{\partial F}{\partial y'} = \frac{\partial y'}{\partial y'} = 1.$$

without $\frac{\partial y'}{\partial y}$.
without $\frac{\partial y}{\partial y'}$.

This is very important \therefore we treat y, y' as independent.

Eventually,

$$\left. \frac{dI(\epsilon)}{d\epsilon} \right|_{\epsilon=0} = \int_{x_1}^{x_2} \left[\frac{\partial F}{\partial y} n(x) + \frac{\partial F}{\partial y'} \cdot n'(x) \right] dx$$

$$\Rightarrow \int_{x_1}^{x_2} \left[\frac{\partial F}{\partial y} n(x) + \frac{\partial F}{\partial y'} n'(x) \right] dx = 0, \quad (*)$$

for this n .

However, n was an arbitrary twice continuously differentiable function on $[x_1, x_2]$, with $n(x_1) = n(x_2) = 0$.

Thus, $(*)$ holds for all such n .

And, for any such n :

$$\int_{x_1}^{x_2} \frac{\partial F}{\partial y'} n'(x) dx \xrightarrow{\text{integration by parts}} \left[\frac{\partial F}{\partial y'} n(x) \right]_{x_1}^{x_2} - \int_{x_1}^{x_2} \frac{d}{dx} \left(\frac{\partial F}{\partial y'} \right) n(x) dx =$$

$\rightarrow 0, \text{ as } n(x_1) = n(x_2) = 0$

Now we remember the dependence of F on x ...

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

(7)

Remember that F was a function $F(x, y(x), y'(x))$.

So, here, to take $\frac{d}{dx} \left(\frac{\partial F}{\partial y'} \right)$ we need to

consider the variables x, y, y' of $\frac{\partial F}{\partial y'}$ as

functions of x . For instance,

$$\text{when } f(x, y, y') = x^2 + y^2 + (y')^2,$$

we have $\frac{\partial F}{\partial y'} = 2y'$, i.e.

$$\frac{\partial F}{\partial y'} \text{ (function of } x) = 2y'(x), \text{ so } \frac{d}{dx} \left(\frac{\partial F}{\partial y'} \right) = 2y''(x).$$

Don't make the mistake here to think of y' as an independent variable from x , like for the partial derivatives before... that would lead here to $\frac{d}{dx} \left(\frac{\partial F}{\partial y'} \right) = \frac{d}{dx} (2y') = 0$, which is very wrong.

We have thus shown that

$$\int_{x_1}^{x_2} \left[\frac{\partial F}{\partial y} n(x) - \frac{d}{dx} \left(\frac{\partial F}{\partial y'} \right) \cdot n(x) \right] dx = 0,$$

i.e. $\int_{x_1}^{x_2} \left[\frac{\partial F}{\partial y} - \frac{d}{dx} \left(\frac{\partial F}{\partial y'} \right) \right] \cdot n(x) dx$, for all

$n: [x_1, x_2] \rightarrow \mathbb{R}$, twice cont. diff., with $n(x_1) = n(x_2) = 0$.

Therefore :

$$\frac{\partial F}{\partial y} - \frac{d}{dx} \left(\frac{\partial F}{\partial y'} \right) = 0 \quad , \quad \text{for our minimiser } y.$$

We thus have:

If $y : [x_1, x_2] \rightarrow \mathbb{R}$ minimises the integral $\int_{x_1}^{x_2} F(x, y(x), y'(x)) dx$, then

$\frac{d}{dx} \left(\frac{\partial F}{\partial y'} \right) - \frac{\partial F}{\partial y} = 0$	<p>Euler Lagrange equation.</p>
--	---

th. if y satisfies it, then y is a minimiser of the integral!

All that we get is that, y is a solution for this y .

i.e., $I(\epsilon)$ stationary at $\epsilon=0$, no matter what η we picked.

$\frac{dI(\epsilon)}{d\epsilon} \Big|_{\epsilon=0} = 0$! This could happen if

y was maximising I instead, or, if $\epsilon=0$ was an inflection point of $I(\epsilon)$; which is

(9)

→ Application to our previous example

(finding the shortest path between two points (x_1, y_1) , (x_2, y_2) in \mathbb{R}^2):

We want to find

$y: [x_1, x_2] \rightarrow \mathbb{R}$, with

$y(x_1) = y_1$, $y(x_2) = y_2$,

that makes $I = \int_{x_1}^{x_2} \sqrt{1 + |y'(x)|^2} dx$ minimal.

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

The y that make I stationary are exactly the y that satisfy the Euler-Lagrange equation

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

Remember, when we calculate $\frac{\partial F}{\partial y'}$, $\frac{\partial F}{\partial y}$, we

forget the dependence of y, y' on x and on

each other: we write $F(x, y, y') = \sqrt{1 + (y')^2}$,

$$\text{so } \frac{\partial F}{\partial y'} = \frac{1}{2\sqrt{1+(y')^2}} \cdot \frac{\partial}{\partial y'} (1 + (y')^2) =$$

$$= \frac{1}{2\sqrt{1+(y')^2}} \cdot 2y' = \frac{y'}{\sqrt{1+(y')^2}}, \quad \text{and}$$

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

Now, we remember how everything depends on x :

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

$$\text{so } \underbrace{\frac{d}{dx} \left(\frac{\partial F}{\partial y'} \right) - \frac{\partial F}{\partial y}}_{\text{LHS of Euler-Lagrange equation}} = \frac{d}{dx} \left(\frac{y'(x)}{\sqrt{1+(y'(x))^2}} \right).$$

LHS of
Euler-Lagrange
equation

Thus, y satisfies the Euler-Lagrange equation

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

$$\Leftrightarrow \frac{y'(x)}{\sqrt{1+(y'(x))^2}} = C, \text{ a constant, } \forall x \in [x_1, x_2].$$

This is the same point we reached when solving the problem from scratch (but it is much faster this way). Just like before, it implies that $y' = \text{constant} \rightarrow \{(x, y(x)) : x \in [x_1, x_2]\}$ is the graph of a line segment.



Notice that, from the way we derived the Euler-Lagrange equation, it doesn't follow that, if y satisfies it, then it is a minimiser of the integral

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

but, if it is, it is one of the solutions of the Euler-Lagrange equation.

All we get is:

y satisfies the Euler-Lagrange equation



$$\frac{dI_y^n(\epsilon)}{d\epsilon} \Big|_{\epsilon=0} = 0, \quad \forall \eta: [x_1, x_2] \rightarrow \mathbb{R} \text{ twice cont. diff., with } \eta(x_1) = \eta(x_2) = 0.$$

we call this situation:

$$\int_{x_1}^{x_2} F(x, y(x), y'(x)) dx \text{ is stationary}$$

This could happen if, for any fixed η as above, $I_y^n(\epsilon)$ has a local minimum at $\epsilon=0$, or a local maximum at $\epsilon=0$, or an inflection point at $\epsilon=0$.

It could even happen that, for some η_1 , $I_y^{\eta_1}(\epsilon)$ is minimised at $\epsilon=0$, while, for another η_2 , $I_y^{\eta_2}(\epsilon)$ is maximised at $\epsilon=0$...

So, if it is not clear from our intuition that the y we found is indeed a minimiser of

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

extra checking is required.

The safest way to go is to show that, for our explicit y ,

$I^n(\epsilon)$ is minimised at $\epsilon=0$, $\forall \eta$ as above.

This will imply that y is a global minimiser of

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

the functions $y + \epsilon \cdot \eta$, over all $\epsilon \in \mathbb{R}$ and all $\eta : [x_1, x_2] \rightarrow \mathbb{R}$

twice cont. diff., with $\eta(x_1) = \eta(x_2) = 0$, are practically all possible perturbations of y with fixed endpoints (more precisely, they are dense in the set of all perturbations of y ...)

→ An observation:

Suppose that we have found a solution y to the Euler-Lagrange equation; i.e., a

y s.t. $\left. \frac{dI_y^n(\epsilon)}{d\epsilon} \right|_{\epsilon=0} = 0 \quad \forall n$ as above.

If we can also show that $\left. \frac{d^2 I_y^n(\epsilon)}{d\epsilon^2} \right|_{\epsilon=0} > 0$

for all n as above,

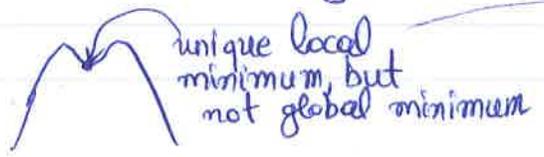
then y is a local minimum

of the integral $\int_{x_1}^{x_2} F(x, \kappa, y'(x)) dx$

(meaning that, amongst all small perturbations of y , y is the function that makes the above integral smallest; but there may exist significant perturbations of y that make the integral even smaller).

If, in addition, the Euler-Lagrange equation has no other solutions, then the above imply that y is the unique local minimiser

of the integral. But, theoretically, it still doesn't have to be a global minimum.

ex: 



In any case, however, all minimisers and all maximisers y of the integral

$$\int_{x_1}^{x_2} f(x, y(x), y'(x)) dx$$

are included in the solutions of the Euler-Lagrange equation.

Lecture 35:

1

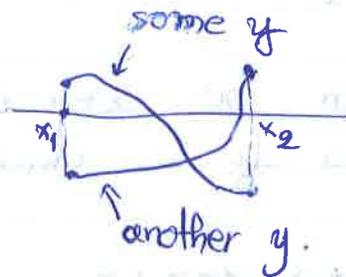
Reformulating the problem with different coordinates

We have developed a method that gives us the graphs $y: [x_1, x_2] \rightarrow \mathbb{R}$ that make the integral

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

stationary.

Note that $y(x_1), y(x_2)$ are not specified in the general case (unlike in the example of finding the shortest path between (x_1, y_1) and (x_2, y_2) in \mathbb{R}^2) i.e., the Euler-Lagrange equation gives all the y that make I stationary, independently of endpoint values:

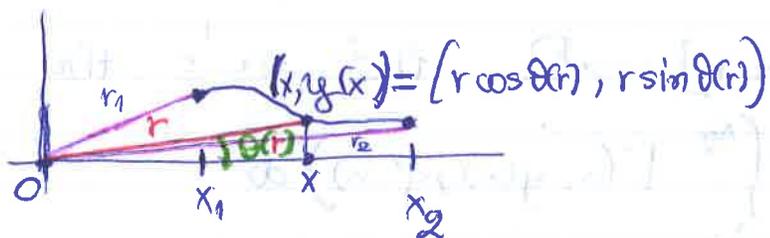


Sometimes, it is very useful, to avoid hard calculations, to switch to different coordinates. For instance:

2

• Polar coordinates:

Suppose that we have a graph $y: [x_1, x_2] \rightarrow \mathbb{R}$.



For each $x \in [x_1, x_2]$, $(x, y(x))$ has some distance r from $(0,0)$, and some angle $\theta(r)$ with the horizontal half-line $[0, +\infty)$

Important: the angle is a function of r ;
as we move across the graph, r changes,
and $\theta(r)$ changes with it. It is not independent
of r .

Thus, y can be seen as a function of r .
So, we can do a change of variables, to write

$$I = \int_{x_1}^{x_2} f(x, y(x), y'(x)) dx = \int_{r_1}^{r_2} \tilde{f}(r, \theta(r), \theta'(r)) dr.$$

change of
variables,
 $x = r \cos \theta(r)$,
 $y = r \sin \theta(r)$

another
function

3

So, to make I stationary, we just need to

make $\int_{r_1}^{r_2} \tilde{F}(r, \theta(r), \theta'(r)) dr$ stationary;

and we do this by solving the Euler-

Lagrange equation that corresponds to this new

function \tilde{F} :

$$\frac{d}{dr} \left(\frac{\partial \tilde{F}}{\partial \theta'} \right) - \frac{\partial \tilde{F}}{\partial \theta} = 0 \quad (*)$$

This will give us all θ as functions of r ,

that make $\int_{r_1}^{r_2} \tilde{F}(r, \theta(r), \theta'(r)) dr = I$ stationary.

Thus, the graphs $\left\{ (x, y(x)) : x \in [r_1, r_2] \right\}$ that make I stationary
are the graphs $\left\{ (r \cos \theta(r), r \sin \theta(r)) : r \in [r_1, r_2] \right\}$,
for all θ that solve $(*)$.

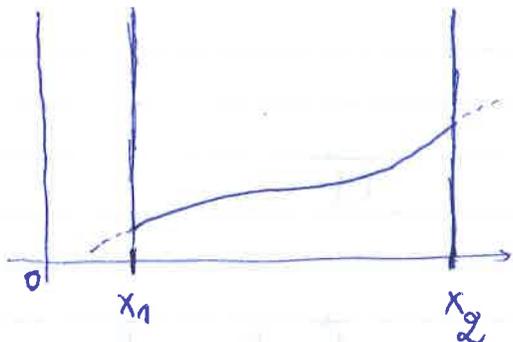
Let us see an explicit example (Example 1, p. 478)
of textbook

(4)

→ Finding the path a light ray follows when it goes through a ^{2-dim} material of varying

density: Suppose that the index of refraction is $n(r) = \frac{1}{r^2}$ at all points at distance r from O .

It holds that, as the light moves between the vertical lines $x=x_1$ and $x=x_2$,



the path $\{(x, y(x)) : x \in [x_1, x_2]\}$

that it follows is the path

$$\text{that makes } I = \int_{x_1}^{x_2} \underbrace{n(r)}_{\frac{1}{r^2} = \frac{1}{x^2 + y^2(x)}} ds = \int_{x_1}^{x_2} \frac{1}{x^2 + y^2(x)} \sqrt{1 + (y'(x))^2} dx$$

the arc-length measure on the path

$$= \int_{x_1}^{x_2} \underbrace{\frac{1}{x^2 + y^2(x)} \cdot \sqrt{1 + (y'(x))^2}}_{f(x, y(x), y'(x))} dx \text{ stationary,}$$

for all x_1, x_2 .

Thus, y satisfies, on its whole domain, the

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

$$f(x, y, y') = \frac{\sqrt{1 + (y')^2}}{x^2 + y^2}, \text{ thus}$$

(5)

$$\frac{\partial F}{\partial y} = \frac{-\sqrt{1+(y')^2}}{(x^2+y^2)^2} \cdot 2y,$$

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

∞ we want to solve for y the following equation:

$$\frac{d}{dx} \left(\frac{y'(x)}{(x^2+y(x)^2)\sqrt{1+(y'(x))^2}} \right) = \frac{-2y(x) \cdot \sqrt{1+(y'(x))^2}}{(x^2+y(x)^2)^2}$$

The integral of the RHS is too hard to calculate...

But changing I into polar coordinates makes things much simpler:

We want to make stationary the integral

$$I = \int_{x_1}^{x_2} \frac{1}{\underbrace{x^2 + y^2(x)}_{=r^2}} \underbrace{\sqrt{1+(y'(x))^2}}_{ds} dx.$$

$$\text{Now, } ds = \sqrt{1+(y'(x))^2} dx = \sqrt{(dx)^2 + \underbrace{(y'(x) dx)^2}_{=dy^2}} = \sqrt{dx^2 + dy^2}.$$

And: for $x = x(r) = r \cos \theta(r)$, $y = y(r) = r \sin \theta(r)$

(as this is the change of variables for $(x, y(x))$ set

(6)

distance r from $(0,0)$, we have:

$$\begin{aligned} dx &= x'(r) dr = \left(r' \cos \theta(r) + r (\cos \theta(r))' \right) dr \\ &= \left(\cos \theta(r) - r \sin \theta(r) \cdot \theta'(r) \right) dr, \end{aligned}$$

$$\begin{aligned} dy &= y'(r) dr = \left(r' \sin \theta(r) + r (\sin \theta(r))' \right) dr \\ &= \left(\sin \theta(r) + r \cos \theta(r) \cdot \theta'(r) \right) dr, \end{aligned}$$

$$\begin{aligned} \text{so } (dx)^2 + (dy)^2 &= \left(\sin \theta(r) \right)^2 + r^2 \left(\sin \theta(r) \right)^2 \left(\theta'(r) \right)^2 - \\ &- \cancel{2 \cos \theta(r) \cdot r \sin \theta(r) \cdot \theta'(r)} + \left(\sin \theta(r) \right)^2 + r^2 \left(\cos \theta(r) \right)^2 \left(\theta'(r) \right)^2 + \\ &+ \cancel{2 \sin \theta(r) \cdot r \cos \theta(r) \cdot \theta'(r)} \Big] (dr)^2 \\ &= \left(1 + r^2 \left(\theta'(r) \right)^2 \right) (dr)^2 \end{aligned}$$

$$\text{thus } \sqrt{(dx)^2 + (dy)^2} = \sqrt{1 + r^2 \left(\theta'(r) \right)^2} dr,$$

This is useful to remember by heart: if ds

is the arc-length measure on a curve, and

we change to polar coordinates, then

$$ds = \sqrt{1 + r^2 \left(\theta'(r) \right)^2} dr$$

(7)

$$So, I = \int_{r_1}^{r_2} \frac{1}{r^2} \sqrt{1 + r^2 (\theta'(r))^2} dr,$$

$\tilde{F}(r, \theta(r), \theta'(r))$

and we want to make it stationary for all r_1, r_2 .

Thus, we solve the Euler-Lagrange equation

$$\frac{d}{dr} \left(\frac{\partial \tilde{F}}{\partial \theta'} \right) - \frac{\partial \tilde{F}}{\partial \theta} = 0 \text{ for } \theta,$$

where $\tilde{F}(r, \theta, \theta') = \frac{\sqrt{1 + r^2 (\theta')^2}}{r^2}$

Now, $\frac{\partial \tilde{F}}{\partial \theta} = 0,$ and

an advantage of using polar coordinates here!

$$\frac{\partial \tilde{F}}{\partial \theta'} = \frac{1}{r^2} \cdot \frac{1}{2\sqrt{1 + r^2 (\theta')^2}} \cdot 2r^2 \theta' = \frac{\theta'}{\sqrt{1 + r^2 (\theta')^2}}$$

So, we need to solve for θ the equation

$$\frac{d}{dr} \left(\frac{\theta'(r)}{\sqrt{1 + r^2 (\theta'(r))^2}} \right) = 0. \text{ It implies that } \exists c, \text{ constant}$$

s.t. $\frac{\theta'(r)}{\sqrt{1 + r^2 (\theta'(r))^2}} = c$, $\forall r \Rightarrow (\theta'(r))^2 = c^2 \cdot (1 + r^2 (\theta'(r))^2)$

$$\Rightarrow (\theta'(r))^2 \cdot (1 - c^2 r^2) = c^2 \quad \forall r$$

$$\Rightarrow (\theta'(r))^2 = \frac{c^2}{1 - c^2 r^2} \quad \forall r$$

$> 0,$
as everything
is > 0

by \oplus ,
 $\theta(r)$ and
 c have the
same sign

$$\theta'(r) = \frac{c}{\sqrt{1 - c^2 r^2}} \quad \forall r$$

$$\Rightarrow \theta(r) = \arcsin(cr) + \text{constant}, \quad \forall r$$

really, $\forall r \in (0, \frac{1}{c})$!
for $r \geq \frac{1}{c}$, the light
moves on the path glued
to this, corresponding to
some other c that has
the same tangent as ours
at the point at $r = \frac{1}{c}$.

depends on
initial direction
of light ray

depends on initial
angle of light ray with x-axis



See how, in this application again, the solitary

derivative $\frac{\partial F}{\partial \theta}$ in the Euler-Lagrange equation

was 0. This simplified matters a lot.

in more detail

We will now see how missing variables

in $F(x, y, y')$ can make our life easier when

solving the Euler-Lagrange equation:

(9)

→ Pursuing simple first integrals of the Euler-Lagrange equation:

The Euler-Lagrange equation gives

$$\frac{d}{dx} \left(\frac{\partial F}{\partial y'} \right) = \frac{\partial F}{\partial y}, \quad \text{for all } y \text{ that make } I = \int_{x_1}^{x_2} F(x, y(x), y'(x)) dx \text{ stationary.}$$

$$\text{So, } \frac{\partial F}{\partial y'} = \int \frac{\partial F}{\partial y} + c \quad \text{i.e. } \frac{\partial F}{\partial y'} \text{ can be}$$

given by an integral (which may, however, be hard to calculate).

The definition of "first integral of the Euler-Lagrange equation" is a little vague:

it is any meaningful integral of $\frac{d}{dx}$ (function depending on F) we can take,

so as to get rid of $\frac{d}{dx}$ in the Euler-Lagrange equation. Two useful cases, which we pursue because of their simplicity, are the following:

→ Case 1:

$f(x, y, y')$ does not depend on y , when we are considering x, y, y' as independent variables.

⚠ This can only be valid if we forget the dependence of x, y, y' on each other. We are really seeing x, y, y' as names for the 3 variables of f , and checking if f depends on its middle variable.

For instance, for $f(x, y(x), y'(x)) = x + y'(x)$,
 where $y(x) = x^2 \forall x$, we have: $f(x, y, y') = x + y'$,
($\rightarrow y(x) = 2x + x$)
 so, when x, y, y' are considered as independent variables, then f doesn't depend on y

Then, $\frac{\partial f}{\partial y} = 0$. So, the Euler-Lagrange equation becomes

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

thus, by integrating both sides,

← simpler, maybe easy to solve.

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

→ we say that this is a first integral of the Euler-Lagrange equation in this case (as it follows from integrating something once w.r.t. x).

Both examples we have seen so far

(shortest path between 2 points in \mathbb{R}^2 ,

and

path of light ray in some ^(specific) material)

fall in this category.

! Notice that polar coordinates in light refraction helped exactly because of this: they made the integrand independent of the middle variable!

Here is another example, where we will force our integrand to be independent of its middle variable

(so that we can take advantage of the simple first integral of the Euler-Lagrange equation that we get in that case) :

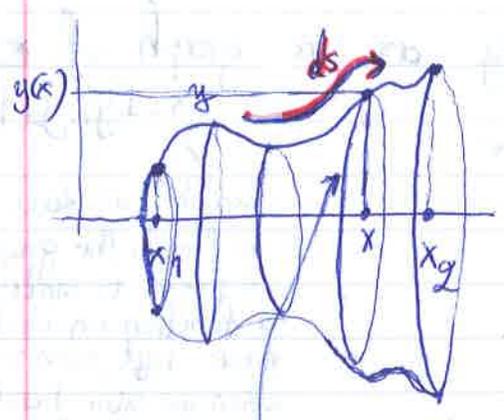
→ Which is the curve connecting two fixed points, such that, when we revolve it around the x-axis, we get the smallest possible surface area?

Let $(x_1, y_1), (x_2, y_2)$ be the two fixed points.

We are looking for $y: [x_1, x_2] \rightarrow \mathbb{R}$,
with $y(x_1) = y_1, y(x_2) = y_2$, s.t.

$$I = \int_{x_1}^{x_2} 2\pi y(x) \underbrace{\sqrt{1+(y'(x))^2}}_{ds} dx$$

is minimised.



length of circle: $2\pi \cdot \text{radius} = 2\pi y(x)$

We thus want to minimise $I = \int_{x_1}^{x_2} F(x, y(x), y'(x)) dx,$

for $f(x, y, y') = 2\pi y \sqrt{1+(y')^2}$. See that this F depends

on its middle variable y ; thus, we do not get

the nice first integral $\frac{\partial F}{\partial y'} = \text{constant}$... This

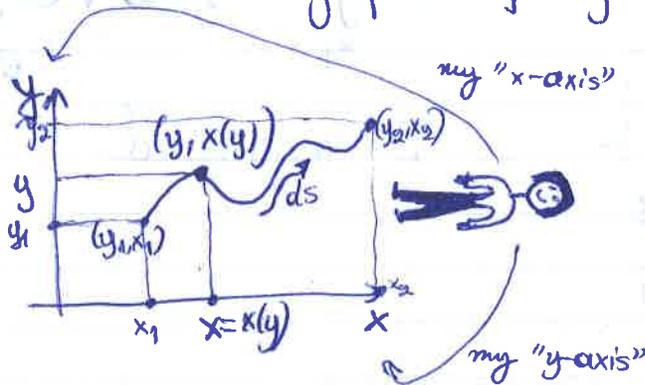
makes things complicated. So, we will try to change
our integrand (and of course the variable of integration)

so that the new integrand doesn't depend any more
on its middle variable:

Important observation:

$$ds = \sqrt{1 + x'(y)} dy \text{ as well,}$$

when we see the graph of y as a graph of $x: [y_1, y_2] \rightarrow \mathbb{R}$:



maybe we have to break the graph in pieces to make this a function on each piece; but, even so, when we glue back we get the above expression for ds .

$$\text{So, } I = \int_{y_1}^{y_2} \underbrace{2\pi y \cdot \sqrt{1 + (x'(y))^2}}_{\parallel}$$

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

doesn't depend on x !

To find the $x: [y_1, y_2] \rightarrow \mathbb{R}$ that minimises this,

we need to solve the Euler-Lagrange equation

$$\frac{d}{dy} \left(\frac{\partial F}{\partial x'} \right) - \frac{\partial F}{\partial x} \overset{0}{=} 0, \text{ i.e. } \frac{d}{dy} \left(\frac{\partial F}{\partial x'} \right) = 0,$$

↑
middle variable

i.e. $\frac{\partial F}{\partial x'} = \text{constant}$: this is a nice first integral.

Note that $\frac{\partial F}{\partial x'} = 2\pi y \cdot \frac{1}{2\sqrt{1+(x')^2}} \cdot 2x' =$

$= \frac{4\pi y x'}{\sqrt{1+(x')^2}}$, thus

$\frac{\partial F}{\partial x'} = \text{constant}$ means that F is const.

$\frac{y x'(y)}{\sqrt{1+(x'(y))^2}} = c, \forall y \in [y_1, y_2]$

$\Rightarrow y^2 (x'(y))^2 = c^2 + c^2 (x'(y))^2$

$\Rightarrow (x'(y))^2 \cdot (y^2 - c^2) = c^2$

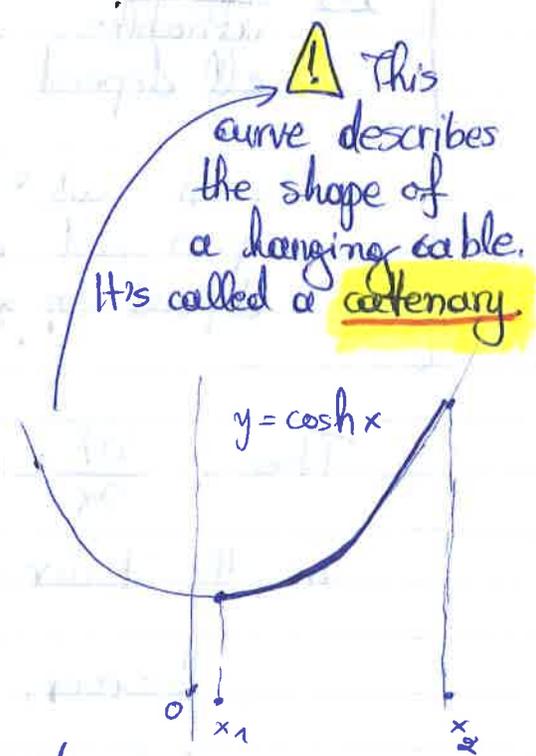
$\Rightarrow (x'(y))^2 = \frac{c^2}{y^2 - c^2}$

$\Rightarrow x'(y) = \frac{c}{\sqrt{y^2 - c^2}}$

$\Rightarrow x(y) = c \cdot \cosh^{-1}\left(\frac{y}{c}\right) + \tilde{c}$ \hookrightarrow another constant

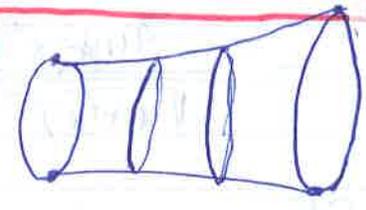
$\Rightarrow \cosh^{-1}\left(\frac{y}{c}\right) = \frac{x(y) - \tilde{c}}{c} \Rightarrow \frac{y(x)}{c} = \cosh\left(\frac{x - \tilde{c}}{c}\right)$

i.e. $y(x) = \cosh\left(\frac{x - \tilde{c}}{c}\right), \forall x \in [x_1, x_2]$. From $y(x_1) = y_1, y(x_2) = y_2$, we get c, \tilde{c}

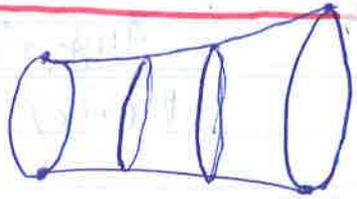


This is the surface that soap forms between two circular hoops

the surface has to be symmetric all around, so it comes from a revolution of a curve



This is the surface that soap forms between two circular hoops



the surface has to be symmetric all around, so it comes from a revolution of a curve

Lecture 36:

1

Case 2: $f(x, y, y')$ does not depend on x

⚠ Again, here x, y, y' are seen as independent variables, forgetting that they actually all depend on x . Such a case is $f(x, y, y') = yty'$: even when $y(x) = x^2 \forall x$, we pretend here that y is just an independent variable, thus yty' doesn't depend on x from this perspective.

Then, $\frac{\partial F}{\partial x} = 0$. Of course, $\frac{\partial F}{\partial x}$ doesn't appear in the Euler-Lagrange equation.

However, the following equality holds thanks to ① the Euler-Lagrange equation, and ② our assumption that $\frac{\partial F}{\partial x} = 0$:

2

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

→ for y that makes $\int_{x_1}^{x_2} f(x, y(x), y'(x)) dx$ stationary.

Proof: By the usual rules of differentiation,

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

here we have dependence on x , and $\frac{d}{dx}$ takes it into account.

NOT 0!

Chain rule for $\frac{d}{dx} F(x)$

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

by Euler-Lagrange

$$y''(x) \cdot \frac{\partial F}{\partial y'}(x) + y'(x) \cdot \frac{\partial F}{\partial y} - \left(\frac{\partial F}{\partial x}(x) + \frac{\partial F}{\partial y} y'(x) + \frac{\partial F}{\partial y'} y''(x) \right) =$$

here, $\frac{\partial F}{\partial x}$, $\frac{\partial F}{\partial y}$, $\frac{\partial F}{\partial y'}$ are taken w.r.t. the variables x, y, y' , considered independent: they really are just the partial derivatives of F w.r.t its 3 variables

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

↓
we are in Case 2.

③

So, in Case 2, we have that our desired y satisfies the first order ODE

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

→ the Beltrami identity.

↓
this is a simple first integral of the Euler-Lagrange equation in this case.

→ Application of Beltrami identity to our previous example, of finding curve whose revolution produces surface of minimal area:

We saw that, if $y: [x_1, x_2] \rightarrow \mathbb{R}$,

with $y(x_1) = y_1, y(x_2) = y_2$,

is the curve connecting $(x_1, y_1), (x_2, y_2)$, whose revolution produces a surface of smallest possible area (amongst all surfaces produced that way),

then I minimises

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

we can forget about this...

$F(x, y(x), y'(x))$

If we hadn't thought of the fact that $\sqrt{1+(y'(x))^2} dx = \sqrt{1+(x'(y))^2} dy$ (or if this led us nowhere),

we could still notice that

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

Thus, the Beltrami identity holds:

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

$$\begin{aligned} \frac{\partial F}{\partial y'} &= y \cdot \frac{1}{2\sqrt{1+(y')^2}} \cdot 2y' \text{ , so } y' \frac{\partial F}{\partial y'} - F = \\ &= y' \cdot y \cdot \frac{1}{\sqrt{1+(y')^2}} \cdot y' - y \sqrt{1+(y')^2} = \frac{y \cdot (y')^2}{\sqrt{1+(y')^2}} - \frac{y \sqrt{1+(y')^2}}{\sqrt{1+(y')^2}} = \\ &= \frac{-y}{\sqrt{1+(y')^2}}. \end{aligned}$$

Thus, $\exists c \in \mathbb{R}$ s.t.

(5)

$$\frac{-y(x)}{\sqrt{1+(y'(x))^2}} = c \quad \forall x$$

$$\rightarrow c^2 + c^2(y'(x))^2 = y^2(x) \quad \forall x$$

$$\rightarrow (c^2 + c^2(y'(x))^2)' = (y^2(x))' \quad \forall x$$

$$\rightarrow c^2 \cdot \cancel{2y'(x)} \cdot y''(x) = \cancel{2y(x)} \cdot y'(x) \quad \forall x$$

$$\rightarrow y''(x) = \frac{1}{c^2} y(x) \quad \forall x.$$

We can see that this last equation (combined with the ones above) leads to the catenary (hyperbolic cosine) solution, as before.

Maybe this way was a bit more involved, but it didn't require us to be too clever, or too lucky.

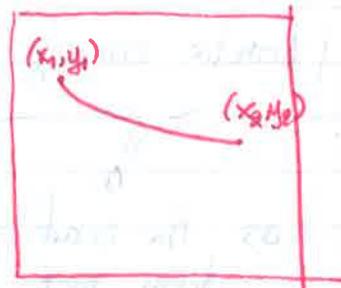
6

→ The brachistochrone problem: We are given a bead.

Given points (x_1, y_1) , (x_2, y_2) , on the same vertical

2-dim board, find the curve connecting the two points, down which the bead will slide in the least time (from rest)

(i.e., st. the bead will take longer to slide down any other curve connecting the two points):

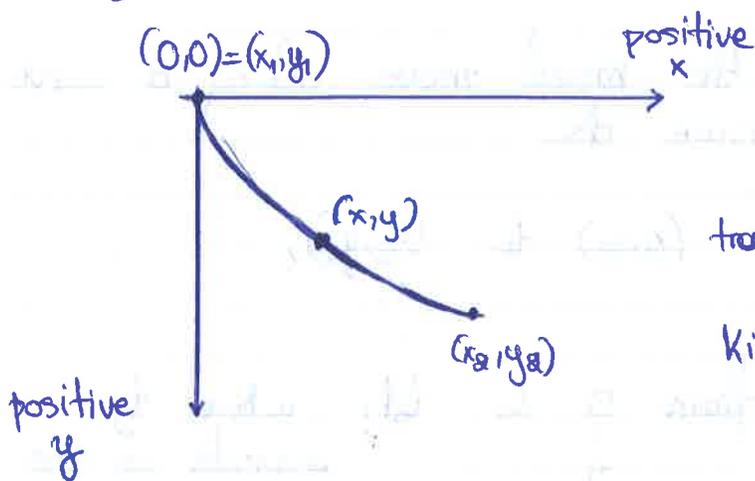


Assume:
no friction.

Let m be the mass of the bead.

We put a coordinate system on the vertical board, so that

$(x_1, y_1) = (0, 0)$, and the positive y -axis points downwards:



When our bead is at arbitrary position (x, y) , then it has:

kinetic energy = $\frac{1}{2} m \cdot v^2$,
↓
its velocity at position (x, y)

and potential energy = $-mg \cdot y$

Due to the lack of friction, energy is preserved, thus:

$$[\text{kinetic energy at } (x,y)] + [\text{potential energy at } (x,y)]$$

$$[\text{kinetic energy at } (0,0)] + [\text{potential energy at } (0,0)] = 0.$$

as the bead starts from rest

$$mg \cdot 0 = 0$$

$$\text{So, } \frac{1}{2} m \cdot v^2 - mgy = 0$$

$$\rightarrow v^2 = 2gy \Rightarrow v = \sqrt{2gy}$$

the velocity at (x,y).

Now, suppose that the bead moves along a curve with arc-length measure ds.

As it moves from (x₁, y₁) to (x₂, y₂),

time changes from 0 to t₂, where t₂ depends on the curve.

We want to find the curve that minimises

$$t_2 = \int_0^{t_2} dt \quad \frac{ds}{dt} = v \Rightarrow dt = \frac{ds}{v}$$

(v) changes depending on (x,y) on curve

$$= \int_{\substack{(x_1, y_1) \\ (0,0)}}^{(x_2, y_2)} \frac{1}{v(x,y)} ds = \int_{(x_1, y_1)}^{(x_2, y_2)} \frac{1}{\sqrt{2gy}} ds = \frac{ds = \sqrt{1+(y'(x))^2} dx}{\sqrt{2gy}}$$

$$= \int_0^{x_2} \frac{1}{\sqrt{2gy(x)}} \sqrt{1+(y'(x))^2} dx =$$

$$= \frac{1}{\sqrt{2g}} \int_0^{x_2} \underbrace{\frac{1}{\sqrt{y(x)}} \cdot \sqrt{1+(y'(x))^2}}_{f(x, y(x), y'(x))} dx$$

We see that $f(x, y, y') = \frac{\sqrt{1+(y')^2}}{\sqrt{y}}$ depends on its middle variable $y \dots$

(but not on its 1st variable, so we can use Beltrami identity. It will give us an answer, but we will follow another way here.)

But we remember that, given that we derived this formula by integrating against ds , we can use the standard trick

$$\sqrt{1+(y'(x))^2} dx = ds = \sqrt{1+(x'(y))^2} dy$$

So, the integral we want to minimise becomes

$$\int_0^{y_2} \underbrace{\frac{1}{\sqrt{y}} \sqrt{1+(x'(y))^2}}_{\tilde{F}(y, x(y), x'(y))} dy$$

Now, $\tilde{F}(y, x, x') = \frac{\sqrt{1+(x')^2}}{\sqrt{y}}$, independent of its middle variable!

(9)

So, the Euler-Lagrange equation gets the nice form

$$\frac{d}{dy} \left(\frac{\partial \tilde{F}}{\partial x'} \right) = 0, \text{ i.e. } \frac{\partial \tilde{F}}{\partial x'}(y) = c \quad \forall y \in [y_1, y_2],$$

for some constant $c \in \mathbb{R}$.

$$\text{Now, } \frac{\partial \tilde{F}}{\partial x'} = \frac{1}{\sqrt{y}} \cdot \frac{1}{\sqrt{1+(x')^2}} \cdot x' = \frac{x'}{\sqrt{y} \cdot \sqrt{1+(x')^2}}$$

$$\text{So, } \frac{x'(y)}{\sqrt{y} \cdot \sqrt{1+(x'(y))^2}} = c, \quad \forall y \in [y_1, y_2]$$

$\hookrightarrow x'$ has always the same sign. Since x changes from $x=0$ to $x_2 > 0$, it must increase at some point, thus $\exists y$ s.t. $x'(y) > 0 \Rightarrow x'(y) > 0 \quad \forall y$

$$\rightarrow (x'(y))^2 = c^2 \cdot y \cdot (1 + (x'(y))^2)$$

$$\rightarrow (x'(y))^2 \cdot (1 - c^2 y) = c^2 y \Rightarrow x'(y) = \sqrt{\frac{c^2 y}{1 - c^2 y}}$$

$$x'(y) > 0 \quad \forall y \Rightarrow dx = \sqrt{\frac{c^2 y}{1 - c^2 y}} dy$$

A trick to simplify this is setting $c^2 y = \sin^2 \left(\frac{\theta}{2} \right)$

$$\left(= \frac{1}{2} (1 - \cos \theta) \right)$$

$$\text{As } \theta \text{ varies, } dy = d \left(\frac{1}{c^2} \sin^2 \frac{\theta}{2} \right) = \frac{1}{c^2} \cdot 2 \sin \frac{\theta}{2} \cdot \left(\sin \frac{\theta}{2} \right)' d\theta =$$

$$= \frac{1}{c^2} \cdot 2 \sin \frac{\theta}{2} \cos \frac{\theta}{2} \cdot \frac{1}{2} d\theta = \frac{1}{c^2} \left(\sin \frac{\theta}{2} \right) \left(\cos \frac{\theta}{2} \right) d\theta$$

$$\text{So, } dx = \sqrt{\frac{\sin^2 \frac{\theta}{2}}{1 - \sin^2 \frac{\theta}{2}} \cdot \frac{1}{c^2} \left(\sin \frac{\theta}{2} \right) \left(\cos \frac{\theta}{2} \right) d\theta =$$

$1 - \sin^2 \frac{\theta}{2} = \cos^2 \frac{\theta}{2}$

(10)

$$= \frac{\sin \frac{\theta}{2}}{\cos \frac{\theta}{2}} \cdot \frac{1}{2c^2} (\sin \frac{\theta}{2}) \cdot (\cos \frac{\theta}{2}) d\theta = \frac{1}{2c^2} \cdot \sin^2 \frac{\theta}{2} d\theta =$$

$$= \frac{1}{2c^2} \cdot (1 - \cos \theta).$$

So, integrating, we get $x(\theta) = \int \frac{1}{2c^2} (1 - \cos \theta) d\theta + c'$

$$= \int \frac{1}{2c^2} d\theta - \int \frac{\cos \theta}{2c^2} d\theta + c' =$$

$$= \frac{1}{2c^2} (\theta - \sin \theta) + c'.$$

Remember that, for $\theta=0$, we get $y=0$: we are at our initial position. So, for $\theta=0$, also $x=0$, thus $c'=0$.

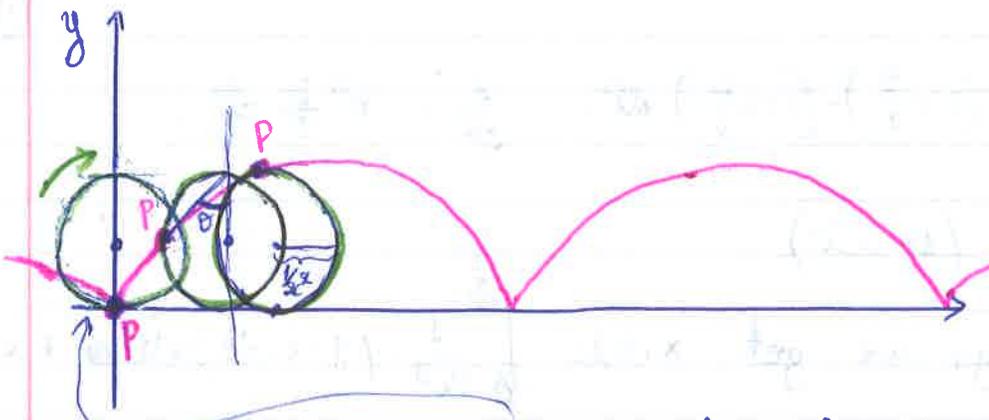
So, the curve that minimises our desired integral is the one consisting of the points (x, y) , st.

$$x = \frac{1}{2c^2} (\theta - \sin \theta)$$

and $y = \frac{1}{2c^2} (1 + \cos \theta).$

, as θ runs in $[0, 2\pi]$

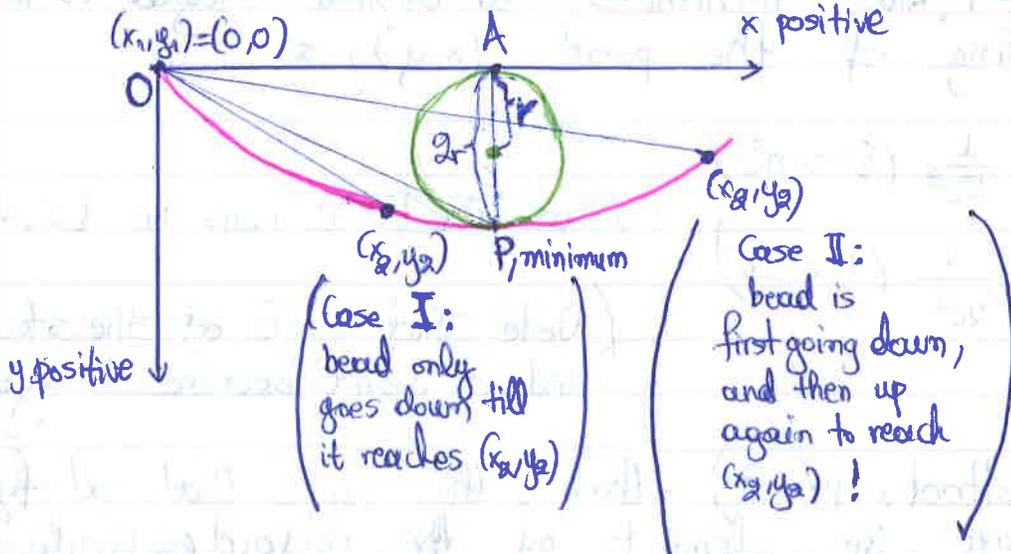
We see (Textbook, p.483) that the (x, y) that satisfy the above are the elements of the cycloid produced by rolling a circle of radius $\frac{1}{2c^2}$:



As we roll this circle, of radius $\frac{1}{2c\epsilon}$, the point P follows the pink orbit: this pink curve is a cycloid.

Note that the cycloid is not differentiable for $\theta = 2\pi, 4\pi, \dots$; while our graph $(x, y(x))$ had to be differentiable.

So, the bead will slide on the unique cycloid connecting $(x_1, y_1), (x_2, y_2)$ within its 1st period (before it reaches 2π):



We find the radius $r = \frac{1}{2g}$ of the cycloid, and the angle θ that achieves (x_2, y_2) , via the system of equations:

$$x_2 = \underbrace{r}_{\frac{1}{2g}} (\theta - \sin\theta)$$

(our (x_2, y_2) is known).

$$y_2 = r \cdot (1 - \cos\theta)$$

We can deduce easily whether the bead "hits" (x_2, y_2) in a downward or upward orbit: Let P be the lowest pt of the cycloid.

• If $\frac{y_2}{x_2} > \frac{PA}{OA} = \frac{2r}{\pi r} = \frac{2}{\pi}$, then $O(x_2, y_2)$ is below the line OP, so the bead "hits" (x_2, y_2) a downward orbit.

• If $\frac{y_2}{x_2} < \frac{PA}{OA} = \frac{2}{\pi}$, then $O(x_2, y_2)$ is above OP, so the bead "hits" (x_2, y_2) in an upward orbit, after first having passed through P.

Lecture 37:

①

→ finding graphs $\{(x, y_1(x)) : x \in [x_1, x_2]\}$,
 $\{(x, y_2(x)) : x \in [x_1, x_2]\}$,
 \vdots
 $\{(x, y_n(x)) : x \in [x_1, x_2]\}$

that simultaneously make

$$\int_{x_1}^{x_2} f(x, y_1(x), y_1'(x), y_2(x), y_2'(x), \dots, y_n(x), y_n'(x)) dx$$

stationary :

Suppose we want to find functions

$$y_1: [x_1, x_2] \rightarrow \mathbb{R}, y_2: [x_1, x_2] \rightarrow \mathbb{R}, \dots, y_n: [x_1, x_2] \rightarrow \mathbb{R},$$

that minimise $I = \int_{x_1}^{x_2} f(x, y_1(x), y_1'(x), y_2(x), y_2'(x), \dots, y_n(x), y_n'(x)) dx$

(doesn't matter how large $n \in \mathbb{N}$ is).

Suppose we have found such functions y_1, \dots, y_n .

Then, if we fix y_2, \dots, y_n , and perturb y_1 ,

then the integral can only get larger. I.e.:

$$I_{y_1}^n(\epsilon) := \int_{x_1}^{x_2} f(x, (y_1 + \epsilon n)(x), (y_1 + \epsilon n)'(x), y_2(x), y_2'(x), \dots, y_n(x), y_n'(x)) dx$$

(2)

$$is \geq I_{y_i}^n(0), \quad \forall \epsilon \in \mathbb{R},$$

for any $n: [x_1, x_2] \rightarrow \mathbb{R}$ twice cont. differentiable,
with $\eta(x_1) = \eta(x_2) = 0$,

$$thus \quad \frac{dI_{y_i}^n(\epsilon)}{d\epsilon} \Big|_{\epsilon=0} = 0, \quad \text{for any } n \text{ as above.}$$

Similarly, perturbations $y_i + \epsilon n$ of y_i , $\forall \epsilon \in \mathbb{R}$,

and all n as above,
can only lead to larger integrals:

$$I_{y_i}^n(\epsilon) := \int_{x_1}^{x_2} f(x, y_1(x), y_1'(x), \dots, y_{i-1}(x), y_{i-1}'(x), (y_i + \epsilon n)(x), (y_i + \epsilon n)'(x), \\ y_{i+1}(x), y_{i+1}'(x), \dots, y_n(x), y_n'(x)) dx$$

$$\geq I_{y_i}^n(0), \quad \forall \epsilon \in \mathbb{R}, \quad \forall n \text{ as above,}$$

$$thus \quad \frac{dI_{y_i}^n(\epsilon)}{d\epsilon} \Big|_{\epsilon=0} = 0, \quad \text{for all } n \text{ as above.}$$

Thus, if y_1, y_2, \dots, y_n minimise $I = \int_{x_1}^{x_2} f(x, y_1(x), y_1'(x), \dots, y_n(x), y_n'(x)) dx$,

then they make I stationary;

meaning that $\frac{dI_{y_i}^n(\epsilon)}{d\epsilon} \Big|_{\epsilon=0} = 0$, for all $i=1, \dots, n$,

(3)

for all $\eta: [x_1, x_2] \rightarrow \mathbb{R}$

twice cont. differentiable,

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

Thus, **each** y_i

makes stationary the integral

$$\int_{x_1}^{x_2} F_i(x, y_i(x), y_i'(x)) dx,$$

where $F_i(x, y_i(x), y_i'(x)) = f(x, y_1(x), y_1'(x), \dots, y_n(x), y_n'(x))$,

where $y_1, \dots, y_{i-1}, y_{i+1}, \dots, y_n$

are the fixed other functions (that we are looking for).

So, y_i satisfies the Euler - Lagrange equation

$$\frac{d}{dx} \left(\frac{\partial F_i}{\partial y_i'} \right) - \frac{\partial F_i}{\partial y_i} = 0, \quad \text{i.e.}$$

\parallel \parallel

$$\frac{\partial f}{\partial y_i'} \quad \quad \quad \frac{\partial f}{\partial y_i}$$

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

Therefore, the functions $y_1, \dots, y_n : [x_1, x_2] \rightarrow \mathbb{R}$ that make the integral

$$I = \int_{x_1}^{x_2} F(x, y_1(x), y_1'(x), \dots, y_n(x), y_n'(x)) dx \text{ stationary.}$$

are those that satisfy

$$\left\{ \begin{array}{l} \frac{d}{dx} \left(\frac{\partial F}{\partial y_1'} \right) - \frac{\partial F}{\partial y_1} = 0, \\ \frac{d}{dx} \left(\frac{\partial F}{\partial y_2'} \right) - \frac{\partial F}{\partial y_2} = 0, \\ \vdots \\ \frac{d}{dx} \left(\frac{\partial F}{\partial y_n'} \right) - \frac{\partial F}{\partial y_n} = 0 \end{array} \right.$$

And of course, if minimisers (or maximisers) of the integral exist, then they are solutions to the above system of Euler-Lagrange equations.

(5)

→ Application: Lagrange's equations:

Hamilton's principle says that a particle in \mathbb{R}^3 always moves in such a way that the integral

$$I = \int_{t_1}^{t_2} L(t) dt$$

is stationary. Here,

$$L(t) = T(t) - V(t)$$

↓
kinetic energy
at time t

↓
potential energy
at time t .

We denote by $(x(t), y(t), z(t))$ the coordinates of the particle at time t . Let m be its mass. Then:

$$T(t) = \frac{1}{2} m \cdot (v(t))^2 = \frac{1}{2} m \cdot \left[\underbrace{(x'(t))^2 + (y'(t))^2 + (z'(t))^2}_{v(t)} \right]^{\frac{1}{2}} =$$

$$= \frac{1}{2} m \cdot \left((x'(t))^2 + (y'(t))^2 + (z'(t))^2 \right) \rightsquigarrow \text{depends only on } x'(t), y'(t), z'(t).$$

In general, $V(t)$ depends on $t, x(t), y(t), z(t), x'(t), y'(t), z'(t)$.

(6)

So, $L(t) = L(t, x(t), x'(t), y(t), y'(t), z(t), z'(t))$

Thus, the functions $x, y, z: [t_1, t_2] \rightarrow \mathbb{R}$

that make $I = \int_{t_1}^{t_2} L(t) dt$ stationary

are those that satisfy the system of Euler - Lagrange equations

$$\frac{d}{dt} \left(\frac{\partial L}{\partial x'} \right) - \frac{\partial L}{\partial x} = 0$$

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

$$\frac{d}{dt} \left(\frac{\partial L}{\partial z'} \right) - \frac{\partial L}{\partial z} = 0$$

These are known as

Lagrange's equations

(but they are the same Euler-Lagrange equations we have been working on all along, just for integrand $L = T - V$).

→ Example: Find the position $(x(t), y(t), z(t))$, at each time $t \in [t_1, t_2]$, of a particle of mass m , moving (close to the earth's surface) under gravity:

Solution:

(7)

At each time t , the particle has:

- kinetic energy = $\frac{1}{2} m (v(t))^2 = \frac{1}{2} m \cdot \left((x'(t))^2 + (y'(t))^2 + (z'(t))^2 \right)$
- potential energy = $m \cdot g \cdot z(t)$.

Thus,

$$L(t) = \frac{1}{2} m \cdot \left((x'(t))^2 + (y'(t))^2 + (z'(t))^2 \right) - m \cdot g \cdot z(t),$$

$$\forall t \in [t_1, t_2].$$

We have explained that the functions $x, y, z: [t_1, t_2] \rightarrow \mathbb{R}$ are solutions to Lagrange's equations. Now:

$$L(t, x, x', y, y', z, z') = \frac{1}{2} m \left((x')^2 + (y')^2 + (z')^2 \right) - mgz. \text{ So:}$$

$$\left. \begin{aligned} \frac{\partial L}{\partial x} = \frac{\partial L}{\partial y} = 0, \text{ so we must have} \\ \frac{d}{dt} \left(\frac{\partial L}{\partial x'} \right) = 0 \\ \frac{d}{dt} \left(\frac{\partial L}{\partial y'} \right) = 0 \end{aligned} \right\} \rightarrow$$

$$\rightarrow \frac{\partial L}{\partial x'} = \text{constant}, \quad \frac{\partial L}{\partial y'} = \text{constant}.$$

$$\text{Since } \frac{\partial L}{\partial x'} = \frac{1}{2} m \cdot 2x' = mx'$$

$$\text{and } \frac{\partial L}{\partial y'} = \frac{1}{2} m \cdot 2y' = my'$$

$$\text{, we have: } x'(t) = c_1 \forall t \in [t_1, t_2]$$

$$\text{and } y'(t) = c_2 \forall t \in [t_1, t_2].$$

As for z : It satisfies

$$\left. \frac{d}{dt} \left(\frac{\partial L}{\partial z'} \right) - \frac{\partial L}{\partial z} = 0. \right] \textcircled{*}$$

$$\frac{\partial L}{\partial z} = -mg, \quad \frac{\partial L}{\partial z'} = \frac{1}{2} m \cdot \cancel{2} z' = m z',$$

$$\text{so } \textcircled{*} \Rightarrow \frac{d}{dt} (m \cdot z'(t)) = -mg$$

$$\Leftrightarrow m \cdot z''(t) = -mg$$

$$\Leftrightarrow z''(t) = -g, \quad \forall t \in [t_1, t_2].$$

the \downarrow vertical
acceleration, as
expected.

So, we have constant horizontal velocity, and vertical acceleration $-g$.

→ Lagrange multipliers:
Finding minima and maxima under constraints:

We know that if $f: \mathbb{R} \rightarrow \mathbb{R}$ is differentiable,
and we want to find $x_0 \in \mathbb{R}$ s.t. $f(x_0)$ is a
local minimum or maximum of f ,

we can first solve $f'(x_0) = 0$; because,
if f has a local ^{or global!} min or max at x_0 , then $f'(x_0) = 0$.

Similarly, if all partial derivatives of

$f: \mathbb{R}^n \rightarrow \mathbb{R}$ are well-defined,
 $(x_1, \dots, x_n) \rightarrow f(x_1, \dots, x_n) \in \mathbb{R}$

it holds that:

if f has a local ^{or global!} min or max at
 $x_0 = (x_1, \dots, x_n)$, then $\frac{\partial f}{\partial x_1}(x_0) = \frac{\partial f}{\partial x_2}(x_0) = \dots = \frac{\partial f}{\partial x_n}(x_0) = 0$;

thus, to find x_0 , we first solve the equations.
 $\frac{\partial f}{\partial x_1} = \dots = \frac{\partial f}{\partial x_n} = 0$.

! But what if we impose an extra constraint, such as:
 x_0 also lies on a specific curve?

A method to attack the problem, that can simplify calculations dramatically

(compared to, for instance, doing the obvious thing of restricting f on the curve/surface, and taking partial derivatives equal to 0)

is the method of Lagrange multipliers:

Lagrange multipliers:

→ Let us first consider the 2-dim case:

We want to find the minimum or maximum of a function

$$f: \mathbb{R}^2 \rightarrow \mathbb{R}$$

$$(x, y) \rightarrow f(x, y) \in \mathbb{R},$$

where x and y are s.t. $\phi(x, y) = \text{constant}$.

Thus, x and y are related to each other;

so, we can see y as a function of x , $y(x)$, according to the constraint.

(technically, we may have to split the curve $\phi(x, y) = \text{constant}$ in pieces for this to be true on each piece ...)

and we need the Jacobian of ϕ to be $\neq 0$ always...

We don't now find the function $y(x)$. We instead use that: (2)

for this function $y(x)$,

- $\phi(x, y(x)) = \text{constant} \Rightarrow \frac{d}{dx} (\phi(x, y(x))) \equiv 0$

$$\Rightarrow \frac{\partial \phi}{\partial x} + \frac{\partial \phi}{\partial y} y'(x) = 0 \quad \forall x, \quad \text{and}$$

$$\Rightarrow \lambda \left(\frac{\partial \phi}{\partial x} + \frac{\partial \phi}{\partial y} y'(x) \right) = 0 \quad \forall x, \quad \forall \lambda \in \mathbb{R}$$

- $f(x, y(x))$ is a function of one variable, so any x_0 at which it takes minimal / maximal values satisfies

$$\left. \frac{d}{dx} (f(x, y(x))) \right|_{x=x_0} = 0$$

$$\Leftrightarrow \frac{\partial f}{\partial x}(x_0, y(x_0)) + \frac{\partial f}{\partial y}(x_0, y(x_0)) \cdot y'(x_0) = 0$$

So, x_0 satisfies both

$$\left. \begin{aligned} \lambda \cdot \frac{\partial \phi}{\partial x}(x_0, y(x_0)) + \lambda \cdot \frac{\partial \phi}{\partial y}(x_0, y(x_0)) \cdot y'(x_0) &= 0 \quad \forall \lambda \in \mathbb{R} \\ \text{and} \quad \frac{\partial f}{\partial x}(x_0, y(x_0)) + \frac{\partial f}{\partial y}(x_0, y(x_0)) \cdot y'(x_0) &= 0. \end{aligned} \right\} \Rightarrow$$

$$\Rightarrow \left[\frac{\partial f}{\partial x}(x_0, y(x_0)) + \lambda \cdot \frac{\partial \phi}{\partial x}(x_0, y(x_0)) \right] + \left[\frac{\partial f}{\partial y}(x_0, y(x_0)) + \lambda \cdot \frac{\partial \phi}{\partial y}(x_0, y(x_0)) \right] \cdot y'(x_0) = 0, \quad \forall \lambda \in \mathbb{R} \quad (*)$$

③

In particular, this holds for λ st.

$$\frac{\partial f}{\partial y}(x_0, y(x_0)) + \lambda \cdot \frac{\partial \phi}{\partial y}(x_0, y(x_0)) = 0 \quad (1)$$

(we don't find this λ at this point, even though we could)

For such λ ,

$$\circledast \rightarrow \frac{\partial f}{\partial x}(x_0, y(x_0)) + \lambda \cdot \frac{\partial \phi}{\partial x}(x_0, y(x_0)) = 0 \quad (2)$$

And, of course, $\phi(x_0, y(x_0)) = \text{constant}.$ (3)

Thus, the three unknowns $x_0, y(x_0), \lambda$ can be found by solving the system of equations (1), (2), (3) (3 equations, 3 unknowns).

What we need are $x_0, y(x_0)$

(λ helps with calculations, and it may help to find its exact value at some point on the way, but it is not our goal).

This λ is called a Lagrange multiplier.

(4)

We have thus shown that, if f has a min/max
at (x_0, y_0) , under the constraint $\phi(x, y) = \text{const}$,
then the following equations are
satisfied at (x_0, y_0) :

$$\frac{\partial f}{\partial x} + \lambda \frac{\partial \phi}{\partial x} = 0$$

$$\frac{\partial f}{\partial y} + \lambda \frac{\partial \phi}{\partial y} = 0$$

$$\phi(x, y) = \text{const}$$

for some $\lambda \in \mathbb{R}$

(which will also be derived by solving the system of these equations).

Notice that, if f has a min/max (local or global),
at (x_0, y_0) , under the constraint $\phi(x, y) = \text{const}$,
then (x_0, y_0) will be given as a solution to the above
system of equations.

However, if we have found a solution (x, y) to this
system, there is no guarantee that it is a min/max
of f . Further checking is required. And our intuition
can help: for instance, if it is clear that a
min must exist, and we have found a unique

(x, y) satisfying the system of equations above, then this (x, y) must be the element of \mathbb{R}^2 where the minimum is attained.

!
 This can be generalised to any

$$f: \mathbb{R}^n \rightarrow \mathbb{R}, \quad n \geq 2 : \\ (x_1, \dots, x_n) \rightarrow f(x_1, \dots, x_n) \in \mathbb{R}.$$

If f has a min/max at $\vec{x}_0 \in \mathbb{R}^n$, n coordinates under the constraint $\phi(\vec{x}_0) = \text{const.}$, then \vec{x}_0 satisfies the system of equations

$$\frac{\partial f}{\partial x_1} + \lambda \frac{\partial \phi}{\partial x_1} = 0 \\ \frac{\partial f}{\partial x_2} + \lambda \frac{\partial \phi}{\partial x_2} = 0 \\ \vdots \\ \frac{\partial f}{\partial x_n} + \lambda \frac{\partial \phi}{\partial x_n} = 0 \\ \phi(x_1, \dots, x_n) = 0$$

for some $\lambda \in \mathbb{R}$ (which can also be derived by solving this system of equations, even though its value is not pursued a priori).

Proof: (a generalisation of the 2-dim case):

The set of points (x_1, \dots, x_n) that satisfy $\phi(x_1, \dots, x_n) = \text{const}$ form an $(n-1)$ -dim surface, on which x_n is a function of x_1, \dots, x_{n-1} .

For this function $x_n(x_1, \dots, x_{n-1})$, we have:

no need to find it explicitly

• $\phi(x_1, x_2, \dots, x_{n-1}, x_n(x_1, \dots, x_{n-1})) = \text{const}, \forall x_1, x_2, \dots, x_{n-1}$.

Thus:

• $\frac{d}{dx_i} \phi(x_1, \dots, x_{n-1}, x_n(x_1, \dots, x_{n-1})) = 0$
 $\iff \frac{\partial \phi}{\partial x_i}(x_1, \dots, x_{n-1}, x_n(x_1, \dots, x_{n-1})) + \frac{\partial \phi}{\partial x_n}(x_1, \dots, x_{n-1}, x_n(x_1, \dots, x_{n-1})) \cdot \frac{d}{dx_i} x_n(x_1, \dots, x_{n-1}) = 0,$

$\forall i=1, \dots, n-1, \forall x_1, \dots, x_n$

$\implies \lambda \frac{\partial \phi}{\partial x_i}(x_1, \dots, x_{n-1}, x_n(x_1, \dots, x_{n-1})) + \lambda \frac{\partial \phi}{\partial x_n}(x_1, \dots, x_{n-1}, x_n(x_1, \dots, x_{n-1})) \cdot \frac{d}{dx_i} x_n(x_1, \dots, x_{n-1}) = 0,$

$\forall i=1, \dots, n-1, \quad \forall x_1, \dots, x_{n-1} \quad \forall n \in \mathbb{R}.$

• $f(x_1, \dots, x_{n-1}, x_n(x_1, \dots, x_{n-1}))$ is a function of the $n-1$ variables x_1, \dots, x_{n-1} .

So, if it has a min/max at

$\vec{x}_0 = \left(\underbrace{\overrightarrow{x_0^{(n-1)}}}_{\in \mathbb{R}^{n-1}}, \underbrace{x_n(x_0^{(n-1)})}_{\in \mathbb{R}} \right)$, then
the arbitrary element of the surface $\{f(x_1, \dots, x_n) = \text{const}\}$

$\frac{df}{dx_i}(\vec{x}_0) = 0, \quad \forall i=1, \dots, n-1, \quad \text{ie.}$

$\frac{\partial f}{\partial x_i}(\vec{x}_0) + \frac{\partial f}{\partial x_n}(\vec{x}_0) \frac{d}{dx_i} x_n(x_0^{(n-1)}) = 0, \quad \forall i=1, \dots, n-1.$

So, \vec{x}_0 satisfies both

②

$$\lambda \frac{\partial \phi}{\partial x_i}(\vec{x}_0) + \lambda \frac{\partial \phi}{\partial x_n}(\vec{x}_0) \cdot \frac{d}{dx_i} x_n(\vec{x}_0^{(n-1)}), \quad \forall \lambda \in \mathbb{R},$$

and

$$\frac{\partial f}{\partial x_i}(\vec{x}_0) + \frac{\partial f}{\partial x_n}(\vec{x}_0) \cdot \frac{d}{dx_i} x_n(\vec{x}_0^{(n-1)}),$$

$\forall i=1, \dots, n-1$

$$\Rightarrow \left[\frac{\partial f}{\partial x_i}(\vec{x}_0) + \lambda \frac{\partial \phi}{\partial x_i}(\vec{x}_0) \right] + \left[\frac{\partial f}{\partial x_n}(\vec{x}_0) + \lambda \frac{\partial \phi}{\partial x_n}(\vec{x}_0) \right] \cdot \frac{d}{dx_i} x_n(\vec{x}_0^{(n-1)}),$$

$\forall \lambda \in \mathbb{R},$
 $\forall i=1, \dots, n-1.$

In particular, for λ such that

$$\frac{\partial f}{\partial x_n}(\vec{x}_0) + \lambda \cdot \frac{\partial \phi}{\partial x_n}(\vec{x}_0) = 0,$$

we have: $\frac{\partial f}{\partial x_i}(\vec{x}_0) + \lambda \cdot \frac{\partial \phi}{\partial x_i}(\vec{x}_0) = 0, \quad \forall i=1, \dots, n-1.$

Thus, \vec{x}_0 satisfies the system of $n+1$ equations

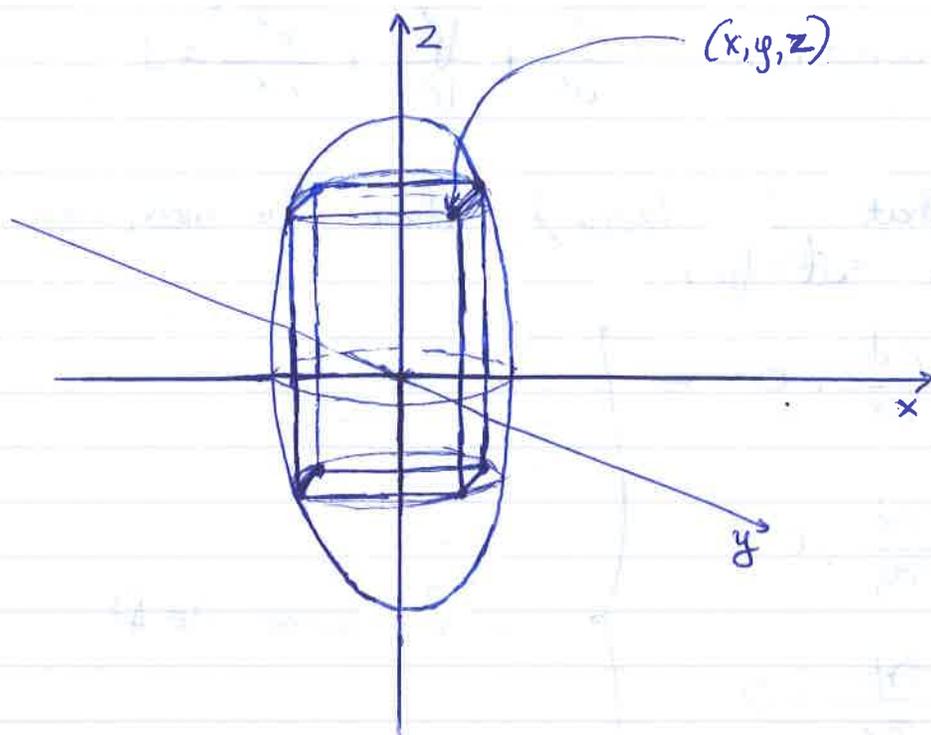
\downarrow n coordinates

$$\left\{ \begin{array}{l} \frac{\partial f}{\partial x_1} + \lambda \frac{\partial \phi}{\partial x_1} = 0 \\ \vdots \\ \frac{\partial f}{\partial x_{n-1}} + \lambda \frac{\partial \phi}{\partial x_{n-1}} = 0 \\ \frac{\partial f}{\partial x_n} + \lambda \frac{\partial \phi}{\partial x_n} = 0 \\ \phi(x_1, \dots, x_n) = \text{const.} \end{array} \right. \quad \left(\begin{array}{l} \text{from this system, we find the } n \\ \text{coordinates of } \vec{x}_0 \text{ (which we needed)} \\ \text{and the } \lambda \text{ (which is not needed,} \\ \text{but helps us with calculations).} \end{array} \right)$$

■

(9)

→ Application:



Find the volume of the largest box, with edges parallel to the axes, that has vertices on

the ellipsoid $\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1$.

Solution: If the vertex shown in the picture is the point (x, y, z) , then the box has volume

$$f(x, y, z) = \underbrace{(2x)}_{\text{length}} \cdot \underbrace{(2y)}_{\text{width}} \cdot \underbrace{(2z)}_{\text{height}} = 8xyz.$$

We thus want to find the maximum of $f(x,y,z)$

under the constraint $\phi(x,y,z) = \frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1$

We know that the (x,y,z) where the maximum is attained satisfies

System of equations: $\frac{\partial f}{\partial x} + \lambda \frac{\partial \phi}{\partial x} = 0$, $\frac{\partial f}{\partial y} + \lambda \frac{\partial \phi}{\partial y} = 0$, $\frac{\partial f}{\partial z} + \lambda \frac{\partial \phi}{\partial z} = 0$, and $\phi(x,y,z) = 1$, for some $\lambda \in \mathbb{R}$

Equivalent system: $\begin{cases} 2yz + \lambda \cdot \frac{2x}{a^2} = 0 & (1) \\ 2xz + \lambda \cdot \frac{2y}{b^2} = 0 & (2) \\ 2xy + \lambda \cdot \frac{2z}{c^2} = 0 & (3) \\ \frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1 & (4) \end{cases}$

(11)



$$\left. \begin{aligned} x \times \textcircled{1} &\Rightarrow 8xyz + \lambda \cdot \frac{2x^2}{a^2} = 0 \\ y \times \textcircled{2} &\Rightarrow 8xyz + \lambda \cdot \frac{2y^2}{b^2} = 0 \\ z \times \textcircled{3} &\Rightarrow 8xyz + \lambda \cdot \frac{2z^2}{c^2} = 0 \end{aligned} \right\} \Rightarrow$$

$$\Rightarrow 24xyz + 2\lambda \cdot \left(\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} \right) = 0 \quad \textcircled{4}$$

$$\Rightarrow 24xyz + 2\lambda \cdot 1 = 0 \Rightarrow \boxed{\lambda = -12xyz} \quad \textcircled{5}$$

Now:

$$\textcircled{1} \textcircled{5} \Rightarrow 8yz - 12xyz \cdot \frac{2x}{a^2} = 0 \Rightarrow yz \left(1 - 3 \frac{x^2}{a^2} \right) = 0$$

$$\xrightarrow{\substack{y \neq 0 \\ z \neq 0}} 1 - 3 \frac{x^2}{a^2} = 0 \Rightarrow x^2 = \frac{1}{3} a^2 \xrightarrow{x > 0} x = \frac{a}{\sqrt{3}}$$

or else we'd
be finding volume
0, when it is clear
that there exist
larger parallelepipeds
in the ellipsoid

$$\text{Similarly, } \textcircled{2} \textcircled{5} \Rightarrow y = \frac{b}{\sqrt{3}}, \quad \textcircled{3} \textcircled{5} \Rightarrow z = \frac{c}{\sqrt{3}}$$

$$\text{So, largest volume} = 8 \cdot \frac{a}{\sqrt{3}} \cdot \frac{b}{\sqrt{3}} \cdot \frac{c}{\sqrt{3}} = \frac{8abc}{3\sqrt{3}}$$

⚠ Note that the (x,y,z) we have found through this method could have been a minimum instead, or a saddle point... However, our intuition tells us that there must exist a parallelepiped, confined in the ellipsoid, with maximal volume $f(x,y,z)$. So, the unique (x,y,z) we found has to be the point where the maximum is attained.

Lecture 39:

①

→ Application of Lagrange multipliers:

Isoperimetric problems.

Isoperimetric problems ask the following:

find the graph $(x, y(x)), x \in [x_1, x_2]$,

that minimises (or maximises, or, in general, makes stationary) the integral

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

under the condition

$$\textcircled{*} \quad \mathcal{I} = \int_{x_1}^{x_2} G(x, y(x), y'(x)) dx = c, \text{ for some constant } c.$$

These problems are called isoperimetric because, historically, the condition $\textcircled{*}$ that $(x, y(x)), x \in [x_1, x_2]$

had to satisfy was that its length should

be equal to c (i.e., $\int_{x_1}^{x_2} \underbrace{\sqrt{1 + (y'(x))^2}}_{G(x, y(x), y'(x))} dx = c$).

Note that the desired graph doesn't have to be a minimiser of I , meaning that there may exist other graphs, that doesn't satisfy $\textcircled{*}$, which make

(9)

I even smaller. So, the desired y doesn't have to satisfy the Euler-Lagrange equation for F ! We will show, however, that:

The desired $y: [x_1, x_2] \rightarrow \mathbb{R}$ satisfies the Euler-Lagrange equation for $F + \lambda G$:

for some specific λ , that can be derived by condition $(*)$ for y .

Proof:

Suppose we have found the desired minimizer $y: [x_1, x_2] \rightarrow \mathbb{R}$.

We perturb y by two ^{fixed} $\eta_1, \eta_2: [x_1, x_2] \rightarrow \mathbb{R}$, twice cont. differentiable, with $\eta_1(x_1) = \eta_1(x_2) = \eta_2(x_1) = \eta_2(x_2) = 0$:

$$y + \varepsilon_1 \eta_1 + \varepsilon_2 \eta_2 : [x_1, x_2] \rightarrow \mathbb{R}$$

is a perturbation of y , $\forall \varepsilon_1, \varepsilon_2 \in \mathbb{R}$.

The idea now is that, when $\varepsilon_1 = \varepsilon_2 = 0$ (i.e., before we start perturbing), $y + \varepsilon_1 \eta_1 + \varepsilon_2 \eta_2 = y$, which satisfies $(*)$, and also minimises I under $(*)$.

(3)

So, amongst all (ϵ_1, ϵ_2) with the property that

$$\underbrace{y + \epsilon_1 \eta_1 + \epsilon_2 \eta_2}_{\substack{\text{perturbation} \\ \text{of } y}} \text{ satisfies } (*),$$

it is $(\epsilon_1, \epsilon_2) = (0, 0)$ that minimises I .

Thus, the point $(\epsilon_1, \epsilon_2) = (0, 0)$ is a minimiser of the quantity

$$I_y^{n_1, n_2}(\epsilon_1, \epsilon_2) := \int_{x_1}^{x_2} f\left(x, \underbrace{y(x) + \epsilon_1 \eta_1(x) + \epsilon_2 \eta_2(x)}_{y + \epsilon_1 \eta_1 + \epsilon_2 \eta_2}, (y(x) + \epsilon_1 \eta_1(x) + \epsilon_2 \eta_2(x))'\right) dx,$$

under the condition

$$(*)': \quad \int_{x_1}^{x_2} G\left(x, \underbrace{y(x) + \epsilon_1 \eta_1(x) + \epsilon_2 \eta_2(x)}_{y + \epsilon_1 \eta_1 + \epsilon_2 \eta_2}, (y(x) + \epsilon_1 \eta_1(x) + \epsilon_2 \eta_2(x))'\right) dx = c.$$

Thus, for these fixed (but arbitrary) n_1, n_2 , the point $(\epsilon_1, \epsilon_2) = (0, 0)$ satisfies the system of equations

$$\left. \begin{array}{l} (1) \quad \frac{\partial}{\partial \epsilon_1} I_y^{n_1, n_2} + \lambda \frac{\partial}{\partial \epsilon_1} \int_{x_1}^{x_2} G = 0 \\ (2) \quad \frac{\partial}{\partial \epsilon_2} I_y^{n_1, n_2} + \lambda \frac{\partial}{\partial \epsilon_2} \int_{x_1}^{x_2} G = 0 \\ (3) \quad \int_{x_1}^{x_2} G = c \end{array} \right\}, \quad \text{for an appropriate Lagrange multiplier } \lambda \text{ (that is derived as a solution of this system of equations).}$$

Amongst these 3 equations, we only use (1):

$$\frac{\partial}{\partial \epsilon_1} I_y^{n_1, n_2} \Big|_{\epsilon_1 = \epsilon_2 = 0} + \lambda \frac{\partial}{\partial \epsilon_2} I_y^{n_1, n_2} \Big|_{\epsilon_1 = \epsilon_2 = 0} = 0,$$

for the λ above.

Now, we remember our old notation of I , with input a perturbation of y by a single n :

$$I_y^\epsilon(n) = \int_x^{x_2} F(x, y + \epsilon n, (y + \epsilon n)') dx.$$

Note that
$$\frac{\partial}{\partial \epsilon_1} I_y^{n_1, n_2} \Big|_{\epsilon_1 = \epsilon_2 = 0} = \frac{d}{d\epsilon_1} I_y^{n_1}(\epsilon_1) \Big|_{\epsilon_1 = 0}$$

Similarly, we define

$$I_y^\epsilon(n) := \int_{x_1}^{x_2} G(x, y + \epsilon n, (y + \epsilon n)') dx,$$

and we notice that

$$\frac{\partial}{\partial \epsilon_2} I_y^{n_1, n_2} \Big|_{\epsilon_1 = \epsilon_2 = 0} = \frac{d}{d\epsilon_2} I_y^{n_2}(\epsilon_2) \Big|_{\epsilon_2 = 0}.$$

thus:

$$(1) \Rightarrow \frac{d}{d\epsilon_1} \left(I_y^{n_1} + \lambda I_y^{n_2} \right) (\epsilon_1) \Big|_{\epsilon_1 = 0} = 0, \text{ for the } \lambda \text{ above.}$$

And this holds for the arbitrary $n_i: [x_1, x_2] \rightarrow \mathbb{R}$, twice cont. diff., with $n_1(x_1) = n_1(x_2) = 0$.

(5)

Thus, y makes stationary the integral

$$\int_{x_1}^{x_2} (F + \lambda G)(x, y(x), y'(x)) dx !$$

So, y satisfies the Euler-Lagrange equation for $F + \lambda G$:

$$\frac{d}{dx} \left(\frac{\partial}{\partial y'} (F + \lambda G) \right) - \frac{\partial}{\partial y} (F + \lambda G) = 0, \text{ for the}$$

So, to find the desired $y: [x_1, x_2] \rightarrow \mathbb{R}$, that minimises I under the condition $(*)$, we just solve the Euler-Lagrange equation for $F + \lambda G$ (for λ not specified).

This gives $y: [x_1, x_2] \rightarrow \mathbb{R}$, that depends on the Lagrange multiplier λ .

To find the value of λ , we now use the fact that this explicit y satisfies $(*)$.

6

We thus plug in this particular y in

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

and solve for A .

This way, the formula for the desired y is complete.

See p. 492, Example 1 (Textbook) for a demonstration of the above technique.