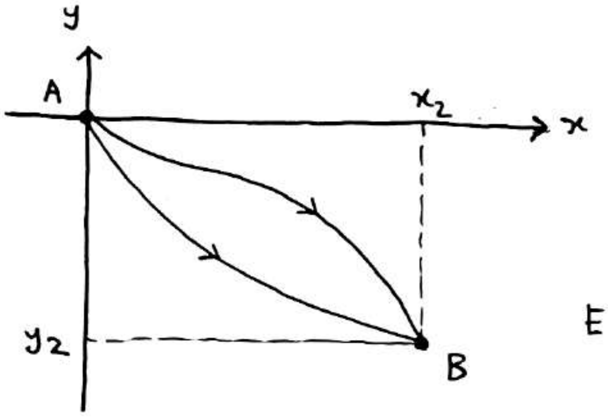
Variational Principles – Lecture 1 21/6/21

Section 0 Motivation

Example 0.1 : Brachistochrone problem (Johann Bernoulli, 1696)

Particle slides on wire under influence of gravity, between 2 fixed points, A and B. What shape should the wire be for the shortest travel time, starting from rest?



Travel time

$$T = \int dt = \int_A^B \frac{dl}{v(x,y)}$$

E conserved: $T + V = \text{const.}$

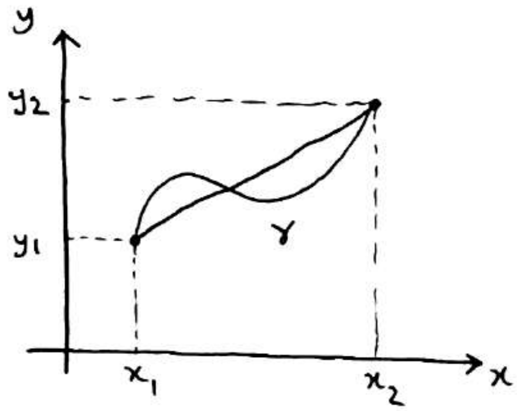
$$\frac{1}{2}mv^2 + mgy = \overset{\text{at A: } y_1 = 0}{mgy_1} = 0 \qquad \text{so} \qquad v = \sqrt{2g}\sqrt{-y}$$

So minimise
$$T(y) = \frac{1}{\sqrt{2g}} \int_0^{x_2} \frac{\sqrt{1+(y')^2}}{\sqrt{-y}}\, dx$$
(line elt. dl)

subject to $y(0) = 0$, $y(x_2) = y_2$.

Example 0.2 : Geodesics (shortest path γ between 2 points on a surface Σ, if they exist)

Take $\Sigma = \mathbb{R}^2$ (a plane, Pythagorean thm holds)



Distance along γ:

$$D(y) = \int_A^B dl = \int_{x_1}^{x_2} \sqrt{1+(y')^2}\, dx$$

Minimise D by varying path.

In general we want to minimise / maximise some function

$$F(y) = \int_{x_1}^{x_2} f(x, y(x), y'(x)) \, dx \qquad (0.1)$$

among all functions with $y(x_1) = y_1$, $y(x_2) = y_2$

Note $y = y(x)$.

(0.1) is a functional (a function on the space of functions).

Functions: numbers $\rightarrow$ numbers

Functionals: functions $\rightarrow$ numbers (example: area under graph)

Calculus of variations: finding extrema of functionals on spaces of functions.

Notation $C(\mathbb{R})$ space of continuous functions on $\mathbb{R}$

functions on R with

$C^k(\mathbb{R})$ space of continuous k^{th}-derivatives

$C^k_{(\alpha,\beta)}(\mathbb{R})$ space of continuous k^{th}-derivatives s.t. $f(\alpha) = f(\beta)$

functions on ^
R with

Need to specify the function space beforehand.

Example 0.3 Fermat's principle

Light between two points travels along paths which require the least time.

Example 0.4 Principle of least (stationary) action

$$S(\gamma) = \int_{t_1}^{t_2} (T - V) \, dt \qquad \text{for motion of particle}$$

(e.g. $m\ddot{\underline{x}} = -\nabla V$ so Newton's eqns should follow)

Variational Principles - Lecture 2

Section 1 Calculus for Functions on $\mathbb{R}^n$

$f \in C^2(\mathbb{R}^n)$, $\quad f: \mathbb{R}^n \to \mathbb{R}$ (cts 2nd derivatives)

The point $\underline{a} \in \mathbb{R}^n$ is stationary if $\underline{\nabla f(\underline{a}) = \underline{0}}$

Expand near $\underline{x} = \underline{a}$:

$$f(\underline{x}) = f(\underline{a}) + \underbrace{(\underline{x} - \underline{a}) \cdot \nabla f|_{\underline{a}}}_{\substack{0 \text{ as } \underline{a} \text{ stationary} \\ \nabla f|_{\underline{a}} = \underline{0}}} + \frac{1}{2}(x_i - a_i)(x_j - a_j)\, \partial^2_{ij} f|_{\underline{a}} + o(|\underline{x} - \underline{a}|^2)$$

The Hessian matrix is given by $\underline{H_{ij} = \partial_i \partial_j f = H_{ji}}$

Shift origin to set $\underline{a} = \underline{0}$. Diagonalise $H(\underline{0})$ by an orthogonal transformation.

$$H' = R^T H(\underline{0}) R = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$$

$$f(\underline{x}') - f(\underline{0}) = \frac{1}{2} \sum_i \lambda_i (x_i')^2 + o(|\underline{x}'|^2)$$

(note 1st order term is 0 as $\underline{0}$ is stationary)

(i) If all $\underline{\lambda_i > 0}$ then $f(\underline{x}') > f(\underline{0})$ in all directions
local minimum

(ii) all $\underline{\lambda_i < 0}$: local maximum

(iii) some $\underline{\lambda_i > 0}$, some $\underline{\lambda_i < 0}$: saddle point

(iv) some $\underline{\lambda_i = 0}$: need higher order terms.

Special case $n = 2$ (may be easier than finding evals)

$\det(H) = \lambda_1 \lambda_2$ $\quad\quad$ $\operatorname{tr}(H) = \lambda_1 + \lambda_2$

- $\det H > 0$ and $\operatorname{tr} H > 0$: local minimum
- $\det H > 0$ and $\operatorname{tr} H < 0$: local maximum
- $\det H < 0$: saddle point
- $\det H = 0$: higher order derivatives

Remarks If we have $f : D \to \mathbb{R}$ where $D \subseteq \mathbb{R}^n$ then may have extremum on boundary of domain

Here we cannot find the global max/min via derivatives.

Remark 2 If f is harmonic on $\mathbb{R}^2$: $f_{xx} + f_{yy} = 0$
with f defined on $D \subseteq \mathbb{R}^2$

Then anywhere in D, $\operatorname{tr} H = 0$ so critical points must be saddle points and the min/max is on the boundary (see reason above).

Example 1.1 $f(x,y) = x^3 + y^3 - 3xy$

$\nabla f = (3x^2 - 3y,\ 3y^2 - 3x) = (0,0)$ at critical pt

$x^2 - y = 0, \quad y^2 - x = 0 \quad \Rightarrow \quad y^4 = y$

So critical points are $(0,0)$ and $(1,1)$.

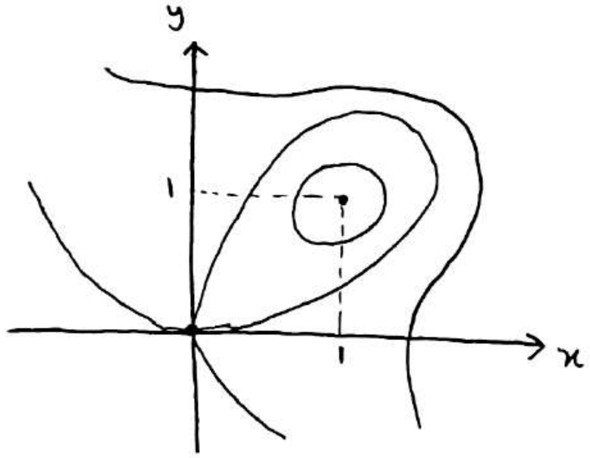
$$H = \begin{pmatrix} 6x & -3 \\ -3 & 6y \end{pmatrix}$$

At $(0,0)$ $\det H = -9 < 0$ $\operatorname{tr} H = 0$

so have a saddle point, where $f = 0$.

At $(1,1)$ $\det H = 27 > 0$ $\operatorname{tr} H = 12 > 0$

so have a local minimum, where $f = -1$.

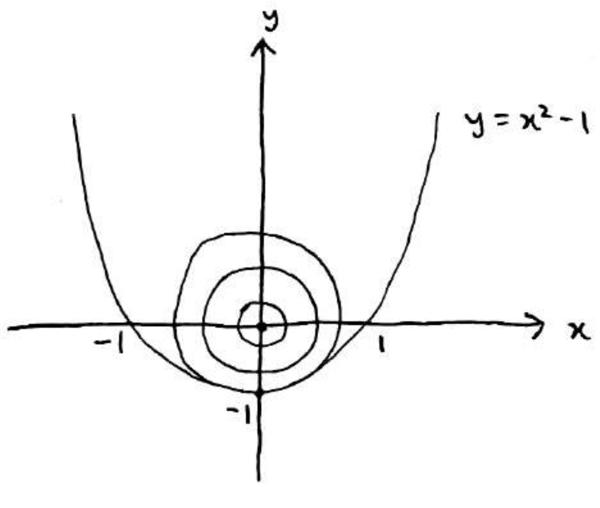


Near $(0,0)$ $f \approx -3xy$
as x^3, y^3 small
so decreases on $y = x$
and increases on $y = -x$
(no global min / max)

(only a rough sketch)

Section 1.1 Constraints and Lagrange multipliers

Example 1.2 Find the circle centred at $(0,0)$ with smallest radius, which intersects the parabola $y = x^2 - 1$.



Direct method: solve the constraints

Minimise $x^2 + y^2$ subject to $y = x^2 - 1$

so minimise

$$x^2 + (x^2-1)^2 = x^4 - x^2 + 1$$

$$\Rightarrow 4x^3 - 2x = 0$$

2 solns: $x = \pm \frac{1}{\sqrt{2}}$, $y = -\frac{1}{2}$ $\Rightarrow$ radius is $\frac{\sqrt{3}}{2}$

$x = 0$, $y = -1$ $\Rightarrow$ radius is 1

so $\sqrt{3}/2$ is smallest.

But what if we can't solve analytically?

Lagrange Multipliers

Define new function $\underline{h(x, y, \lambda) = f(x,y) - \lambda g(x, y)}$

where f is the function we want to min / max

and $\underline{g(x, y) = 0}$ is the constraint. λ is Lagrange multiplier.

Here $h = x^2 + y^2 - \lambda(y - x^2 + 1)$

Then <u>extremise</u> over <u>3</u> variables with no constraint.

$$\frac{\partial h}{\partial x} = 2x + 2\lambda x = 0 \qquad \frac{\partial h}{\partial y} = 2y - \lambda = 0 \qquad \frac{\partial h}{\partial \lambda} = y - x^2 + 1 = 0$$

(constraint)

Combining first 2: $2x + 4xy = 0 \Rightarrow x = 0$ or $y = -\frac{1}{2}$

so using $\frac{\partial h}{\partial \lambda}$ have either $x = 0,\ y = -1$

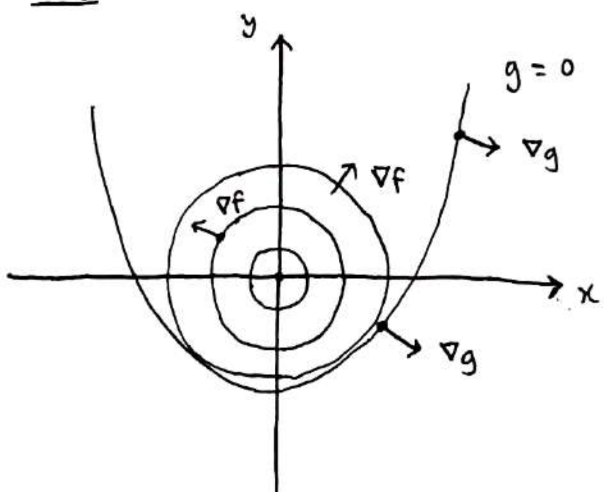
or $x = \pm\frac{1}{\sqrt{2}},\ y = -\frac{1}{2}$ as before.

$(0, -1) \rightarrow f = 1 \quad (\lambda = -2)$

$\left(\pm\frac{1}{\sqrt{2}}, -\frac{1}{2}\right) \rightarrow f = \frac{3}{4} \quad (\lambda = -1)$

Finding the critical points without solving analytically.

<u>Why</u> does this work?



∇g perpendicular to "surface" $g = 0$.

Representing circles by $f = \text{constant}$,

have also ∇f perpendicular to $f = \text{constant}$.

At the extremum, ∇f and ∇g are <u>parallel</u>

i.e. $\nabla f = \lambda \nabla g$ i.e. $\underline{\nabla(f - \lambda g) = \underline{0}}$

so we consider the extrema of

$\underline{h = f - \lambda g}$ to find the solution(s).

For multiple constraints:

Extremise $f : \mathbb{R}^n \to \mathbb{R}$ subject to $g_\alpha(\underline{x}) = 0$

$(g_\alpha : \mathbb{R}^n \to \mathbb{R}, \ \alpha = 1, \dots, k)$

Then define $\underline{h(x_1, \dots, x_n, \lambda_1, \dots, \lambda_k) = f - \sum_{\alpha=1}^{k} \lambda_\alpha g_\alpha}$

a function of $n + k$ vars, with k Lagrange multipliers.

So we work with

$\frac{\partial h}{\partial x_i} = 0$, $\frac{\partial h}{\partial \lambda_\alpha} = 0$, eliminate λ_α and solve for $\underline{x}$.

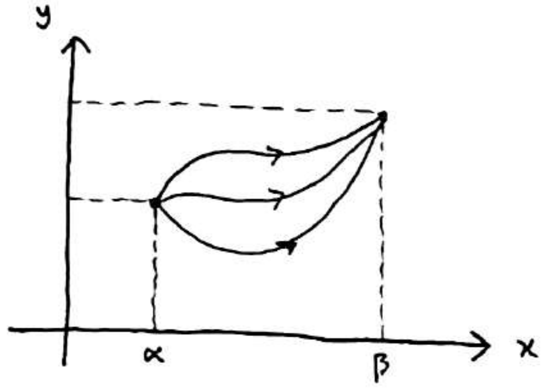
The method works also if the constraints cannot be eliminated algebraically.

Variational Principles - Lecture 3

22/6/21

Section 2 Euler-Lagrange equations

Extremise functional 0.1
$$F(y) = \int_\alpha^\beta f(x, y, y')\,dx \qquad (2.1)$$



f is given, α, β are fixed
Functional depends on y.

Small perturbation $y \to y + \varepsilon\eta(x)$ in (2.1)
Compute $\underline{F(y + \varepsilon\eta(x))}$ with $\eta(\alpha) = \eta(\beta) = 0$
(so perturbed function goes through <u>same</u> fixed pts)

<u>Lemma</u> If $g : [\alpha, \beta] \to \mathbb{R}$ is continuous on $[\alpha, \beta]$ <u>and</u>
$\int_\alpha^\beta g(x)\eta(x)\,dx = 0$ for <u>all</u> continuous $\eta(x)$ on $[\alpha, \beta]$
s.t. $\eta(\alpha) = \eta(\beta) = 0$, then $g(x) = 0 \quad \forall x \in [\alpha, \beta]$.

<u>Proof</u> Suppose $\exists\, \bar{x} \in (\alpha, \beta)$ s.t. $g(\bar{x}) \neq 0$. WLOG suppose that $g(\bar{x}) > 0$. Then $\exists$ interval $[x_1, x_2] \subseteq (\alpha, \beta)$ s.t. $g(x) > c$ on $[x_1, x_2]$ for some $c > 0$.

Set
$$\eta(x) = \begin{cases} (x - x_1)(x_2 - x) & x \in [x_1, x_2] \\ 0 & \text{otherwise} \end{cases} \qquad (2.2)$$

Then $\eta(x)$ is continuous and we see $\int_{x_1}^{x_2} g(x)\eta(x)\,dx > 0$
so $\int_\alpha^\beta g(x)\eta(x)\,dx > 0$ (as η is 0 elsewhere). ⋇

(This was on IA Analysis I ES4).

Remark η is a bump function (in 2.2)

A general form for C^k bump functions ($x \in [x_1, x_2]$ bump interval)

is
$$\eta = \begin{cases} \left((x-x_1)(x_2-x)\right)^{k+1} & x \in [x_1, x_2] \\ 0 & \text{otherwise} \end{cases}$$

Now back to 2.1. Expand in ε:

$$F(y+\varepsilon\eta) = \int_\alpha^\beta f(x, y+\varepsilon\eta, y'+\varepsilon\eta')\,dx$$

$$= F(y) + \varepsilon \int_\alpha^\beta \left(\frac{\partial f}{\partial y}\eta + \frac{\partial f}{\partial y'}\eta' \right) dx + O(\varepsilon^2)$$

(Note: we get this by expanding the $y+\varepsilon\eta$ and $y'+\varepsilon\eta'$ as these are the only dependencies of f that depend on ε. Can check this works and the $F(y)$ comes from the 1st order term)

At extremum, have $F(y+\varepsilon\eta) = F(y) + O(\varepsilon^2)$

i.e. $\left.\dfrac{df}{d\varepsilon}\right|_{\varepsilon=0} = 0$. (1st derivative term must be 0 at extremum).

Integrating by parts on the ε-term, we want

$$0 = \int_\alpha^\beta \left(\frac{df}{dy}\eta - \frac{d}{dx}\left(\frac{\partial f}{\partial y'}\right)\eta \right) dx + \cancel{\left[\frac{\partial f}{\partial y'}\eta \right]_\alpha^\beta} \to 0$$

($\eta(\alpha) = \eta(\beta) = 0$)

$$= \int_\alpha^\beta \underbrace{\left(\frac{\partial f}{\partial y} - \frac{d}{dx}\left(\frac{\partial f}{\partial y'}\right)\right)}_{\text{apply lemma with this } \underline{0}} \eta\, dx$$

as $\eta(\alpha) = \eta(\beta) = 0$.

So $$\boxed{\frac{d}{dx}\left(\frac{\partial f}{\partial y'}\right) = \frac{\partial f}{\partial y}} \quad (2.3)$$

(necessary condition for extremum)

Euler-Lagrange equation

Remarks

- (2.3) is a 2nd order ODE for $y(x)$ with boundary conditions $y(\alpha) = y_1, \quad y(\beta) = y_2$.

- $\dfrac{d}{dx}\left(\dfrac{\partial f}{\partial y'}\right) - \dfrac{\partial f}{\partial y}$ is called a functional derivative

  denoted $\dfrac{\delta F(y)}{\delta y(x)}$

- Sometimes $\delta y = \varepsilon \eta(x)$ is written, called a "small variation" and they write $F(y + \delta y) = F(y) + \delta F(y)$ where
$$\delta F = \int_\alpha^\beta \left(\frac{\delta F(y)}{\delta y(x)} \, \delta y(x) \right) dx .$$

- Other boundary conditions are possible e.g. $\left.\dfrac{\partial f}{\partial y'}\right|_{\alpha, \beta} = 0$

- We consider x, y, y' to be independent vars when taking partial derivatives.

  Total derivative: $\dfrac{dh}{dx} = \dfrac{\partial h}{\partial x} + \dfrac{\partial h}{\partial y}\dfrac{dy}{dx} + \dfrac{\partial h}{\partial y'}\dfrac{d^2 y}{dx^2}$

Section 2.1 First integrals of the Euler-Lagrange equation

In some cases 2.3 (E-L eqn) can be integrated once to get a 1st order ODE - "first integral"

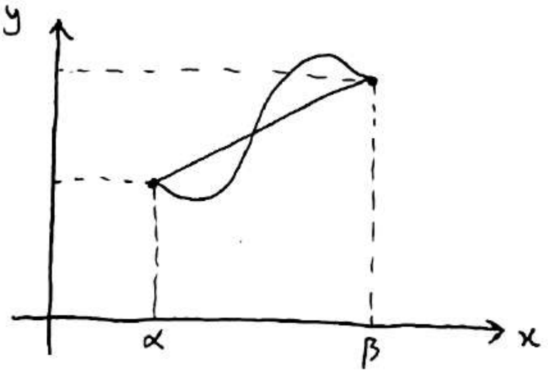
Cases:

(a) f does not explicitly depend on y, so $\dfrac{\partial f}{\partial y} = 0$

Then we get $\dfrac{d}{dx}\left(\dfrac{\partial f}{\partial y'}\right) = 0$ i.e. $\dfrac{\partial f}{\partial y'} = c$ (constant)

(2.4)

Example 2.1 Geodesics on Euclidean plane



$$F(y) = \int_\alpha^\beta \sqrt{dx^2+dy^2}$$

arc length functional

so $$F(y) = \int_\alpha^\beta \underbrace{\sqrt{1+(y')^2}}\, dx$$

$f(y')$ doesn't depend on y explicitly

So use 2.4: $\frac{\partial f}{\partial y} = 0$ so $\frac{y'}{\sqrt{1+(y')^2}}$ = constant

so $(y')^2 = c + c(y')^2$ for some c

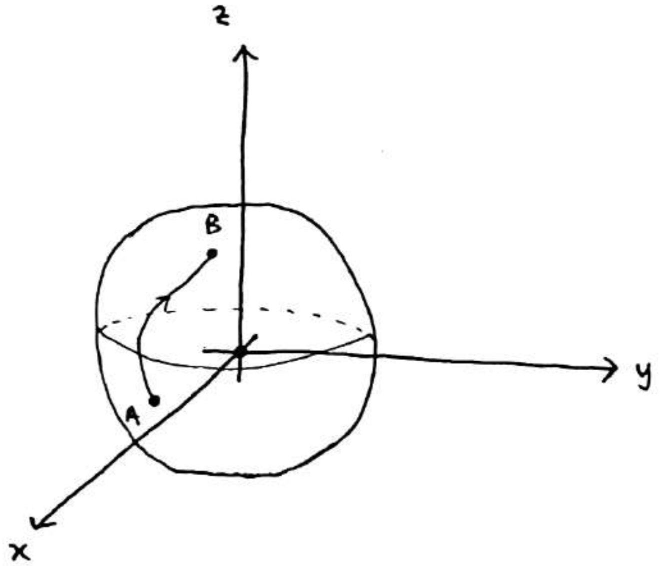
$\Rightarrow \quad (y')^2 = \frac{c}{1-c} \quad \Rightarrow \quad y' = \pm\sqrt{\frac{c}{1-c}}$

So y' must be constant so have $y' = m$ for some m

$\Rightarrow$ $y = mx + k$ a straight line.

Variational Principles - Lecture 4   22/6/21

Example 2.2 Geodesics on a sphere $S^2 \subseteq \mathbb{R}^3$



Parametrise sphere:

$x = \sin\theta \sin\phi$

$y = \sin\theta \cos\phi$

$z = \cos\theta$

$0 \leq \theta \leq \pi, \quad 0 \leq \phi \leq 2\pi$

take as unit sphere

$$ds^2 = dx^2 + dy^2 + dz^2 = d\theta^2 + \sin^2\theta \, d\phi^2$$

Then we parametrise as $\underline{\phi = \phi(\theta)}$ and can write

$$ds = \sqrt{1 + \sin^2\theta (\phi')^2} \, d\theta$$

Then $$F(\phi) = \int_{\theta_1 = \alpha}^{\theta_2 = \beta} \sqrt{1 + \sin^2\theta (\phi')^2} \, d\theta$$

Integrand $f = f(\theta, \phi')$ doesn't depend on ϕ itself

$$\frac{\partial f}{\partial \phi} = 0 \implies \frac{\partial f}{\partial \phi'} = K \quad \text{(constant)}$$

Evaluating $\frac{\partial f}{\partial \phi'}$ we have

$$\frac{\phi' \sin^2\theta}{\sqrt{1 + \sin^2\theta (\phi')^2}} = K$$ We square and solve for $(\phi')^2$:

$$(\phi')^2 = \frac{K^2}{\sin^2\theta (\sin^2\theta - K^2)}$$

So $\phi = \pm \int \frac{K}{\sin\theta \sqrt{\sin^2\theta - K^2}} \, d\theta$ (2 solns, each going one way around sphere)

Make substitution $u = \cot\theta$

Then get $\pm \frac{\sqrt{1-K^2}}{K} \cos(\phi - \phi_0) = \cot\theta$

a <u>great circle</u>. (Geodesics are segments of great circles)

(b) Consider for general $f(x, y, y')$:

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

$$= y' \underbrace{\cancel{\left(\frac{\partial f}{\partial y} - \frac{d}{dx}\frac{\partial f}{\partial y'}\right)}}_{\text{(by E-L)}}^{0} + \frac{\partial f}{\partial x} = \underline{\frac{\partial f}{\partial x}}$$

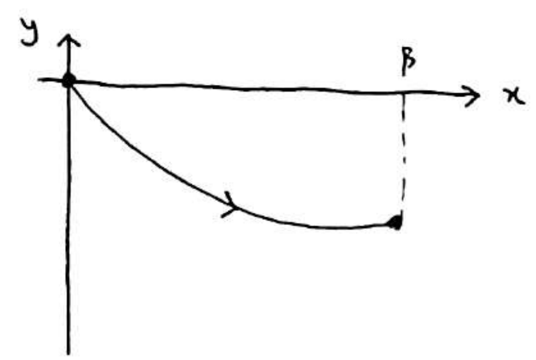
So if f does <u>not</u> explicitly depend on $\underline{x}$, then $\underline{\frac{\partial f}{\partial x} = 0}$

so we have $\frac{d}{dx}\left(f - y' \frac{\partial f}{\partial y'}\right) = 0$

so $\underline{f - y' \frac{\partial f}{\partial y'} = c}$ (2.5)

and we can reduce the order to get a 1st order ODE.

<u>Example 2.3</u> Brachistochrone problem



Recall from section 0 that the functional is

$$F(y) = \frac{1}{\sqrt{2g}} \int_0^\beta \underbrace{\frac{\sqrt{1+(y')^2}}{\sqrt{-y}}}_{\text{depends on } y,\, y' \text{ but not } x \,:\, \text{use } 2.5.} \, dx$$

From $f - y' \frac{\partial f}{\partial y'} = c$ we get

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

Rearranging gives $\frac{1}{\sqrt{1+(y')^2}} = c\sqrt{-y}$

$$\Rightarrow \quad y' = \pm \frac{\sqrt{1+c^2y^2}}{c\sqrt{-y}} \quad \Rightarrow \quad x = \pm c \int \frac{\sqrt{-y}}{\sqrt{1+c^2y}} \, dy$$

Set $y = -\frac{1}{c^2}\sin^2\frac{\theta}{2}$, $dy = -\frac{1}{c^2}\sin\frac{\theta}{2}\cos\frac{\theta}{2}$

Then have

$$x = \pm c \int (-1)\frac{1}{c^2} \frac{\sin^2\frac{\theta}{2}\cos\frac{\theta}{2}}{\sqrt{1-\sin^2\frac{\theta}{2}}} \, d\theta$$

$$= \mp \frac{1}{2c^2} \int (1-\cos\theta) \, d\theta$$

$$= \mp \frac{1}{2c^2}(\theta - \sin\theta)$$ (curve goes through $(0,0)$ so constant of integration is 0)

Then we get $x = \frac{\theta - \sin\theta}{2c^2}$, $y = -\frac{1}{c^2}\sin^2\frac{\theta}{2}$

The parametrisation of a cycloid (curve traced out by a point on a circle as it rolls without slipping)

Section 2.2 Fermat's principle

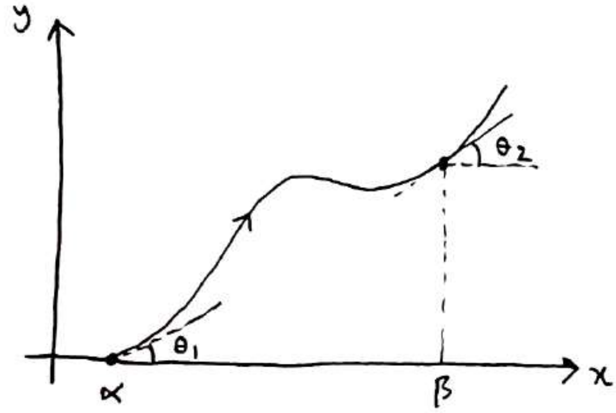
Light / sound travels along the path between two points that takes the least time.

Ray : path $y = y(x)$ speed $c(x, y)$

$$F(y) = \int \frac{dl}{c} = \int_{\alpha}^{\beta} \frac{\sqrt{1+(y')^2}}{c(x,y)}\, dx$$

Assume $\underline{c = c(x)}$ only. Then $\frac{\partial f}{\partial y} = 0 \Rightarrow \frac{\partial f}{\partial y'} = \text{constant}$

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

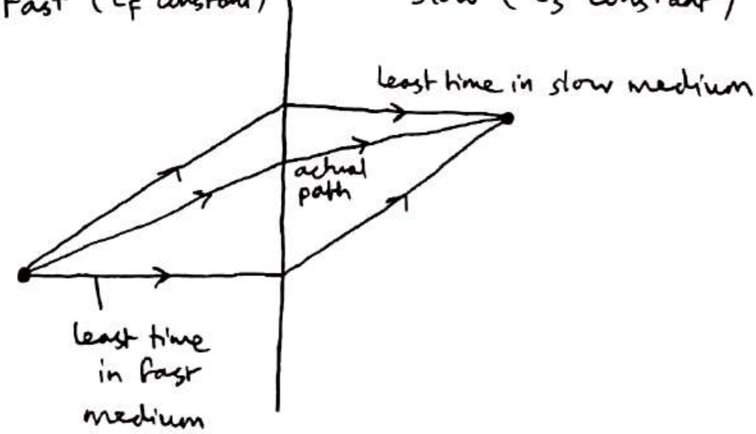


$\tan\theta = y'$ so making this sub, we get

$$\frac{\sin\theta_1}{c(x_1)} = \frac{\sin\theta}{c(x)} \qquad (2.6)$$

Snell's law in optics

Fast (c_F constant) Slow (c_S constant) $c_S < c_F$



Variational Principles - Lecture 5

30/6/21

Section 3 Extensions of the Euler-Lagrange equations

3.1 Euler-Lagrange equations with constraints

Extremise $$F(y) = \int_\alpha^\beta f(x, y, y')\, dx$$ subject to

$$G(y) = \int_\alpha^\beta g(x, y, y')\, dx = k \quad (\text{constant}).$$

Use Lagrange multipliers: extremise

$$\Phi(y; \lambda) = F(y) - \lambda G(y)$$

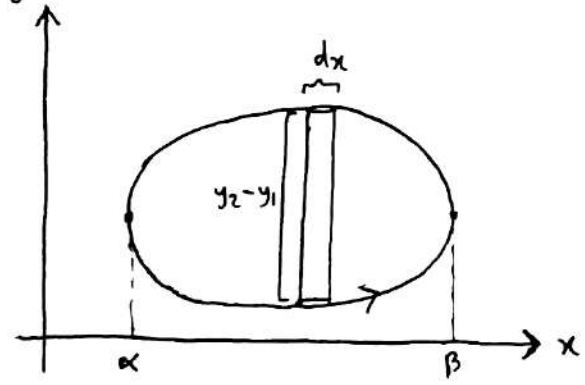
Replace f in E-L eqn by $f - \lambda g$:

$$\frac{d}{dx}\left(\frac{\partial}{\partial y'}(f - \lambda g)\right) = \frac{\partial}{\partial y}(f - \lambda g) = 0 \qquad (3.1)$$

Example 3.1 Dido problem / isoparametric problem

What simple, closed plane curve of fixed length L maximises the enclosed area?

WLOG we assume convexity.



x monotonically increases from α to β, decreases from β to α.

(not coords)

Given x, there exists (y_1, y_2) on the curve with $y_1(x) = y_2(x)$ and $y_2 > y_1$

so $dA = \left[y(x)\right]_{x_1}^{x_2} dx$

Find area functional:

Area functional

$$A(y) = \int_\alpha^\beta (y_2(x) - y_1(x))\, dx = \oint_C y(x)\, dx$$

Constraint

$$L(y) = \oint dl = \oint_C \sqrt{1+(y')^2}\, dx = L$$

Let $\underline{h = y - \lambda\sqrt{1+(y')^2}}$ (Lagrange multiplier function)

Using $\frac{d}{dx}\left(\frac{\partial h}{\partial y'}\right) = \frac{\partial h}{\partial y}$ note $\frac{\partial h}{\partial x} = 0$ so can use <u>first integral</u> of E-L. (see L4)

So $h - y' \frac{\partial h}{\partial y'} = k$ for some constant k

Writing this out <u>explicitly</u>, we get

$$k = y - \frac{\lambda}{\sqrt{1+(y')^2}} \implies \underline{(y')^2 = \frac{\lambda^2}{(y-k)^2} - 1}$$

Solving this gives $(x-x_0)^2 + (y-y_0)^2 = \lambda^2$

And circumference $2\pi\lambda = L \implies \lambda = \frac{L}{2\pi}$

This gives a <u>circle</u> of radius $\frac{L}{2\pi}$.

<u>Example 3.2</u> The Sturm-Liouville problem

Let function $\rho(x) > 0$ for $x \in [\alpha, \beta]$. Define also $\sigma = \sigma(x)$.

Then define

$$F(y) = \int_\alpha^\beta \left(\rho \cdot (y')^2 + \sigma \cdot (y^2)\right) dx$$

and <u>extremise</u> F subject to

$$G(y) = \int_\alpha^\beta y^2 \, dx = 1 .$$

$G(y) - 1$

rewrite 1 using an integrand :

$$\int_\alpha^\beta \frac{1}{\beta - \alpha} \, dx = 1$$

$$\Phi(y; \lambda) = F(y) - \lambda \left(G(y) - 1 \right)$$

Write $\quad h = \rho \cdot (y')^2 + \sigma \cdot (y^2) - \lambda \left(y^2 - \frac{1}{\beta - \alpha} \right)$

$$\frac{\partial h}{\partial y'} = 2 \rho y' , \qquad \frac{\partial h}{\partial y} = 2 \sigma y - 2 \lambda y$$

Applying E-L, rearranging gives

$$\underbrace{- \frac{d}{dx} (\rho y') + \sigma y}_{\mathcal{L}(y)} = \lambda y \qquad (3.2)$$

$\mathcal{L}$ is called the Sturm - Liouville operator

Then $\underline{\mathcal{L}(y) = \lambda y}$ is an eigenvalue problem

(note: similar form to TISE)

Note: if $\sigma > 0$ then $\underline{F(y) > 0}$.

<u>Claim</u> The smallest ~~positive~~ eigenvalue is equal to the positive minimum.

<u>Proof</u> Take (3.2) $\times$ y and IBP from α to β :

$$F(y) - \underbrace{\left[y \cdot y' \rho \right]_\alpha^\beta}_{0} = \lambda \underbrace{G(y)}_{1} = \lambda$$

(boundary term is 0 - fixed end problem)

So lowest eigenvalue = minimum of $\underline{\dfrac{F(y)}{G(y)}}$. $\square$

3.2 Several dependent variables

$$\underline{y}(x) = (y_1(x), \ y_2(x), \ \dots, \ y_n(x)) \qquad \text{Extremise}$$

$$F(\underline{y}) = \int_\alpha^\beta f(x, y_1, \dots, y_n, y_1', \dots, y_n') \, dx$$

Perturbation $\quad y_i \to y_i(x) + \varepsilon \eta_i(x)$

with $i = 1, \dots, n$ and $\eta_i(\alpha) = \eta_i(\beta) = 0$.

Using same derivation as that of E-L eqn, we get

$$F(\underline{y} + \varepsilon \underline{\eta}) - F(\underline{y}) = \int_\alpha^\beta \sum_{i=1}^n \eta_i \left(\frac{d}{dx}\left(\frac{\partial f}{\partial y_i'} \right) - \frac{\partial f}{\partial y_i} \right) dx$$

$$+ \text{ boundary terms } + O(\varepsilon^2)$$

Recall the "fundamental lemma" from L3: by setting all the η_i's but one to zero in turn, we get again

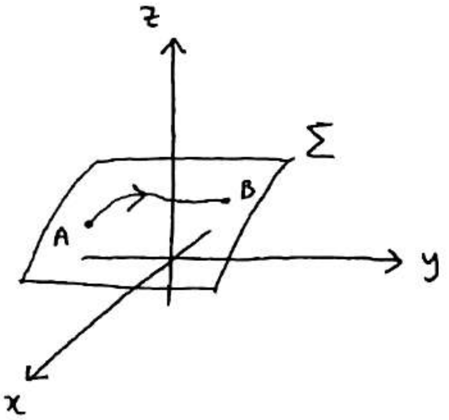
$$\frac{d}{dx}\left(\frac{\partial f}{\partial y_i'} \right) = \frac{\partial f}{\partial y_i} \qquad (3.3)$$

a system of n 2^{nd} order ODEs.

First integrals of 3.3

* if $\dfrac{\partial f}{\partial y_j} = 0$ for some $1 \le j \le n$ then $\dfrac{\partial f}{\partial y_j'} = \text{constant}$
* if $\dfrac{\partial f}{\partial x} = 0$ then $f - \sum_i y_i' \dfrac{\partial f}{\partial y_i'} = \text{constant}$

Example Geodesics on surfaces



$\Sigma \subseteq \mathbb{R}^3$ is a surface given by $g(x, y, z) = 0$.

Find shortest path on surface between 2 points, if one exists.

Take t to be a parameter on the curve:

$A = \underline{x}(0)$, $\quad B = \underline{x}(1)$ $\qquad \underline{x} = (x, y, z) = \underline{x}(t)$

$$\Phi(\underline{x}, \lambda) = \int_0^1 \sqrt{\dot{x}^2 + \dot{y}^2 + \dot{z}^2} - \lambda(t)\, g(x, y, z) \; dt$$

(Note $\lambda = \lambda(t)$ as we need the curve to lie on the surface i.e. $g(x, y, z) = 0$ everywhere - if don't have $\lambda = \lambda(t)$ then we are just saying $g(x, y, z)$ integrates to 0 over curve, not necessarily $g(x, y, z) = 0$)

Write integrand $\sqrt{\dot{x}^2 + \dot{y}^2 + \dot{z}^2} - \lambda(t)\, g = \underline{h(x, y, z, \dot{x}, \dot{y}, \dot{z}, \lambda)}$

Use E-L with h

- Variation wrt λ: $\underbrace{\frac{d}{dt}\left(\frac{\partial h}{\partial \dot{\lambda}}\right)}_{0 \text{ (} h \text{ does not depend on } \dot{\lambda}\text{)}} = \underbrace{\frac{\partial h}{\partial \lambda}}_{g(x,y,z)} \Rightarrow \underline{g(x, y, z) = 0} \quad \forall t$

Variation wrt x_i (x, y, z) :

$$\frac{d}{dt}\left(\frac{\dot{x}_i}{\sqrt{\dot{x}_1^2+\dot{x}_2^2+\dot{x}_3^2}}\right) + \lambda \frac{\partial g}{\partial x_i} = 0 \qquad i = 1, 2, 3$$

a system of ODEs
to be solved.

Alternatively, solve the constraint $\underline{g = 0}$, as we did in Ex 2.2 (geodesics on sphere).

3.3 Several independent variables

In general with $\Phi : \mathbb{R}^n \to \mathbb{R}^m$: for $n > 1$, E-L eqns become PDEs.

Consider case where $n = 3$, $m = 1$ so $\Phi : \mathbb{R}^3 \to \mathbb{R}$

$$F(\phi) = \iiint_D f(x, y, z, \phi, \phi_x, \phi_y, \phi_z) \, dx\, dy\, dz$$

(note $\phi_x := \frac{\partial \phi}{\partial x}$)

This is a volume integral over a domain $D \subseteq \mathbb{R}^3$

Assume there is an extremum, consider perturbation

$\phi \to \phi(x, y, z) + \varepsilon \eta(x, y, z)$ where $\underline{\eta = 0 \text{ on } \partial D}$

$$\underset{\text{(expanding)}}{F(\phi + \varepsilon\eta) - F(\phi)} = \varepsilon \int_D \left(\eta \frac{\partial f}{\partial \phi} + \eta_x \frac{\partial f}{\partial \phi_x} + \eta_y \frac{\partial f}{\partial \phi_y} + \eta_z \frac{\partial f}{\partial \phi_z}\right) dx\,dy\,dz + O(\varepsilon^2)$$

$$= \varepsilon \int_D \left(\eta \frac{\partial f}{\partial \phi} + \underbrace{\nabla \cdot \left(\eta \left(\frac{\partial f}{\partial \phi_x}, \frac{\partial f}{\partial \phi_y}, \frac{\partial f}{\partial \phi_z}\right)\right)}_{\substack{\text{use div. thm: } \eta = 0 \text{ on } \partial D \\ \text{so this becomes } 0}} - \eta \nabla \cdot \left(\frac{\partial f}{\partial \phi_x}, \frac{\partial f}{\partial \phi_y}, \frac{\partial f}{\partial \phi_z}\right)\right) dx\,dy\,dz + O(\varepsilon^2)$$

Then we get

$$F(\phi+\varepsilon\eta) - F(\phi) = \varepsilon \int \eta \underbrace{\left(\frac{\partial f}{\partial \phi} - \nabla \cdot \left(\frac{\partial f}{\partial \phi_x}, \frac{\partial f}{\partial \phi_y}, \frac{\partial f}{\partial \phi_z} \right) \right)}_{\text{then apply lemma:}} dx\,dy\,dz + o(\varepsilon^2)$$

$$\frac{\partial f}{\partial \phi} - \sum_{i=1}^{3} \frac{\partial}{\partial x_i} \left(\frac{\partial \phi}{\partial(\partial_i \phi)} \right) = 0 \qquad (3.4)$$

In general we have this for n rather than 3 specifically:

$$\underline{\frac{\partial f}{\partial \phi} - \sum_{i=1}^{n} \frac{\partial}{\partial x_i} \left(\frac{\partial \phi}{\partial(\partial_i \phi)} \right) = 0} \qquad (3.4.1)$$

Example Extremise "potential energy" $n=2$

$$F(\phi) = \iint_{D \subseteq \mathbb{R}^2} \frac{1}{2}\left(\phi_x^2 + \phi_y^2 \right) dx\,dy \qquad f = \text{integrand}$$

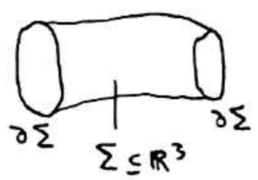
$\frac{\partial f}{\partial \phi} = 0$, $\frac{\partial f}{\partial \phi_x} = \phi_x$, $\frac{\partial f}{\partial \phi_y} = \phi_y$ so (3.4.1) gives

$$\frac{\partial}{\partial x}\phi_x + \frac{\partial}{\partial y}\phi_y = 0 \qquad \text{so} \qquad \underline{\phi_{xx} + \phi_{yy} = 0}$$

Laplace equation

Example Minimal surfaces

Minimise area of $\Sigma \subseteq \mathbb{R}^3$ subject to boundary conditions of 2 curves given as the boundary of the surface.



Take $\Sigma = \{ \underline{x} \in \mathbb{R}^3 \text{ s.t. } g(x,y,z) = 0 \}$

Assume (note: can use implicit function theorem) that we solved $g = 0$ to give $\underline{z = \phi(x, y)}$

$$ds^2 = dx^2 + dy^2 + dz^2, \qquad dz = \phi_x \, dx + \phi_y \, dy$$

so $$ds^2 = (1 + \phi_x^2) \, dx^2 + (1 + \phi_y^2) \, dy^2 + 2\phi_x \phi_y \, dx \, dy$$

This is called the <u>first fundamental form</u> or <u>Riemannian metric</u>

$$ds^2 = \sum_{i,j=1}^{2} \bar{g}_{ij}(x, y) \, dx_i \, dx_j \qquad \bar{g} = \begin{pmatrix} 1 + \phi_x^2 & \phi_x \phi_y \\ \phi_x \phi_y & 1 + \phi_y^2 \end{pmatrix}$$

Area element $\sqrt{\det \bar{g}} \, dx \, dy$

Then we get the area functional

$$A(\phi) = \int_D \underbrace{\sqrt{1 + \phi_x^2 + \phi_y^2}}_{h} \, dx \, dy$$ apply E-L to h (3.4.1)

$$\frac{\partial h}{\partial \phi_x} = \frac{\phi_x}{\sqrt{1 + \phi_x^2 + \phi_y^2}} \qquad \frac{\partial h}{\partial \phi_y} = \frac{\phi_y}{\sqrt{1 + \phi_x^2 + \phi_y^2}}$$

Since $\frac{\partial h}{\partial \phi} = 0$, applying 3.4.1 gives

$$\partial_x \left(\frac{\partial h}{\partial \phi_x} \right) + \partial_y \left(\frac{\partial h}{\partial \phi_y} \right) = 0$$ and expanding derivatives

gives $$\underline{(1 + \phi_y^2) \phi_{xx} + (1 + \phi_x^2) \phi_{yy} - 2 \phi_x \phi_y \phi_{xy} = 0} \qquad (3.5)$$

the minimal surface equation.

Assume <u>circular symmetry</u> $\underline{z = \phi(r)}$, $r = \sqrt{x^2 + y^2}$

$\Rightarrow$ 3.5 becomes an <u>ODE</u>

$$\underline{r z'' + z' + (z')^3 = 0}$$ (check by chain rule - find derivatives of ϕ).

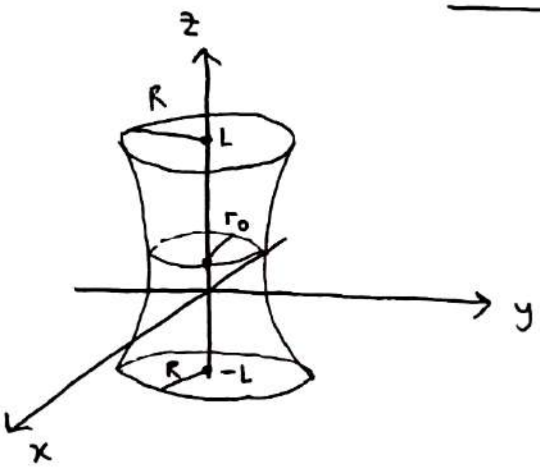
Recall ODE from area functional from L6:

$r z'' + z' + (z')^3 = 0$ Set $z' = w$:

Multiply by z' to give

$$\frac{1}{2} r \frac{d}{dr}(w^2) + w^2 + w^4 = 0$$

This has solution $\underline{r = r_0 \cosh\left(\frac{z - z_0}{r_0}\right)}$ corresponding to a <u>catenoid</u> (a catenary rotated about an axis)



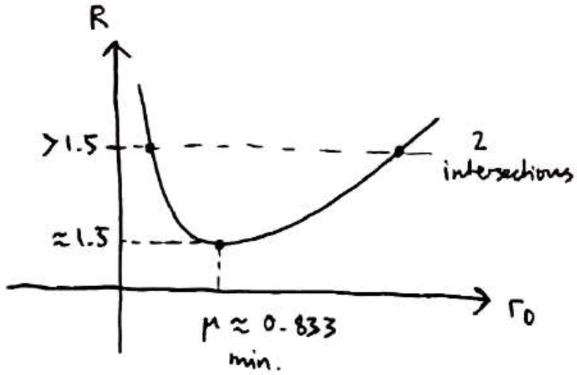
$r(L) = r(-L)$

If $L \neq 0$ then $z_0 = 0$

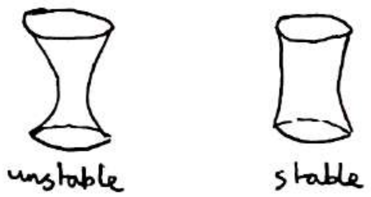
Set $r = R$, divide by L:

$$\frac{R}{L} = \frac{r_0}{L} \cosh\left(\frac{L}{r_0}\right)$$

Set $L = 1$ WLOG and get $R = r_0 \cosh\left(\frac{1}{r_0}\right)$



If $R > \approx 1.5$ have 2 minimal surfaces



(none exists for $R < \approx 1.5$)

<u>Higher derivatives</u>

Extremise

$$F(y) = \int_\alpha^\beta f(x, y, y', \dots, y^{(n)})\, dx.$$

Suppose y exists and as before, perturb $y \to y + \varepsilon \eta$.

Have $\eta = \eta' = \dots = \eta^{(n-1)} = 0$ at α, β.

$$F(y+\varepsilon\eta) - F(y) = \varepsilon \int_\alpha^\beta \Big(\frac{\partial f}{\partial y}\eta + \underbrace{\frac{\partial f}{\partial y'}\eta'}_{\text{IBP once}} + \dots + \underbrace{\frac{\partial f}{\partial y^{(n)}}\eta^{(n)}}_{\text{IBP } n \text{ times}} \Big)\, dx + o(\varepsilon^2)$$

After doing the IBP and applying the lemma, we get

$$\underline{\frac{\partial f}{\partial y} - \frac{d}{dx}\frac{\partial f}{\partial y'} + \frac{d^2}{dx^2}\frac{\partial f}{\partial y''} + \dots + (-1)^n \frac{d^n}{dx^n}\frac{\partial f}{\partial y^{(n)}} = 0} \qquad (3.6)$$

(Euler-Lagrange for higher derivatives).

<u>Example</u> $n=2$ and $\frac{\partial f}{\partial y} = 0$

Then (3.6) $\Rightarrow$ $\frac{d}{dx}\left(\frac{\partial f}{\partial y'} - \frac{d}{dx}\frac{\partial f}{\partial y''} \right) = 0$

so $\frac{\partial f}{\partial y'} - \frac{d}{dx}\frac{\partial f}{\partial y''} = c$ (1st <u>integral</u>)

<u>Example</u> Extremise $F(y) = \int_0^1 (y'')^2\, dx$ subject to

$y(0) = y'(0) = 0$ and $y(1) = 0$, $y'(1) = 1$.

The integrand satisfies $\frac{\partial f}{\partial y} = 0$ so applying above 1st integral,

we get $\frac{d}{dx}(2y'') = c \Rightarrow \underline{y''' = k}$ for some constant k.

Then just <u>integrate</u> and use BCs to give $\underline{y = x^3 - x^2}$.

$\square$

<u>Note</u>: This is an absolute minimum.

$y_0 = x^3 - x^2$, $\quad \eta(0) = \eta'(0) = \eta(1) = \eta'(1) = 0$

(do not assume η small).

$$F(y_0+\eta) - F(y_0) = \underbrace{\int_0^1 (\eta'')^2\,dx}_{>0} + 2\int_0^1 (y_0''\,\eta'')\,dx$$

(else if BCs hold have $\eta = 0$)

(We get this using difference of two squares as we have $(y'')^2$ term in the integrand).

Then $$F(y_0+\eta) - F(y_0) > 4\int_0^1 (3x-1)\,\eta''$$

$$= 4\left(\underbrace{[-\eta']_0^1}_{0} + \int_0^1 \left(\frac{d}{dx}(3x\eta') - \eta'\right)dx\right)$$

$$= 4\left([3x\eta']_0^1 - [3\eta]_0^1\right) = \underline{0} \quad \text{(by BCs)}$$

So y_0 is an <u>absolute minimum</u>.

(To be clear, this is because no matter what η we perturb y_0 by, we have $F(y_0+\eta) > F(y_0)$ as shown above). $\square$

<u>Section 4</u> Least action principle and Noether's theorem

Consider a particle moving in $\mathbb{R}^3$, kinetic energy T, potential energy V.

Lagrangian $\quad \underline{L = L(\underline{x}, \dot{\underline{x}}, t) = T - V} \quad (4.1)$

t = independent variable

$\underline{x} = (x, y, z)$ dependent variables

Action functional

$$S(\underline{x}) = \int_{t_1}^{t_2} L \, dt \qquad (4.2)$$

sometimes called least action, but should be stationary

Hamilton's principle / principle of stationary action:

The motion of the particle is such that S is made stationary, so the particle's motion satisfies the Euler-Lagrange equations.

Example 4.1 $\quad T = \frac{1}{2} m |\dot{\underline{x}}|^2, \quad V = V(\underline{x})$

Using E-L, we get

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}_i}\right) = \frac{\partial L}{\partial x_i} \quad \Rightarrow \quad m\ddot{x}_i = -\frac{\partial V}{\partial x_i}$$

i.e. $\underline{m\ddot{\underline{x}} = -\nabla V}$

recovering Newton's 2nd Law.

1/7/21

Variational Principles - Lecture 8

Action $S(\underline{x}) = \int_{t_1}^{t_2} L \, dt$, $\quad L = T - V$

Example 4.2 Central force in 2 dimensions

$$L = \frac{1}{2} m \left(\dot{r}^2 + r^2 \dot{\theta}^2 \right) - V(r)$$

Using E-L :
$$\frac{d}{dt} \frac{\partial L}{\partial \dot{r}} - \frac{\partial L}{\partial r} = 0 \qquad (1)$$
$$\frac{d}{dt} \frac{\partial L}{\partial \dot{\theta}} - \underbrace{\frac{\partial L}{\partial \theta}}_{0} = 0 \qquad (2)$$

(2) $\Rightarrow$ $\frac{\partial L}{\partial \dot{\theta}} = \text{constant}$, so $\underline{m r^2 \dot{\theta} = \text{constant}}$

conservation of angular momentum

Also $\frac{\partial L}{\partial t} = 0$ so can use the other first integral (2.5)

to give $\dot{r} \frac{\partial L}{\partial \dot{r}} + \dot{\theta} \frac{\partial L}{\partial \dot{\theta}} - L = \text{constant}$

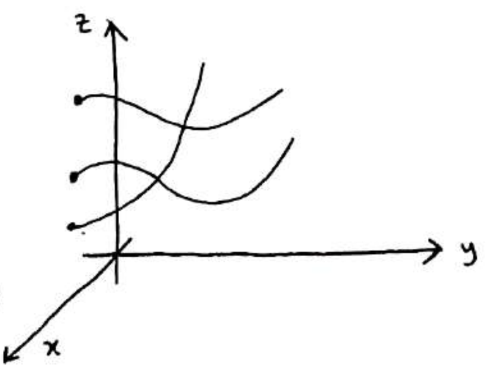
$$\Rightarrow \dot{r} m \dot{r} + \dot{\theta} m r^2 \dot{\theta} - \frac{1}{2} m \dot{r}^2 - \frac{1}{2} m r^2 \dot{\theta}^2 + V(r) = \text{const.}$$

$$\Rightarrow \frac{1}{2} m \dot{r}^2 + \frac{1}{2} m r^2 \dot{\theta}^2 + V(r) = E = \text{constant}$$

conservation of total energy

Example 4.3 (Configuration space, and generalised coordinates)

N particles in $\mathbb{R}^3$ Configuration space $\mathbb{R}^{3N}$



$t \to (q_i(t), \dot{q}_i(t), t)$

$i = 1, \dots, 3N$

$L = L(q_i, \dot{q}_i, t)$

4.2 Noether's Theorem

$$F(y) = \int_\alpha^\beta f(y_i, y_i', x)\, dx \qquad i = 1, \dots, n$$

Suppose there exists a 1-parameter family of transformations

$y_i(x) \to Y_i(x, s)$ where $Y_i(x, 0) = y_i(x)$

$s \in \mathbb{R}$ a continuous parameter.

This is a continuous symmetry of a Lagrangian f if

$$\frac{d}{ds} f\big(Y_i(x,s),\ Y_i'(x,s),\ x\big) = 0.$$

Noether's theorem Given a continuous symmetry $Y_i(x,s)$ of f, the quantity

$$\sum_i \frac{\partial f}{\partial y_i'} \frac{\partial Y_i}{\partial s}\bigg|_{s=0} \qquad (4.3)$$

is a first integral of the Euler-Lagrange equation with $Y_i(x, 0) = y_i(x)$ $\forall i$.

Proof $\frac{d}{ds} f = 0$ so certainly $\frac{d}{ds} f\big|_{s=0} = 0$.

(cts symmetry)
of f, true by defn above

Use chain rule:

$$\frac{d}{ds} f\bigg|_{s=0} = \frac{\partial f}{\partial y_i} \frac{dY_i}{ds}\bigg|_{s=0} + \frac{\partial f}{\partial y_i'} \frac{dY_i'}{ds}\bigg|_{s=0}$$

$$= \left[\underbrace{\frac{d}{dx}\left(\frac{\partial f}{\partial y_i'}\right)}_{\text{by E-L}} \frac{dY_i}{ds} + \frac{\partial f}{\partial y_i'} \frac{d}{dx} \frac{dY_i}{ds} \right]\Bigg|_{s=0}$$

note: using Σ conv

$$\Rightarrow \quad \frac{d}{dx}\Bigg(\underbrace{\frac{\partial f}{\partial y_i'} \frac{\partial Y_i}{\partial s}}_{\Sigma \text{ conv}} \Bigg)\Bigg|_{s=0} = 0 \qquad \text{as required.} \qquad \square$$

<u>Example 4.4</u> $\quad f = \frac{1}{2}(y')^2 + \frac{1}{2}(z')^2 - V(y-z)$

Lagrangian of a particle moving on a plane in a potential

Define $\quad Y = y+s, \quad Z = z+s, \quad Y' = y', \quad Z' = z'$

note $\quad V(Y-Z) = V(y-z)$

so $\quad \frac{d}{ds} f = 0$

and (4.3) $\Rightarrow \quad \frac{\partial f}{\partial y'} \frac{dY}{ds} + \frac{\partial f}{\partial z'} \frac{dZ}{ds} \Big|_{\text{at } s=0} = \underline{y' + z'}$ is

<u>conserved</u> "momentum in $y+z$ direction".

<u>Example 4.5</u> $\quad$ Return to example 4.2

$(r, \theta) \quad \Theta = \theta + s, \quad R = r$

$\frac{dL}{ds} = 0 \quad$ and (4.3) $\Rightarrow \left(\frac{\partial L}{\partial \dot\theta} \frac{d\Theta}{ds} + \frac{\partial L}{\partial \dot r} \frac{\partial R}{\partial s} \right)\Big|_{s=0}$

$\quad = \underline{mr^2 \dot\theta}$ conserved angular momentum

Isotropy of space $\Rightarrow$ rotational invariance of L $\Rightarrow$ conserved angular momentum.

<u>Convex Functions</u>

A class of functions for which it is easy to classify stationary points.

PTO

Definition A set $S \subseteq \mathbb{R}^n$ is convex if $\forall \underline{x}, \underline{y} \in S$,

$(1-t)\,\underline{x} + t\underline{y} \in S \quad \forall\ 0 \leq t \leq 1$.

Definition A graph of a function $f: \mathbb{R}^n \to \mathbb{R}$ is a surface $\{ z - f(\underline{x}) = 0 \}$ in $\mathbb{R}^{n+1}$.

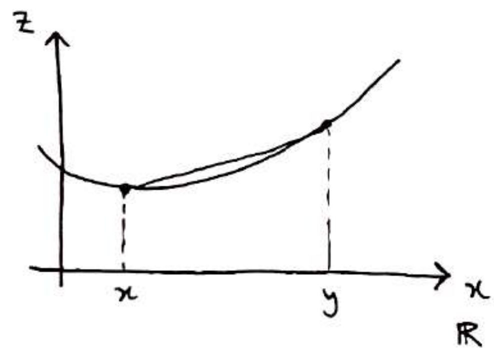
A chord of f is a line segment joining two points on the graph.

Definition A function $f: \mathbb{R}^n \to \mathbb{R}$ is convex if

(i) the domain of f is a convex set, and

(ii) $$f\left((1-t)\underline{x} + t\underline{y}\right) \leq (1-t) f(\underline{x}) + t f(\underline{y}) \qquad (5.1)$$
$\forall\ 0 < t < 1$

Illustration $f: \mathbb{R} \to \mathbb{R}$:



between x and y, the function lies below the line segment / chord joining $(x, f(x))$ and $(y, f(y))$.

Variational Principles - Lecture 9 2/7/21

Remarks on convexity:

(i) f is concave when we replace $\leq$ by $\geq$

(ii) f is convex iff $-f$ is concave

(iii) f is strictly convex if we replace $\leq$ by $<$

Example $f: \mathbb{R} \to \mathbb{R}$ $f(x) = x^2$ domain $\mathbb{R}$ convex

$$f((1-t)x + ty) - (1-t)f(x) - tf(y)$$

$$= ((1-t)x + ty)^2 - (1-t)x^2 - ty^2$$

$$= x^2(1-t)(-t) + ty^2(1-t) + 2(1-t)txy$$

$$= -(1-t)t(x-y)^2 < 0 \quad \forall\, 0<t<1$$ strictly convex

Example $f(x) = \frac{1}{x}$ domain $\mathbb{R} \setminus \{0\}$ not a convex set

But on restricted domain $\mathbb{R} > 0$ f is convex

5.1 Conditions for convexity

3 tests for convexity

(a) If f is once differentiable, then f is convex iff

$$f(\underline{y}) \geq f(\underline{x}) + (\underline{y} - \underline{x}) \cdot \nabla f(\underline{x}) \qquad (5.2)$$

Proof Assume (5.2) and apply twice: $\Rightarrow$ back implication

$$f(\underline{x}) \geq f(\underline{z}) + (\underline{x} - \underline{z}) \cdot \nabla f(\underline{z}) \qquad (i)$$

$$f(\underline{y}) \geq f(\underline{z}) + (\underline{y} - \underline{z}) \cdot \nabla f(\underline{z}) \qquad (ii)$$

Choose $\underline{z} = (1-t)\underline{x} + t\underline{y} \in S$ (the domain of f)

with $0 < t < 1$.

Compute $(1-t)\cdot(i) + t\cdot(ii)$: $\nabla f(\underline{z})$ cancels and we get (5.1) the defn of convexity. ✓

<u>Reverse</u> implication: assume convexity (5.1).

Set $h(t) = (1-t) f(\underline{x}) + t f(\underline{y}) - f((1-t)\underline{x} + t\underline{y}) \geq 0$ for $0 < t < 1$.

$$h'(0) = -f(\underline{x}) + f(\underline{y}) - (\underline{y}-\underline{x})\cdot \nabla f(\underline{x})$$

So (5.2) is equivalent to $\underline{h'(0) \geq 0}$.

Note $h(0) = 0$ so $\underline{\dfrac{h(t)-h(0)}{t}} \geq 0$ because $h(t) \geq 0$ for $0<t<1$

Taking <u>limit</u> as $t \to 0$, we get indeed that $h'(0) \geq 0$.

So (5.2) holds. ✓ $\square$

<u>Corollary</u> If f is convex and has a stationary point, then the stationary point is a global minimum.

<u>Proof</u> Given $\nabla f(\underline{x}_0) = 0$, (5.2) implies

$f(\underline{y}) \geq f(\underline{x}_0)$ $\forall \underline{y}$ in domain $\square$

(b) If $(\nabla f(\underline{y}) - \nabla f(\underline{x}))\cdot(\underline{y}-\underline{x}) \geq 0$ (5.3)

then f is convex.

<u>Proof</u> Exercise $\square$

(c) Second order conditions: Assume f twice differentiable. Then f is convex iff the Hessian $\dfrac{\partial^2 f}{\partial x_i \partial x_j}$ has <u>all eigenvalues non-negative</u>.

Proof Assume convex, apply (5.3) by taking $\underline{y} = \underline{x} + \underline{h}$

Then $\underline{h} \cdot (\nabla f(\underline{x}+\underline{h}) - \nabla f(\underline{y})) \geq 0$

For small $\underline{h}$ we have

$$\partial_i f(\underline{x}+\underline{h}) = \partial_i f(\underline{x}) + \sum_j h_j \underbrace{H_{ij}(\underline{x})}_{\text{Hessian } \frac{\partial^2 f}{\partial x_i \partial x_j}} + O(|\underline{h}|^2)$$

So by dotting with $\underline{h}$ (taking $\partial_i f(\underline{x})$ to LHS first) get

$$\sum_{j,i} h_i h_j H_{ij}(\underline{x}) + O(|\underline{h}|^2) \geq 0$$

Hessian is symmetric, so diagonalisable and must be true in all bases, so eigenvalues must be non-negative. $\square$

Example $f(x,y) = \dfrac{1}{xy}$ $\quad x, y > 0$

$$H = \frac{1}{xy}\begin{pmatrix} 2/x^2 & 1/xy \\ 1/xy & 2/y^2 \end{pmatrix} \qquad \begin{array}{l} \det H = 3/x^3y^3 > 0 \\ \operatorname{tr} H > 0 \end{array}$$

so evals ≥ 0, f is convex.

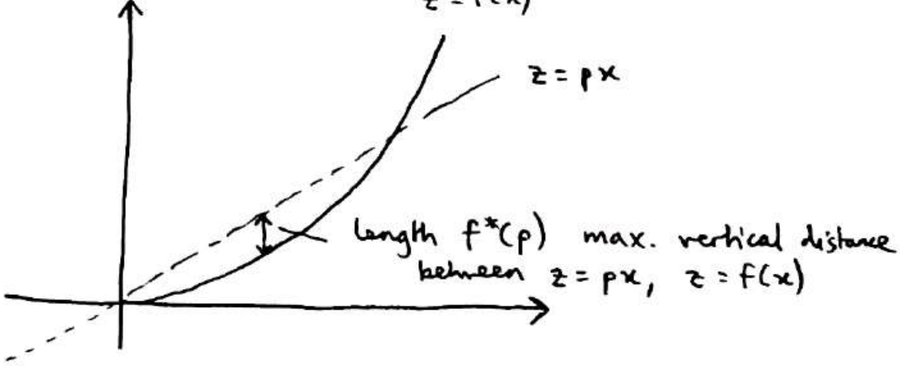
Section 6 Legendre transform

Definition The Legendre transform of $f : \mathbb{R}^n \to \mathbb{R}$ is

$$\underline{f^*(\underline{p}) = \sup_{\underline{x}} (\underline{p} \cdot \underline{x} - f(\underline{x}))}$$

The domain of f^* consists of all vectors $\underline{p} \in \mathbb{R}^n$ s.t. the sup is finite.

Example $n = 1$



<u>Example</u> $n=1$, $f(x) = ax^2$, $a>0$

$f^*(p) = \sup_x (px - ax^2)$ $\quad \frac{\partial}{\partial x}(px - ax^2) = 0 \Rightarrow p = 2xa$

so $x = \frac{p}{2a}$ sub into $f^*(p)$ to get

$f^*(p) = \underline{\frac{p^2}{4a}}$ the Legendre transform.

$(f^*)^*(s) = \sup_p \left(sp - \frac{p^2}{4a} \right) \Rightarrow p = 2as,\ (f^*)^*(s) = as^2$

In fact, $\underline{(f^*)^* = f}$ for $\underline{f \text{ convex}}$.

<u>Proposition</u> The domain of f^* is a convex set, and f^* is convex.

<u>Proof</u> $f^*((1-t)\underline{p} + t\underline{q}) = \sup_{\underline{x}} \left((1-t)\underline{p}\cdot\underline{x} + t\underline{q}\cdot\underline{x} - f(\underline{x}) \right)$

$= \sup_{\underline{x}} \left((1-t)(\underline{p}\cdot\underline{x} - f(\underline{x})) + t(\underline{q}\cdot\underline{x} - f(\underline{x})) \right)$

$\leq (1-t) f^*(p) + t f^*(q)$

(using $\sup(A+B) \leq \sup A + \sup B$)

and we are done, as (i) $(1-t)\underline{p} + t\underline{q} \in D(f^*)$
(convex domain)

and (ii) f^* satisfies the convexity definition (5.1). $\square$

If f is convex, differentiable then $f^*(p)$ is the global <u>minimum</u> over $\underline{x}$

We have $\nabla(\underline{p}\cdot\underline{x} - f(\underline{x})) = 0 \Rightarrow \underline{p} = \nabla f$ $\quad (*)$

(substitute to defn of $f^*(p)$).

If f is strictly convex, there is a unique inversion of $(*)$

$\underline{x} = \underline{x}(\underline{p})$ with $\underline{f^*(\underline{p}) = \underline{p}\cdot\underline{x}(\underline{p}) - f(\underline{x}(\underline{p}))}$. (6.2)

(cf x^2 example). This generalises that method to find f^*.

Variational Principles - Lecture 10 2/7/21

6.1 Applications to Thermodynamics

Many particles → few variables - pressure, volume, temperature, entropy

Internal energy $U(S, V)$

Helmholtz free energy $F(T, V) = \min_S (U(S, V) - TS)$

$$= -\max_S (TS - U(S, V))$$

$= - \underline{U^*(T, V)}$ Legendre transform of U wrt S (V fixed).

$$\frac{\partial}{\partial S}(TS - U(S, V))\Big|_{T,V} = 0 \Rightarrow \underline{T = \frac{\partial U}{\partial S}\Big|_V}$$

Enthalpy $H(S, P) = \min_V (U(S, V) + PV) = -\underline{U^*(-P, S)}$

and at minimum, as above, $\underline{P = -\frac{\partial U}{\partial V}\Big|_S}$.

Legendre transform is a way to swap from (S, V) dependence to dependence on other variables.

Section 7 Hamilton's Equations

$L(q, \dot{q}, t) = T - V$ function on configuration space

Define Hamiltonian as Legendre transform of L wrt $\dot{\underline{q}} = v$

$$H(\underline{q}, \underline{p}, t) = \sup_{\underline{v}} (\underline{p} \cdot \underline{v} - L) = \underline{\underline{p} \cdot \underline{v} - L(\underline{q}, \underline{v}, t)}$$

where $\underline{v} = \underline{v}(\underline{p})$ is soln to $p_i = \frac{\partial L}{\partial \dot{q}_i}$

↘ recall the $\underline{p} = \nabla f$ from end of L 9 to find Legendre transform assuming convexity

Example $T = \frac{1}{2} m |\dot{\underline{q}}|^2$, $V = V(\underline{q})$

$$\underline{p} = \frac{\partial L}{\partial \dot{\underline{q}}} = m \dot{\underline{q}} \Rightarrow \dot{\underline{q}} = \frac{\underline{p}}{m}$$

$$H(\underline{q}, \underline{p}, t) = \underline{p} \cdot \frac{\underline{p}}{m} - \left(\frac{1}{2} m \frac{|\underline{p}|^2}{m^2} - V(\underline{q}) \right)$$

$$= \underline{\frac{1}{2m} |\underline{p}|^2 + V(\underline{q})} \qquad \text{total energy.}$$

What happens to the Euler-Lagrange equations?

$$H = H(\underline{q}, \underline{p}, t) = p_i \dot{q}^i - L(q^i, \dot{q}^i, t) \qquad (\Sigma \text{ conv})$$

$$dH = \frac{\partial H}{\partial q^i} dq^i + \frac{\partial H}{\partial p_i} dp_i + \frac{\partial H}{\partial t} dt$$

$$= \cancel{p_i d\dot{q}^i} + \dot{q}^i dp_i - \underbrace{\frac{\partial L}{\partial q^i}}_{\text{by E-L } \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}^i}\right) = \dot{p}_i} dq^i - \underbrace{\cancel{\frac{\partial L}{\partial \dot{q}^i} d\dot{q}^i}}_{p_i} - \frac{\partial L}{\partial t} dt$$

$$= \dot{q}^i dp_i - \dot{p}_i dq^i - \frac{\partial L}{\partial t}$$

Compare differentials:

$$\underline{\dot{q}^i = \frac{\partial H}{\partial p_i}, \qquad \dot{p}_i = -\frac{\partial H}{\partial q^i}, \qquad \frac{\partial H}{\partial t} = -\frac{\partial L}{\partial t}}$$

Hamilton's equations (7.2).

Warning: $\left.\frac{\partial}{\partial t}\right|_{p,q} \neq \left.\frac{\partial}{\partial t}\right|_{q,\dot{q}}$ to avoid confusion we assume $\frac{\partial L}{\partial t} = 0$.

Then (7.2) is a system of $\underline{2n}$ first order ODEs.

Need to specify $q^i(0)$, $p_i(0)$, $i = 1, \ldots, n$

Solutions are trajectories in $2n$-dimensional phase space.

<u>Remark</u> Hamilton's equations can be derived from extremising a functional in phase space.

$$S(\underline{q}, \underline{p}) = \int_{t_1}^{t_2} \underbrace{\left(\dot{q}^i p_i - H(\underline{q}, \underline{p}, t) \right)}_{f(\underline{q}, \underline{p}, \underline{\dot{q}}, \underline{\dot{p}}, t)} \, dt$$

E-L for S $\quad \dfrac{\partial f}{\partial p_i} - \underbrace{\dfrac{d}{dt}\left(\dfrac{\partial f}{\partial \dot{p}_i} \right)}_{0 \text{ as } \frac{\partial f}{\partial \dot{p}_i} = 0} = 0$ $\quad$ variation wrt p_i

$$\Rightarrow \quad \frac{\partial f}{\partial p_i} = 0 \quad \Rightarrow \quad \underline{q^i = \frac{\partial H}{\partial p_i}}$$

wrt q_i we have $\quad \dfrac{\partial f}{\partial q^i} - \dfrac{d}{dt}\left(\dfrac{\partial f}{\partial \dot{p}_i} \right) = 0$

$$\Rightarrow \quad -\frac{\partial H}{\partial q^i} - \frac{d}{dt} p_i = 0 \quad \text{so} \quad \underline{\dot{p}_i = -\frac{\partial H}{\partial q^i}}$$

recovering (7.2).

Newton's, Lagrange's, Hamilton's formulations are equivalent.

Section 8 The second variation

Euler-Lagrange may lead to a min/max/saddle point. How do we examine the nature of these critical points?

$$F(y) = \int_\alpha^\beta f(x, y, y')\,dx$$

Expand $F(y+\varepsilon\eta)$ to 2nd order in ε, around a solution to E-L equation (stationary pt).

$$F(y+\varepsilon\eta) - F(y) = \int_\alpha^\beta \left(f(x, y+\varepsilon\eta, y'+\varepsilon\eta') - f\right) dx$$

$$= \varepsilon \int_\alpha^\beta \eta \, \overbrace{\cancel{\left(\frac{\partial f}{\partial y} - \frac{d}{dx}\frac{\partial f}{\partial y'}\right)}}^{0 \text{ by E-L}} dx$$

(note: 0th order term is 0, omitted)

$$+ \frac{\varepsilon^2}{2}\int_\alpha^\beta \left(\eta^2 \frac{\partial^2 f}{\partial y^2} + (\eta')^2 \frac{\partial^2 f}{\partial (y')^2} + 2\frac{\partial^2 f}{\partial y \partial y'}\eta\eta'\right) dx$$

$$+ O(\varepsilon^3)$$

Consider 2nd order term, the 2nd variation: take

$$\underset{\substack{\text{2nd}\\ \text{variation}}}{\delta^2 F(y)} = \frac{1}{2}\int_\alpha^\beta \left(\eta^2 \frac{\partial^2 f}{\partial y^2} + (\eta')^2 \frac{\partial^2 f}{\partial (y')^2} + \underbrace{2\frac{\partial^2 f}{\partial y \partial y'}\eta\eta'}_{\text{IBP, use } \eta = 0 \text{ at } \alpha, \beta}\right) dx$$

$$= \frac{1}{2}\int_\alpha^\beta \left(Q\eta^2 + P(\eta')^2\right) dx$$

where $P = \dfrac{\partial^2 f}{\partial (y')^2}$, $\quad Q = \dfrac{\partial^2 f}{\partial y^2} - \dfrac{d}{dx}\dfrac{\partial^2 f}{\partial y \partial y'}$ (8.1)

So we have just proved:

<u>Proposition</u> If y solves the Euler-Lagrange equation <u>and</u> $Q\eta^2 + P(\eta')^2 > 0$ $\forall \eta$ vanishing at α, β, then $y(x)$ is a local <u>minimiser</u> of $F(y)$.

<u>Example</u> Geodesics on a plane

$$f = \sqrt{1+(y')^2} \qquad P = \frac{1}{(1+(y')^2)^{3/2}} > 0, \quad Q = 0$$

If $\eta' = 0$ then $\eta = 0$ (as $\eta(\alpha) = \eta(\beta) = 0$), so have $\eta' \neq 0$ so $P(\eta')^2 > 0$ $\forall \eta$.

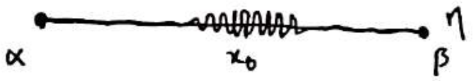
So straight lines are local length <u>minimisers</u> of $\mathbb{R}^2$.

<u>Proposition</u> If $y_0(x)$ is a local minimiser, then we have

$$P = \left.\frac{\partial^2 f}{\partial (y')^2}\right|_{y_0} \geq 0 \qquad \text{(Legendre condition necessary).} \qquad (8.2)$$

<u>Sketch proof</u> If η' small, then η can't be too large

Assume $\exists x_0$ s.t. $P(x_0, y_0, y_0') < 0$.



η oscillates around x_0 and vanishes everywhere else so η' gets very large around x_0

(8.2) <u>not sufficient</u>

But <u>$P > 0$</u> <u>and</u> <u>$Q \geq 0$</u> is sufficient, as if $\eta \not\equiv 0$ on $[\alpha, \beta]$ then $\exists x_0 \in (\alpha, \beta)$ s.t. $\eta'(x_0) \neq 0$.

Example Brachistochrone problem

$$f = \sqrt{\frac{1+(y')^2}{-y}} \qquad \text{is the cycloid a minimiser?}$$

$$P = \frac{1}{(1+(y')^2)^{3/2}\sqrt{-y}} > 0 \;, \qquad Q = \frac{1}{2\sqrt{1+(y^2)^2}\, y^2 \sqrt{-y}} > 0$$

so sufficient condition holds, and this is a local minimum.

8.1 Associated eigenvalue problem

Go back to eqn (8.1) and rewrite:

$$Q\eta^2 + P(\eta')^2 = Q\eta^2 + \underbrace{\frac{d}{dx}(P\eta\eta')}_{\substack{\text{boundary term} \\ \text{drop: } \eta(\alpha)=\eta(\beta)=0}} - \eta(P\eta')'$$

Integrate:

$$\delta^2 F(\underset{\text{soln to E-L eqn}}{y_0}) = \frac{1}{2}\int_\alpha^\beta \eta \underbrace{\left(-(P\eta')' + Q\eta\right)}_{\mathcal{L}(\eta) \;\; \text{Sturm-Liouville operator}} dx \qquad (8.3)$$

If $\exists\, \eta$ such that $\mathcal{L}(\eta) = -\omega^2\eta \quad (\omega \in \mathbb{R})$

and $\eta(\alpha) = \eta(\beta) = 0$ (8.4)

If the eval problem has a solution, then y_0 is <u>not</u> a minimiser because

$$\underset{\text{2nd variation}}{\delta^2 F(y_0)} = -\frac{1}{2}\omega^2 \int_\alpha^\beta \eta^2\, dx \;\underline{<0}.$$

Note (8.4) the eval problem can have solutions even if $P>0$, so Legendre condition (8.2) is <u>not sufficient</u>.

<u>Example</u> $F(y) = \int_0^\beta \underbrace{\left((y')^2 - y^2\right)}_{f} dx$ $\quad y(0) = y(\beta) = 0$, $\beta \neq n\pi$ for $n \in \mathbb{Z}$

Euler - Lagrange: $y'' + y = 0$

$y = y_0 = 0$ is the <u>only</u> solution satisfying the BCs.

2nd variation:

$$\delta^2 F(0) = \frac{1}{2}\int_0^\beta \left((\eta')^2 - \eta^2\right) dx \qquad \underline{P = 1 > 0, \quad Q < 0}$$

So we examine the eigenvalue (S-L) problem:

Problem (8.4) is $\underline{-\eta'' - \eta = -\omega^2 \eta}$ with $\underline{\eta(0) = \eta(\beta) = 0}$

Take $\underset{\text{(soln)}}{\eta = A \sin\left(\frac{\pi x}{\beta}\right)} \Rightarrow \underline{\left(\frac{\pi}{\beta}\right)^2 = 1 - \omega^2}$

This is possible for $\omega \in \mathbb{R}$ only if $\underline{\beta > \pi}$.

So if $P > 0$ a problem may arise if $\underline{\beta > \pi}$

(i.e. the interval is "<u>too large</u>", and the eval problem may have a solution, meaning a minimiser may <u>not</u> exist).

In L12 we will formalise this notion of "too large".

8.2 Jacobi condition

Legendre tried to prove that $P > 0$ is sufficient for a local minimum. This would not work, but the idea was good...

Let $\phi = \phi(x)$ be differentiable on $[\alpha, \beta]$.

Note since $\eta(\alpha) = \eta(\beta) = 0$, using FTC we have

$$0 = \int_\alpha^\beta (\phi \eta^2)' \, dx = \int_\alpha^\beta (\phi' \eta^2 + 2\eta\eta'\phi) \, dx$$

Add this to (8.1) and then we can rewrite (8.1) as

$$\delta^2 F(y) = \frac{1}{2} \int_\alpha^\beta \left(P(\eta')^2 + 2\eta\eta'\phi + (Q + \phi')\eta^2 \right) dx$$

Assume that $\underline{P|_y > 0}$ and complete the square on integrand.

Then we get

$$\delta^2 F(y) = \frac{1}{2} \int_\alpha^\beta \left(P\left(\eta' + \frac{\phi}{P}\eta\right)^2 + \left(Q + \phi' - \frac{\phi^2}{P}\right)\eta^2 \right) dx$$

This is positive if we can choose some ϕ such that

$$\underline{\phi^2 = P(Q + \phi')} \qquad (8.3)$$

(Note $Q + \phi' - \frac{\phi^2}{P} = 0$ if (8.3) holds.)

If (8.3) holds, then $\delta^2 F(y) > 0$ unless $\eta' + \frac{\phi}{P}\eta = 0$ $\ast\ast$ on $[\alpha, \beta]$. But $\eta = 0$ at α so $\eta'(\alpha) = 0$ if $\ast\ast$ holds.

But then $\eta \equiv 0$ on $[\alpha, \beta]$ (uniqueness of soln to 1st order ODE) so $** \neq 0$.

<u>Does</u> a solution to (8.3) exist on $[\alpha, \beta]$?

Transform (8.3) into 2nd order ODE by setting

$\underline{\phi = -P \dfrac{u'}{u}}$ with $u \neq 0$ on $[\alpha, \beta]$.

Then we get $P\left(\dfrac{u'}{u}\right)^2 = Q - \left(\dfrac{(Pu')}{u}\right)'$

$$= Q - \frac{(Pu')'}{u} + P\left(\frac{u'}{u}\right)^2$$

$$\Rightarrow \quad \underline{-(Pu') + Qu = 0} \qquad (8.4)$$

This is the Jacobi accessory condition.
(Remember: we are looking to find out when $\delta^2 F > 0$).

We need a solution to (8.4) $\mathcal{L}(u) = 0$ with $u \neq 0$ on $[\alpha, \beta]$.

But there may <u>not</u> be a solution if $[\alpha, \beta]$ is <u>large</u>.

To be clear: <u>if</u> there is a nonzero solution to (8.4), then we have a suitable ϕ, so given $P > 0$ for this y, we have $\delta^2 F(y) > 0$.

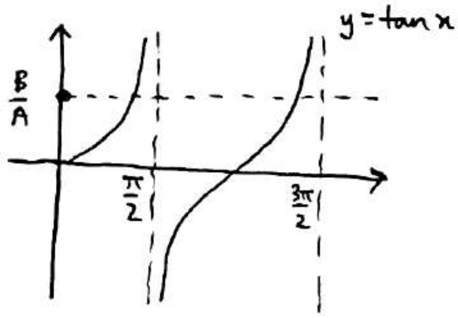
<u>Example 1</u> $\quad F(y) = \dfrac{1}{2}\displaystyle\int_\alpha^\beta \left((y')^2 - y^2\right) dx$

$$\delta^2 F(y) = \frac{1}{2}\int_\alpha^\beta \left((\eta')^2 - \eta^2\right) dx \qquad \begin{array}{l} P = 1 \\ Q = -1 \end{array}$$

(8.4) general solution to $u'' + u = 0$ is

$$u = A \sin x - B \cos x \qquad (-\text{ sign for convenience}).$$

When we find u, we need $u \not\equiv 0$ for $[\alpha, \beta]$ i.e.
$\tan x \neq \frac{B}{A}$ on $[\alpha, \beta]$

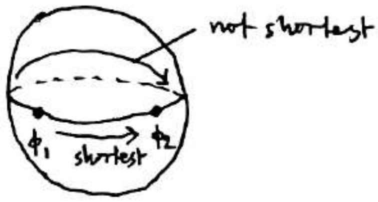


We can avoid $\frac{B}{A}$ only on an interval smaller than π.
So if $|\beta - \alpha| < \pi$ then $\delta^2 F(y_0) > 0$.

Example 2 Geodesics on a sphere

$$\sqrt{d\theta^2 + \sin^2\theta \, d\phi^2} = \underbrace{\sqrt{(\theta')^2 + \sin^2\theta}}_{f} \, d\phi \qquad \theta = \theta(\phi)$$

We showed earlier that geodesics are segments of great circles. This great circle can be made equatorial by rotation:



Great circle: θ constant
$\theta_0 = \frac{\pi}{2}$ (equatorial)

$$\left.\frac{\partial^2 f}{\partial \theta'^2}\right|_{\theta_0} = 1 = P, \qquad Q = -1$$

$$\delta^2 F\left(\theta_0 = \frac{\pi}{2} ; \eta\right) = \frac{1}{2}\int_{\phi_1}^{\phi_2} \left((\eta')^2 - (\eta^2)\right) d\phi$$

The same as Example 1. This can be made positive if $\phi_2 - \phi_1 < \pi$ (have a "unique minimum".)

(note if $\phi_2 - \phi_1 = \pi$ then have two equal geodesics.