Variational Principles Lecture Notes

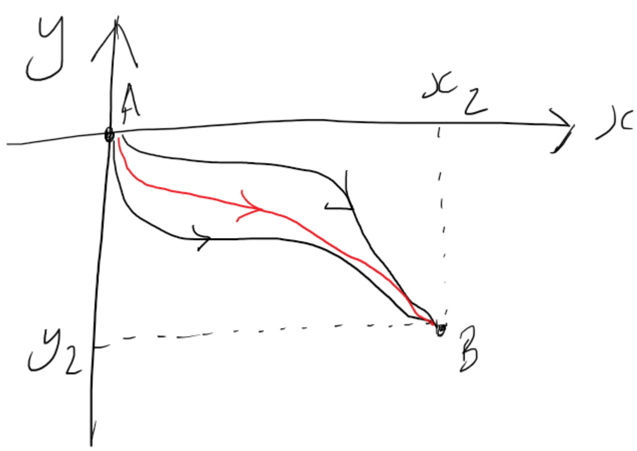
Conor Rajan

Easter 2021

0 Motivations

Example 0.1 (The Brachistochrome Problem)**.** Particle slides on a wire, under influence of gravity between two fixed points A,B. Which shape of the wire gives the shortest travel time, starting from rest?

The problem was proposed by Johann Bernoulli in 1696.



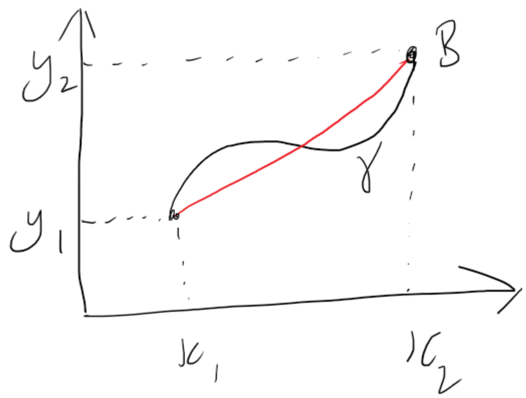
The travel time is given by

$$T = \int \mathrm{d}t = \int_A^B \frac{\mathrm{d}t}{v(x,y)}$$

$$KE + V = const \text{ (conservation of energy)}$$
$$\frac{1}{2}mv^2 + mgy = mgy_1 = 0,\ v = \sqrt{2}g\sqrt{-y}$$

$$\text{minimise } T[y] = \frac{1}{\sqrt{2g}}\int_0^{x_2} \frac{\sqrt{1+(y')^2}}{\sqrt{-y}}\mathrm{d}x \text{ subject to } y(0) = 0, y(x_2) = y_2$$

Example 0.2 (Geodesic)**.** A geodesic is the shortest path γ between 2 points on a surface Σ (if one exists). Take $\Sigma = \mathbb{R}^2$



Distance along γ:

$$D[y] = \int_A^B \mathrm{d}t = \int_{x_1}^{x_2} \sqrt{1 + (y')^2}\mathrm{d}x$$

Seek to minimise D by varying γ.

We want to be able to minimise functions of the form

$$F[y] = \int_{x_1}^{x_2} f(x, y(x), y'(x))\mathrm{d}x \tag{0.1}$$

among all functions such that $y(x_1) = y_1, y(x_2) = y_2$.

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

function: numbers $\rightarrow$ numbers
functionals: functions $\rightarrow$ functions

Examples of functionals are

- Area under a graph: $f(x, y, y') = y$
- Length of a curve: $f(x, y, y') = \sqrt{1 + (y')^2}$

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

Notation:

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

- $C^k(\mathbb{R})$ = space of functions with continuous k-th derivatives
- $C^k_{(\alpha,\beta)}(\mathbb{R})$ =space of functions with continuous k-th derivatives such that $f(\alpha) = f(\beta)$

Need to specify the function space beforehand

(Calculus of variations is a branch of Functional Analysis - part III - analysis on the space of functions).

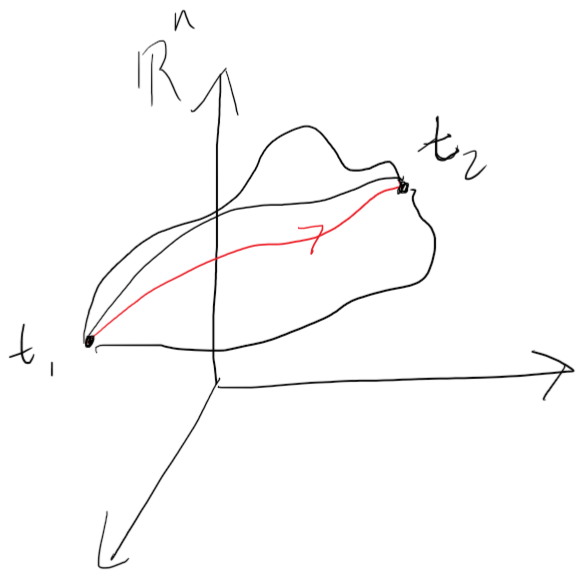
Analysis I was about analysis on the number line.

Variational Principles - principles in nature, where the laws follow from extremizing functions.

Example 0.3 (Fermat's Principle)**.** "Light between two points travels along paths which require least time"

Example 0.4 (Principle of least action)**.** T = kinetic energy, V = potential energy

$$S[\gamma] = \int_{t_1}^{t_2} (T - V)\mathrm{d}t$$



"Action is minimised along paths of motion"

Leibnitz's take on Example 0.4: "We live in the best of all possible worlds"

Feynmann's take on Example 0.4: "This is wrong. In quantum theory the motion takes place along all possible paths with different probabilities." (see Part

III Quantum Field Theory)

This course:

- Necessary conditions of extremum of (0.1) given by Euler-Lagrange equation.
- Lots of examples (geometry, physics, problems with constraints - e.g maximise the area given a fixed length of perimeter)
- Second variation: some sufficient conditions for minima/maxima

Books:

1. Gelfand-Fomin "Calculus of Variations"
2. DAMPT notes online (e.g P.Townsend)

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

$f \in C^2(\mathbb{R}^n)$, $f : \mathbb{R}^n \to \mathbb{R}$, with continuous 2nd partial derivatives.

The point $\mathbf{a} \in \mathbb{R}^n$ is stationary if

$$\nabla f(\mathbf{a}) = (\partial_1 f, \dots, \partial_n f)|_{\mathbf{x}=\mathbf{a}} = 0,\ \partial_i f = \frac{\partial f}{\partial x_i}$$

Expanding near $\mathbf{x} = \mathbf{a}$:

$$f(\mathbf{x}) = f(\mathbf{a}) + \underbrace{(\mathbf{x}-\mathbf{a}) \cdot \nabla f|_{\mathbf{a}}}_{0, \text{ since } \mathbf{a} \text{ is stationary}} + \frac{1}{2}(x_i - a_i)(x_j - a_j)^2 \partial^2_{ij} f|_{\mathbf{a}} + O(|\mathbf{x}-\mathbf{a}|^2)$$

The Hessian matrix $H_{ij} = \partial_i \partial_j f = H_{ji}$

Shift the origin to set $\mathbf{a} = \mathbf{0}$, and diagonalise $H(\mathbf{0})$ by an orthogonal transformation

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

$$f(\mathbf{x}') - f(\mathbf{0}) = \frac{1}{2} \sum \lambda_i (x_i)^2 + O(|\mathbf{x}'|^2)$$

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

(ii) All $\lambda_i < 0$ - local maximum

(iii) Somce $\lambda_i > 0$, some $\lambda_i < 0$ so f increases in some directions and decreases in others - saddle point

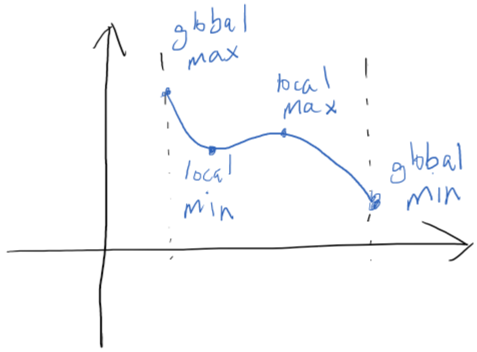
(iv) Some $\lambda_i = 0$. Need to consider higher order terms in Taylor expansion.

For the special case $n = 2$, $\det(H) = \lambda_1 \lambda_2$, $\mathrm{Tr}(H) = \lambda_1 + \lambda_2$

- $\det H > 0, \mathrm{Tr}(H) > 0 \implies$ local maximum
- $\det H > 0, \mathrm{Tr}(H) < 0 \implies$ local minimum
- $\det H < 0 \implies$ saddle point
- $\det H = 0 \implies$ need to look at higher order terms

Remarks:

1. $f : D \to \mathbb{R}$, may not be able to determine global maxima/minima just by looking at derivatives:



2. f harmonic $f_{xx} + f_{yy} = 0$, $D \subset \mathbb{R}^2$ on $\mathbb{R}^2$. Then $\mathrm{Tr}(H) = 0$ so the turning point is a sadddle so the min/max is on the boundary

Example 1.1.

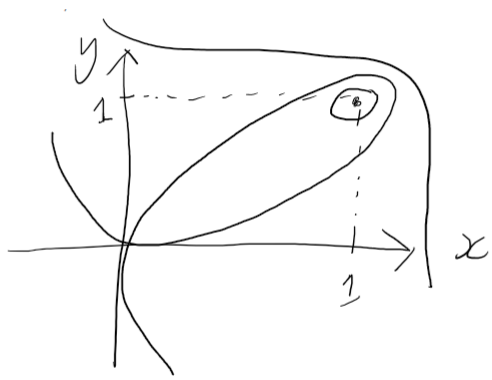
$$f(x, y) = x^3 + y^3 - 3xy$$

$\nabla f = (2x^2 - 3y, 3y^2 - 3x) = (0, 0)$ for critical points. $x^2 - y = 0, y^2 - x = 0 \Rightarrow y^4 = y \Rightarrow$ solutions are $(0, 0)$ and $(1, 1)$.

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

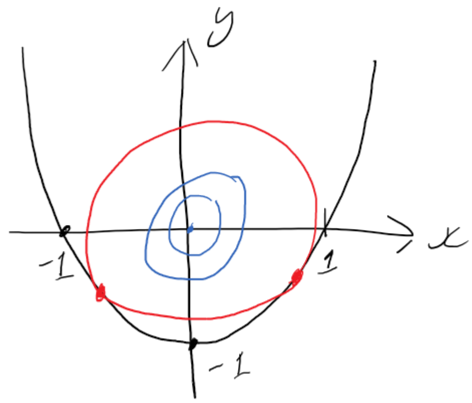
At $(0, 0)$, $\det H = -9 < 0$, so saddle point $f = 0$

At $(1, 1)$, $\det H = 26 > 0$, $\mathrm{Tr}(H) = 12 > 0$, so local minimum $f = -1$.



1.1 Constraints and Lagrange multipliers

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



Two approaches:

- Direct method. Solve the constraints

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

$$\partial_x f = 0 = 4x^3 - 2x = 0$$

So $(x, y) = (\pm\frac{1}{\sqrt{2}}, -\frac{1}{2}),$ radius $= \frac{\sqrt{3}}{2}$ or $(x, y) = (0, 1),$ radius $= 1$

- Lagrange multipliers. Define new function $h(x, y, \lambda) = f(x, y) - \lambda g(x, y)$ where $g(x, y) = 0$ is the constraint. λ is called the Lagrange Multiplier.

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

Extremise over 3 variables with no contraints:

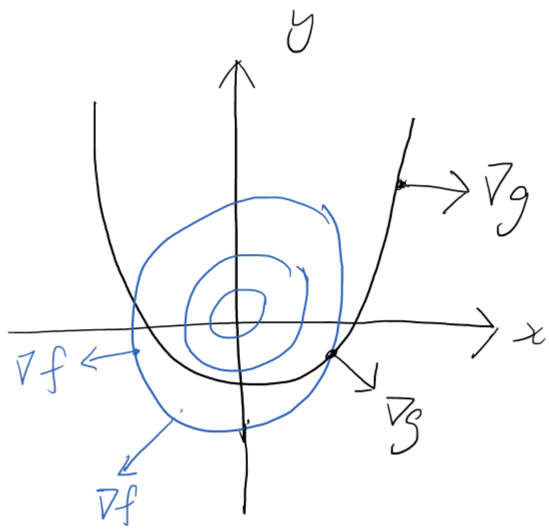
$$\frac{\partial h}{\partial x} = 2x + 2\lambda x = 0$$

$$\frac{\partial h}{\partial y} = 2y - \lambda = 0$$

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

Solving these gives $(0, -1, -2) \Rightarrow f = 1$ or $(\pm\frac{1}{\sqrt{2}}, -\frac{1}{2}, -1) \Rightarrow f = \frac{3}{4}$.

Why does it work (geometrically)?



∇g perpendicular to $g = 0$, ∇f perpendicular to $f = const$. At the extremum ∇f is parallel to ∇g, or $\nabla f = \lambda \nabla g$, i.e $\nabla(f - \lambda g) = 0$.

Multiple constraints:

$$\text{Extremize } f : \mathbb{R}^n \to \mathbb{R} \text{ subject to } g_\alpha(\mathbf{x}) = 0$$

$$g_\alpha : \mathbb{R}^n \to \mathbb{R}, \alpha = 1, 2, \ldots, k$$

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

Then solve

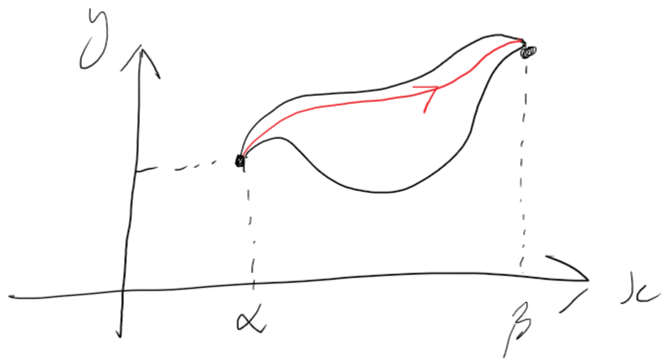
$$\frac{\partial h}{\partial x_i} = 0, \ \frac{\partial h}{\partial \lambda_\alpha} = 0$$

by eliminating λ_α and solving for $\mathbf{x}$.

2 The Euler-Lagrange equations

We wish to extremize the functional (0.1)

$$F[y] = \int_\alpha^\beta f(x, y, y')\mathrm{d}x \tag{2.1}$$



With f given, depending on y, with fixed ends. For a given y consider a small pertubation $y \mapsto y + \varepsilon\eta(x)$ in (2.1) with $\eta(\alpha) = \eta(\beta) = 0$.

We will need

Lemma 2.1. *If $g : [\alpha, \beta] \to \mathbb{R}$ is continuous on $[\alpha, \beta]$ and*

$$\int_\alpha^\beta g(x)\eta(x)\mathrm{d}x = 0$$

For all η continuous on $[\alpha, \beta]$ with $\eta(\alpha) = \eta(\beta) = 0$, then $g(x) = 0$ for all $x \in [\alpha, \beta]$.

Proof. For contradiction, suppose $\exists \bar{x} \in (\alpha, \beta)$ with $g(\bar{x}) = 0$. Say $g(\bar{x}) > 0$. Then there is an interval $[x_1, x_2] \subset (\alpha, \beta)$ with $g(x) > c$ on $[x_1, x_2]$ for some $c > 0$. Set

$$\eta(x) = \begin{cases} (x - x_1)(x_2 - x) & x \in [x_1, x_2] \\ 0 & x \notin [x_1, x_2] \end{cases} \tag{2.2}$$

Then

$$\int_\alpha^\beta g(x)\eta(x) > c\int_{x_1}^{x_2} (x - x_1)(x_2 - x)\mathrm{d}x > 0$$

Which is a contradiction. □

Remark: then η given by (2.2) is a bump function. Can give a bump function in C^k by considering

$$\eta(x) = \begin{cases} ((x - x_1)(x_2 - x))^{k+1} & x \in [x_1, x_2] \\ 0 & x \notin [x_1, x_2] \end{cases} \tag{2.2}$$

Now we return to considering (2.1)

$$\begin{aligned}F[y+\varepsilon\eta] &= \int_\alpha^\beta f(x, y+\varepsilon\eta, y'+\varepsilon\eta')\mathrm{d}x \\ &= F[y] + \varepsilon\int_\alpha^\beta \left(\frac{\partial f}{\partial y}\eta + \frac{\partial f}{\partial y'}\eta'\right)\mathrm{d}x + O(\varepsilon^2)\end{aligned}$$

At the extremum, $\frac{\mathrm{d}F}{\mathrm{d}\varepsilon}|_{\varepsilon=0} = 0$. Integrating the ε term by parts

$$\begin{aligned}0 &= \int_\alpha^\beta \left(\frac{\partial f}{\partial y}\eta - \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\partial f}{\partial y'}\right)\eta\right)\mathrm{d}x + \left[\frac{\partial f}{\partial y'}\eta\right]_\alpha^\beta \\ &= \int_\alpha^\beta \left(\frac{\partial f}{\partial y}\eta - \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\partial f}{\partial y'}\right)\eta\right)\mathrm{d}x\end{aligned}$$

Now let $g = \frac{\partial f}{\partial y}\eta - \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\partial f}{\partial y'}\right)\eta$ and apply the lemma to see that we must have g identically 0 if it is extremiszed for all η. Hence a necessary condition for an extremum is

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\partial f}{\partial y'}\right) - \frac{\partial f}{\partial y}\eta = 0 \tag{2.3}$$

This is called the Euler-Lagrange equation.

Remarks:

- (2.3) is a 2nd order ODE for $y(x)$ with boundary conditions $y(\alpha) = y_1$ and $y(\beta) = y_2$.
- Notation: the LHS of (2.3) denoted by $\frac{\delta F[y]}{\delta y(x)}$ and is called the functional derivative
- Some books (e.g Townsend's notes) use $\delta y = \varepsilon\eta(x)$

  $$F[y+\delta y] = F[y] + \delta F[y]$$

  where

  $$\delta F = \int_\alpha^\beta \left[\frac{\delta F[y]}{\delta y(x)}\delta y(x)\right]\mathrm{d}x$$

- Other boundary conditions possible, e.g $\frac{\partial f}{\partial y'}|_{\alpha,\beta} = 0$
- Careful with derivatives, e.g $\frac{\partial f}{\partial y}$ means $\left(\frac{\partial f}{\partial y}\right)|_{x,y'}$ and x, y, y' are independent so

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

  and

  $$\frac{\mathrm{d}}{\mathrm{d}x} = \partial_x + y'\partial_y + y''\partial_{y'}$$

For example, for $f(x, y, y') = x((y')^2 - y^2)$, we have

$$\partial_x f = (y')^2 - y^2,\ \partial_y f = -2xy,\ \partial_{y'} f = 2xy'$$

$$\frac{\mathrm{d}f}{\mathrm{d}x} = (y')^2 - y^2 - 2xyy' + 2xy'y''$$

2.1 First integrals of the Euler-Lagrange equation

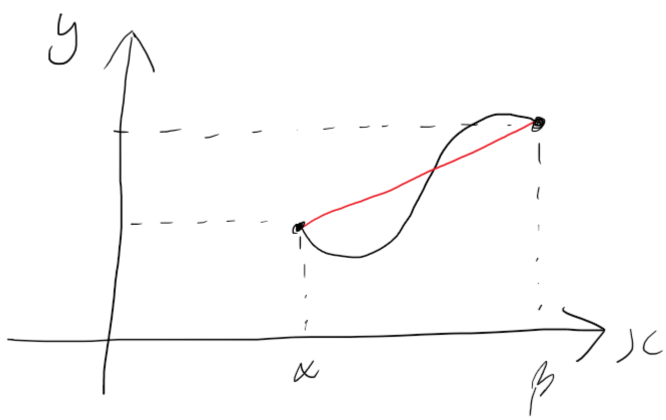
In some cases (2.3) can be integrated once to a 1st order ODE:

(a) f does not explicitly depend on y, i.e $\frac{\mathrm{d}f}{\mathrm{d}y} = 0$. Then (2.3) implies

$$\frac{\partial f}{\partial y'} = const \tag{2.4}$$

Example 2.1. Geodesics on the Euclidean plane

$$F[y] = \int_\alpha^\beta \sqrt{\mathrm{d}x^2 + \mathrm{d}y^2} = \int_\alpha^\beta \sqrt{1 + (y')^2}\mathrm{d}x$$



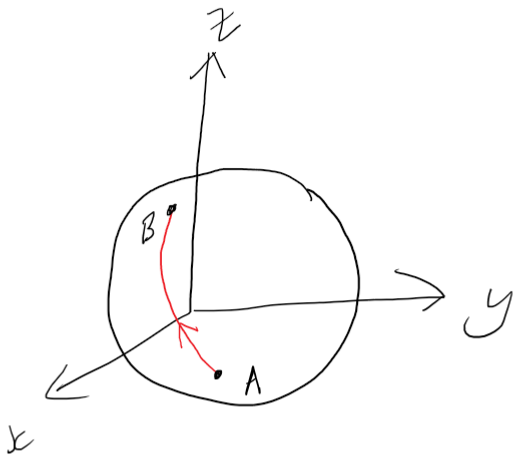
have $f(y') = \sqrt{1 + (y')^2}$ and $\frac{\partial f}{\partial y} = 0$ so by (2.4):

$$\frac{y'}{\sqrt{1 + (y')^2}} = const$$

and therefore $y' = const$ and $y = mx + c$ (straight line)

Example 2.2 (Geodesics on a sphere)**.** $S^2 \subset \mathbb{R}^3$. Using spherical coordinates:

$$\begin{aligned} x &= \sin\theta\sin\phi \\ y &= \sin\theta\cos\phi \\ z &= \cos\theta \end{aligned}$$



for $0 \leq \theta \leq \pi, 0 \leq \phi < 2\pi$. Also $\mathrm{d}s^2 = \mathrm{d}x^2 + \mathrm{d}y^2 + \mathrm{d}z^2 = \mathrm{d}\theta^2 + \sin^2\theta \mathrm{d}\phi^2$.

Parameterise as $\phi = \phi(\theta)$ so

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

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

Note $pdvf\phi = 0$, so $\frac{\partial f}{\partial \phi'} = \kappa$ where κ is a constant. Hence

$$\frac{\sin^2\theta\phi'}{\sqrt{1 + \sin^2\theta(\phi')^2}} = \kappa$$

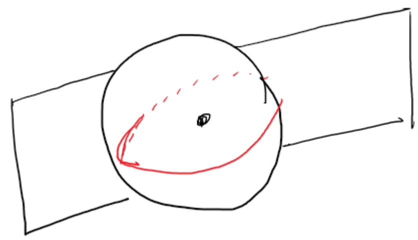
$$(\phi')^2 = \frac{\kappa^2}{\sin^2\theta(\sin^2\theta - \kappa^2)}$$

$$\phi = \pm\int \frac{\kappa}{\sin\theta\sqrt{\sin^2\theta - \kappa^2}}\mathrm{d}\theta$$

Two solutions, each corresponding to opposite orientations. Using the substitution $u = \cos\theta$ we get

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

This is the great circle, so the geodesics are segments of great circles.



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

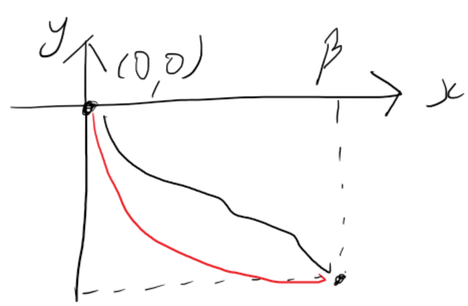
$$\frac{\mathrm{d}}{\mathrm{d}x}\left(f - y'\frac{\partial f}{\partial y'}\right) = \frac{\partial f}{\partial x} + y'\frac{\partial f}{\partial y} + y''\frac{\partial f}{\partial y'} - y''\frac{\partial f}{\partial y'} - y'\frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\partial f}{\partial y'}\right)$$
$$=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

So if f does not explicitly depend on x, i.e $\frac{\partial f}{\partial x}$ then

$$f - y'\frac{\partial f}{\partial y'} = const \tag{2.5}$$

Example 2.3 (Brachistochrome)**.**

$$F[y] = \frac{1}{\sqrt{2g}}\int_0^\beta \underbrace{\frac{\sqrt{1+(y')^2}}{\sqrt{-y}}}_{f(y,y')}\,\mathrm{d}x$$



$\frac{\partial f}{\partial x} = 0$ so use (2.5)

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

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

$$x = \pm\kappa \int \frac{\sqrt{-y}}{\sqrt{1+\kappa^2 y}} \mathrm{d}y$$

Set

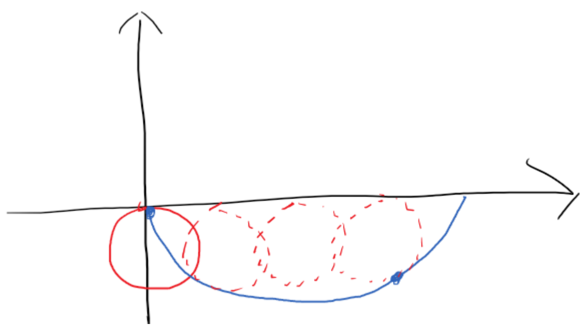
$$y = -\frac{1}{\kappa^2}\sin^2\frac{\theta}{2}, \ \mathrm{d}y = -\frac{1}{\kappa^2}\sin\frac{\theta}{2}\cos\frac{\theta}{2}$$

$$x = \mp\frac{1}{2\kappa^2}\int (1-\cos\theta)\mathrm{d}\theta = \mp\frac{1}{2\kappa^2}(\theta - \sin\theta) + C$$

initial condition: $\theta_0 = 0 \Rightarrow C = 0$ so

$$\begin{cases} x = \frac{\theta - \sin\theta}{2\kappa^2} \\ y = -\frac{1}{\kappa^2}\sin^2\frac{\theta}{2} \end{cases}$$

this is a cycloid - the curve traced by a point on the rim of a wheel, as the wheel rolls along a straight line.



2.2 Fermat's principle

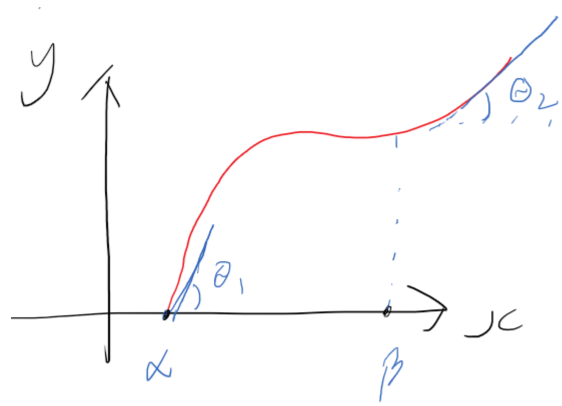
Light/sound travels along the paths between two points which require the least time.

Ray = path $y = y(x)$. Speed of light $c(x, y)$

$$F[y] = \int \frac{\mathrm{d}t}{c} = \int_\alpha^\beta \underbrace{\frac{\sqrt{1+(y')^2}}{c(x,y)}}_{f(x,y,y')} \mathrm{d}x$$

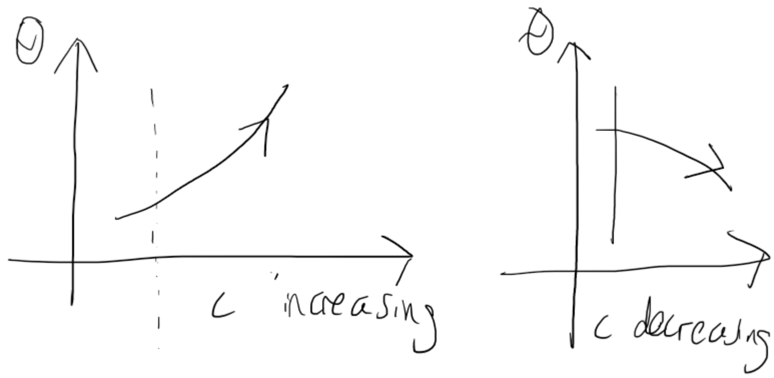
assume $c = c(x) \implies \frac{\partial f}{\partial y} = 0$, so (2,4) implies $\frac{\partial f}{\partial y'} = const$

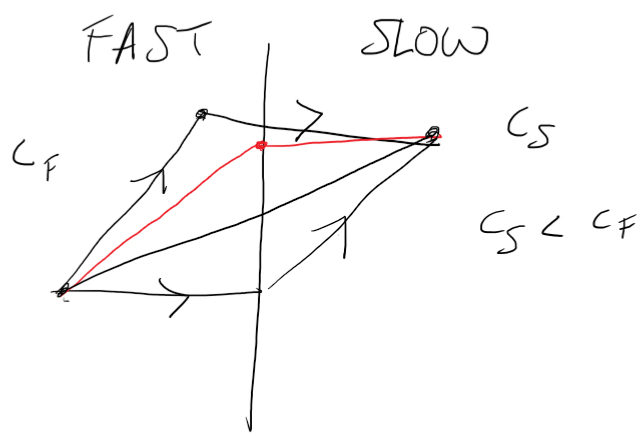
$$\frac{y'}{\sqrt{1+(y')^2}c(x)} = const$$



ray launched at θ_1 and $\tan\theta = y'$, which gives us

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





3 Extensions of the Euler-Lagrange equations

3.1 Euler-Lagrange equations with constraints

We wish to extremize

$$F[y] = \int_{\alpha}^{\beta} f(x, y, y')\mathrm{d}x$$

subject to

$$G[y] = \int_{\alpha}^{\beta} g(x, y, y')\mathrm{d}x = \kappa$$

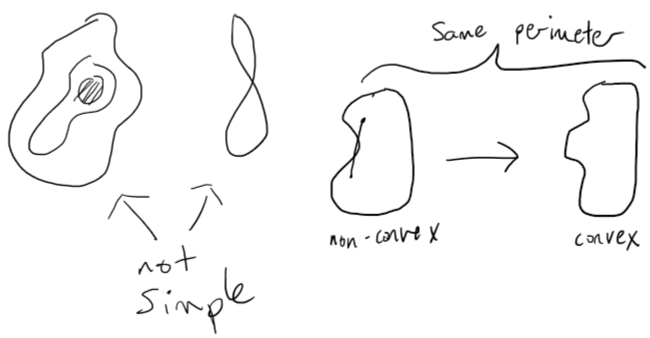
for some constant κ. Using lagrange multipliers, extremize

$$\Phi[y; \lambda] = F[y] - \lambda G[y]$$

so replace f in $E - L$ with $f - \lambda g$

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\partial}{\partial y'}(f - \lambda g)\right) - \frac{\partial}{\partial y}(f - \lambda g) = 0$$

Example 3.1 (Dido problem)**.** What simple and closed plane curve of fixed lenth L maximizes the enclosed area?



Assume convexity. x monotonically increases from $\alpha \to \beta$ and decreases $\beta \to \alpha$. Given $x \in (\alpha, \beta)$, there exists points $(x, y_1), (x, y_2)$ on the curve with $y_2 > y_1$ and $\mathrm{d}A = y(x)|_{x_1}^{x_2} \cdot \mathrm{d}x$.

$$A[y] = \int_{\alpha}^{\beta} (y_2(x) - y_1(x))\mathrm{d}x = \oint_C y(x)\mathrm{d}x$$

Constraint

$$L[y] = \oint_C \mathrm{d}l = \oint_C \sqrt{1 + (y')^2}\mathrm{d}x = L$$

Define

$$h = y - \lambda\sqrt{1 + (y')^2}$$

[Note: do not need to worry about the boundary term in the derivation if the E-L equations, as C has no boundary]

Note $\frac{\partial h}{\partial x} = 0$ so we apply (2.5):

$$\kappa = h - y' \frac{\partial h}{\partial y'} = y - \lambda\sqrt{1+(y')^2} + y'\lambda \frac{y'}{\sqrt{1+(y')^2}}$$

$$= y - \frac{\lambda}{\sqrt{1+(y')^2}} \Rightarrow (y')^2 = \frac{\lambda^2}{(y-\kappa)^2} - 1$$

Solution: $(x-x_0)^2 + (y-y_0)^2 = \lambda^2$ (circle of radius λ).

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

Example 3.2 (The Sturm-Liouville problem)**.**

$$\rho(x) > 0 \text{ for } x \in [\alpha, \beta], \sigma = \sigma(x)$$

$$F[y] = \int_\alpha^\beta \left[\rho \cdot (y')^2 + \sigma \cdot y^2\right] \mathrm{d}x, G[y] = \int_\alpha^\beta y^2 \mathrm{d}x$$

Minimize F subject to $G = 1$ (fixed ends).

$$\Phi[y;\lambda] = F[y] - \lambda(G[y] - 1)$$

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

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

$$\underbrace{-\frac{\mathrm{d}}{\mathrm{d}x}(\rho \cdot y') + \sigma \cdot y}_{\mathcal{L}(y)=\lambda y,\ \mathcal{L}=\text{ Sturm-Liouville operator}} = \lambda y \tag{3.2}$$

(3.2) is an eigenvalue problem, e.g if $\rho = 1$, $\sigma(x) =$ 'potential' in time-independent Schrödinger equation (IB Quantum).

Proposition. *If $\sigma > 0$ then $F[y] > 0$. Positive minimum equal to the lowest eigenvalue.*

Proof. (3.2)$\times y$ and integrate $|_\beta^\alpha$ by parts:

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

lowest eigenvalue = minimum of $\frac{F[y]}{G[y]}$. □

3.2 Several dependent variables

$$y(x) = (y_1(x), y_2(x), \ldots, y_n(x))$$

$$F[y] = \int_\alpha^\beta f(x, y_1, \ldots, y_n, y_1', \ldots, y_n')\mathrm{d}x$$

$$y_i \to y_i(x) + \varepsilon\eta_i(x), i = 1, \ldots, n, \eta_i(\alpha) = \eta_i(\beta) = 0$$

$$F[y+\varepsilon\eta] - F[y] = \int_\alpha^\beta \sum_{i=1}^n \eta_i \left\{ \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\partial f}{\partial y_i'}\right) - \frac{\partial f}{\partial y_i} \right\} \mathrm{d}x + \text{ boundary terms } + O(\varepsilon^2)$$

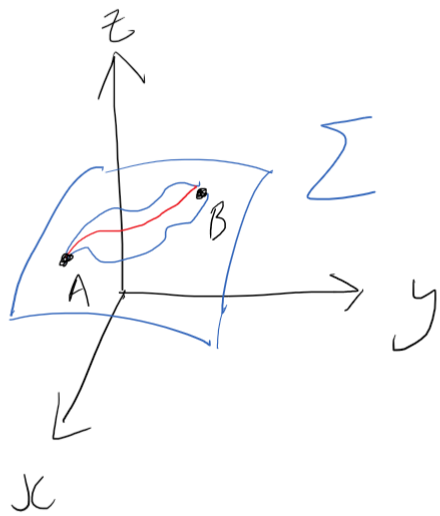
Use the Lemma proved at the start:

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\partial f}{\partial y_i'}\right) = \frac{\partial f}{\partial y_i} \tag{3.3}$$

This is a system of n 2nd order ODEs. Taking the first integrals of (3.3):

- If $\frac{\partial f}{\partial y_j} = 0$ for some $1 \leq j \leq n$, then by (3.3) $\frac{\partial f}{\partial y_j'} = const$
- If $\frac{\partial f}{\partial x} = 0$, then $f - \sum_i y_i' \frac{\partial f}{\partial y_i'} = const$

Example 3.3 (Geodesics on surfaces)**.** $\Sigma \subset \mathbb{R}^3$ (surface) given by $g(x, y, z) = 0$. A geodesic is the shortest path on the suface between $A, B \in \Sigma$ (if one exists).



Parameterise Σ with t, $A = \mathbf{x}(0), B = \mathbf{x}(1)$ where $\mathbf{x} = (x, y, z)$.

$$\Phi[\mathbf{x}, \lambda] = \int_0^1 \sqrt{\dot{x}^2 + \dot{y}^2 + \dot{z}^2} - \lambda(t) \cdot g(x, y, z)\mathrm{d}t$$

Note: the lagrange multiplier λ is now a function of t as we want the entire curve to lie on Σ.

E-L equations with h

- Variation w.r.t λ:

$$\underbrace{\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial h}{\partial \dot{\lambda}}\right)}_{0} - \frac{\partial h}{\partial \lambda} = 0 \Rightarrow g(x, y, z) = 0 \ \forall t$$

- Variation w.r.t x_i:

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\dot{x}_i}{\sqrt{\dot{x}_1^2 + \dot{x}_2^2 + \dot{x}_3^2}}\right) + \lambda \frac{\partial g}{\partial x_i} = 0$$

Alternatively, solve the constraint $g = 0$, as we did in example 2.2 ($\Sigma =$ sphere).

3.3 Several independent variables

In general $\Phi : \mathbb{R}^n \to \mathbb{R}^m$. In $n > 1$, E-L becomes PDEs. Assume that $n = 3, m = 1$.

$$F[\Phi] = \iiint_D f(\underbrace{x, y, z}_{\text{indpendent}}, \Phi, \Phi_x, \Phi_y, \Phi_z)\mathrm{d}x\mathrm{d}y\mathrm{d}z$$

Notation: $\Phi_x = \frac{\partial \Phi}{\partial x}$ etc.

Assume Φ is extremised, consider pertubations $\Phi \to \Phi(x,y,z) + \varepsilon\eta(x,y,z)$ s.t $\eta = 0$ on ∂D.

$$F[\Phi+\varepsilon\eta] - F[\Phi] = \varepsilon \int_D \left\{ \eta \frac{\partial f}{\partial \Phi} + \eta_x \frac{\partial f}{\partial \Phi_x} + \eta_y \frac{\partial f}{\partial \Phi_y} + \eta_z \frac{\partial f}{\partial \Phi_z} \right\} \mathrm{d}x\mathrm{d}y\mathrm{d}z + O(\varepsilon^2)$$

$$= \varepsilon \int_D \left\{ \eta \frac{\partial f}{\partial \Phi} + \nabla \cdot \left(\eta \left(\frac{\partial f}{\partial \Phi_x}, \frac{\partial f}{\partial \Phi_y}, \frac{\partial f}{\partial \Phi_z} \right) \right) - \eta \nabla \cdot \left(\frac{\partial f}{\partial \Phi_x}, \frac{\partial f}{\partial \Phi_y}, \frac{\partial f}{\partial \Phi_z} \right) \right\} \mathrm{d}x\mathrm{d}y\mathrm{d}z + O(\varepsilon^2)$$

Can apply divergence theorem to the middle term, and use

$$\int_{\partial D} \eta \left(\frac{\partial f}{\partial \Phi_x}, \frac{\partial f}{\partial \Phi_y}, \frac{\partial f}{\partial \Phi_z} \right) \cdot \mathrm{d}\mathbf{S} = 0$$

$$F[\Phi+\varepsilon\eta] - F[\Phi] = \varepsilon \int_D \left\{ \eta \frac{\partial f}{\partial \Phi} - \eta \nabla \cdot \left(\frac{\partial f}{\partial \Phi_x}, \frac{\partial f}{\partial \Phi_y}, \frac{\partial f}{\partial \Phi_z} \right) \right\} \mathrm{d}x\mathrm{d}y\mathrm{d}z + O(\varepsilon^2)$$

E-L equation: single 2nd order PDE for 1 function Φ.

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

(3.4) is also true for general n instead of 3.

Example 3.4 (Exremize 'potential energy')**.** Using $n = 2$

$$F[\Phi] = \iint_{D \subset \mathbb{R}^2} \frac{1}{2} \left[\Phi_x^2 + \Phi_y^2 \right] \mathrm{d}x\mathrm{d}y$$

$$\frac{\partial F}{\partial \Phi} = 0, \frac{\partial f}{\partial \Phi_x} = \Phi_x, \frac{\partial f}{\partial \Phi_y} = \Phi_y$$

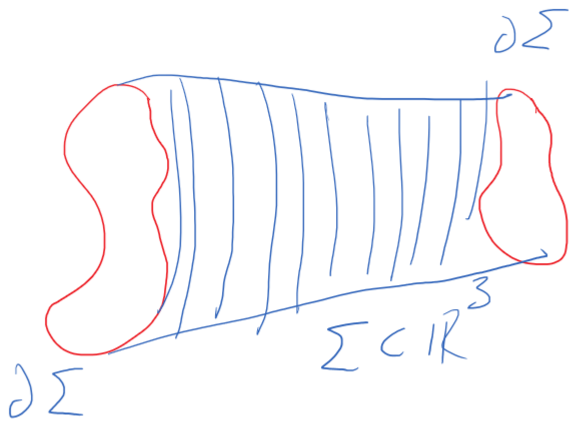
(3.4) implies

$$\frac{\partial}{\partial x}\Phi_x + \frac{\partial}{\partial y}\Phi_y = 0 = \Phi_{xx} + \Phi_{yy}$$

This is the Laplace equation.

Example 3.5 (Minimal surfaces)**.** Minimise the area of $\Sigma \subset \mathbb{R}^3$ subject to boundary conditions

$$\Sigma = \{ \mathbf{x} \in \mathbb{R}^3 : k(x,y,z) = 0 \}$$



Assume (can do it by implicit function theorem) that we solved $k = 0$ to give $z = \phi(x, y)$.

$$\mathrm{d}s^2 = \mathrm{d}x^2 + \mathrm{d}y^2 + \mathrm{d}z^2,\ \mathrm{d}z = \phi_x \mathrm{d}x + \phi_y \mathrm{d}y$$

$$\mathrm{d}s^2 = (1 + \phi_x^2)\mathrm{d}x^2 + (1 + \phi_y^2)\mathrm{d}y^2 + 2\phi_x\phi_y \mathrm{d}x\mathrm{d}y$$

[IB geometry, this is called the 1st fundamental form, or Rimemannian metric]

$$\mathrm{d}s^2 = \sum_{i,j=1}^{2} g_{ij}(x, y)\mathrm{d}x_i \mathrm{d}x_j$$

$$g = \begin{pmatrix} 1 + \phi_x^2 & \phi_x\phi_y \\ \phi_x\phi_y & 1 + \phi_y^2 \end{pmatrix}$$

Area element $= \sqrt{\det g}\mathrm{d}x\mathrm{d}y$

Area functional

$$A[\phi] = \int_D \underbrace{\sqrt{1 + \phi_x^2 + \phi_y^2}}\, h\mathrm{d}x\mathrm{d}y$$

Apply (3.4) to h:

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

$$\partial_x \left(\frac{\phi_x}{\sqrt{1 + \phi_x^2 + \phi_y^2}} \right) + \partial_y \left(\frac{\phi_y}{\sqrt{1 + \phi_x^2 + \phi_y^2}} \right) = 0$$

Expanding derivatives:

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

This is the minimal surface equation.

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

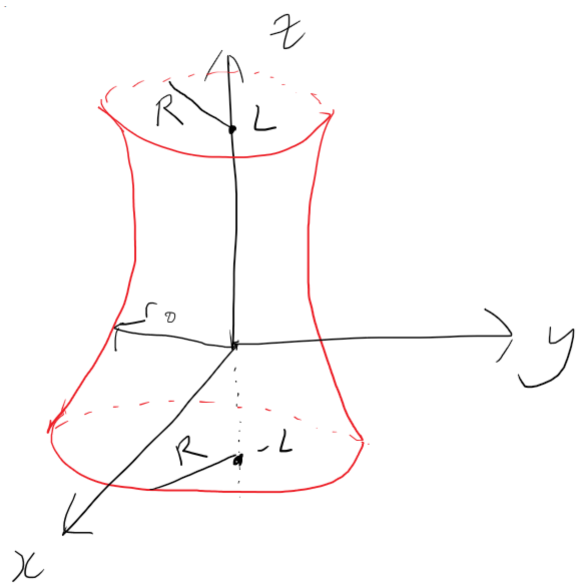
$$\phi_x = \frac{\mathrm{d}z}{\mathrm{d}r}\frac{\partial r}{\partial x} = \ \frac{x}{r}, \phi_y = z' z \frac{y}{r}$$

Then (3.5) implies $rz'' + z' + (z')^3 = 0$.
Set $z' = w$, then

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

This has solution

$$r = r_0 \cosh\left(\frac{z - z_0}{r_0}\right)$$

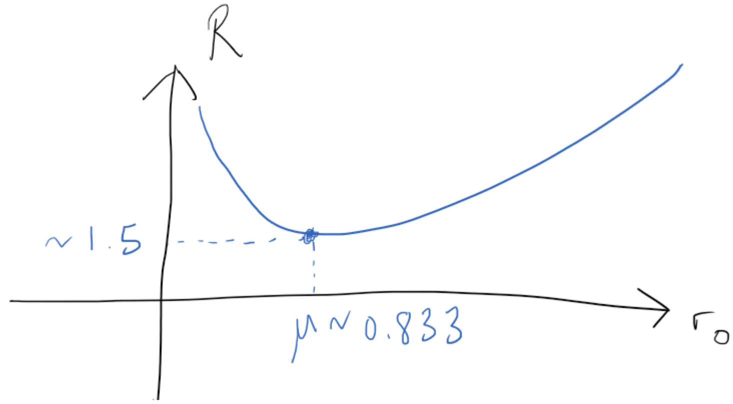


This is a catenoid: minimal surface of revolution

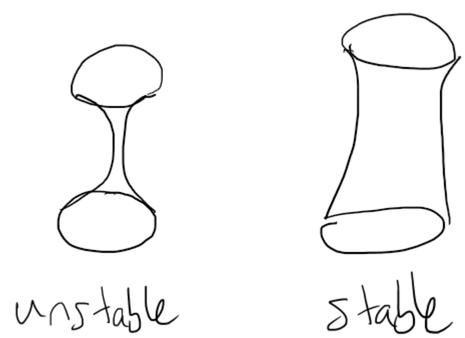
$r(L) = r(-L)$. If $L \neq -$, then $z_0 = 0$. Set $r = R$, and divide by L:

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

Set $L = 1$. $R = r_0 \cosh(1/r_0)$.



If $R > 1.5$, then there are two possible minimal surfaces with different r_0:



3.4 Higher Derivatives

$$F[y] = \int_\alpha^\beta f(x, y, y', \dots, y^{(n)})\mathrm{d}x$$

Proceed as in section 2. Assume a maximising y exists, $y \to y + \varepsilon\eta$, where $\eta = \eta' = \ldots = \eta^{(n-1)} = 0$ at α, β.

$$F[y+\varepsilon\eta] - F[y] = \varepsilon \int_\alpha^\beta \left(\frac{\partial f}{\partial y}\eta + \frac{\partial f}{\partial y'}\eta' + \ldots + \frac{\partial f}{\partial y^{(n)}}\eta^{(n)} \right) \mathrm{d}x + O(\varepsilon^2)$$

Integrating by parts and applying fundamental lemma:

$$\frac{\partial f}{\partial y} - \frac{\mathrm{d}}{\mathrm{d}x}\frac{\partial f}{\partial y'} + \frac{\mathrm{d}^2}{\mathrm{d}x^2}\frac{\partial f}{\partial y''} + \ldots + (-1)^n \frac{\mathrm{d}^n}{\mathrm{d}x^n}\frac{\partial f}{\partial y^{(n)}} = 0 \tag{3.6}$$

Euler-Lagrange equation.

Example 3.6. If $n = 2$ and $\frac{\partial f}{\partial y} = 0$, (3.6) implies

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\partial f}{\partial y'} - \frac{\mathrm{d}}{\mathrm{d}x}\frac{\partial f}{\partial y''}\right) = 0$$

Example 3.7. Extremize

$$F[y] = \int_0^1 (y'')^2 \mathrm{d}x$$

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

$$\frac{\mathrm{d}}{\mathrm{d}x}(2y'') = const \implies y''' = k = const$$

Imposing boundary conditions gives $y = x^3 - x^2$.
Note: this is an absolute minimum. Consider $y_0 = x^3 - x^2$, $\eta(0) = \eta'(0) = \eta(1) = \eta'(1) = 0$ [do not assume η small].

$$F[y_0 + \eta] - F[y_0] = \int_0^1 (\eta'')^2 \mathrm{d}x + 2\int_0^1 (y_0''\eta'')\mathrm{d}x$$

$$> 4\int_0^1 (3x-1)\eta'' \mathrm{d}x = 4\left([-\eta']_0^1 + \int_0^1 \left(\frac{\mathrm{d}}{\mathrm{d}x}(2x\eta') - \eta'\right)\mathrm{d}x\right)$$

$$= 4\left([3x\eta']_0^1 - 3\eta|_0^1\right) = 0$$

So y_0 is the absolute minimizer of F.

4 Least action principle & Noether's Theorem

Particle in $\mathbb{R}^3$, T = kinetic energy, V = potential energy.

$$L(\mathbf{x}, \dot{\mathbf{x}}, t) = T - V \text{ (Lagrangian)} \tag{4.1}$$

t = independent variable, $\mathbf{x} = (x, y, t)$ dependent variable.

$$\text{Action } S[\mathbf{x}] = \int_{t_1}^{t_2} L\mathrm{d}t \tag{4.2}$$

Hamilton's principle (least action principle):

The motion is such that $S[\mathbf{x}]$ is stationary, i.e L satisfies the E-L equations.

Example 4.1.

$$T = \frac{1}{2}m|\dot{\mathbf{x}}|^2, \ V = V(\mathbf{x})$$

E-L:

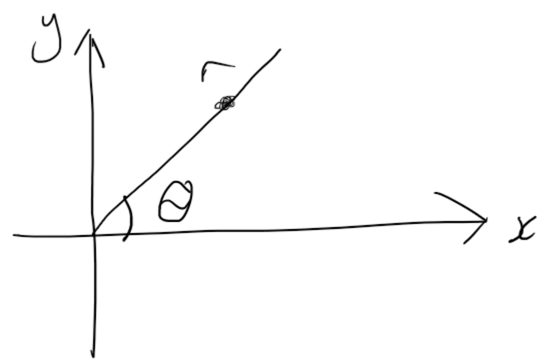
$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{x}_i} = \frac{\partial L}{\partial x_i}$$

$$m\ddot{x}_i = -\frac{\partial V}{\partial x_i}, \text{ or } m\ddot{\mathbf{x}} = -\nabla V$$

This is Newton's 2nd Law.

Example 4.2 (Central force in 2 dimensions)**.**

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



E-L gives

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{r}} - \frac{\partial L}{\partial r} = 0$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{\theta}} - \underbrace{\frac{\partial L}{\partial \theta}}_{0} = 0 \implies \frac{\partial L}{\partial \dot{\theta}} = mr^2\dot{\theta} = const$$

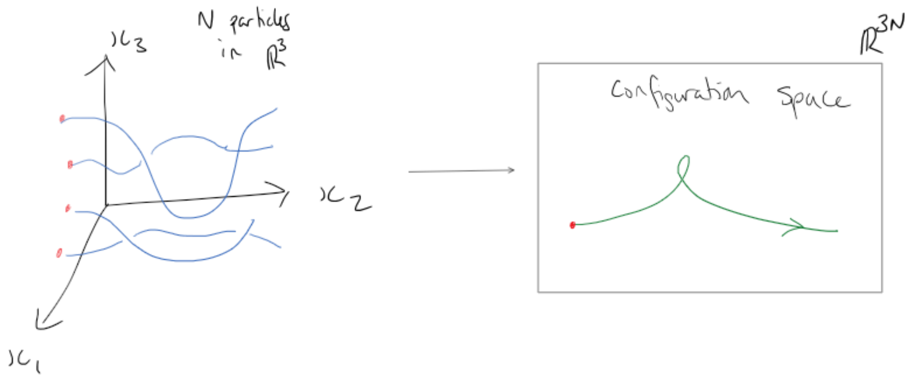
i.e conservation of angular momentum. Since $\frac{\partial L}{\partial t} = 0$, use (2.5) to give

$$\dot{r}\frac{\partial L}{\partial \dot{r}} + \dot{\theta}\frac{\partial L}{\partial \dot{\theta}} - L = const$$

$$\implies \underbrace{\frac{1}{2}m\dot{r}^2 + \frac{1}{2}mr^2\dot{\theta}^2}_{T} + V(r) = E = const$$

i.e conservation of total energy.

Example 4.3 (Configuration space and generalised coordinates)**.** N particles in $\mathbb{R}^3$



The configuration map is a map $t \mapsto \{q_i(t), \dot{q}_i(t), t\}$ for $i = 1, \ldots, 3N$ where q_i is the position (This is motivation for Part II Classical Dynamics).

4.1 Noether's Theorem

$$F[y] = \int_\alpha^\beta f(y_i, y_i', x)\mathrm{d}x, \ i = 1, \ldots, n$$

Suppose there exists a 1-parameter family of transformations $y_i(x) \mapsto Y_i(x, s)$ with $Y_i(x, 0) = y_i(x)$ and $s \in \mathbb{R}$.

This is called a continuous symmetry (or symmetry) of a Lagrangian f, if

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

Theorem 4.1 (Noether's Theorem)**.** *Given a continuous summetry $Y_i(x, s)$ of f, the quantity*

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

is a first integral of the E-L equation with $Y_i(x, 0) = y_i(x)$.

Proof. Using summation convention on the i's:

$$\begin{aligned}
0 = \frac{\mathrm{d}}{\mathrm{d}s} f\bigg|_{s=0} &= \frac{\partial f}{\partial y_i} \frac{\mathrm{d}Y_i}{\mathrm{d}s}\bigg|_{s=0} + \frac{\partial f}{\partial y_i'} \frac{\mathrm{d}Y_i'}{\mathrm{d}s}\bigg|_{s=0} \\
&= \left[\frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\partial f}{\partial y_i'}\right) \frac{\mathrm{d}Y_i}{\mathrm{d}s} + \frac{\partial f}{\partial y_i'} \frac{\mathrm{d}}{\mathrm{d}x} \frac{\mathrm{d}Y_i}{\mathrm{d}s}\right]\bigg|_{s=0} \\
&= \frac{\mathrm{d}}{\mathrm{d}x}\left[\frac{\partial f}{\partial y_i'} \frac{\partial Y_i}{\partial s}\bigg|_{s=0}\right.
\end{aligned}$$

□

Example 4.4.

$$f = \frac{1}{2}(y')^2 + \frac{1}{2}(z')^2 - V(y-z),\ \mathbf{y} = (y,z)$$

Lagrangian of a particle moving on a plane, in a potential.

$$Y = y+s, Z = z+s, Y' = y', Z' = z', V(Y-Z) = V(y-z)$$

so $\frac{\mathrm{d}}{\mathrm{d}s} f = 0$. (4.3) implies

$$\left(\frac{\partial f}{\partial y'}\frac{\mathrm{d}Y}{\mathrm{d}s} + \frac{\partial f}{\partial z'}\frac{\mathrm{d}Z}{\mathrm{d}s}\right) = y' + z'$$

i.e conservation of momentum in the $y+z$ direction.

Example 4.5. Back to example 4.2, $\Theta = \theta + s, R = r$

$\frac{\mathrm{d}L}{\mathrm{d}s} = 0$, (4.3) implies

$$\left.\frac{\partial L}{\partial \dot\theta}\frac{\partial \theta}{\partial s} + \frac{\partial L}{\partial \dot r}\frac{\partial R}{\partial s}\right|_{s=0} = mr^2\dot\theta$$

i.e conservation of angular momentum. Hence

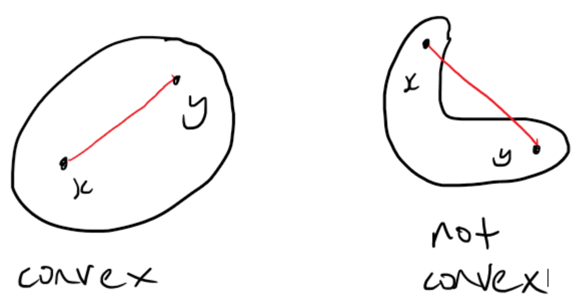
Isotropy of space $\implies$ rotational invariance of L $\implies$ conservation of angular momentum

5 Convex functions

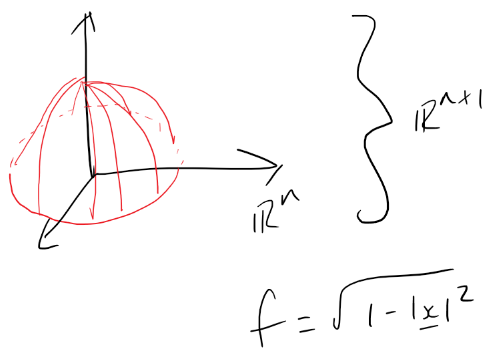
Going back to calculus on $\mathbb{R}^n$, we look at a class of functions for which it is easy to classify stationary points.

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

$$(1-t)\mathbf{x} + t\mathbf{y} \in S,\ \forall t \in [0,1]$$



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



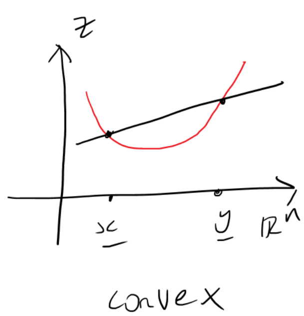
Chord of f = line segment in $\mathbb{R}^{n+1}$ joining two points on the graph.

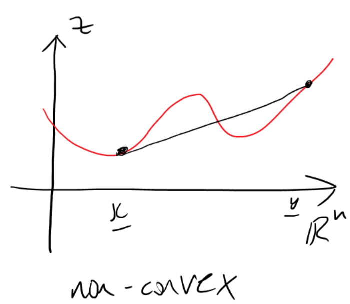
Definition. A function $f : \mathbb{R}^{n+1} \to \mathbb{R}$ is convex if

(i) The domain of f is a convex set

(ii)
$$f((1-t)\mathbf{x} + t\mathbf{y}) \leq (1-t)f(\mathbf{x}) + tf(\mathbf{y}), \ \forall t \in [0,1] \tag{5.1}$$

f is convex if the graph of f lies below or on its chords.





Remarks

1. f is called concave if (5.1) holds with $\leq$ replaced by $\geq$
2. f convex $\iff$ $-f$ concave
3. f is called strictly convex if (5.1) holds with $\leq$ replaced by $<$

Example 5.1. $f\mathbb{R} \to \mathbb{R}$, $f(x) = x^2$. Domain $\mathbb{R}$ is convex and

$$\begin{aligned}
&f((1-t)x + ty) - (1-t)f(x) - tf(y) \\
&= [(1-t)x + ty]^2 - (1-t)x^2 - ty^2 \\
&= x^2(1-t)(-t) + t(1-t)y^2 + 2(1-t)txy \\
&= -(1-t)t(x-y)^2 < 0
\end{aligned}$$

For all $t \in (0,1)$. So f is strictly convex.

Example 5.2. $f(x) = \frac{1}{x}$, domain $\mathbb{R} \setminus \{0\}$ which is not convex. However, on the restricted domain $\mathbb{R}_{x \geq 0}$, f is convex.

5.1 Conditions for convexity

3 tests for f to be convex:

1. If f is once differentiable, then f is convex iff

$$f(\mathbf{y}) \geq f(\mathbf{x}) + (\mathbf{y} - \mathbf{x}) \cdot \nabla f(\mathbf{x}) \tag{5.2}$$

Proof. Assume (5.2) holds, and apply it twice:

$$f(\mathbf{x}) \geq f(\mathbf{z}) + (\mathbf{x} - \mathbf{z}) \cdot \nabla f(\mathbf{z})$$

$$f(\mathbf{y}) \geq f(\mathbf{z}) + (\mathbf{y} - \mathbf{z}) \cdot \nabla f(\mathbf{z})$$

Take $\mathbf{z} = (1-t)\mathbf{x} + t\mathbf{y} \in S$ (the domain of f) for $t \in (0,1)$. Now adding $(1-t)$ times the first equation to t times the second we get (5.1).

Now to prove the converse, assume (5.1) and set

$$h(t) = (1-t)f(\mathbf{x}) + t(f(\mathbf{y})) - f((1-t)\mathbf{x} + t\mathbf{y}) \geq 0$$

for $t \in (0,1)$

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

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

Note $h(0) = 0$, so $\frac{h(t) - h(0)}{t} \geq 0$ for $t \in (0,1)$. Now take the limit $t \to 0$. $\square$

A quick corollary of this is

Corollary 5.1. *If f is convex and has a stationary point, then it is a global minimum*

Proof. If $\nabla f(\mathbf{x}_0) = 0$ then (5.2) implies $f(\mathbf{y}) \geq f(\mathbf{x}_0)$ for all $\mathbf{y}$. □

2. If
$$(\nabla f(\mathbf{y}) - \nabla f(\mathbf{x})) \cdot (\mathbf{y} - \mathbf{x}) \geq 0 \tag{5.3}$$
Then f is convex

Proof. Exercise □

3. (Second order conditions) Assume f is twice differentiable, then f is convex iff the Hessian $\frac{\partial^2 f}{\partial x_i \partial x_j}$ has all eigenvalues non-negative. If all eigenvalues are positive, f is strictly convex.

Proof. Assume convex and apply (5.3) by taking $\mathbf{y} = \mathbf{x} + \mathbf{h}$
$$\mathbf{h} \cdot (\nabla f(\mathbf{x} + \mathbf{h}) - \nabla f(\mathbf{y})) \geq 0$$
For small $\mathbf{h}$
$$\partial_i f(\mathbf{x} + \mathbf{h}) = \partial_i f(\mathbf{x}) + \sum_j h_j H_{ij}(\mathbf{x}) + O(|\mathbf{h}|^2)$$
so (by dotting with $\mathbf{h}$)
$$\sum_{i,j} h_i h_j H_{ij}(\mathbf{x}) + O(|\mathbf{h}|^2) \geq 0$$
The converse is left as an exercise. □

Example 5.3. $f(x, y) = \frac{1}{xy}$, domain $\{(x, y) : x, y > 0\}$.
$$H = \frac{1}{xy} \begin{pmatrix} \frac{2}{x^2} & \frac{1}{xy} \\ \frac{1}{xy} & \frac{2}{y^2} \end{pmatrix}$$
$\det(H) = \frac{3}{x^3 y^3} > 0$, $\mathrm{Tr}(H) > 0$ so f is strictly convex.

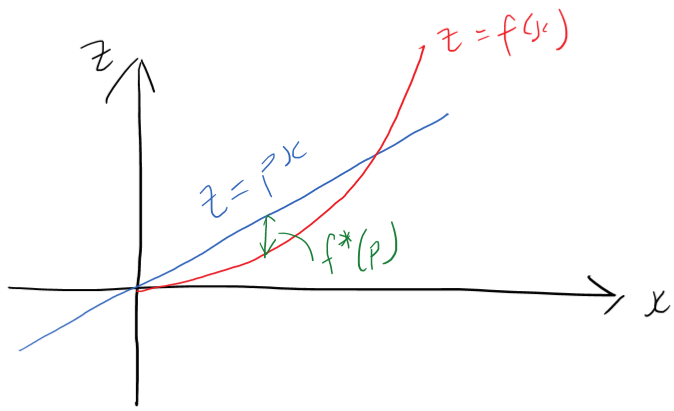
6 Legendre transform

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

$$f^*(\mathbf{p}) = \sup_{\mathbf{x}}(\mathbf{p} \cdot \mathbf{x} - f(\mathbf{x})) \tag{6.1}$$

The domain of f^* consists of all vectors $\mathbf{p} \in \mathbb{R}^n$ such that the supremum is finite.

Example 6.1. $n = 1$



Example 6.2. $n = 1$, $f(x) = ax^2$, $a > 0$

$$f^*(p) = \sup_x(px - ax^2)$$

$$\frac{\partial}{\partial x}(px - ax^2) = 0 \implies p = 2xa$$

So $x = \frac{p}{2a}$ and substitute so

$$f^*(p) = \frac{p^2}{4a}$$

Can compute

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

$$f^{**}(s) = as^2$$

Turns out that $f^{**} = f$ is true for any convex f.

If $a < 0$, then $\sup(px - ax^2) = \infty$ for all p, so f^* has empty domain.

Proposition. *Domain of f^* is a convex set, and f^* is convex.*

Proof.

$$f^*((1-t)\mathbf{p}+t\mathbf{q}) = \sup_{\mathbf{x}}\left[(1-t)\mathbf{p}\cdot\mathbf{x}+t\mathbf{q}\cdot\mathbf{x}-f(\mathbf{x})\right]$$

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

Use $\sup(A+B) \leq \sup(A)+\sup(B)$ so

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

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

Hence $(1-t)\mathbf{p}+t\mathbf{q}$ is in the domain of f^* and f^* satisfies convexity definition (5.1).

□

In practice, if f is convex and differentiable

$$f^*(p) = \text{ global minimum over } \mathbf{x}$$

$$\nabla(\mathbf{p}\cdot\mathbf{x}-f(\mathbf{x})) = 0 \implies \mathbf{p} = \nabla f$$

(Substitute into the definition of $f^*(p)$). If f strictly convex, then there is a unique inversion $\mathbf{x} = \mathbf{x}(\mathbf{p})$ of $\mathbf{p} = \nabla f$ so that

$$f^*(\mathbf{p}) = \mathbf{p}\cdot\mathbf{x}(\mathbf{p}) - f(\mathbf{x}(\mathbf{p})) \tag{6.2}$$

6.1 Applications to Thermodynamics

Many particles (gas $\sim 10^{23}$ particles) with a few macroscopic variables: p (pressure), V (volume), T (temperature), S (entropy) [Part II Statistical Physics].

Internal energy $U(S,V)$. Hermholtz free energy is given by

$$\begin{aligned} F(T,V) &= \min_S (U(S,V) - TS) \\ &= -\max_S (TS - U(S,V)) = -u^*(T,V) \end{aligned}$$

Legendre transform of U w.r.t S, with V held fixed as a parameter.

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

Other quantities as Legendre transform: e.g Enthalpy

$$H(S,p) = \min_V (U(S,V) + pv) = -U^*(-p,S)$$

At min $p = -\left(\frac{\partial U}{\partial V}\right)\big|_S$

Legendre transform: a way to swap from (S,V) dependence, to dependence of other variables e.g $U(S,V)$ can become $F(T,V)$ or $H(S,p)$.

7 Hamilton's Equations

Recall (from section 4.1) the Lagrangian $L = T - V = L(\mathbf{q}, \dot{\mathbf{q}}, t)$ which is a function on the configuration space.

Hamiltonian = Legendre transform of L w.r.t $\dot{\mathbf{q}} = \mathbf{v}$ (velocity)

$$H(\mathbf{q},\mathbf{p},t) = \sup_{\mathbf{v}}(\mathbf{p}\cdot\mathbf{v} - L) = \mathbf{p}\cdot\mathbf{v} - L(\mathbf{q},\mathbf{v},t) \tag{7.1}$$

Where $\mathbf{v} = \mathbf{v}(\mathbf{p})$ is the solution to $p_i = \frac{\partial L}{\partial \dot{q}_i}$ (assuming convexity of L in $\mathbf{v}$)

$\mathbf{p}$ = generalised momenta.

Example 7.1.

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

$$\mathbf{p} = \frac{\partial L}{\partial \dot{\mathbf{q}}} = m\dot{\mathbf{q}} \implies \dot{\mathbf{q}} = \frac{\mathbf{p}}{m}$$

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

What happened to the Euler-Lagrange equations?

$$H = H(\mathbf{q}, \mathbf{p}, t) = p_i \dot{q}_i - L(q_i, \dot{q}_i, t)$$

$$\mathrm{d}H = \frac{\partial H}{\partial q_i}\mathrm{d}q_i + \frac{\partial H}{\partial p_i}\mathrm{d}p_i + \frac{\partial H}{\partial t}\mathrm{d}t$$

$$= p_i \mathrm{d}\dot{q}_i + \dot{q}_i \mathrm{d}p_i - \frac{\partial L}{\partial q_i}\mathrm{d}q_i - \frac{\partial L}{\partial \dot{q}_i}\mathrm{d}\dot{q}_i - \frac{\partial L}{\partial t}\mathrm{d}t$$

Since $\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{q}_i} = \dot{p}_i$, $\frac{\partial L}{\partial \dot{q}_i} = p_i$ and so

$$\mathrm{d}H = \dot{q}_i \mathrm{d}p_i - \dot{p}_i \mathrm{d}q_i - \frac{\partial L}{\partial t}\mathrm{d}t$$

Comparing differentials

$$\dot{q}_i = \frac{\partial H}{\partial p_i},\ \dot{p}_i = -\frac{\partial H}{\partial q_i},\ \frac{\partial H}{\partial t} = -\frac{\partial L}{\partial t} \tag{7.2}$$

Careful: $\frac{\partial}{\partial t}\Big|_{p,q} \neq \frac{\partial}{\partial t}\Big|_{q,\dot{q}}$

Assume no explicit t-dependence in L. Then (7.2) is a system of $2n$ 1st order ODEs. Need to specify $\mathbf{q}(0), \mathbf{p}(0)$.

Solution curves to (7.2) = trajectory in $2n$-dimensional phase space.

Remark: Hamilton's equations also arise from extremizing a functional in phase space:

$$S[\mathbf{q}, \mathbf{p}] = \int_{t_1}^{t_2} \underbrace{\{\dot{q}_i p_i - H(\mathbf{q}, \mathbf{p}, t)\}}_{f(\mathbf{q}, \mathbf{p}, \dot{\mathbf{q}}, \dot{\mathbf{p}}, t)} \mathrm{d}t$$

E-L for S

Variations w.r.t p_i:

$$\frac{\partial f}{\partial p_i} - \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial f}{\partial \dot{p}_i} = 0 \implies \dot{q}_i = \frac{\partial H}{\partial p_i}$$

Variations w.r.t q_i:

$$\frac{\partial f}{\partial q_i} - \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial f}{\partial \dot{p}_i} = 0 \implies -\frac{\partial H}{\partial q_i} - \frac{\mathrm{d}}{\mathrm{d}t}(p_i) = 0 \implies \dot{q}_i = -\frac{\partial H}{\partial q_i}$$

So we have recovered (7.2).

So Newtons eq, Lagrange's eq, Hamilton's eq are all equivalent. So far, (7.2) is just another formalism.

$\sim$ 1926. Paul Dirac, walking to Granchester, noticed the Hamiltonian was a bridge between classical and quantum physics [Part II PQM].

8 The Second Variation

E-L equation: minimum/maximum/saddle point?

Look at the nature of the stationary points of

$$F[y] = \int_\alpha^\beta f(x, y, y')\mathrm{d}x$$

Expanding $F[y + \varepsilon\eta]$ to 2nd order in ε, around a solution to the E-L equation:

$$F[y + \varepsilon\eta] - F[y] = \int_\alpha^\beta \left[f(x, y + \varepsilon\eta, y' + \varepsilon\eta') - f\right]\mathrm{d}x$$

$$= 0 + \varepsilon \underbrace{\int_\alpha^\beta \eta\left(\frac{\partial f}{\partial y} - \frac{\mathrm{d}}{\mathrm{d}x}\frac{\partial f}{\partial y'}\right)\mathrm{d}x}_{0 \text{ by E-L}} + \frac{\varepsilon^2}{2}\int_\alpha^\beta \left[\eta^2\frac{\partial^2 f}{\partial y^2} + (\eta')^2\frac{\partial f}{\partial (y')^2} + 2\frac{\partial^2 f}{\partial y \partial y'}\eta\eta'\right]\mathrm{d}x$$

2nd variation:

$$\delta^2 F[y] \equiv \frac{1}{2}\int_\alpha^\beta \left[\eta^2\frac{\partial^2 f}{\partial y^2} + (\eta')^2\frac{\partial^2 f}{\partial (y')^2} + \frac{\mathrm{d}}{\mathrm{d}x}(\eta^2)\frac{\partial^2 f}{\partial y \partial y'}\right]\mathrm{d}x$$

Integrating the last term by parts and using $\eta = 0$ at α and β:

$$\delta^2 F[y] = \frac{1}{2}\int_\alpha^\beta \left\{Q\eta^2 + P(\eta')^2\right\}\mathrm{d}x \tag{8.1}$$

Where $P = \frac{\partial^2 f}{\partial (y')^2}$, $Q = \frac{\partial^2 f}{\partial y^2} - \frac{\mathrm{d}}{\mathrm{d}x}\frac{\partial^2 f}{\partial y \partial y'}$

We proved:

Proposition. *If $y(x)$ is a solution to the E-L equation (2.3) and $Q\eta^2 + P(\eta')^2 > 0$ for all η vanishing at α, β then $y(x)$ is a local minimizer of $F[y]$.*

Example 8.1 (Geodesics on a plane)**.**

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

If $\eta' = 0$ then $\eta = 0$, so $\eta \neq 0$ and $P(\eta')^2 > 0$ for all η. So straight lines are local length minimizers on $\mathbb{R}^2$.

Proposition. *If $y_0(x)$ is a local minimum, then*

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

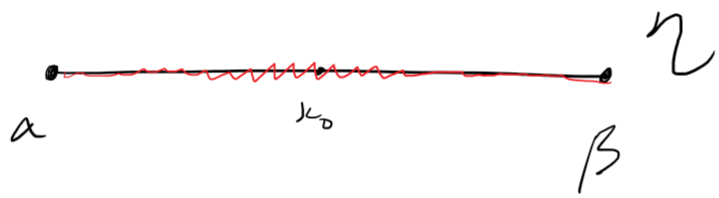
This is called the Legendre condition.

Proof. See Gelfund-Fomin for details.

Idea: if η' is small, then η can't be too large.

The converse is not true: η can be small, η' large.

Assume $\exists x_0$ such that $P(x_0, y_0, y_0') < 0$



$\square$

Note: (8.2) not sufficient for local minimum but $P > 0$, $Q \geq 0$ is sufficient as if $\eta \neq 0$ on (α, β), then $\exists x_0 \in (\alpha, \beta)$ such that $\eta'(x_0) \neq 0$.

Example 8.2. Go back to brachistochrome

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

Is the cycloid a local minimizer?

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

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

$$Q = \ldots = \frac{1}{2\sqrt{1 + (y')^2} y^2 \sqrt{-y}} > 0$$

So the cycloid is indeed a local minimizer.

8.1 Associated Eigenvalue Problem

Go back to (8.1)

$$Q\eta^2 + P(\eta')^2 = Q\eta^2 + \underbrace{\frac{\mathrm{d}}{\mathrm{d}x}(P\eta\eta')}_{\text{boundary term}} - \eta(P\eta')'$$

Integrate, drop the boundary term as $\eta = 0$ at α, β:

$$\delta^2 F[y_0] = \frac{1}{2}\int_\alpha^\beta \eta \underbrace{\left[-(P\eta')' + Q\eta\right]}_{\mathcal{L}(\eta)} \mathrm{d}x \tag{8.3}$$

Where $\mathcal{L}(\eta)$ is the Sturm-Liouville operator (compare with 3.2).

If $\exists \eta$ such that

$$\mathcal{L}(\eta) = -\omega^2\eta \text{ (ω real) and } \eta(\alpha) = \eta(\beta) = 0 \tag{8.4}$$

then y_0 is not a minimizer as

$$\delta^2 F[y_0] = -\frac{1}{2}\omega^2 \int_\alpha^\beta \eta^2 \mathrm{d}x < 0$$

There can be such η even if $P > 0$, so the Legendre condition (8.2) is not sufficient for y_0 to be a minimizer.

Example 8.3.

$$F[y] = \int_0^\beta \left\{(y')^2 - y^2\right\} \mathrm{d}x$$

Where $y(0) = y(\beta) = 0$ and $\beta \neq N\pi$ for any $N \in \mathbb{N}$

(2.3) implies $y'' + y = 0$ so $y = y_0 \equiv 0$ is the stationary point of $F[y]$.

2nd variation:

$$\delta^2 F[0] = \frac{1}{2}\int_0^\beta \left[(\eta')^2 - \eta^2\right] \mathrm{d}x$$

$P = 1 > 0$, but $Q < 0$. Examine (8.4):

$$-\eta'' - \eta = -\omega^2\eta$$

Take $\eta = A\sin\left(\frac{\pi x}{\beta}\right) \Rightarrow \left(\frac{\pi}{\beta}\right)^2 = 1 - \omega^2$ which is possible if $\beta > \pi$.

So, if $P > 0$ a problem may arise if the interval is ”too large”.

8.2 The Jacobi Condition

Legendre tried to prove that $P > 0$ is sufficient for $y = y_0$ to be a minimum. This couldn't have worked (see the last example), but the idea was good.

Let $\Phi = \Phi(x)$ be any differentiable function of x on $[\alpha, \beta]$. As $\eta(\alpha) = \eta(\beta) = 0$:

$$0 = \int_\alpha^\beta (\Phi\eta^2)' \mathrm{d}x = \int_\alpha^\beta \left\{\Phi'\eta^2 + 2\eta\eta'\Phi\right\} \mathrm{d}x$$

Rewrite (8.1) as

$$\delta^2 F[y] = \frac{1}{2}\int_\alpha^\beta \left\{P(\eta')^2 + 2\eta\eta'\Phi + (Q + \Phi')\eta^2\right\} \mathrm{d}x$$

Assume $P_y > 0$, and complete the square:

$$\delta^2 F[y] = \frac{1}{2}\int_\alpha^\beta \left[P\left(\eta' + \frac{\Phi}{P}\eta\right)^2 + \left(Q + \Phi' - \frac{\Phi^2}{P}\right)\eta^2\right] \mathrm{d}x$$

Which is positive* if we can choose Φ such that

$$\Phi^2 = P(Q + \Phi') \tag{8.3}$$

Does a solution to (8.3) exist on $[\alpha, \beta]$? We can transform it into a linear 2nd order ODE. Set $\Phi = -P\frac{u'}{u}$, where $u \neq 0$ on $[\alpha, \beta]$.

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

Or

$$-(Pu')' + Qu = 0 \tag{8.4}$$

This is the Jacobi accessory condition.

Need a solution to (8.4) [which can be written as $\mathcal{L}(u) = 0$] such that $u \neq 0$ on $[\alpha, \beta]$. This may not exist on a large enough interval.

Example 8.4.

$$F[y] = \frac{1}{2}\int_\alpha^\beta \left[(y')^2 - y^2\right] \mathrm{d}x$$

$y \to y + \varepsilon\eta$ gives:

$$\delta^2 F[y] = \frac{1}{2}\int_\alpha^\beta \left[(\eta')^2 - \eta^2\right] \mathrm{d}x$$

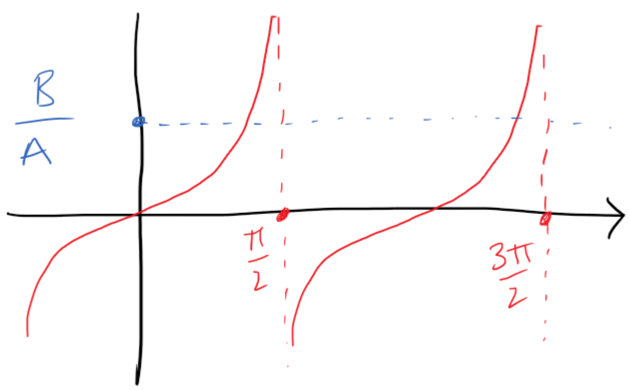
*If (8.3) holds then $\delta^2 F > 0$ unless $\eta' + \frac{\Phi}{P}\eta = 0$ on $[\alpha, \beta]$. But $\eta = 0$ at α, so $\eta'(\alpha) = 0$, and then $\eta \equiv 0$ (by uniqueness of solution to 1st order ODEs).

$P = 1$ and $Q = -1$ so (8.4) is $u'' + u = 0$, which has general solution

$$u = A\sin(x) - B\cos(x)$$

We want u to be non-zero on $[\alpha, \beta]$, i.e

$$\tan(x) \neq \frac{B}{A}$$

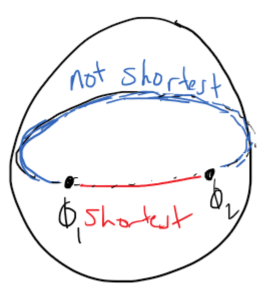


So it is possible to 'avoid' B/A on an interval smaller than π. So if $|\beta - \alpha| < \pi$ then y is a positive 2nd minimizer.

Example 8.5. Back to geodesics on the sphere.

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

We saw earlier that critical points are segments of great circles.



$\theta = const$, $\theta_0 = \frac{\pi}{2}$ (this is any great circle up to rotation).

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

$$\delta^2 F[\theta - 0 = \frac{\pi}{2}; \eta] = \frac{1}{2} \int_{\phi_1}^{\phi_2} \left[(\eta')^2 - \eta^2 \right] \mathrm{d}\phi$$

Which is positive if $\phi_2 - \phi_1 < \pi$.