Variational Principles

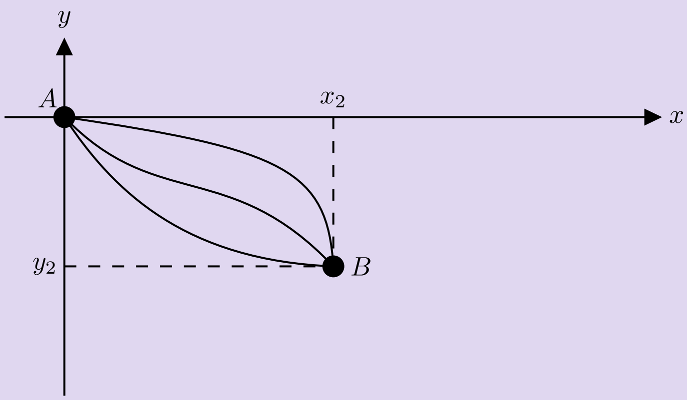
Hasan Baig

Easter 2021

Contents

0 Motivation

Example (The Brechistochrome Problem)**.** A 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?



Johenn Bernoulli proposed the problem of finding the optimal shape, in 1696. Travel time:

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

$$K.E. + V = \text{ const. (energy conservation)}$$

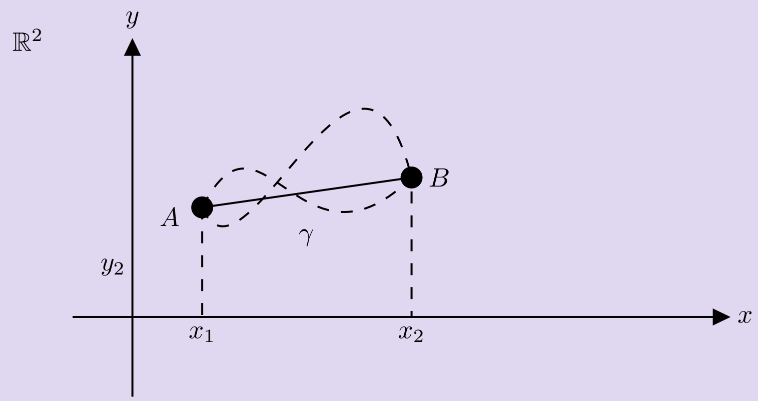
$$\frac{1}{2}mv^2 + mgy = mgy_1 = 0 \qquad\qquad v = \sqrt{2g}\sqrt{-y}$$

Minimise

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

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

Example (Geodesic)**.** Finding the shortest path γ between 2 points on a surface Σ (if one exists). Take $\Sigma = \mathbb{R}^2$ (a plane, Pythagorean theorem holds).



Distance along γ:

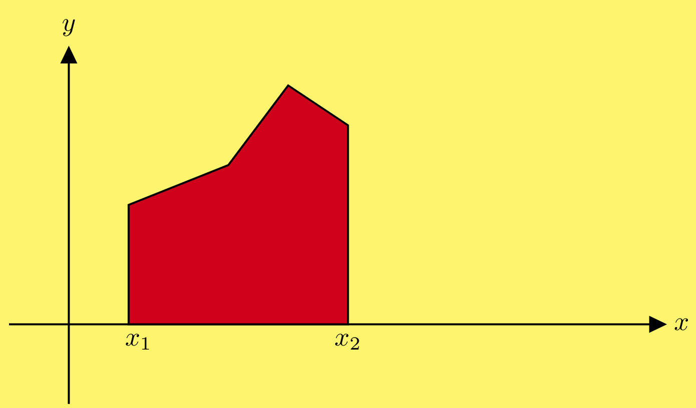
$$D[y] = \int_A^B \mathrm{d}l = \int_{x_1}^{x_2} \sqrt{1 + (y')^2}\,\mathrm{d}x$$

Seek to minimise D by varying γ.

Remark. Generally, we are trying to minimise (maximise)

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

among all functions s.t. $y(x_1) = y_1,\ y(x_2) = y_2$.
(0.1) is a **functional** (a function on the space of functions)
Functions map numbers to numbers. Functionals map functions to numbers e.g.

- 

  $y(x) \to$area under the graph

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

- curve $\to$ length

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

Calculus of variations is finding extrema (min/max/stable) of functions on spaces of functions.

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

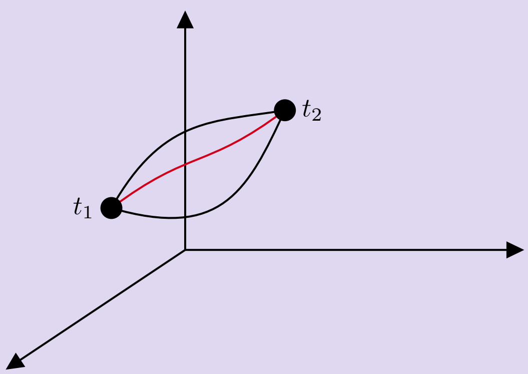
Warning. NEED to specify the function space beforehand (a branch of Functional Analysis – Part III – analysis on the space of functions)

Variational Principles are principles in nature where the laws follow from extremising Functionals

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

Example (Principle of least action)**.** T = kinetic energy (e.g. $m|\dot{\mathbf{x}}|^2/2$)
V = potential energy (e.g. $V(\mathbf{x}),\ \mathbf{x} \in \mathbb{R}^n$)

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



"Action is minimised along paths of motion"

Moral. Leibnitz's take: We live in "the best of all possible worlds".
Science → Theology.
Feynman's take: "This is wrong. In quantum theory, the motion takes place along all possible paths with different probabilities." (see Part III QFT)

In this course

- We consider necessary conditions of extremum of (0.1). Euler-Lagrange equation.
- Lots of examples (geometry, physics, problems with constraints – e.g. maximise area given a fixed length of perimeter)
- Second variation: some sufficient conditions for min/ max

Books:
(i) Gelfend- Fomin 'Calculus of Variations.'
(ii) DAMTP notes online (e.g. P. Townsend)

Note. Lectures have a different order but similar content to (ii).

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

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

Definition. The postition $\mathbf{a} \in \mathbb{R}^n$ is **stationary** if

$$\nabla f(\mathbf{a}) = (\partial_1 f, \ldots, \partial_n f) \mid_{\mathbf{x}=\mathbf{a}} = 0,\ \text{where } \partial_i f = \frac{\partial f}{\partial x_i}$$

Method. Expanding near $\mathbf{x} = \mathbf{a}$

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

using the summation convention. The **Hessian** matrix is

$$H_{ij} = \partial_i \partial_j f = H_{ji}$$

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

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

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

(i) If all $\lambda_i > 0$, $f(\mathbf{x}') > f(\mathbf{0})$ in all directions so we have a local minimum
(ii) If all $\lambda_i < 0$, then we have a local maximum
(iii) If some $\lambda_i > 0$, and some $\lambda_i < 0$, then f increases in some directions and decreases in others. We have a **saddle point** in this case.
(iv) If some $\lambda_i = 0$, then we need to consider higher order derivatives in Taylor's expansion.

Method. Special case $n = 2$:

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

- $\det > 0,\ \operatorname{tr} > 0$ gives local minimum
- $\det > 0,\ \operatorname{tr} < 0$ gives local maximum
- $\det < 0$ gives saddle point
- $\det = 0$ requires us to look at 3^{rd}/ higher derivatives

Remarks.

(i) For $f : D \to \mathbb{R}$ (domain) we can have local minimum, local maximum or global minimum

(ii) For f harmonic, $f_{xx} + f_{yy} = 0,\ D \subseteq \mathbb{R}^2$ gives $\mathrm{tr}(H) = 0$ so our turning point is a saddle point and the min/ max is on the boundary

Example.

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

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

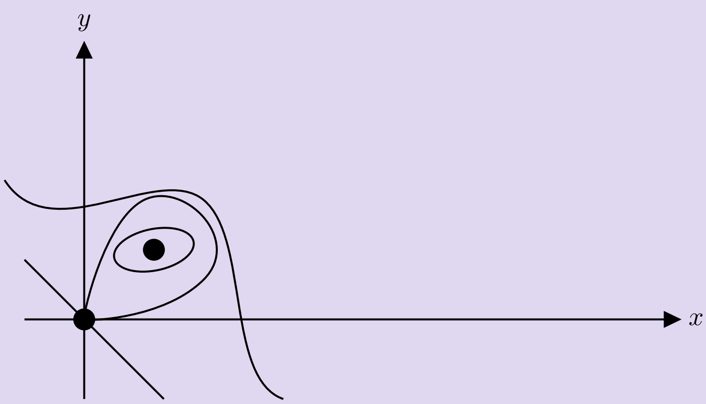
for critical points.

$$x^2 - y = 0,\ y^2 = 0 \implies y^4 = y \implies \begin{cases} y = 0,\ x = 0 \\ y = 1,\ x = 1 \end{cases}$$

Stationary points $(0, 0)$ and $(1, 1)$

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

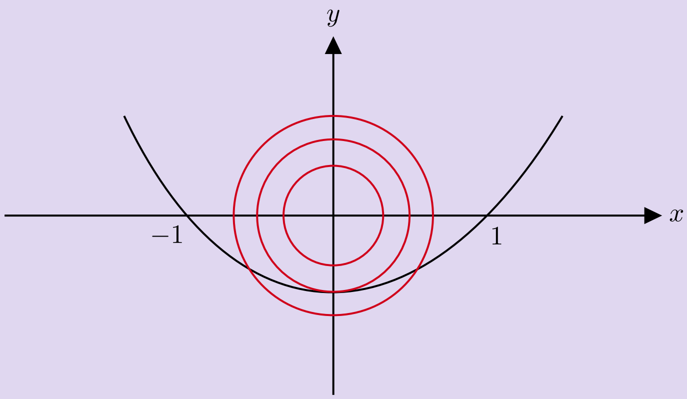
$(0, 0)$ has $\det H = -9 < 0$, saddle point $f = 0$.
$(1, 1)$ has $\det H = 27 > 0$, $\mathrm{tr}(H) = 12 > 0$ so is local minimum with $f = -1$



For $(0, 0)$, near $f = 0$, $f \simeq -3xy$ which decreases on the line $y = x$ but increases on $y = -x$.
This function has no global min/ max

1.1 Constraints and Lagrange Multipliers

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



Two approaches:

(i) Direct method. Solve the constraintsL

$$f = x^2 + y^2 = x^2 + (X^2 - 1)^2 = \underbrace{x^4 - x^2 + 1}_{f(x)}$$

We have

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

Giving two solutions

- $x = \pm 1/\sqrt{2},\ y = -1/2$, radius $\sqrt{3}/2$
- $x = 0, y = -1$, radius 1

Example. (ii) Lagrange Multipliers. Define new function $h(x, y, \lambda) = f(x, y) - \lambda g(x, y)$ with $g(x, y) = 0$ the constraint. λ = Lagrange multiplier.

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

Extremising over 3 variables with no constraints:

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

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

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

The first two give:

$$2x + 4xy = 0 \implies x = 0 \text{ or } y = -\frac{1}{2}$$

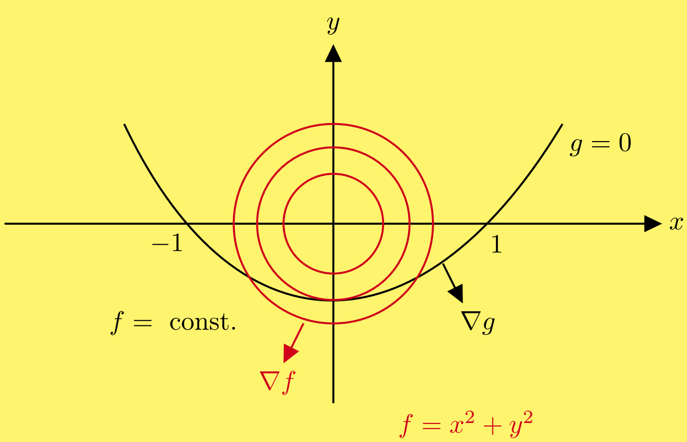
Subbing these in the final equation gives solutions:

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

$$(0, 1) \to f = 1 \text{ so } (\lambda = 2)$$

$$(\pm\frac{1}{\sqrt{2}}, -\frac{1}{2}) \to f = \frac{3}{4},\ \lambda = -1$$

Moral. Why does it work (geometry):



$\nabla g \perp$ to $g = 0$

At the extrema, $\nabla f \parallel \nabla g$, so

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

Extremum of $h = f - \lambda g$

Method. For multiple constraints, extremise $f : \mathbb{R}^n \to \mathbb{R}$, subject to $g_\alpha(\mathbf{x}) = 0$

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

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

We have $n + k$ variables, k Lagrange Multipliers

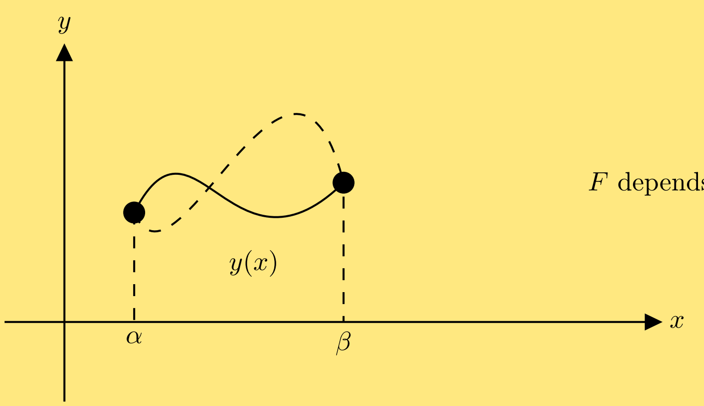
$$\frac{\partial h}{\partial x_i} = 0, \ \frac{\partial h}{\partial \lambda_\alpha} = 0$$

Eliminate λ_α and solve for $\mathbf{x}$
This method works also if constraints can't be eliminated

2 Euler-Lagrange Equations

Method. Our task is to extremise functional (0.1)

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



F depends on y with $y(\alpha), y(\beta)$ fixed

f given, depends on y with fixed ends.
Consider a small pertubation $y \to y + \epsilon\eta(x)$ in (2.1)
Compute $F[y + \varepsilon\eta]$, $\eta(\alpha) = \eta(\beta) = 0$.
We will need the lemma below

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

$$\int_\alpha^\beta g(x)\eta(x)\,\mathrm{d}x = 0 \text{ for all } \eta \text{ continuous on } [\alpha, \beta], \text{ s.t. } \eta(\alpha) = \eta(\beta) = 0$$

Then $g(x) \equiv 0,\ \forall x \in [\alpha, \beta]$

Proof. We show $\exists \bar{x} \in (\alpha, \beta)$ s.t. $g(\bar{x}) = 0$. Suppose $g(\bar{x}) > 0$. 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

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

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

Remark. η given by (2.2) is a bump function. A C^k bump function:

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

Method. Back to (2.1):

$$\begin{aligned} F[y + \varepsilon\eta] &= \int_\alpha^\beta f(x, y + \varepsilon\eta, y' + e\eta')\,\mathrm{d}x \\ &= F[y] + \varepsilon \int_\alpha^\beta \left(\frac{\partial f}{\partial y}\eta + \frac{\partial f}{\partial y'}y'\right)\mathrm{d}x + \underbrace{O(\varepsilon^2)}_{\text{return in section 8}} \\ &= F[y] + O(\varepsilon^2) \text{ at extremum, i.e. } \left.\frac{\mathrm{d}F}{\mathrm{d}\varepsilon}\right|_{\varepsilon=0} = 0 \end{aligned}$$

Integrating the ε-term by parts

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

Applying the Lemma with g as above, we must have $g \equiv 0$.

Equation. We have proved a nexessary condition for an extremum is:

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

This is called the Euler-Lagrange equation

Remarks.

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

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

  where

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

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

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

  $$\frac{\mathrm{d}}{\mathrm{d}x} = \delta_x + y'\delta_y + y''\delta_{y'}$$

  is the total derivative.

Example.

$$f(x, y, y') = x \cdot ((y')^2 - y^2)$$
$$\delta_x f = (y')^2 - y^2 \quad \delta_y f = -2xy \quad \delta_{y'} f = 2xy'$$
$$\frac{\mathrm{d}f}{\mathrm{d}x} = (y')^2 - y^2 - 2xyy' + 2xy'y''$$

2.1 First Integrls of the E-L equation

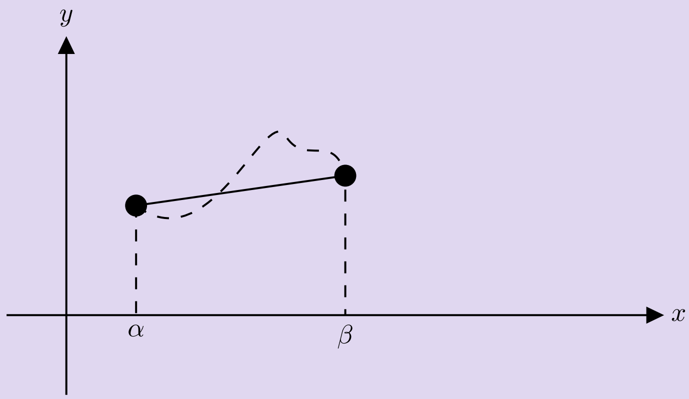
In some cases (2.3) (2nd order ODE) can be integrated once to a 1st order ODE "first integral".

(i) f does not explicitly depend on y, $\frac{\mathrm{d}f}{\mathrm{d}y} = 0$

$$\frac{\partial f}{\partial y} = 0$$

$$(2.3) \to \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\partial f}{\partial y'}\right) = 0 \implies \frac{\partial f}{\partial y'} = \text{ constant}$$

Examples. Geodesics on the Euclidean plane



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

$$\frac{\partial f}{\partial y} = 0 \implies \frac{y'}{\sqrt{1+(y')^2}} = \text{ const.}$$

So

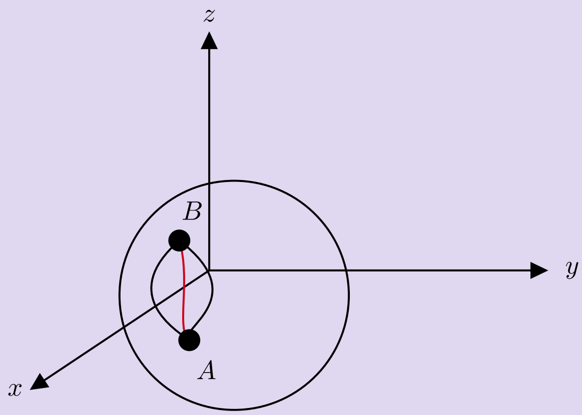
$$y' = m$$

for some constant m, and so

$$y = mx + c$$

straight line

Example. Geodesics on a sphere $S^2 \subset \mathbb{R}^3$



$$\begin{aligned} x &= \sin\theta\sin\phi \quad 0 \le \theta \le \pi \\ y &= \sin\theta\cos\phi \quad 0 \le \phi \le 2\pi \\ z &= \cos\theta \end{aligned}$$

$$\mathrm{d}s^2 = \mathrm{d}x^2 + \mathrm{d}y^2 + \mathrm{d}z^2 = \mathrm{d}\theta^2 + \sin^2\mathrm{d}\phi^2$$

Parametrise as $\phi = \phi(\theta)$

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

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

$$\frac{\partial f}{\partial\phi} = 0 \implies \frac{\partial f}{\partial\phi} = \kappa \text{ (constant)}$$

first integral.

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

Example (continued)**.** Squaring to solve for $(\phi')^2$

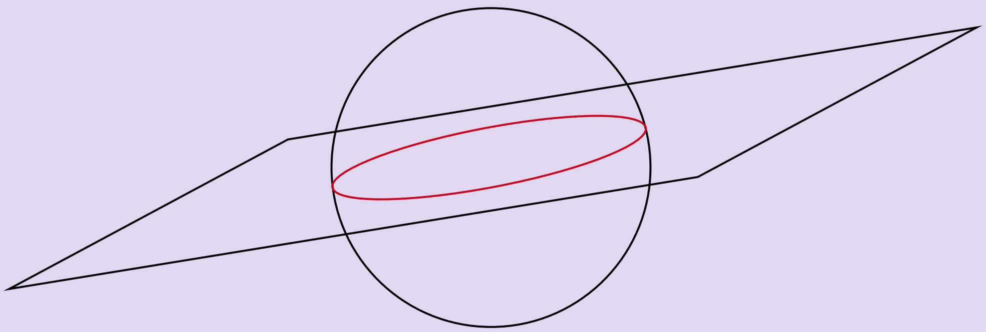
$$(\phi')^2 = \frac{\kappa^2}{\sin^2\theta \cdot (\sin^2 - \kappa^2)}$$

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

Two solutions, each going one way around the sphere. Using substitution $\cot(\theta) = u$

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

for $\phi_0 = \text{const}$. Great circle



(Geodesics are segments of great circles)

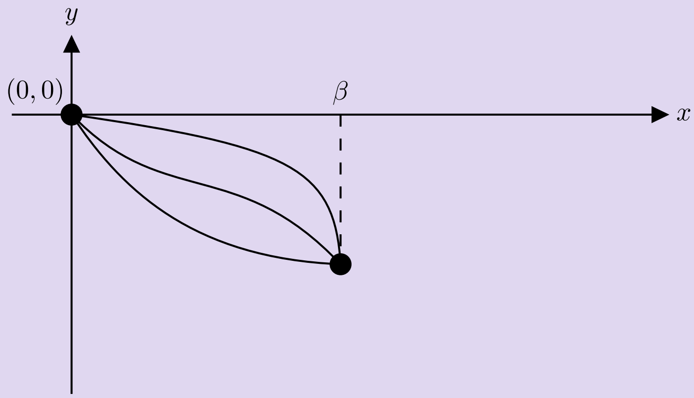
(ii) Consider, for general $f(x, y, y')$

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

If f does not explicitly depend on x, i.e. $\frac{\partial f}{\partial x} = 0$ then

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

Example (Brachistochrome).



Going back to section 0,

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

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

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

$$\frac{1}{\sqrt{1+(y')^2}} = K\sqrt{-y} \implies y' = \pm \frac{\sqrt{1_K^2 y^2}}{K\sqrt{-y}}$$

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

Set

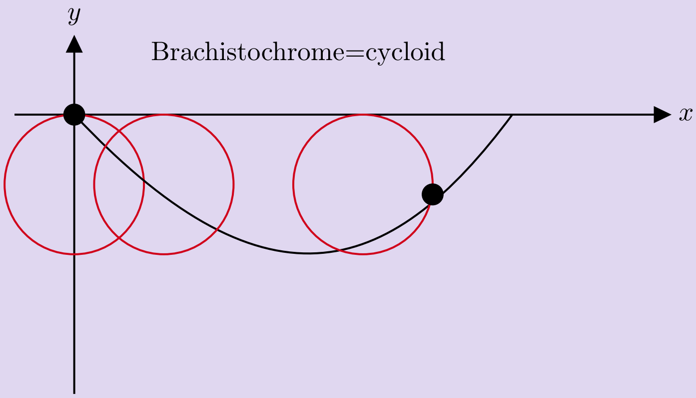
$$y = -\frac{1}{K^2}\sin^2\frac{\theta}{2} \quad \mathrm{d}y = -\frac{1}{K^2}\sin\left(\frac{\theta}{2}\right)\cos\frac{\theta}{2}$$

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

Initial condition $(0,0) \to \theta_0 = 0 \to C = 0$, take positive root

$$x = \frac{\theta - \sin\theta}{2K^2}$$
$$y = -\frac{1}{K^2}\sin^2\frac{\theta}{2}$$

Example (continued)**.**



Parametrised equation of a cycloid. Brachistochrome = cycloid.
The curve traced by a point on the rim of a wheel, as the wheel rolls along a straight line (Galileo)

2.2 Fermat's Principle

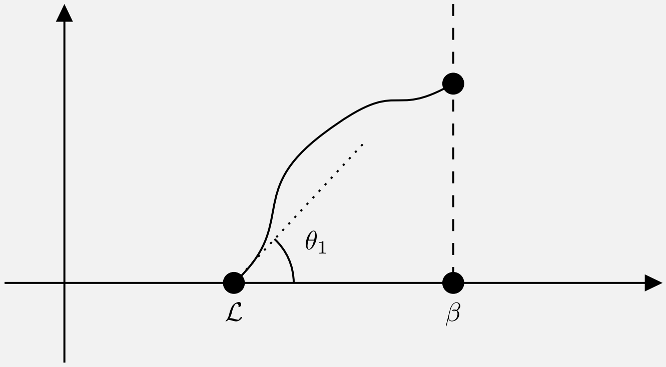
Light/sound travels along paths between two points which requires least time. Rays are represented by path $y = y(x)$. Speed of light $c(x, y)$

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

assume $c = c(x) \to \frac{\partial f}{\partial y} = 0$ so (2.4) gives

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

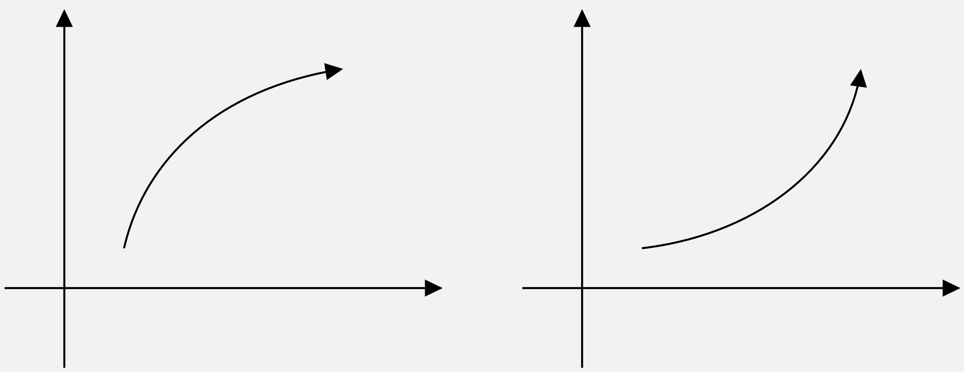
$$\frac{y'}{\sqrt{1+(y')^2}c(x)} = \text{ const.}$$



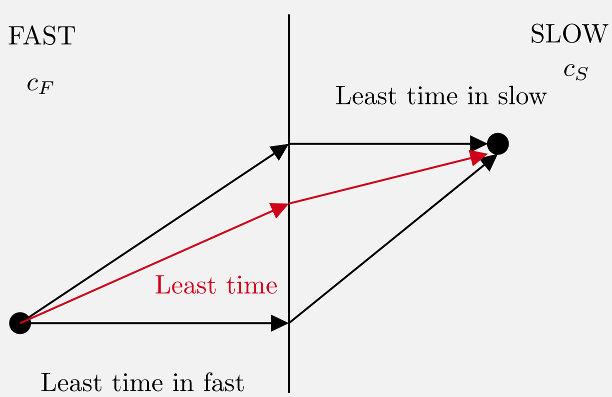
Ray launched at θ_1, $\tan\theta = y'$

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

Snell's law.



c increasing $\quad$ c decreasing



3 Extensions of the Euler-Lagrange Equations

3.1 Euler-Lagrange Equations with Constraints

Extremise

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

subject to

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

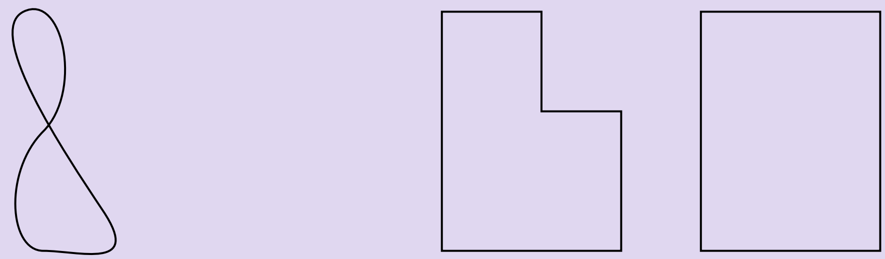
Lagrange multiplier, extremise

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

replace f in E-L by $f - \lambda g$

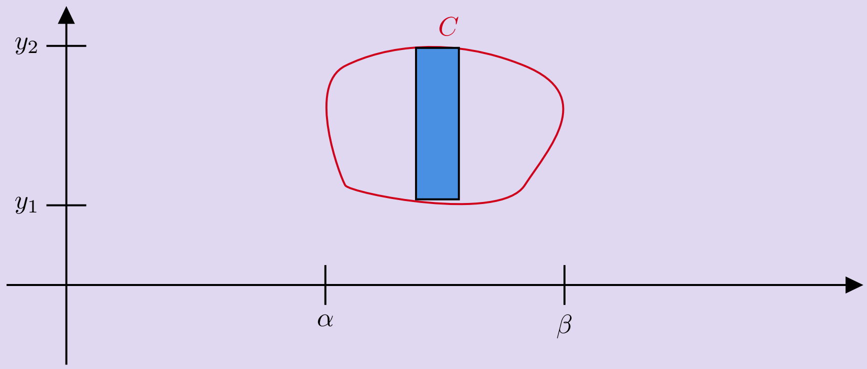
$$\frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\partial}{\partial y'}(f - \lambda g)\right) - \frac{\partial}{\partial y}(f - \lambda g) = 0 \tag{3.1}$$

Example. Dido problem (a.k.a. isoperemetric problem). What simple and closed plane curve of fixed length L maximises the enclosed area?



Not simple Not convex →convex with the same perimeter

Assume convexity.



$$\mathrm{d}A = y(x)\,|_{x_1}^{x_2}\,\mathrm{d}x$$

x monotonically increases from $\alpha \to \beta$ and decreases from $\beta \to \alpha$. Given $x, \exists(y_1, y_2)$ on the curve with

$$y_1(x) = y_2(x),\ y_2 > y_1,\ \mathrm{d}A = y(a)]_{x_1}^{x_2} \cdot \mathrm{d}x$$

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

Constraint

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

$$K = y0\lambda\sqrt{1+(y')^2}$$

(Note: do not worry about the boundary trm in the derivation of the E-L, as C has no boundary) $\frac{\partial h}{\partial x} = 0$ so we use (2.5)

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

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

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. The Sturn-Liouville problem.
$\rho(x) > 0$ for $x \in [\alpha, \beta]$, $\sigma = \sigma(x)$

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

Minimise f subject to $G = 1$ (fixed ends)

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

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

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

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

$\mathcal{L}$ is the Sturm-Liouville operator. (3.2) is an eigenvalue probleme.g. if $\rho = 1$, $\sigma(x) =$ 'potential' in time-independent Shrodinger equation (IB Quantum Mechanics).
If $\sigma > 0$, then $F[y] > 0$. Positive minimum equal to the lowest eigenvalue

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

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

Lowest eigenvalue is the minimum of $F[y]/G[y]$

3.2 Several dependent variables

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

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

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

Following the derivation of the E-L equation:

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

Use Lemma

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

A system of n 2nd order ODEs.
First integrals of 3.3

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

Example. Geodesics on surfaces
$\Sigma \subset \mathbb{R}^3$ (surface) given by

$$g(x, y, z) = 0$$

Geodesic = shortest path on the surface between $A, B \in \Sigma$ (if one exists). t = parameter on the curve

$$\begin{aligned} A &= \mathbf{x}(0) \\ B &= \mathbf{x}(1) \quad \mathbf{x} = (x, y, z) \end{aligned}$$

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

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:

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

- Variation w.r.t. $x_i = (x, y, z)$

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\dot{x}_i}{\sqrt{\dot{x}^2 + \dot{y}^2 + \dot{z}^2}}\right) + \lambda \frac{\partial g}{\partial x_i} = 0 \quad i = 1, 2, 3$$

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

3.3 Several Independent Variables

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

$$F[\phi] = \iiint_D f(\underbrace{x, y, z}_{\text{indep}}, \phi, \phi_x, \phi_y, \phi_z), \,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z$$

notation $\phi_x = \frac{\partial \phi}{\partial x}$ etc. Volume integral over a domain $D \subset \mathbb{R}^3$. Assume ϕ extremum, consider perturbations

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

$$\begin{aligned} F[\phi + \varepsilon\eta] - F[\phi] &= \varepsilon \int_D (\eta\frac{\partial f}{\partial \phi} + \eta_x \frac{\partial f}{\partial \phi_x} + \eta_y \frac{\partial f}{\partial \phi_y} + \eta_z \frac{\partial f}{\partial \phi_z}) \,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z + O(\varepsilon^2) \\ &= \varepsilon \int_D \eta\frac{\partial f}{\partial \phi} + \nabla \cdot (\eta(\frac{\partial f}{\partial \phi_x}, \frac{\partial f}{\partial \phi_y}, \frac{\partial f}{\partial \phi_z})) - \eta\nabla \cdot (\frac{\partial f}{\partial \phi_x}, \frac{\partial f}{\partial \phi_y}, \frac{\partial f}{\partial \phi_z}) \,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z + O(\varepsilon^2) \end{aligned}$$

Apply divergence theorem to first div term and use

$$\int_{\partial D} \eta(\frac{\partial f}{\partial \phi_x}, \frac{\partial f}{\partial \phi_y}, \frac{\partial f}{\partial \phi_z}) \cdot \mathrm{d}\mathbf{s} = 0$$

as $\eta = 0$ on ∂D

$$F[\phi + \varepsilon\eta] - F[\phi] = \varepsilon \int \eta(\frac{\partial f}{\partial \phi} - \nabla \cdot (\frac{\partial f}{\partial \phi_x}, \frac{\partial f}{\partial \phi_y}, \frac{\partial f}{\partial \phi_z})) \,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z + O(\varepsilon^2)$$

E-L equation: single 2nd order PDE for one function ϕ

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

remains valid with $3 \to n$

Example. Extremise 'potential energy' $n = 2$

$$F[\phi] = \iint_D \frac{1}{2}[\phi_x^2 + \psi_y^2] \,\mathrm{d}x\,\mathrm{d}y$$

$$\frac{\partial f}{\partial \phi} = 0 \quad \frac{\partial f}{\partial \phi_x} = \phi_x \quad \frac{\partial f}{\partial \phi_y} = \phi_y$$

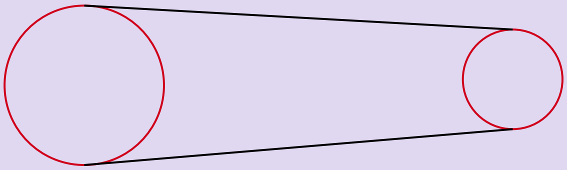
$$(3.4) \to \frac{\partial \phi_x}{\partial \phi_x} + \frac{\partial \phi_y}{\partial y} = 0$$

i.e.

$$\phi_{xx} + \phi_{yy} = 0$$

(Laplace equation)

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



e.g. soap forms

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

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

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

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

(IB geometry, this is called the 1st fundamental form or Riemannian metric)

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

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

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

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

Apply E-L (3.4) to h

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

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

Expand derivatives (exercise)

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

The minimal surface equation. Assume circular symmetry

$$z = \phi(r) \quad r = \sqrt{x^2+y^2}$$

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

by calculating 2nd derviatives, we get from (3.5) the ODE

$$rz'' + z' + (z')^3 = 0$$

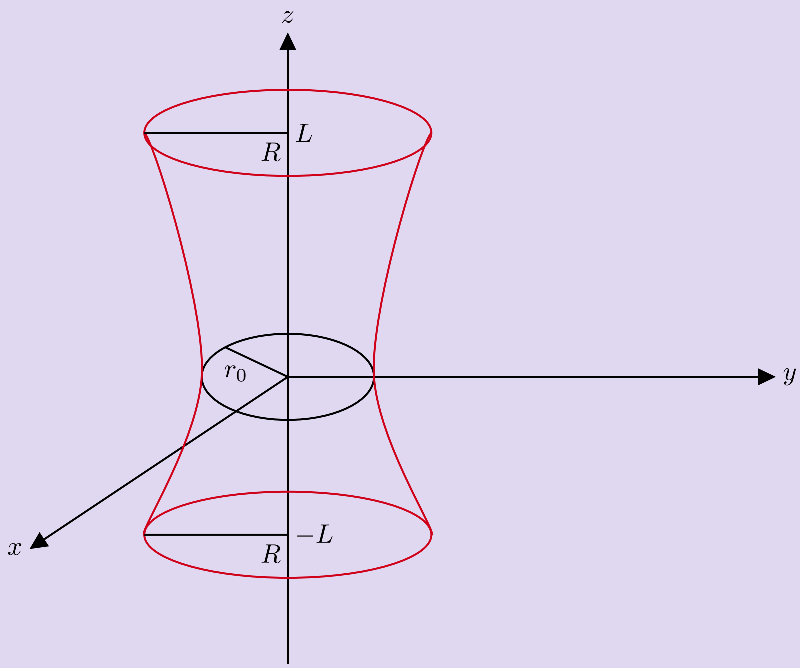
Set $z' = w$ to get

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

Solution

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

Example (continued)**.** Catenoid: minimal surface of revolution (Euler 1744)

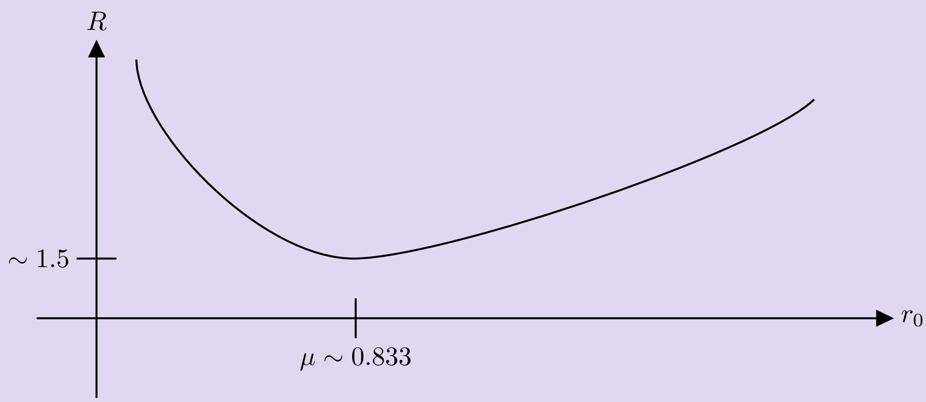


$r(L) = r(-L)$. If $L \neq 0$ then $z_0 = 0$. Set $r = R$, and divide by L

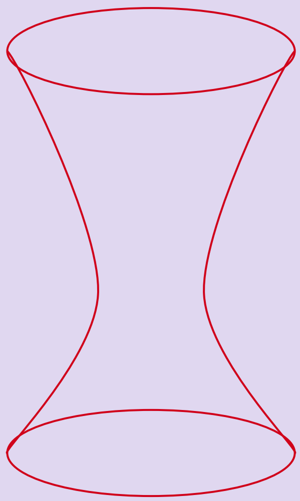
$$\frac{R}{L} = \frac{r_0}{L} \cosh\left(\frac{L}{r_0}\right)$$

Set $L = 1$. Algebraic relation

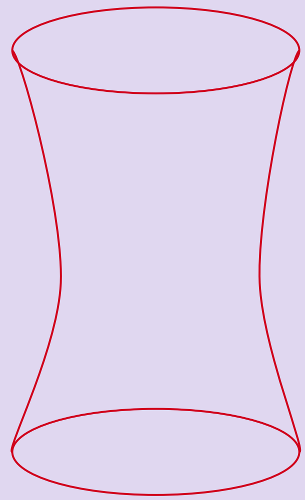
$$R = r_0 \cosh(1/r_0)$$



Example (continued)**.** If $R >\sim 1.5$, $\exists$ 2 minimal surfaces



unstable



stable

3.4 Higher Derivatives

Equation.

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

Proceed as in section 2. Assume y exists, $y \to t + \varepsilon\eta$ where

$$\eta = \eta' = \cdots = \eta^{(n-1)} = 0 \text{ at } \alpha, \beta$$

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

Apply Lemma

$$\frac{\partial f}{\partial y} - \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\partial f}{\partial y'}\right) + \frac{\mathrm{d}^2}{\mathrm{d}x^2}\left(\frac{\partial f}{\partial y''}\right) + \cdots + (-1)^n \frac{\mathrm{d}^n}{\mathrm{d}x^n}\left(\frac{\partial f}{\partial y^{(n)}}\right) \tag{3.6}$$

Euler-Lagrange equation

Example. If $n = 2$ and if $\frac{\partial f}{\partial y} = 0$

$$(3.6) \to \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\partial f}{\partial y'} - \frac{\mathrm{d}}{\mathrm{d}x}\frac{\partial f}{\partial y''}\right) = 0$$

so

$$\frac{\partial f}{\partial y'} - \frac{\mathrm{d}}{\mathrm{d}x}\frac{\partial f}{\partial y''} = \text{const.}$$

Example. Extremise $F[y] = \int_0^1 (y'')^2 \, \mathrm{d}x$ where $y(0) = y'(0) = 0$ and $y(1) = 0,\ y'(1) = 1$

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

Impose boundary conditions to get $y = x^3 - x^2$

Note. This is an absolute minimum. $Y_0 = x^3 - x^2$

$$\eta(0) = \eta'(0) = \eta(1) = \eta'(1) = 0$$

(do not assume η small)

$$\begin{aligned} F[y_0 + \eta] - F[y_0] &= \int_0^1 (\eta'')^2 \, \mathrm{d}x + 2 \cdot \int_0^1 (y_0''\eta'') \, \mathrm{d}x > 4 \int_0^1 (3x - 1)\eta'' \\ &= 4([-\eta]_0^1 + \int_0^1 \frac{\mathrm{d}}{\mathrm{d}x}(3x\eta') - \eta)\mathrm{d}x \\ &= 4([3x\eta']_0^1 - 3\eta]_0^1) = 0 \end{aligned}$$

y_0 absolute minimiser of F

4 Least Action Principle and Noether's Theorem

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

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

is the Langranian. t is the independent variable, $\mathbf{x} = (x, y, z)$ are dependent variables. Action

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

Hamilton's principle (Least action principle, or principle of stationary action). The motion is such that $S[\mathbf{x}]$ is stationary, i.e. L satisfies the E-L equations

Example.

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

Euler-Lagrange equations give

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{x}_i}\right) = \frac{\partial L}{\partial \dot{x}_i}$$

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

Newton's 2nd Law

Example. Central force in 2 dimensions:

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

E-L

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial l}{\partial \dot{r}}\right) - \frac{\partial L}{\partial r} = 0$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{\theta}}\right) - \underbrace{\frac{\partial L}{\partial \theta}}_{0} = 0$$

$$\implies \frac{\partial L}{\partial \dot{\theta}} = mr^2\dot{\theta} = \text{const.}$$

$\frac{\partial L}{\partial t} = 0$, use (2.5)

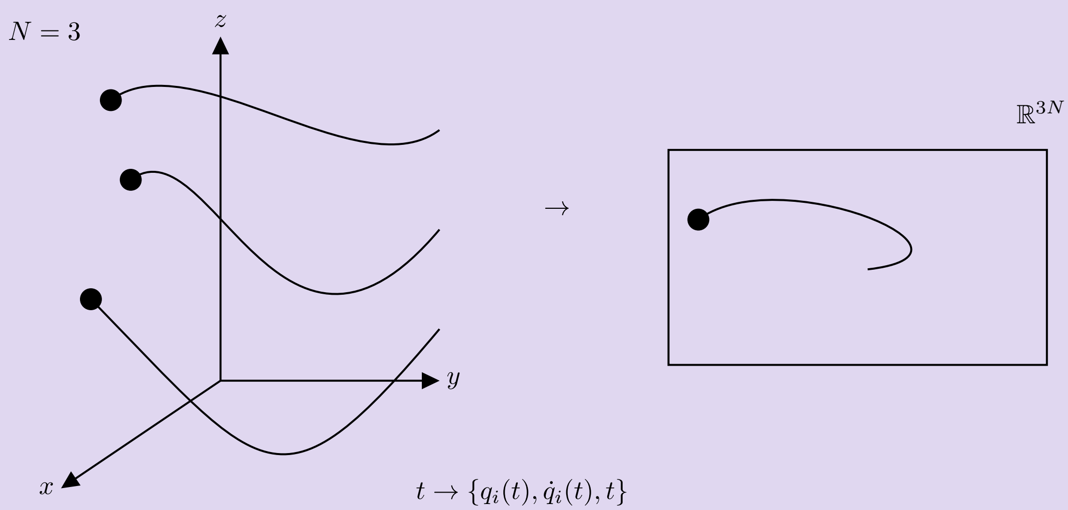
$$\dot{r}\frac{\partial L}{\partial \dot{r}} + \dot{\theta}\frac{\partial L}{\partial \dot{\theta}} - L = \text{const}$$

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

which is constant. Conservation of total energy

Example (Configuration space, and general coordinates)**.**

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



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

$$q_i = \text{generalised coordinates, } i = 1, \ldots, 3N$$

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

(Part II Classical Dynamics)

4.1 Noether's Theorem

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

Suppose $\exists$ a 1-parameter family of transformations $y_i(x) \to Y_i(x, s)$ s.t. $Y_i(x, 0) = y_i(x)$. This is a continuous symmetry of a Lagrangian f, if

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

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

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

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

Proof.

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

Example.

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

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

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

so

$$\frac{\mathrm{d}f}{\mathrm{d}s} = 0$$

$$(4.3) \to (\frac{\partial f}{\partial y'}\frac{\mathrm{d}Y}{\mathrm{d}s} + \frac{\partial f}{\partial z'}\frac{\mathrm{d}Z}{\mathrm{d}y}) = y' + z'$$

(conserved momentum in $y + z$ direction)

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

$$\frac{\mathrm{d}L}{\mathrm{d}s} = 0$$

$$(4.3) \to (\frac{\partial L}{\partial \dot\theta}\frac{\partial \theta}{\partial s} + \frac{\partial L}{\partial \dot r}\underbrace{\frac{\partial R}{\partial s}}_{0})|_{s=0} = mr^2\dot\theta$$

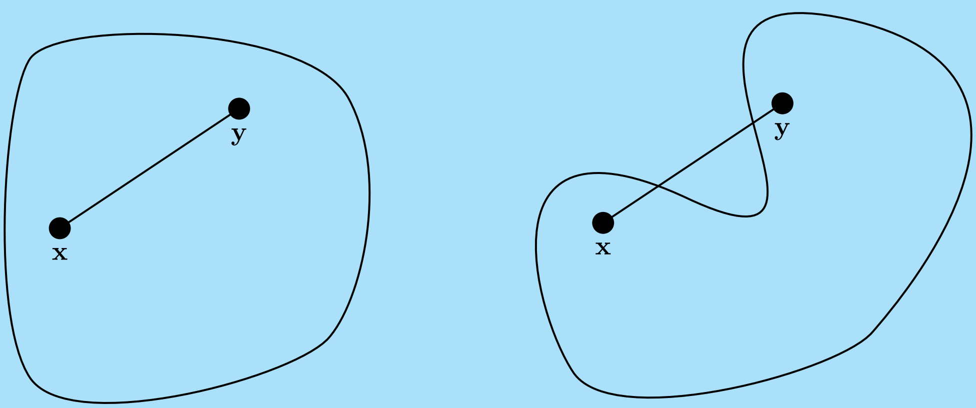
(conserved angular momentum). Isotropy of space gives rotational invariance of L

5 Convex Functions

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

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

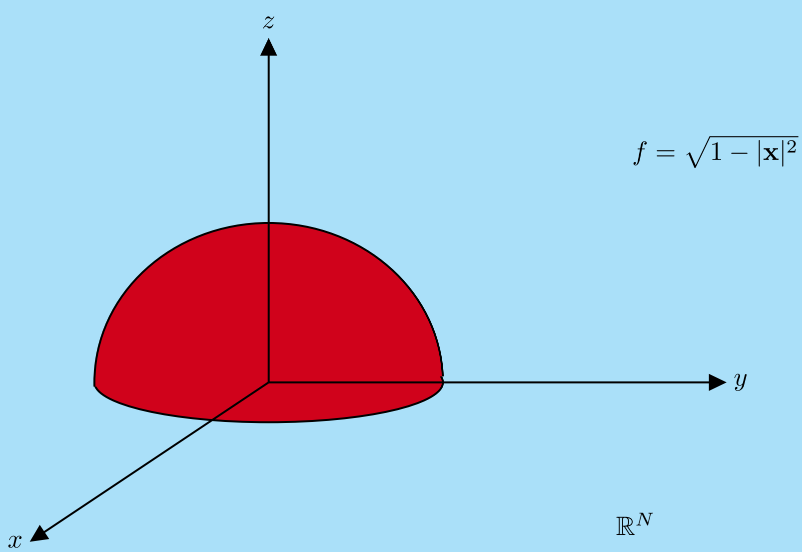
$$(1-t)\mathbf{x} + t\mathbf{y} \in S \quad 0 \leq t \leq 1$$



convex non-convex

Definition. A **graph of a function** $f : \mathbb{R}^n \to \mathbb{R}$ is a surface

$$\{z - f(\mathbf{x}) = 0\} \text{ in } \mathbb{R}^{n+1}$$



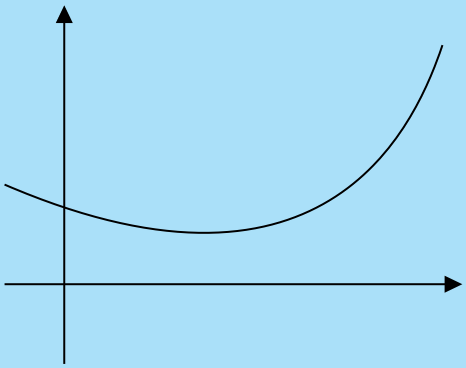
A **chord** of f is a line segment in $\mathbb{R}^{n+1}$ joining two points on the graph

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

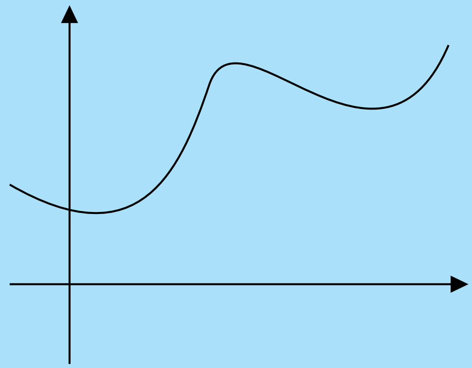
(i) The domain of f is a convex set

(ii)

$$f((1-t)\mathbf{x} + t\mathbf{y}) \leq (1-t)f(\mathbf{x}) + tf(\mathbf{y}) \quad 0 < t < 1 \tag{5.1}$$



convex



non-convex

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

Remarks.

(i) f is concave if we replace $\leq$ by $\geq$ in (5.1)

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

(iii) f strictly convex if we replace $\leq$ by $<$ in (5.1)

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

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

strictly convex

Example. $f(x) = 1/x$, domain $\mathbb{R}\backslash\{0\}$, not a convex set. On restricted domain $\mathbb{R} > 0$, f is convex

5.1 Conditions for Convexity

3 tests for f to be convex

(i) 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

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

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

Take $\mathbf{z} = (1-t)\mathbf{x} + t\mathbf{y} \in S$ (the domain of f), $0 < t < 1$

$$(1-t) \cdot (i) + t \cdot (ii) \to \nabla f(\mathbf{z}) \text{ cancel. get } (5.1)$$

Converse: assume convexity (5.1) and set

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

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

So (5.2) is equivalent to $h'(0) \geq 0$. Note $h(0) = 0$, so

$$\frac{h(t) - h(0)}{t} \geq 0 \quad 0 < t < 1$$

Now take the limit $t \to 0$

Corollary. If f is convex and have a stationary point, then it is a global minimum

Proof. Given $\nabla f(\mathbf{x}_0) = 0$, we get from (5.2) that $f(\mathbf{y}) \geq f(\mathbf{x}_0)$ $\forall \mathbf{y}$

(ii) If

$$(\nabla f(\mathbf{y}) - \nabla f(\mathbf{x})) \cdot (\mathbf{y}-\mathbf{x}) \geq 0 \tag{5.3}$$

then f is convex (f' monotonically increasing if $n = 1$)

Proof. exercise

(iii) (Second order conditions): Assume f twice differentiable, then f convex iff the Hessian $\frac{\partial^2 f}{\partial x^i \partial x^j}$ has all eigenvalues non-negative. If all eigenvalues positive, then 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 $\mathbf{h}$)

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

Example.

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

$$H = \frac{1}{xy}\begin{bmatrix} \frac{2}{x^2} & \frac{1}{xy} \\ \frac{1}{xy} & \frac{2}{y^2} \end{bmatrix} \quad \det(H) = \frac{3}{x^3 y^3} > 0 \quad \mathrm{tr}(H) > 0$$

so f is strictly convex

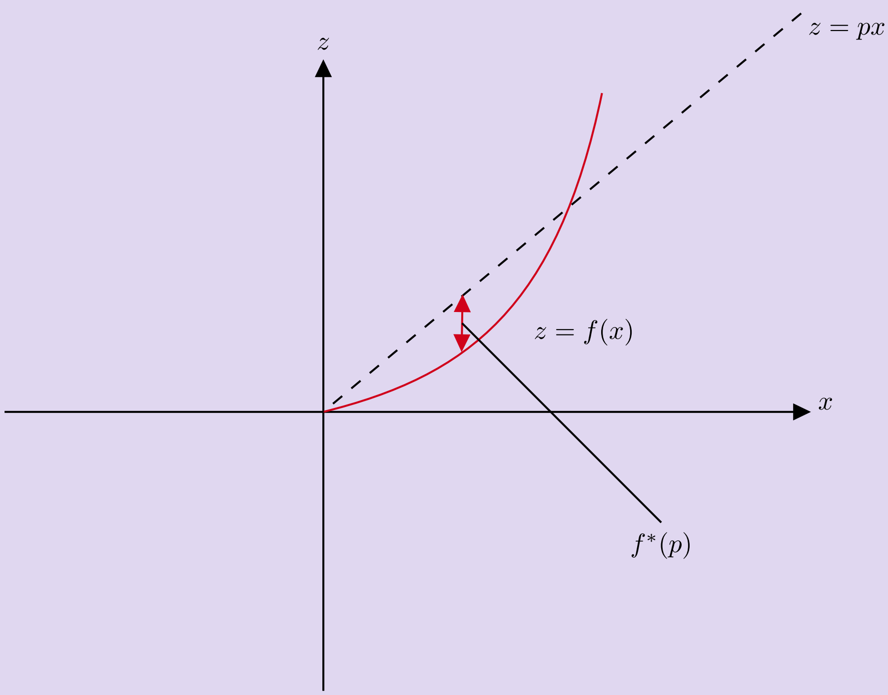
6 Legendre Transform

Definition. The Legendre transform of $f : \mathbb{R}^n \to \mathbb{R}$ 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 $\mathbf{p} \in \mathbb{R}^n$ s.t. the sup is finite

Example. $n = 1$



Maximum vertical distance between graphs of $z = f(x)$ and $z = px$

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

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

So $x = p/2a$ and substitute

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

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

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

so $f^{**} = f$ (always true if f convex)
If $a < 0$, $\sup_x (px - ax^2) = \infty$ $\forall p$ so f^* has empty domain

Prop. Domain of f^* is a convex set, find f^* convex

Proof.

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

Use $\sup(A + B) \le \sup(A) + \sup(B)$ to get

$$LHS \le (1-t)f^*(\mathbf{p}) + tf^*(\mathbf{q})$$

(i)
$$(1-t)\mathbf{p} + t\mathbf{q} \in D(f^*)$$

(ii) f^* satisfies convextiy definition (5.1)

Note. In practice, if f convex and diffrentiable,

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

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

(substitute to definition of $f^*(p)$)
If f is strictly convex, then $\exists$ unique inversion $\mathbf{x} = \mathbf{x}(\mathbf{p})$ 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) so we use a few macroscopic variables: p (pressure), V (volume), T (temperature), S (entropy). (Part II Statistical Physics)
Internal energy $U(S,V)$. Hermholtz free is defined

$$F(T,V) = \min_S (U(S,V) - TS) = \max_S (TS - U(S,V)) = -U^*(T,V)$$

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

$$\frac{\partial}{\partial S}\left(TS - U(S,V)\right)|_{T,V} = 0 \to R = \frac{\partial U}{\partial S}|_V$$

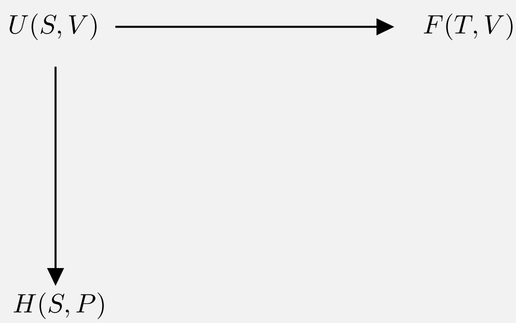
Other quantities as Legendre transform e.g. Entropy

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

at min

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

Entropy is a fixed parameter. The Legendre transform is a way to swap from (S,V) dependence to dependence of other variables



7 Hamilton's Equations

Remark. Recall (section 4.1) Lagranian $L = T - V = L(\mathbf{q}, \dot{\mathbf{q}}, t)$ function on the configuration space

Definition. The Hamiltonian is the Legendre transform of h w.r.t. $\dot{\mathbf{q}} = \mathbf{v}$

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

where $\mathbf{v} = \mathbf{v}(\mathbf{p})$ is the solution to

$$p_i = \frac{\partial}{\partial L \dot{q}_i}$$

(assume convexity of L in $\mathbf{v}$). p is the generalised momentum

Example.

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

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

$$\begin{aligned} H(\mathbf{q}, \mathbf{p}, t) &= \mathbf{p} \cdot \frac{\mathbf{p}}{m} - (\frac{1}{2}m\frac{|\mathbf{p}|^2}{m^2} - V(\mathbf{q})) \\ &= \frac{1}{2m}|\mathbf{p}|^2 + V(\mathbf{q}) \text{ (the total energy)} \end{aligned}$$

What happened to the Euler-Lagrange equations?

$$H = H(\mathbf{q}, \mathbf{p}, t) = p_i\dot{q}^i = L(q^i, \dot{q}^i, t)$$

$$\begin{aligned} \mathrm{d}H &= \frac{\partial H}{\partial q^i}\mathrm{d}q^i + \frac{\partial H}{\partial p_i}\mathrm{d}p_i + \frac{\partial H}{\partial t}\mathrm{d}t \\ &= p_i\mathrm{d}\dot{q}^i + \dot{q}^i\mathrm{d}p_i - \frac{\partial h}{\partial q^i}\mathrm{d}q^i - \frac{\partial L}{\partial \dot{q}^i}\mathrm{d}\dot{q}^i - \frac{\partial L}{\partial t}\mathrm{d}t \\ &= \dot{q}^i\mathrm{d}p_i - \dot{p}_i\mathrm{d}q^i - \frac{\partial L}{\partial t} \end{aligned}$$

by E-L. Compare differentials

$$\dot{q}^i = \frac{\partial H}{\partial p_i} \quad \dot{p}_i = -\frac{\partial H}{\partial q^i} \quad \frac{\partial H}{\partial t} = -\frac{\partial L}{\partial t} \tag{7.2}$$

Warning.

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

Assume no explicit t-dependence in L. Then (7.2) is a system of $2n$ 1st order ODEs. Need to specify $q^i(0), p_i(0),\ i = 1, \ldots, n$. Solution curves to (7.2) are trajectories in $2n$-dimensional phase space

Remark. Hailton's equations also arise from extremizing a functional in phase space

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

E-L for S

- Variation w.r.t. p_i

$$\frac{\partial f}{\partial p_i} - \underbrace{\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial f}{\partial \dot{p}_i}\right)}_{0} = 0 \implies \dot{q}^i = \frac{\partial H}{\partial \dot{p}_i}$$

- Variation w.r.t. q^i

$$\frac{\partial f}{\partial q^i} - \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial f}{\partial \dot{q}^i}\right) = 0 \implies \dot{q}^i = -\frac{\partial H}{\partial \dot{q}^i} - \frac{\mathrm{d}p_i}{\mathrm{d}t} = 0$$

$$\dot{p}_i = -\frac{\partial H}{\partial q^i}$$

Recovered (7.2), Newton's equation, Lagrange's equation, Hamiltons equation so far (7.2) is just another formulation

8 The Second Variation

E-L equation gives us necessary conditon so we could get a minimum, maximum or a saddle point. And so we look at the nature of stationary points of

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

Expand $F[y+\varepsilon y]$ to 2nd order in ε around a solution to E-L equation

$$\begin{aligned} F[y+\varepsilon\eta] - F[y] &= \int_{\alpha}^{\beta} [f(x, y+\varepsilon\eta, y'+\varepsilon'\eta') - f]\,\mathrm{d}x \\ &= 0 + \varepsilon \int_{\alpha}^{\beta} \eta \underbrace{(\frac{\partial f}{\partial y} - \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\partial f}{\partial y'}\right))}_{0}\,\mathrm{d}x + \frac{\varepsilon^2}{2}\int_{\alpha}^{\beta} [\eta^2 \frac{\partial^2 f}{\partial y^2} + (\eta')^2 \frac{\partial f}{\partial (y')^2} + 2\frac{\partial^2 f}{\partial y \partial y'}\eta\eta']\,\mathrm{d}x \\ &+ O(\varepsilon^3) \end{aligned}$$

2nd variation is

$$\begin{aligned} \delta^2 F[y] &\equiv \frac{1}{2}\int_{\alpha}^{\beta} [\eta^2 \frac{\partial^2 f}{\partial y^2} + (\eta')^2 \frac{\partial f}{\partial (y')^2} + \frac{\mathrm{d}}{\mathrm{d}x}\left(\eta^2\right)\frac{\partial^2 f}{\partial y' \partial y}]\,\mathrm{d}x \\ &= \frac{1}{2}\int_{\alpha}^{\beta} Q\eta^2 + P(\eta')^2\,\mathrm{d}x \end{aligned}$$

where

$$P = \frac{\partial f}{\partial (y')^2} \quad Q = \frac{\partial^2 f}{\partial y^2} - \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\partial^2 f}{\partial y' \partial y}\right) \tag{8.1}$$

we have proved

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

Example. Geodesics on a plane (in section 2)

$$f = \sqrt{1+(y')^2} : \begin{cases} P &= \frac{\partial}{\partial y'}\left(\frac{y'}{\sqrt{1+(y')^2}}\right) \to \frac{1}{(1+(y')^2)^{3/2}} > 0 \\ Q &= 0 \end{cases}$$

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

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

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

so the Legendre condition is necessary for local min.
"P is more important than Q in (8.1)"

Proof. See Gelfend-Fomin for details. Idea: if η' small, then η canbe too lare. Converse not true: η can be small, η' large. Assume $\exists x_0$ s.t. $P(x_0, y_0, y_0') < 0$

Note. (8.2) not sufficient for local minimum see section 8.1 but $P > 0, Q \geq 0$ is sufficient as if $\eta \neq 0$ on (α, β) then $\exists x_0 \in (\alpha, \beta)$ s.t. $\eta'(x_0) \neq 0$

Example. Go back to Brachistochrome

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

Is cycloid a minimizer?

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

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

$$Q = \cdots = \frac{1}{2\sqrt{1+(y^2)^2} y^2 \sqrt{-y}}$$

8.1 Associated Eigenvalue Problem

Go back to (8.1)

$$Q\eta^2 + P(\eta')^2 = Q\eta^2 + \frac{\mathrm{d}}{\mathrm{d}x}(P\eta\eta') - \eta(P\eta')'$$

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

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

Sturm-Liouville operator. If $\exists\eta$ s.t.

$$\begin{cases}\mathcal{L}(\eta) = -\omega\eta \ (\omega \text{ real}) \\ \eta(\alpha) = \eta(\beta) = 0\end{cases} \tag{8.4}$$

Then y_0 is not a minimizer as

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

(8.4) can have solutions even if $P > 0$, so the Legendre condition (8.2) is not sufficient for y_0 to be a minimizer

Example.

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

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

$$(2.3) \to y'' + y = 0 \implies y = y_0 = 0$$

is the stationary point of $F[y]$. 2nd variation:

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

but $Q < 0$. Examine (8.4):

$$-\eta'' - \eta = -\omega^2\eta \quad \eta(0) = \eta(\beta) = 0$$

Take

$$\eta = A\cdot\sin\left(\frac{\pi x}{\beta}\right) \to \left(\frac{\pi}{\beta}\right)^2 = 1 - \omega^2$$

Possible if $\beta > \pi$. So, if $P > 0$ a problem mat arise if the interval is "too large'.

8.2 The Jacobi Condtion

Legendre tried to prove that $P > 0$ is sufficient for $y = y_0$ to be a local minimum. This couldn't have worked (last example), but the idea was good.
Let $\phi = \phi(x)$ be a any differentiable function of x on $[\alpha, \beta]$

$$0 = \int_\alpha^\beta (\phi\eta^2)' \,\mathrm{d}x = \int_\alpha^\beta \phi'\eta^2 + 2\eta\eta'\phi \,\mathrm{d}x$$

(as $\eta(\alpha) = \eta(\beta) = 0$). Adding to (8.1), we can rewrite

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

Assume $P|_y > 0$ and complete the square

$$\delta^2 F[y] = \frac{1}{2}\int_\alpha^\beta [P(\eta' + \frac{\phi}{P}\eta)^2 + \underbrace{(Q + \phi - \frac{\phi}{p})\eta^2}_{=0 \text{ if (8.3) holds}}] \,\mathrm{d}x$$

which is positive if we can choose ϕ s.t.

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

If (8.3) holds, then $\delta^2 F > 0$ unless

$$\eta' + \frac{\phi}{P}\eta = 0 \tag{**}$$

on $[\alpha, \beta]$. But $\eta = 0$ at α, so $\eta'(\alpha) = 0$ if (**) holds but then $\eta \equiv 0$ on $[\alpha, \beta]$ (uniqueness of solution to 1st order ODEs), so $(**) \neq 0$.

Method. Does a solution to (8.3) exist on $[\alpha, \beta]^2$
Transform (8.3) into a linear 2nd order ODE by setting $\phi = -Pu'/u$ where $u \neq 0$ on $[\alpha, \beta]$

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

or

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

This is the Jacobi accessory condition.
Need a solution to (8.4) (which is $\mathcal{L}(u) = 0$) s.t. $u \neq 0$ on $[\alpha, \beta]$. This may not exist on a large enough interval

Example.

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

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

(8.4) is $u'' + u = 0$, general solution $u = A\sin x + B\cos x$. Want u to be non-zero on $[\alpha, \beta]$, i.e.

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

possible to avoid B/A on interval smaller than π

$$|\beta - \alpha| < \pi \to \text{ positive nd variation}$$

Example. Back to geodesics on the sphere

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

Found earlier that critical points are segments of great circles
$\theta =$ const, $\theta_0 = \pi/2$ (any great circle is this after a rotation)

$$\frac{\partial^2 f}{\partial(\theta')^2}|_{\theta_0} = 1 = P \quad Q = \cdots = -1$$

$$\delta^2 F[\theta_0 = \frac{\pi}{2} l\eta] = \frac{1}{2}\int_{\phi_1}^{\phi_2} [(\eta')^2 - \eta^2]\,\mathrm{d}\phi$$

positive if $\phi_2 - \phi_1 < \pi$