Introduction.

Calculus of variations.

X: a family of objects (configurations, shapes, ...)

$X \xrightarrow{\Phi} \mathbb{R}$: a "functional", i.e. a function defined on X.

Problem. Find $x_0 \in X$ s.t. $\Phi(x_0) \leq \Phi(x)$ $\forall x \in X$.

i.e. find the "admissible" $x_0 \in X$ which minimizes Φ

"Calculus of Variations" consists in attacking the problem by means of extensions of the usual differential calculus.

Review of facts about minima of functions.

(A) **Theorem of Weierstrass.** Let $K \subseteq \mathbb{R}^n$ be compact (which means: closed and bounded) and let $K \xrightarrow{f} \mathbb{R}$ be a continuous function.

Then $\exists x_0 \in K$: $\forall x \in K \Rightarrow f(x) \geq f(x_0)$.

In this case minima exist.

(~~B~~ Can we do better?

Definition. $K \xrightarrow{f} \mathbb{R}$ is lower-semicontinuous if $\forall x \in K$: $\underline{\lim}_{y \to x} f(y) \geq f(x)$.

This means that for all sequences $\{a_n\}_{n=0}^{\infty}$ in K,

$$\lim_{n\to\infty} a_n = x_0 \Rightarrow \underline{\lim}_{n\to\infty} f(a_n) \geq f(x_0).$$

Theorem of Weierstrass for l.s.c. functions.

$$\underset{\text{compact}}{\mathbb{R}^n \supseteq K} \xrightarrow{f} \mathbb{R} \text{ is lower semicontinuous}$$

$$\Downarrow$$

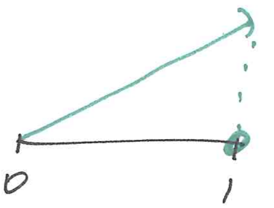
$$\exists x_0 \in K \quad \forall x \in K \Rightarrow f(x) \geq f(x_0)$$

Proof. Open the calculus book where Weierstrass' Theorem is proved for $f \in C([a,b], \mathbb{R})$ and ~~modify~~ check that the proof works for our Thm.

The assumption l.s.c. does not guarantee the existence of maxima:

$$f(x) = \begin{cases} x & \text{if } 0 \le x < 1 \\ 0 & \text{if } x = 1 \end{cases}$$

is l.s.c. but has no maximum in $[0,1]$.



(B) Fermet' Theorem. If $A \subseteq \mathbb{R}^n$ is open and $f \in C^1(A, \mathbb{R})$, if $x_0 \in A$ is a point of (local) minimum for f,

then $\nabla f(x_0) = 0$.

x_0 is a local minimum if $\exists \delta > 0 \ \forall x \in A : |x - x_0| < \delta \Rightarrow f(x) \ge f(x_0)$

The condition in the Theorem is necessary only, but it tipically reduces the number of candidates to be a minimum to a few points.

(C) Criteria based on the Hessian.

Thm. Let $A \subseteq \mathbb{R}^n$ be open and $f \in C^2(A, \mathbb{R})$. If $x_0 \in A$,

$\nabla f(x_0) = 0$ and

$\operatorname{Hess} f(x_0)$ is positive definite,

then

x_0 is a local minimum for f.

Here

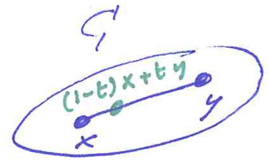
$$\operatorname{Hess} f(x_0) = \begin{bmatrix} \partial_{11} f(x_0) & \partial_{21} f(x_0) & \cdots & \partial_{n1} f(x_0) \\ \partial_{21} f(x_0) & \partial_{22} f(x_0) & & \vdots \\ \vdots & & \ddots & \vdots \\ \partial_{n1} f(x_0) & \cdots & \cdots & \partial_{nn} f(x_0) \end{bmatrix}$$

is the $n \times n$ matrix of the 2^{nd} partial derivatives of f at x_0 and, if A is an $n \times n$ matrix, we say that A is positive definite if and only if

$$\forall h \in \mathbb{R}^n : h \ne 0 \Rightarrow \langle Ah, h \rangle > 0.$$

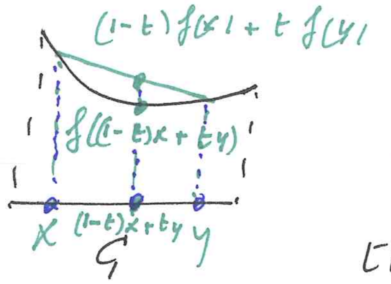
(D) Convexity. Let $G \subseteq \mathbb{R}^n$ be convex:

$\forall x, y \in G$ and $\forall t \in [0,1] \Rightarrow (1-t)x + ty \in G$.



A function $G \xrightarrow{f} \mathbb{R}$ is strongly convex if

$\forall x, y \in G$ $\forall t \in (0,1) \Rightarrow f((1-t)x + ty) < (1-t) f(x) + t f(y)$



Theorem. If $G \subseteq \mathbb{R}^n$ is convex and $G \xrightarrow{f} \mathbb{R}$ is strongly convex, then f is continuous on G and if x_0 is a minimum for f, it is unique: if $x_1 \in G$ and $f(x_1) \le f(x)$ $\forall x \in G$, then $x_1 = x_0$.

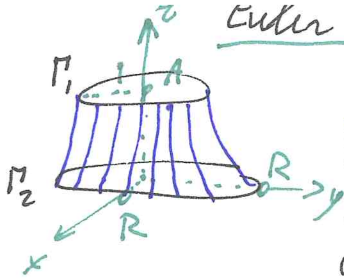
(E) Constrained extreme.

Let $\Omega \xrightarrow{F, G} \mathbb{R}$ be $C^1(\Omega, \mathbb{R})$, $\Omega \subseteq \mathbb{R}^n$ open.
If $x_0 \in \Omega$ is a local minimum for $F(x)$ constrained to $G(x) = 0$, then $\exists \lambda \in \mathbb{R}$ s.t.

$$\nabla F(x_0) = \lambda \cdot \nabla G(x_0).$$

Def. local minimum under the constraint $G(x) = 0$ means that $\exists \delta > 0$ s.t. $\forall x \in \Omega$, if $|x - x_0| < \delta$ and $G(x) = 0$, then $F(x) \ge F(x_0)$; and obviously $G(x_0) = 0$.

Euler – Lagrange equations.



Problem. Find the surface having as boundaries two concentric circles in space and having minimal surface area.

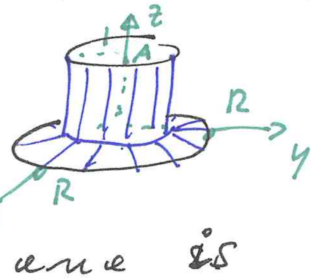
Step 1. Analytic reformulation of the problem.

Consider circles $\Gamma_1 (x^2+y^2=1,\ z=A)$ and $\Gamma_2 (x^2+y^2=R^2>1,\ z=0)$ in $\mathbb{R}^3$.

Suppose the surface is the graph of (smooth) $f:\Omega\to\mathbb{R}$,

$z=f(x,y)$ and $\Omega=\{(x,y):\ 1\le (x^2+y^2)^{1/2}\le R\}$.

Obs. Not all surfaces which are admissible for the problem have this form; e.g.



is a candidate for which no f exists.

The formula for the surface area is

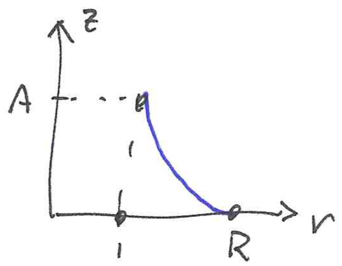
$$S(f)=\iint_\Omega \sqrt{(\partial_x f)^2+(\partial_y f)^2+1}\ dx\,dy$$

Problem: find (if there is one) $f_{min}\in C^1(\Omega,\mathbb{R})$ s.t.

$$S(f_{min})\le S(f)\quad \forall f\in C^1(\Omega,\mathbb{R}),\ \begin{cases} f(x,y)=A>0 & \text{if } x^2+y^2=1\\ f(x,y)=0 & \text{if } x^2+y^2=R^2\end{cases}$$

Step 2. If a minimal f_{min} exists and it is unique, then it is radially symmetric: there is $\varphi\in C^1([1,R],\mathbb{R})$ such that

$$f(x,y)=\varphi(\sqrt{x^2+y^2})\qquad \begin{cases}\varphi(1)=A>0\\ \varphi(R)=0\end{cases}$$



This means that graph(f) is a surface of rotation in $\mathbb{R}$

$\|$

$\{(x,y,z):\ z=f(x,y);\ (x,y)\in\Omega\}$.

Sketch of the proof. Given $f\in C^1(\Omega,\mathbb{R})$ which is admissible for the problem and $\theta\in\mathbb{R}$, let

$$R_\theta f(x,y)=f\left(\begin{pmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{pmatrix}\begin{pmatrix}x\\ y\end{pmatrix}\right)=f(\cos\theta\cdot x-\sin\theta\cdot y,\ \sin\theta\cdot x+\cos\theta\cdot y)$$

Then $R_\theta f\in C^1(\Omega,\mathbb{R})$ and it is admissible for the problem, and $S(R_\theta f)=S(f)$ (EXERCISE).

If $f = f_{min}$ minimizes $\mathcal{J}(f)$ over all admissible functions, so it does $R_\theta f_{min}$. But the minimizer is unique by hypothesis, hence

$$R_\theta f_{min} = f_{min} \quad \forall \theta \in \mathbb{R}.$$

This is equivalent to asking that

$$f_{min}(x,y) = \varphi_{min}(\sqrt{x^2+y^2})$$

for some $\varphi_{min} \in C^1([1,R],\mathbb{R})$, $\begin{cases} \varphi(1) = A > 0 \\ \varphi(R) = 0 \end{cases}$

(EXERCISE: show this).

Step 3. For admissible $\varphi \in C^1([1,R],\mathbb{R})$ (i.e. $\varphi(1)=A$, $\varphi(R)=0$), and $f(x,y) = \varphi(\sqrt{x^2+y^2})$,

$$\mathcal{J}(f) = \iint_{1\le\sqrt{x^2+y^2}\le R} \left\{1+\left[\partial_x \varphi(\sqrt{x^2+y^2})\right]^2+\left[\partial_y\varphi(\sqrt{x^2+y^2})\right]^2\right\}^{1/2} dx\,dy$$

$$= \iint_{1\le\sqrt{x^2+y^2}\le R} \left\{1+\left(\frac{x}{\sqrt{x^2+y^2}}\right)^2\cdot\left(\dot\varphi(\sqrt{x^2+y^2})\right)^2+\left(\frac{y}{\sqrt{x^2+y^2}}\right)^2\cdot\left(\dot\varphi(\sqrt{x^2+y^2})\right)^2\right\}^{1/2} dx\,dy$$

$$= \int_0^{2\pi} d\theta \cdot \int_1^R \left\{1+(\dot\varphi(r))^2\right\}^{1/2} r\,dr \quad \text{in polar coordinates } \begin{cases} x = r\cos\theta \\ y = r\sin\theta \end{cases}$$

$$= 2\pi\cdot\int_1^R \left\{1+\dot\varphi(r)^2\right\}^{1/2}\cdot r\,dr \;:=2\pi\,\mathcal{S}(\varphi)$$

where $\mathcal{S}$ is a "functional" defined on admissible φ's.

We want to find φ_{min} admissible s.t.

$$\mathcal{S}(\varphi_{min}) \le \mathcal{S}(\varphi) \quad \text{for all admissible } \varphi\text{'s}.$$

If the solution f_{min} of the original problem is unique, we have so found it.

Having so reduced the problem ~~from~~ two to one dimensions, we make it easier to understand by generalizing it.

Let $L : \mathbb{R} \times \mathbb{R} \times [\alpha, \beta] \longrightarrow \mathbb{R}$ be C^2.

Let $a, b \in \mathbb{R}$.

For $x \in C^1([\alpha, \beta], \mathbb{R})$, consider

$$\Lambda(x) = \int_\alpha^\beta L(x(t), \dot{x}(t), t)\, dt$$

Problem: find "local" minima / maxima for Λ under the constraints $x(\alpha) = a$, $x(\beta) = b$.

What's the meaning of "local"? To be discussed later.

Suppose we just look for minima (necessary conditions).

In our real-life problem: $\alpha = 1$, $\beta = R$, $a = A$, $b = 0$,

$$L(x, \dot{x}, t) = \sqrt{1 + \dot{x}^2} \cdot t$$

If $x \in C^1([\alpha, \beta], \mathbb{R})$ is admissible (i.e. $x(\alpha) = a$, $x(\beta) = b$), then $\forall h \in C^1([\alpha, \beta], \mathbb{R})$ s.t. $h(\alpha) = 0 = h(\beta)$, $x + h$ is admissible.

$$\Lambda(x+h) - \Lambda(x) = \int_\alpha^\beta \left\{ L((x+h)(t), (\dot{x} + \dot{h})(t), t) - L(x(t), \dot{x}(t), t) \right\} dt$$

$$= \int_\alpha^\beta \left[\partial_x L(x(t), \dot{x}(t), t)\, h(t) + \partial_{\dot{x}} L(x(t), \dot{x}(t), t) \cdot \dot{h}(t) \right] dt + \text{Error}$$

$$\text{Error} = \int_\alpha^\beta \left[A(t)\, h(t)^2 + 2B(t)\, h(t) \dot{h}(t) + C(t)\, \dot{h}(t)^2 \right] dt$$

where A, B, C involve $\partial_{xx} L$, $\partial_{x\dot{x}} L$, $\partial_{\dot{x}\dot{x}} L$ computed in $y(t) = x(t) + \Theta(t) h(t)$ $\quad (|\Theta(t)| \leq 1)$ and

$\dot{y}(t) = \dot{x}(t) + \Psi(t)\, \dot{h}(t) \quad (|\Psi(t)| \leq 1)$

by Taylor's formula for L with Lagrange's remainder; hence, if $|h(t)| \leq 1$ $\forall t \in [\alpha, \beta]$ and $|\dot{h}(t)| \leq 1$ $\forall t \in [\alpha, \beta]$,

$$\text{Error} \leq E \cdot \left\{ \int_\alpha^\beta h(t)^2\, dt + \int_\alpha^\beta \dot{h}(t)^2\, dt \right\}$$

by easy estimates,

$$\leq F \cdot \left\{ \max |h(t)| + \max |\dot{h}(t)| \right\}^2.$$

For $h \in C^1([\alpha,\beta],\mathbb{R})$, let

$$\|h\|_{C^1} := \max_{\alpha\le t\le\beta} |h(t)| + \max_{\alpha\le t\le\beta} |\dot h(t)| :$$

this is our measure of how $\left\{\begin{array}{l}\text{small}\\ \text{large}\end{array}\right.$ h is in $C^1([\alpha,\beta],\mathbb{R})$

$$\text{Error} \le F \cdot \|h\|^2_{C^1}.$$

$x_0 \in C^1([\alpha,\beta],\mathbb{R})$ is a local minimum (constrained to $x(\alpha)=a$, $x(\beta)=b$) for Λ if $\exists \delta>0$ s.t. $\forall h \in C^1([\alpha,\beta],\mathbb{R})$ with $h(\alpha)=h(\beta)=0$ we have

$$\|h\|_{C^1([\alpha,\beta],\mathbb{R})} \le \delta \Rightarrow \Lambda(x_0+h) \ge \Lambda(x_0)$$

<u>Fermat-type Theorem.</u> x_0 is a local minimum for $\Lambda \Rightarrow \forall h \in C^1([\alpha,\beta],\mathbb{R})$ s.t. $h(\alpha)=h(\beta)=0$ we have

$$D\Lambda(x_0)(h) := \int_\alpha^\beta \left[\partial_x L(x(t),\dot x(t),t)\cdot h(t) + \partial_{\dot x} L(x(t),\dot x(t),t)\,\dot h(t)\right] dt = 0$$

Pf. similar to that of Fermat's Theorem.
Fix h and let $s \in \mathbb{R}$. If x_0 is a (local) minimum, then

$$0 \le \Lambda(x_0+sh) - \Lambda(x_0) = D\Lambda(x_0)(sh) + \text{Error}(sh)$$

$\uparrow$ if s is s.t. $|s|\cdot\|h\|_{C^1} \le \delta$ $\qquad$ $= s\cdot D\Lambda(x_0)(h) + \text{Error}(sh)$

Dividing by $s>0$: $\quad 0 \le D\Lambda(x_0)(h) + \frac{1}{s}\cdot\text{Error}(sh)$

But $\left|\frac{1}{s}\text{Error}(sh)\right| \le F\cdot s\cdot \|h\|^2_{C^1}$,

hence $\quad 0 \le D\Lambda(x_0)(h) + F\cdot s\cdot\|h\|^2_{C^1} \xrightarrow[s\to 0^+]{} D\Lambda(x_0)(h)$

so that $\quad D\Lambda(x_0)(h) \ge 0$.

Dividing by $s<0$, we see that $D\Lambda(x_0)(h) \le 0$,

hence $\quad D\Lambda(x_0)(h) = 0 \quad \forall h \in C^1$ with $\left\{\begin{array}{l}h(\alpha)=0\\ h(\beta)=0\end{array}\right.$

Obs. that the same reasoning works for (local) maxima.

Obs. that $h \mapsto D\Lambda(x_0)(h)$ is linear:

$$D\Lambda(x_0)(\lambda h + \mu k) = \lambda\cdot D\Lambda(x_0)(h) + \mu\cdot D\Lambda(x_0)(k),$$

$\forall\, h, k \in C^1$ and $\lambda, \mu \in \mathbb{R}$.

In fact, $h \mapsto D\Lambda(x_0)(h)$ plays the rôle of the (1^{st} order) differential in this infinite dimensional context.

$$D\Lambda(x_0)(h) \overset{\text{"metaphor"}}{\longleftrightarrow} df(x_0)(h) = \langle \nabla f(x_0), h \rangle$$
for $f \in C^1(\mathbb{R}^n, \mathbb{R})$

The problem has become: find (admissible) $x \in C^1([\alpha, \beta], \mathbb{R})$ s.t. $\forall h \in C^1([\alpha,\beta], \mathbb{R}) : h(\alpha) = h(\beta) = 0$

$$\Rightarrow \quad D\Lambda(x_0)(h) := \int_\alpha^\beta \left[\partial_x L(x, \dot{x}, t) \cdot h + \partial_{\dot{x}} L(x, \dot{x}, t) \cdot \dot{h} \right] dt = 0.$$

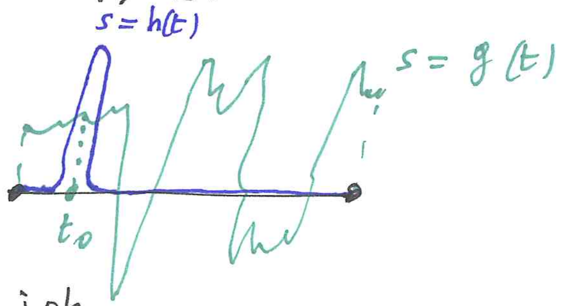
If $x \in C^2([\alpha, \beta], \mathbb{R})$ I could integrate by parts ⟶ MIRACLE: I can even if $x \notin C^2$ by a deep Theorem of C.O.V.

$$0 = \int_\alpha^\beta \partial_x L(x, \dot{x}, t) \cdot h \, dt + \Big[\partial_{\dot{x}} L(x(t), \dot{x}(t), h(t)) \cdot h(t) \Big]_{t=\alpha}^{\beta}$$
$$- \int_\alpha^\beta \left\{ \frac{d}{dt} \partial_{\dot{x}} L(x(t), \dot{x}(t), t) \right\} h(t) \, dt$$
$$= \int_\alpha^\beta \left[\partial_x L(x, \dot{x}, t) - \frac{d}{dt} \partial_{\dot{x}} L(x, \dot{x}, t) \right] h(t)\, dt, \text{ because } h(\alpha) = h(\beta) = 0.$$

Fundamental Lemma of C.O.V.

If $g \in C^1([\alpha, \beta], \mathbb{R})$ and $\int_\alpha^\beta g(t)\, h(t)\, dt = 0$
$\forall h \in C^1([\alpha, \beta], \mathbb{R})$ s.t. $h(\alpha) = h(\beta) = 0$,
then $g = 0$ on $[\alpha, \beta]$.

Pf. by pictures.



If $g(t_0) > 0$, then
$g(t) > 0$ near t_0, and
I can design h to do the job.

Applying the Lemma to the last equality, we obtain the Euler-Lagrange equation for L.

Theorem. Let $L \in C^2(\mathbb{R}\times\mathbb{R}\times[\alpha,\beta], \mathbb{R})$; $a, b \in \mathbb{R}$.
If $x_0 \in C^1([\alpha,\beta], \mathbb{R})$ is a local minimum for

$$\Lambda(x) = \int_\alpha^\beta L(x(t), \dot{x}(t), t)\, dt$$

under the constraints $x(\alpha) = a$, $x(\beta) = b$,

then

$x = x_0$ satisfies the 2^{nd} order differential equation

$$\text{(E-L)} \qquad \frac{d}{dt}\, \partial_{\dot{x}} L(x(t), \dot{x}(t), t) - \partial_x L(x(t), \dot{x}(t), t) = 0$$

(E-L) is the Euler-Lagrange equation for the Lagrangian L.

Back to our real-life problem:

$$L(x, \dot{x}, t) = t(\dot{x}^2+1)^{1/2} \Rightarrow \partial_{\dot{x}} L = \frac{t\dot{x}}{(\dot{x}^2+1)^{1/2}}$$

$$x(1) = A \;;\; x(R) = 0$$

produces (E-L) $$\frac{d}{dt}\left[\frac{t\dot{x}}{(\dot{x}^2+1)^{1/2}}\right] = 0$$

which means $$\frac{t\dot{x}}{(\dot{x}^2+1)^{1/2}} = -K \text{ constant.}$$

Since $\dot{x} \le 0$ for $t > 0$ (this is clear from the problem), $K > 0$.

$$t^2\dot{x}^2 = K^2(\dot{x}^2+1)$$

$$(t^2 - K^2)\dot{x}^2 = K^2 \quad (\Rightarrow t \geq K)$$

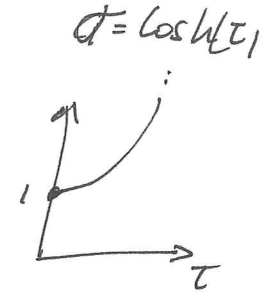
$$\dot{x} = -\frac{K}{(t^2-K^2)^{1/2}}$$

Thus,

$$-x(t) = x(R) - x(t) = \int_t^R \dot{x}(s)\, ds = -\int_t^R \frac{K}{(s^2-K^2)^{1/2}}\, ds$$

$$= -\int_{K\cdot\cosh(v)=t}^{K\cdot\cosh(v)=R} \frac{K\cdot\sinh(v)\cdot dv}{\sinh(v)} = -K\left[\cosh^{-1}(R/K) - \cosh^{-1}(t/K)\right]$$

$s = K\cdot\cosh(v)$, $\cosh^2(v) - 1 = \sinh^2(v)$



$$x(t) = K\left[\cosh^{-1}(R/K) - \cosh^{-1}(t/K)\right]$$

We impose

$$A = x(1) = K\cdot\left[\cosh^{-1}(R/K) - \cosh^{-1}(1/K)\right]$$

Since $\cosh(\tau) = \frac{e^{\tau}+e^{-\tau}}{2} = \sigma$, we have

$$\text{or } \tau = \cosh^{-1}(\sigma) = \log\left(1+\sqrt{\sigma^2-1}\right)$$

The solution to the original problem provided by (E-L) is

$$x(t) = K\cdot\log\frac{K+\sqrt{R^2-K^2}}{K+\sqrt{t^2-K^2}}$$

with $A = x(1) = K.\ \log\frac{K+\sqrt{R^2-K^2}}{K+\sqrt{1-K^2}}$

and $0 \le K \le 1$.

Obs. that $K=0 \Rightarrow A=0$

$K=1 \Rightarrow A = \log\left(1+\sqrt{R^2-1}\right)$

It can be computed $\frac{dA}{dK} > 0$.

Problem: when $A > \log\left(1+\sqrt{R^2-1}\right)$ the equation (E-L) does not provide solutions anymore!!

What has happened?

This analysis should be done inverting functions.