Minima of functionals.

$\mathfrak{F} \xrightarrow{\Lambda} \mathbb{R}$ $\quad m = \inf\{\Lambda(\varphi) : \varphi \in \mathfrak{F}\}$

- Is $\inf = \min$?
- Are there $\varphi \in \mathfrak{F}$: $\Lambda(\varphi) = m$?
- How many?
- Efficient ways to find them?
- Do they have special properties?
- Suppose $\inf \in \mathbb{R}$, but inf is not min. Can we extend in such a way that

$$\begin{array}{ccc} \mathfrak{F} & \xrightarrow{\Lambda} & \mathbb{R} \\ \downarrow i & \nearrow_{\bar{\Lambda}} & \\ \bar{\mathfrak{F}} & & \end{array}$$

$\exists \bar{\varphi} \in \bar{\mathfrak{F}}$ s.t.

(i) $\bar{\Lambda}(\bar{\varphi}) = m$

(ii) $\exists \{\varphi_n\}$ in $\mathfrak{F}$: $\varphi_n \xrightarrow[\text{in } \bar{\mathfrak{F}}]{} \bar{\varphi}$

(the latter requires some topology on $\mathfrak{F}, \bar{\mathfrak{F}}$).

Some inspiring examples from calculus.

(i) Fermat' Theorem. If $\mathbb{R}^n \supseteq \Omega \xrightarrow{f} \mathbb{R}$ is C^1 and $x_0 \in \Omega$ (open) is a local max/min for f, then $\nabla f(x_0) = 0$.

(ii) If $\mathbb{R}^n \xrightarrow{f} \mathbb{R}$ is convex, ~~f(tx + ...)~~ $f(tx + (1-t)y) \le t f(x) + (1-t) f(y)$ $\forall t \in [0,1]$ $\forall x, y \in \mathbb{R}^n$, then f has at most one point of minimum.

(iii) If (K, d) is a compact metric space and $K \xrightarrow{f} \mathbb{R}$ is continuous, then f has a ~~point of~~ minimum in K. (Weierstrass).

Exercise. $(X,d) \xrightarrow{f} \mathbb{R}$ is lower semi continuous if

$$\underline{\lim}_{x \to x_0} f(x) \geq f(x_0).$$

Prove that Weierstrass holds under the weaker assumption that $K \xrightarrow{f} \mathbb{R}$ be l.s.c.

How do we do such things in an infinite dimensional context?

Why should we care?

Classical mechanics is a starting point.

Euler - Lagrange equation.

$\mathcal{F} = \{ \varphi \in C^1([a,b], \mathbb{R}) : \varphi(a) = \alpha,\ \varphi(b) = \beta \}$

$L \in C^2([a,b] \times \mathbb{R} \times \mathbb{R}, \mathbb{R})$; $L = L(\tau, x, \dot{x})$

(1) $$\Lambda(\varphi) = \int_a^b L(\tau, \varphi(\tau), \dot{\varphi}(\tau))\, d\tau$$

Suppose that $\varphi_0 \in \mathcal{F}$ and that $\Lambda(\varphi_0) = m = \inf\{\Lambda(\varphi) : \varphi \in \mathcal{F}\}$.

Let $\mathcal{F}_0 = \{ h \in C^1([a,b], \mathbb{R}) : h(a) = h(b) = 0 \}$.

Then $\forall h \in \mathcal{F}_0$:

(2) $$0 = \int_a^b \left[\frac{\partial L}{\partial x}(\tau, \varphi_0(\tau), \dot{\varphi}_0(\tau))\, h(\tau) + \frac{\partial L}{\partial \dot{x}}(\tau, \varphi_0(\tau), \dot{\varphi}_0(\tau))\, \dot{h}(\tau) \right] d\tau$$

Proof. Recall Taylor's formula for $f \in C^2(\Omega, \mathbb{R})$, $\Omega \subseteq \mathbb{R}^n$ open, $x_0 \in \Omega$:

(T) $$f(x_0 + \varepsilon v) = f(x_0) + \nabla f(x_0) \cdot v\, \varepsilon + \frac{1}{2} v^T \operatorname{Hess} f(x_0 + \varepsilon \Theta v) v \cdot \varepsilon^2$$

where $\Theta = \Theta(x_0, v, \varepsilon) \in [0,1]$, $v \in \mathbb{R}^n$ fixed.

Exercise. Prove (T) using Taylor's formula from 1-variable calculus, Lagrangian remainder.

For all $h \in \mathcal{F}_0$ and $\varepsilon \in \mathbb{R}$:

$0 \leq \Lambda(\varphi_0 + \varepsilon h) - \Lambda(\varphi_0)$ because φ_0 is a minimum

$$= \int_a^b \left[\frac{\partial L}{\partial x}(\tau, \varphi_0, \dot{\varphi}_0) h(\tau) + \frac{\partial L}{\partial \dot{x}}(\tau, \varphi_0, \dot{\varphi}_0) \dot{h}(\tau) \right] d\tau \cdot \varepsilon$$
$$+ \frac{1}{2} \int_a^b \left(\partial_{xx} L \cdot h^2 + 2 \partial_{x\dot{x}} L \cdot h \dot{h} + \partial_{\dot{x}\dot{x}} L\, \dot{h}^2 \right) d\tau \cdot \varepsilon^2$$

where the derivatives in the 2nd integral are computed in $(\tau, \varphi_0(\tau) + \varepsilon \Theta h(\tau), \dot{\varphi}_0(\tau) + \varepsilon \Theta \dot{h}(\tau))$ and $\Theta = \Theta(\tau, \varepsilon, h) \in [0,1]$

$$= \int_a^b \left[\frac{\partial L}{\partial x}(\tau, \varphi_0, \dot{\varphi}_0) h + \frac{\partial L}{\partial \dot{x}}(\tau, \varphi_0, \dot{\varphi}_0) \dot{h} \right] d\tau \cdot \varepsilon$$
$$+ O(\varepsilon^2) \text{ because } L \in C^2;\ h, \varphi_0 \in C^1:$$

Divide by ε and let $\varepsilon \to 0$ to obtain

$$\underset{\text{if } \varepsilon > 0}{0 \leq} \int_a^b \left(\partial_x L \cdot h + \partial_{\dot{x}} L \cdot \dot{h} \right) dt \underset{\text{if } \varepsilon < 0}{\leq} 0$$

■

<u>Fundamental lemma of the calculus of variations.</u>

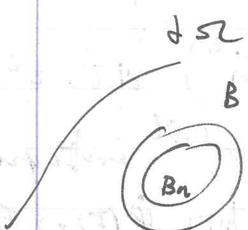
$\Omega \subseteq \mathbb{R}^n$ open and $v \in L^1_{loc}(\Omega)$.

If $\int_\Omega v(x)\, \psi(x)\, dx = 0 \quad \forall \psi \in C_0^\infty(\Omega)$

then $v = 0$ a.e. in Ω.

Proof. The Lemma is proved if we show the following. For each ball $\overline{B} \subset \Omega$ we have $\int_{\overline{B}} |v(x)|\, dx = 0$.

Fix $\eta > 0$ and $\overline{B} \subseteq \Omega$ and find $B_\eta \subsetneq B$ concentric with it s.t. $\int_{B \setminus B_\eta} |v|\, dx < \eta$.



Let $\varepsilon_0 = \text{radius}(B) - \text{radius}(B_\eta)$.

Let $\varphi \in C_c^\infty(\mathbb{R}^n)$; $\varphi \geq 0$;

$\varphi(x) = 0$ if $|x| \geq 1$; $\int \varphi(x)\, dx = 1$

and for $0 < \varepsilon \leq \varepsilon_0$ let

$$\varphi_\varepsilon(x) = \frac{1}{\varepsilon} \varphi\left(\frac{x}{\varepsilon}\right)$$

Then
$$\int_B |v(x)|\, dx = \int_{B \setminus B_\eta} |v(x)|\, dx + \int_{B_\eta} |v(x)|\, dx$$

$$\leq \eta + \int_{B_\eta} |v(x) - v * \varphi_\varepsilon(x)|\, dx = \eta + \text{Int}.$$

where $v * \varphi_\varepsilon(x) = \int v(x-y)\, \varphi_\varepsilon(y)\, dy = 0$ for $x \in B_\eta$ by hypothesis.

$$\text{Int} = \int_{B_\eta} \left| v(x) - \int_{|y| \leq \varepsilon} v(x-y)\, \varphi_\varepsilon(y)\, dy \right| dx$$

$$\leq \int_{B_\eta} \int_{|y| \leq \varepsilon} |v(x) - v(x-y)|\, \varphi_\varepsilon(y)\, dy \cdot dx$$

$$= \int_{B_\eta} \int_{|y| \leq \varepsilon} |v(x) - v(x-y)| \cdot \varphi\left(\frac{y}{\varepsilon}\right) \frac{dy}{\varepsilon} \cdot dx$$

$$= \int_{B_\eta} \int_{|z| \leq 1} |v(x) - v(x - \varepsilon z)|\, \varphi(z)\, dz \cdot dx$$

$$= \int_{|z| \leq 1} \varphi(z) \cdot \int_{B_\eta} |v(x) - v(x - \varepsilon z)|\, dx \cdot dz \qquad (*)$$

<u>Lemma.</u> $\int_{B_\eta} |v(x) - v(x - \varepsilon z)|\, dx \xrightarrow[\varepsilon \to 0]{} 0$ for each z.

Pf. of the Lemma.

The limit holds if v is continuous on $\overline{B}$ (Exercise: use uniform continuity).

In the general case fix $\delta > 0$ and find $v \in C(\bar{B}, \mathbb{R})$ s.t.

$$\int |u - v|\, dx \le \delta. \text{ Then,}$$

$$\int_{B_\eta} |u(x) - u(x - \varepsilon z)|\, dx \le$$

$$\le \int_{B_\eta} |u(x) - v(x)| + |v(x - \varepsilon z) - u(x - \varepsilon z)|\, dx + \int_{B_\eta} |v(x) - v(x - \varepsilon z)|\, dx$$

$$\le 2\delta + \int_{B_\eta} |v(x) - v(x - \varepsilon z)|\, dx \xrightarrow[\varepsilon \to 0]{} 2\delta$$

because v is continuous.

This proves the Lemma being δ arbitrary. ∎

Back to the expression (#) we have:

(a) $g_\varepsilon(z) := \int_{B_\eta} |u(x) - u(x - \varepsilon z)|\, dx \xrightarrow[\varepsilon \to 0]{} 0$

for all z by the Lemma;

(b) $0 \le g_\varepsilon(z) \le 2 \int_B |u(x)|\, dx = 2 \cdot \|u\|_{L^1(B)}$

By Lebesgue Dominated convergence Thm.

$$\int_{|z| \le 1} \varphi(z) \cdot \int_{B_\eta} |u(x) - u(x - \varepsilon z)|\, dx \cdot dz =$$

$$= \int_{|z| \le 1} \varphi(z)\, g_\varepsilon(z)\, dz \xrightarrow[\varepsilon \to 0]{} 0.$$

Thus $\int_B |u(x)|\, dx \le \eta \quad \forall \eta \Rightarrow u = 0$ a.e. on B. ∎

Suppose that the minimizer φ_0 in (2) is $\varphi_0 \in C^2([a,b])$. We can integrate by parts the second summand and using $h(a) = h(b) = 0$:

$$0 = \int_a^b \left[\frac{\partial L}{\partial x}(\tau, \varphi_0, \dot{\varphi}_0) - \frac{d}{d\tau} \frac{\partial L}{\partial \dot{x}}(\tau, \varphi_0, \dot{\varphi}_0) \right] h(\tau)\, d\tau \qquad (2\tfrac{1}{2})$$

$$\forall h \in \mathcal{F}_0$$

hence, by the F.L. of C.o.V. we have that φ_0 satisfies:

$$\frac{d}{d\tau}\left\{ \frac{\partial L}{\partial \dot{x}}(\tau, x, \dot{x}) \right\} - \frac{\partial L}{\partial x}(\tau, x, \dot{x}) = 0. \qquad (3)$$

(3) is called the Euler - Lagrange equation for the "action" Λ.

Theorem. If $\Lambda(\varphi_0) = \min\{\Lambda(\varphi) : \varphi \in \mathcal{F}\}$ with $L \in C^2$ and if $\varphi_0 \in C^2$, then $x = \varphi_0$ satisfies (3) on $[a, b]$.

- We have no existence result
- (3) might hold without φ_0 being minimal
- minima might exist, but not in C^2 (or C^1).

Suppose we do not want to assume that $\varphi_0 \in C^2$. Then we are left with $\varphi_0 \in C^1$ and

$$(\star)\quad \int_a^b \left(f(\tau)\, h(\tau) + g(\tau)\, \dot h(\tau) \right) d\tau \qquad \forall h \in C^1([a,b]),\ h(a) = h(b) = 0.$$

where $f(\tau) = \partial_x L(\tau, \varphi_0(\tau), \dot\varphi_0(\tau))$

and $g(\tau) = \partial_{\dot x} L(\tau, \varphi_0(\tau), \dot\varphi_0(\tau))$

so that $f, g \in C([a,b], \mathbb{R})$ if $\varphi_0 \in C^1$, $L \in C^1$.

Something can be said under these weaker assumptions.

Proposition. If $L \in C^1$ and φ_0 satisfies (2) then:

(i) $\tau \mapsto g(\tau) = \partial_{\dot x} L(\tau, \varphi_0(\tau), \dot\varphi_0(\tau)) \in C^1([a,b])$

and

(ii) $\frac{d}{d\tau} \partial_{\dot x} L(\tau, \varphi_0(\tau), \dot\varphi_0(\tau)) = \partial_x L(\tau, \varphi_0(\tau), \dot\varphi_0(\tau))$

Sketch of the proof.

(A) Let $F(\tau) = \int_a^\tau f(\sigma)\, d\sigma$, $F \in C^1$.

(B) Integrating by parts we have

$$\int_a^b f(\tau)\, h(\tau) + F(\tau)\, \dot h(\tau)\, d\tau = 0 \quad \forall h \dots$$

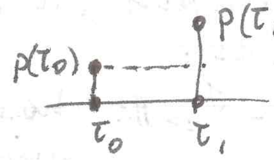
(C) From (B) and $(\star)$ we deduce that

$$\int_a^b \left(F(\tau) - g(\tau)\right) \dot h(\tau)\, d\tau = 0 \quad \forall h \dots$$

(D) Lemma. If $\int_a^b p(\tau)\, q(\tau)\, d\tau = 0$ $\forall q \in C([a,b])$ with $\int_a^b q(\tau)\, d\tau = 0$ and $p \in C([a,b])$ fixed, then

$p = \text{const}$.

Pf. Exercise. Try by contradiction starting with



(E) From (D) and (C) we have

$$F(\tau) = g(\tau) + \text{const},$$

hence $g \in C^1$ and $g' = f$. ∎

A few things we have learned.

(1) Necessary conditions for minima lead to equations. Equation (2),

$$0 = \int_a^b (\partial_x L \cdot h + \partial_{\dot{x}} \cdot \dot{h})\, dt \quad \forall h \in \mathcal{F}_0$$

corresponds to Fermat condition in $\mathbb{R}^n$, written in terms of differentials:

if $\underset{\substack{\Downarrow \\ x_0}}{\Omega} \xrightarrow{f} \mathbb{R}$ has a (local) minimum at x_0

then $df(x_0)h := \nabla f(x_0) \cdot h = 0 \quad \forall h \in \mathbb{R}^n$.

(2) ~~The~~ Integration by parts can be done if $\varphi_0 \in C^2$ and leads to a better expression for the differential:

$$0 = \int_a^b \left(\partial_x L - \frac{d}{dt}\partial_{\dot{x}} L\right) h \, dt$$

In higher dimensions this issue is very subtle.

The advantage of the new formulation is that the differential is now in terms of the $L^2((a,b))$ scalar product.

(3) Moreover, by measure theoretic results, the equation in (2) holds iff $\boxed{\partial_x L - \frac{d}{dt}\partial_{\dot{x}} L = 0}$

which is a bona fide O.D.E., corresponding to $\nabla f(x_0) = 0$.

We want to find an easier way to deal with the integration by parts and to get to a (possibly abstract and "weak") version of the differential equation.

This will take us through a path:

- some functional analysis
- weak derivatives and Sobolev spaces
- direct method in C.o.V.

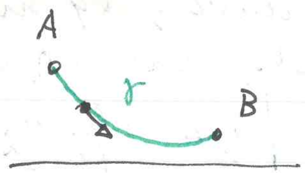
Note on minima/maxima. The Euler-Lagrange eq. is satisfied alike by minima, maxima and other stationary (saddle) points. To have better information, we might extend the notion of (positive definite) Hessian. This would lead to the 2^{nd} variation of Λ.

See DACOROGNA for more on this.

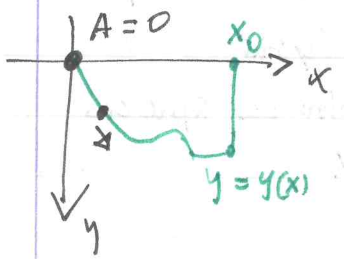
Examples.

Brachistochrone.

Among the curves γ joining A, B fixed in the plane, find the one with the property that: a point constrained on γ and falling under the influence of gravity alone gets from A to B in the ~~xxxxx~~ least time.



m: point mass
g: gravitational acc.



Potential energy:

$V(y) = -mgy$

Kinetic energy:

$T(y) = \frac{1}{2} m v^2$

$V(y) + T(y) = V(0) + T(0) = 0$

(conservation of energy).

Equating: $v = \sqrt{2gy}$ is the speed.

The time required to cover the distance

$ds = (dx^2 + dy^2)^{1/2}$

is $dt = \frac{ds}{v} = \frac{(1 + (dy/dx)^2)^{1/2}}{\sqrt{2gy}} dx$

The total time of descent is

$$T = \int_0^{x_0} \frac{(1 + y'(x)^2)^2}{\sqrt{2g \cdot y(x)}} dx$$

which is the action functional for the Lagrangian $L(x, y, y') = \frac{(1 + y'^2)^{1/2}}{\sqrt{2gy}} = \frac{1}{\sqrt{2g}} \sqrt{\frac{1 + y'^2}{y}}$

Trick of general use if $L = L(y, y')$.

$\frac{d}{dx} L = L_y \cdot y' + L_{y'} \cdot y''$ and

$\frac{d}{dx}(y' L_{y'}) = y'' L_{y'} + y' \frac{d}{dx} L_{y'}$ imply

$0 = \left(\frac{d}{dx} L_{y'} - L_y\right) y' = \frac{d}{dx}(y' L_{y'} - L)$

$\Rightarrow y' L_{y'} - L = \text{const.}$

For the brachistochrone equation:

$c_0 = y' \cdot \frac{y'}{y^{1/2}(1 + y'^2)^{1/2}} - \frac{(1 + y'^2)^{1/2}}{y^{1/2}}$

$= \frac{1}{y^{1/2} \cdot (1 + y'^2)^{1/2}}$

i.e. $y(1 + y'^2) = c_1$

i.e. $x = \int \left(\frac{y}{c_1 - y}\right)^{1/2} dy$

which can be solved using

$y = c_1 \cdot \sin^2 \tau$; $dy = 2 c_1 \cdot \sin \tau \cdot \cos \tau \cdot d\tau$

$x = 2c_1 \int \sin^2 \tau \, d\tau$

$$x = 2c_1 \cdot \int \sin^2\tau \, d\tau$$
$$= 2c_1 \cdot \int \frac{1-\cos(2\tau)}{2}\, d\tau$$
$$= c_1 \cdot \left(\tau - \frac{\sin(2\tau)}{2}\right) + c_2$$

Finally :

$$\begin{cases} x(\tau) = c_1 \cdot \left(\tau - \frac{\sin(2\tau)}{2}\right) + c_2 \\ y(\tau) = c_1 \cdot \sin^2(\tau) = c_1 \cdot \frac{1-\cos(2\tau)}{2} \end{cases}$$

$\tau = 0 \Rightarrow (x, y) = (0, 0) \Rightarrow c_2 = 0$

$x(\tau) = x_0 \Rightarrow y(\tau) = y_0$

i.e. c_1 is determined by

$$\begin{cases} x_0 = c_1 \cdot \left(2\tau_0 - \frac{\sin(2\tau_0)}{2}\right) \\ y_0 = c_1 \cdot \frac{1 - \cos(2\tau_0)}{2} \end{cases}$$

and the solution becomes

$x(\tau) = c_1 \cdot (2\tau - \sin(2\tau))/2$

$y(\tau) = c_1 \cdot (1 - \cos(2\tau_0))/2$.

Of the three main problems we'd like to solve:

1. Existence of minima
2. Uniqueness of minima
3. Checkable conditions for perspective minima;

we have just tackled 3 and we have run into problems (what if $\varphi \notin C^3$?).

About 1, to have something like Weierstrass' Thm. we would like to have compactness.

Let $C_0^1([a,b]) = \{\varphi \in C^1([a,b]) : \varphi(a) = \varphi(b) = 0\}$(*)

with norm:

$$\|\varphi\|_{C^1} = \sup_{[a,b]} |\varphi(t)| + \sup_{[a,b]} |\dot\varphi(t)|.$$

Suppose $\Lambda(\varphi_n) \xrightarrow[n\to\infty]{} \inf\{\Lambda(\varphi) : \varphi \in C_0^1;\ \|\varphi\|_{C^1} \le R\}$ with $R > 0$. If (a) $\{\varphi \in C_0^1 : \|\varphi\|_{C^1} \le R\}$ were compact in C^1 and (b) if Λ were continuous, I could use Weierstrass proof and find an extremal φ_R under the condition $\|\varphi_R\|_{C^1} \le R$.
(This would be something!).
(b) is true (Exercise),
but (a) is not (Exercise).
This turns out to be a major problem.

(*) I'm assuming that $\mathcal{F} = \mathcal{F}_0$.