LESSON 1 - 3 MARCH 2021

CALCULUS OF VARIATIONS: The study of minimization problems

X = set, $F : X \to \mathbb{R}$ (often $\mathbb{R} \cup \{\pm\infty\}$) function

We want to solve:

$$\min \{F(u) \mid u \in X\} \quad , \quad \operatorname{argmin} \{F(u) \mid u \in X\}$$

This is a real number, called MINIMUM

These are elements of X, called MINIMIZERS

We will look at the following classes of methods to study minimization problems:

- INDIRECT METHODS
- DIRECT METHODS
- RELAXATION
- Γ - CONVERGENCE

Let us see some basic examples to see what the above words mean:

EXAMPLE 1 $F : \mathbb{R} \to \mathbb{R}$, $F(x) = x^2 - 4x$. What is $\min\{F(x) \mid x \in \mathbb{R}\}$?

INDIRECT METHOD: Find candidate minimizers. If $\hat{x}$ is min. then $F'(\hat{x}) = 0$. Now $F'(x) = 2x - 4 = 0$ iff $x = 2$.
So $\hat{x} = 2$ is a candidate minimizer and minimum value is $F(2) = -4$

CLAIM $\hat{x} = 2$ is the UNIQUE minimizer of F. Thus

$$\min_{\mathbb{R}} F = -4 \quad , \quad \operatorname{argmin}_{\mathbb{R}} F = \{2\} \subseteq \mathbb{R}$$

Proof We need to show that

1) $F(x) \geq F(2) \quad \forall x \in \mathbb{R} \qquad (\hat{x} = 2 \text{ is minimizer})$
2) $F(x) > F(2) \quad \forall x \in \mathbb{R} \setminus \{2\} \qquad (\hat{x} = 2 \text{ is unique min.})$

1) $F(x) \geq F(2) \iff x^2 - 4x \geq -4 \iff (x-2)^2 \geq 0 \iff x = 2$
2) $F(x) > F(2) \iff (x-2)^2 > 0 \iff x \neq 2$ $\square$

EXAMPLE 2 DIRECT METHOD: proving existence of a minimizer by general results

Ex: $F: \mathbb{R} \to \mathbb{R}$ continuous and coercive, i.e.,

$$\lim_{|x| \to +\infty} F(x) = +\infty$$

Then $\exists$ minimizer by Weierstrass Theorem

Example 3 RELAXATION: This technique is relevant when a minimizer does not exist, e.g.,

$$(*) \quad \min\{(x^2-2)^2 \mid x \in \mathbb{Q}\}$$

Solution of $(*)$ would be $\hat{x} = \pm\sqrt{2}$ which is not in $\mathbb{Q}$. So in this case there is no minimum. But we are left with 2 questions

1) What is

$$\inf \{ (x^2-2)^2 \mid x \in \mathbb{Q} \} \quad ?$$

2) If $\{x_n\}$ is MINIMIZING SEQUENCE, i.e.

$$F(x_n) \to \inf \{ F(x) \mid x \in \mathbb{Q} \}, \quad F(x) = (x^2-2)^2$$

what can we say about accumulation points of $\{x_n\}$?

Answer: 1) As we guessed, min over $\mathbb{R}$ would be $x' = \pm\sqrt{2}$, so

$$\inf \{ (x^2-2)^2 \mid x \in \mathbb{Q} \} = F(\pm\sqrt{2}) = 0$$

2) $x_n \to \sqrt{2}$ OR $x_n \to -\sqrt{2}$ (up to subsequences)

Relaxation is useful to treat problems such as (*). To ensure that a minimizer exists one could, for example,

- Extend F over some set $\hat{X}$ with $\hat{X} \supseteq X$ ($\hat{X} = \mathbb{R}$ for (*))
- Change F so that a minimizer is more likely to exist

EXAMPLE 4 Γ-CONVERGENCE: We have a family of problems

$$\min \{ F_n(x) \mid x \in X \}, \quad F_n : X \to \mathbb{R}, \quad n \in \mathbb{N}$$

What happens as $n \to +\infty$? We hope to find $F_\infty : X \to \mathbb{R}$ such that

1) $\min \{ F_n \mid x \in X \} \to \min \{ F_\infty \mid x \in \mathbb{R} \}$ as $n \to +\infty$

2) If $x_n \in \operatorname{argmin} \{ F_n \mid x \in X \}$ then $x_n \to x_\infty$ with

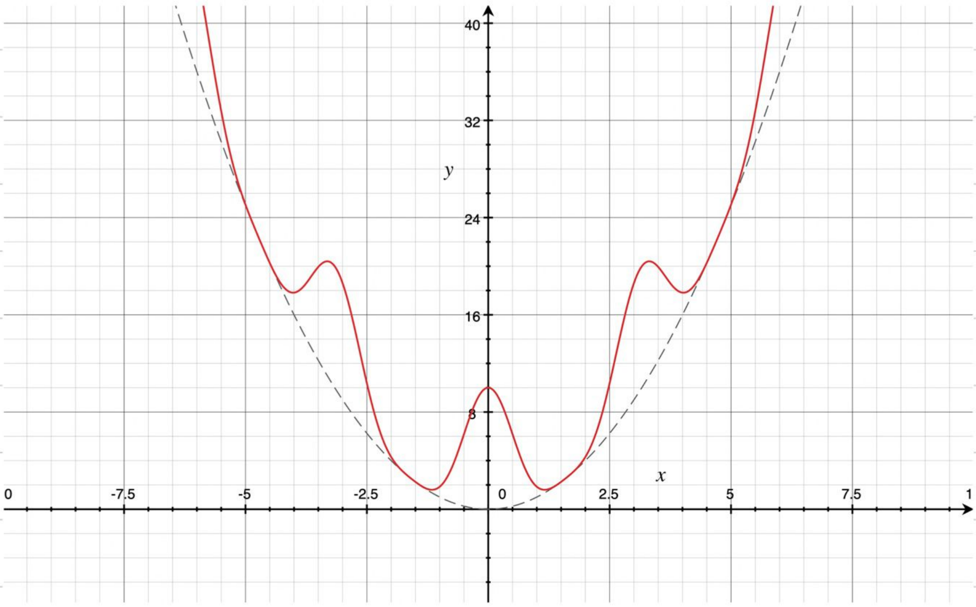
$$x_\infty \in \operatorname{argmin} \{ F_\infty \mid x \in X \}$$

F_∞ is the Γ-limit of $\{F_n\}$ as $n \to +\infty$

For example consider

$$m_n = \min\{F_n(x) \mid x \in \mathbb{R}\}, \quad F_n(x) = x^2 + n\cos^4 x$$

What is the limit of m_n?



Dashed $y = x^2$

Red F_n, $n = 10$

- F_n is sum of two positive terms
- x^2 small $\Leftrightarrow$ $x \simeq 0$
- $n\cos^4 x$ small $\Leftrightarrow$ $\cos x \simeq 0$

True when $x = \pm \frac{\pi}{2}$

Indeed one has

1) $m_n \to \left(\frac{\pi}{2}\right)^2$ as $n \to +\infty$

2) $\{x_n\}$ minimizing sequence converges (up to subsequencies) to $\pm\pi/2$.

3) The Γ-limit is

$$F_\infty(x) = \begin{cases} x^2 & \text{if } \cos x = 0 \\ +\infty & \text{otherwise} \end{cases}$$

INTEGRAL FUNCTIONALS

This course mainly focusses on integral functionals

X = some functions space, e.g.,

$$X = C^k[a,b] = \{ u : [a,b] \to \mathbb{R} \mid u \text{ } k\text{-times continuously differentiable} \}$$

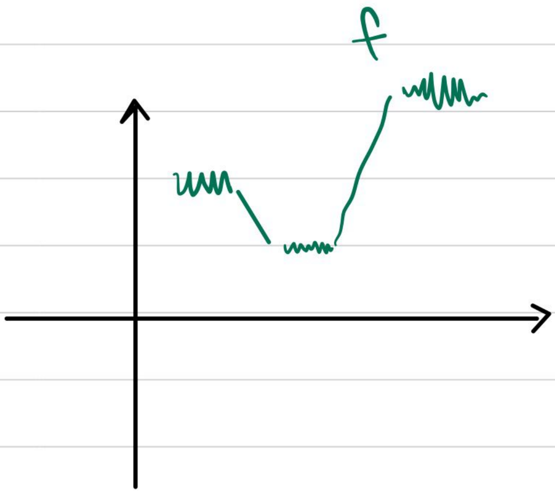
and $F : X \to \mathbb{R}$ is of the form

$$F(u) := \int_a^b L(x, u(x), u'(x), \dots, u^{(k)}(x)) \, dx$$

with $L : [a,b] \times \mathbb{R}^k \to \mathbb{R}$ **LAGRANGIAN**

Typically $L : [a,b] \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}$, $L = L(x, s, p)$

EXAMPLE 1 (DENOISING)



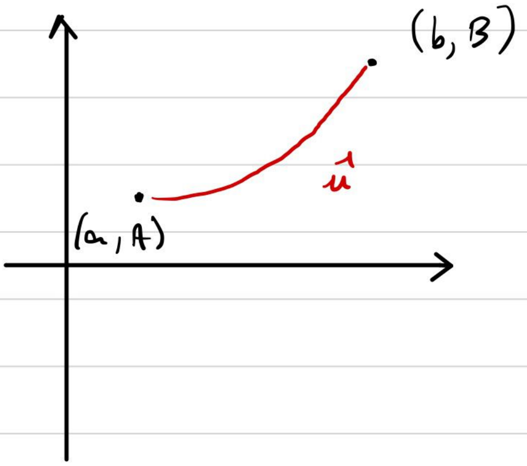
We receive a signal $f : [0,1] \to \mathbb{R}$ which we want to denoise

This task is achieved by solving

$$\hat{u} \in \operatorname{argmin} \left\{ \int_0^1 \dot{u}^2 + (u-f)^2 \, dx \;\middle|\; u \in C^1[0,1] \right\}$$

Note
- $\dot{u}$ penalizes oscillations
- $u - f$ penalizes discrepancy from the noisy signal f

Example 2 (Hanging Rope) Find the profile of a rope hanging at (a, A), (b, B)



The energy of a profile $u : [a,b] \to \mathbb{R}$ is modelled by

$$E(u) = \int_a^b \underbrace{\dot{u}^2}_{\text{elastic energy}} + \underbrace{u}_{\text{potential energy}} \, dx \quad , \quad u(a)=A, u(b)=B$$

Note 1) u can be negative which lowers E

2) Due to boundary conditions, if $u \ll 0$ then $\dot{u} \gg 0 \Rightarrow E$ higher

The solution

$$\hat{u} \in \operatorname{argmin} \{ E(u) \mid u \in C^1[a,b], u(a)=A, u(b)=B \}$$

will be a balance between (1) and (2)

PROBLEMS WE WILL NOT TALK ABOUT

1) GEOMETRIC FUNCTIONALS:

- DIDO'S PROBLEM: $\min \{ \text{Area}(\partial V) \mid V \subseteq \mathbb{R}^3, \text{Vol}(V) = 1 \}$

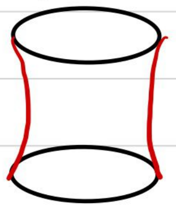
Intuitively the sol is a sphere. However proving it when not requiring regularity on V requires advanced tools (Geometric Measure Theory)

• PLATEAU'S PROBLEM: Given γ collection of curves in $\mathbb{R}^3$ find

$$\min \{ \text{Area}(S) \mid S \subseteq \mathbb{R}^3 \text{ surface}, \ \partial S = \gamma \}$$

Again this requires GMT

γ



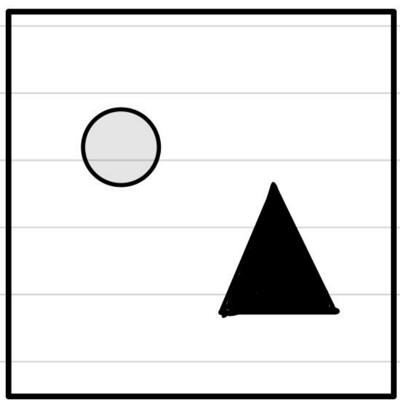
$\overset{1}{S}$ CATENOID is solution

2) IMAGING FUNCTIONALS: used for tasks such as DEnoising, segmentation, reconstruction of medical data. Usually

$$u : \Omega \to \mathbb{R} \quad , \ \Omega \subseteq \mathbb{R}^2, \mathbb{R}^3$$

u encodes gray-scale value of pixels of a picture in the frame Ω

Ω



Example: To segment the image at the left, i.e. find contours of shapes within it, one could minimize the MUMFORD-SHAH functional:

$$F(u, K) = \underbrace{\int_{\Omega \setminus K} |\nabla u|^2 dx}_{(I)} + \underbrace{\int_{\Omega} |u-f|^2 dx}_{(II)} + \underbrace{\text{Length}(K)}_{(III)}$$

• f is the original picture, generally noisy. Want to clean f and detect edges K within it

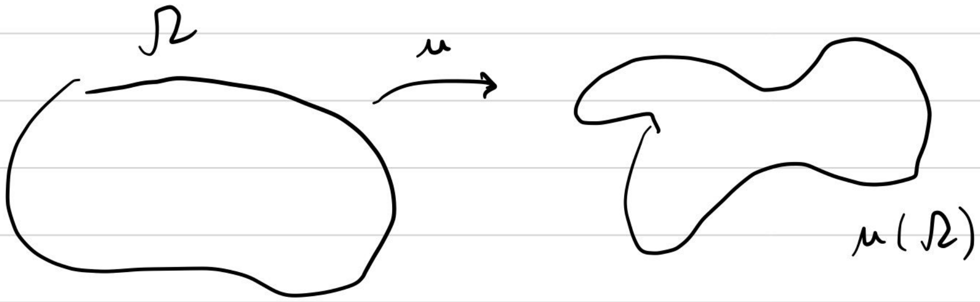
• (I): enforces smoothness of u outside of K (we don't want to pay energy for the natural transitions)

- (II) : Enforces the clean image u to be close to the original f
- (III) : Forces short contours

A solution is then

$$(\hat{u}, \hat{K}) \in \operatorname{argmin} \{ F(u, K) \mid K \subseteq \Omega \text{ compact}, \ u \in C^1(\Omega \setminus K) \}$$

3. VECTORIAL PROBLEMS : For example in materials science $\Omega \subseteq \mathbb{R}^3$ represents the reference configuration of an elastic body, $u : \Omega \to \mathbb{R}^3$ is a deformation



The elastic energy of the deformed configuration is

$$E(u) = \int_\Omega W(\nabla u)\, dx \quad , \quad W : \mathbb{R}^{3\times 3} \to [0, +\infty]$$

In this case the problem of minimizing E is vectorial, and the analysis requires advanced tools (quasi-convexity, etc)

An equilibrium configuration given boundary data $g : \partial\Omega \to \mathbb{R}^3$ is

$$\min \{ E(u) \mid u \in C^1(\Omega; \mathbb{R}^3), \ u = g \text{ on } \partial\Omega \}$$

BASIC FUNCTIONAL ANALYSIS (Revision)

REFERENCE: J. B. CONWAY
"A COURSE IN FUNCTIONAL ANALYSIS"
SECOND EDITION, SPRINGER, 1997

METRIC SPACE X set, $d: X \times X \to [0, +\infty)$. We say that d is a **METRIC** over X if

- $d(x,y) = 0 \iff x = y$
- $d(x,y) = d(y,x)$, $\forall x, y \in X$ (symmetric)
- $d(x,y) \leq d(x,z) + d(y,z)$, $\forall x,y,z \in X$ (Triangle inequality)

The pair (X, d) is called a **Metric Space**

CONVERGENCE For $\{x_n\} \subseteq X$ we say that $x_n \to x_0$ if

$\forall \varepsilon > 0, \exists N_\varepsilon \in \mathbb{N}$ s.t. $d(x_n, x_0) < \varepsilon$ if $n \geq N_\varepsilon$

CAUCHY SEQUENCE $\{x_n\} \subseteq X$ is a Cauchy sequence if

$\forall \varepsilon > 0, \exists N_\varepsilon \in \mathbb{N}$ s.t. $d(x_n, x_m) < \varepsilon$ if $n, m \geq N_\varepsilon$

COMPLETENESS A metric space (X, d) is complete if every Cauchy sequence $\{x_n\} \subseteq X$ converges to some $x_0 \in X$.

Topology generated by d (X, d) metric space. Define

$$\tau := \{ A \subseteq X \mid \forall x \in A, \exists \varepsilon > 0 \text{ s.t. } B_\varepsilon(x) \subseteq A \}$$

with $B_\varepsilon(x) := \{ y \in X \mid d(x,y) < \varepsilon \}$. Then τ is a **TOPOLOGY** over X. The sets in τ are called OPEN. $A \subseteq X$ is CLOSED if $A^c := X \setminus A$ is open.

NOTATION (X, τ) topological space, $A \subseteq X$. We denote by

- $\mathring{A}$ the **INTERIOR** of A: $\mathring{A} = \bigcup \{ O \mid O \subseteq A,\ O \text{ open} \}$
- $\bar{A}$ the **CLOSURE** of A: $\bar{A} = \bigcap \{ C \mid A \subseteq C,\ C \text{ closed} \}$

In other words:

- $\mathring{A}$ is the largest open set contained in A
- $\bar{A}$ is the smallest closed set which contains A

DENSITY (X,d) metric space. $D\subseteq X$ is DENSE in X if $\bar{D}=X$

SEPARABILITY (X,d) metric space is SEPARABLE if $\exists$ a COUNTABLE set $D\subseteq X$ which is dense, i.e., $\bar{D}=X$

LIMITS (X,d_X), (Y,d_Y) metric spaces, $U\subseteq X$ open, $F: U\to Y$, $x_0\in U$. We say that $F(x)\to L$ as $x\to x_0$, in symbols

$$\lim_{x\to x_0} F(x)=L,$$

if $\forall\,\varepsilon>0$, $\exists\,\delta>0$ s.t. $d_Y(F(x),L)<\varepsilon$ if $d_X(x,x_0)<\delta$

CONTINUITY (X,d_X), (Y,d_Y) metric spaces, $U\subseteq X$ open, $F: U\to Y$. We say that F is continuous at $x_0\in U$ if $F(x)\to F(x_0)$ as $x\to x_0$. F is continuous in U if it is continuous $\forall\, x_0\in U$.

NORMED SPACE X vector space over $\mathbb{R}$, $\|\cdot\|: X\to[0,+\infty)$. We say that $\|\cdot\|$ is a norm over X if

- $\|x\|=0 \iff x=0$
- $\|\lambda x\|=|\lambda|\,\|x\|$, $\forall\lambda\in\mathbb{R}$, $x\in X$ (1-homogeneous)
- $\|x+y\|\le\|x\|+\|y\|$, $\forall\, x,y\in X$ (subadditive)

The pair $(X,\|\cdot\|)$ is called normed space

REMARK $(X,\|\cdot\|)$ normed space. Then $d(x,y)=\|x-y\|$ is a metric over X. In particular X is a topological space with τ induced by d. By convention all the topological notions in X are given WRT τ.

BANACH SPACE $(X, \|\cdot\|)$ normed space is BANACH if (X, d) with $d(x,y) = \|x-y\|$ is complete.

LINEAR OPERATORS X, Y normed spaces, $T: X \to Y$. We say that

- T is LINEAR if $T(\lambda x + y) = \lambda Tx + Ty$, $\forall \lambda \in \mathbb{R}$, $x, y \in X$
- T is BOUNDED if

$$\sup_{\|x\|_X \leq 1} \|Tx\|_Y < +\infty$$

FACT Let $T: X \to Y$ be linear. Then

$$T \text{ is continuous} \iff T \text{ is bounded}$$

NOTATION $\mathcal{L}(X,Y) := \{ T: X \to Y \mid T \text{ linear bounded} \}$

$X^* := \mathcal{L}(X, \mathbb{R})$ DUAL SPACE of X

REMARK 1) $\mathcal{L}(X,Y)$ is a vector space over $\mathbb{R}$, with operations

$$(\alpha T_1 + T_2)(x) := \lambda T_1 x + T_2 x, \quad \forall \alpha \in \mathbb{R}, T_1, T_2 \in \mathcal{L}(X,Y)$$

2) $\mathcal{L}(X,Y)$ is a normed space with norm

$$\|T\|_{\mathcal{L}(X,Y)} := \sup_{\|x\|_X \leq 1} \|Tx\|_Y$$

3) If Y is Banach then $\mathcal{L}(X,Y)$ is Banach

4) X normed space $\Rightarrow$ X^* Banach space

CONVERGENCES $(X, \|\cdot\|)$ normed space $\{x_n\} \subseteq X$, $x_0 \in X$, $\{\varphi_n\} \subseteq X^*$ $\varphi_0 \in X^*$,

1) $x_n \to x_0$ **STRONGLY** if $\|x_n - x_0\|_X \to 0$ as $n \to +\infty$
2) $x_n \rightharpoonup x_0$ **WEAKLY** if $\varphi(x_n) \to \varphi(x_0)$, $\forall \varphi \in X^*$
3) $\varphi_n \overset{*}{\rightharpoonup} \varphi_0$ **WEAKLY*** if $\varphi_n(x) \to \varphi_0(x)$, $\forall x \in X$
4) $\varphi_n \to \varphi_0$ **STRONGLY** if $\|\varphi_n - \varphi_0\|_{X^*} \to 0$

NOTE

- $x_n \to x_0 \Rightarrow x_n \rightharpoonup x_0$
- The reverse implication is not true. For example let

$$X = \ell^p := \left\{ (x_1, x_2, \dots, x_n, \dots) \;\middle|\; \sum_{j=1}^{+\infty} |x_j|^p < +\infty \right\}$$

with $1 < p < +\infty$. Recall that $(X, \|\cdot\|)$ is a normed space with

$$\|x\| := \left(\sum_{j=1}^{+\infty} |x_j|^p \right)^{1/p}$$

Let $e_j := (0, \dots, 0, \underset{\substack{\uparrow \\ j\text{-th position}}}{1}, 0, \dots)$. Then $e_j \rightharpoonup 0$ but

$\|e_j\| = 1 \not\to 0$.

DEFINITION $(X, \|\cdot\|)$ normed space, $K \subseteq X$, $\tilde{K} \subseteq X^*$

1) K is **COMPACT** if $\forall \{x_n\} \subseteq K$, $\exists x_0 \in K$ s.t. $x_{n_k} \to x_0$ along some subsequence
2) K is **SEQUENTIALLY WEAKLY COMPACT** if $\forall \{x_n\} \subseteq K$, $\exists x_0 \in K$ s.t. $x_{n_k} \rightharpoonup x_0$ along some subsequence
3) $\tilde{K}$ is **SEQUENTIALLY WEAKLY* COMPACT** if $\forall \{\varphi_n\} \subseteq \tilde{K}$, $\exists \varphi_0 \in \tilde{K}$ s.t. $\varphi_{n_k} \overset{*}{\rightharpoonup} \varphi_0$ along some subsequence.

WARNING

If (X, τ) is a topological space then $K \subseteq X$ is compact if any OPEN COVER of K admits a FINITE SUBCOVER.

If the topology τ is metrizable (e.g. metric or normed spaces) then SEQUENTIAL COMPACTNESS is equivalent to COMPACTNESS.

However, if X is normed space, then the weak topology on X and weak* topology on X^* are NOT metrizable in general. Thus, in general WEAK (WEAK*) COMPACTNESS and WEAK (WEAK*) SEQUENTIAL COMPACTNESS are not equivalent. With additional assumptions, however, they are the same:

1) If X is Banach then WEAK SEQUENTIAL COMPACTNESS and WEAK COMPACTNESS are equivalent (EBERLEIN - SMULIAN THEOREM)

2) If X is a SEPARABLE BANACH space then WEAK* SEQUENTIAL COMPACTNESS and WEAK* COMPACTNESS are equivalent

THEOREM (BANACH - ALAOGLU) $(X, \|\cdot\|)$ normed space. Denote by $B := \{\varphi \in X^* \mid \|\varphi\|_{X^*} \leq 1\}$ the closed unit ball of X^*:

1) Then B is WEAKLY* COMPACT

2) If in addition X is BANACH and SEPARABLE then B is also SEQUENTIALLY WEAKLY* COMPACT

There is a corollary of Banach-Alaoglu concering the weak compactness of the unit ball of X. For that we need the following definition

REFLEXIVITY $(X, \|\cdot\|)$ normed space. Consider X^* and its dual WRT to the strong norm of X^*, i.e., $X^{**} := (X^*, \|\cdot\|_{X^*})^*$

Define the CANONICAL EMBEDDING

$$J : X \to X^{**} \text{ s.t. } J(x)(\varphi) := \varphi(x), \ x \in X, \varphi \in X^*$$

We have $\|J(x)\|_{X^{**}} = \|x\|_X$. We say that X is REFLEXIVE if J is surjective, i.e., if

$$X^{**} = \{ J(x), \ x \in X \}$$

COROLLARY (of BANACH-ALAOGLU)

$(X, \|\cdot\|)$ normed space. Define $B := \{ x \in X \mid \|x\| \leq 1 \}$.

1) If X is Reflexive then B is WEAKLY COMPACT
2) If X is reflexive and Banach then B is WEAKLY SEQUENTIALLY COMPACT

As a consequence of the PRINCIPLE OF UNIFORM BOUNDEDNESS (PUB) (see book of Conway), we have.

PROPOSITION $(X, \|\cdot\|)$ Banach space

1) If $\{x_n\} \subseteq X$ is s.t. $x_n \rightharpoonup x_0$ then $\sup_n \|x_n\| < +\infty$
2) If $\{\varphi_n\} \subseteq X^*$ is s.t. $\varphi_n \overset{*}{\rightharpoonup} \varphi_0$ then $\sup_n \|\varphi_n\|_{X^*} < +\infty$

Another important notion needed throughout the course is the one of lower semicontinuity.

DEFINITION (X, d) metric space, $F: X \to \mathbb{R}$. We say that F is LOWER SEMICONTINOUS at $x_0 \in X$ if

$$F(x_0) \leq \liminf_{n \to +\infty} F(x_n),$$

for all $\{x_n\} \subseteq X$ s.t. $x_n \to x_0$.

DEFINITION $(X, \|\cdot\|)$ normed space, $F: X \to \mathbb{R}$, $G: X^* \to \mathbb{R}$.

1) F is (SEQUENTIALLY) WEAKLY LOWER SEMICONTINOUS at $x_0 \in X$ if

$$F(x_0) \leq \liminf_{n \to +\infty} F(x_n)$$

for all $\{x_n\} \subseteq X$ s.t. $x_n \rightharpoonup x_0$.

2) G is (SEQUENTIALLY) WEAKLY* LOWER SEMICONTINOUS at $\varphi_0 \in X^*$ if

$$G(\varphi_0) \leq \liminf_{n \to +\infty} G(\varphi_n)$$

for all $\{\varphi_n\} \subseteq X^*$ s.t. $\varphi_n \overset{*}{\rightharpoonup} \varphi_0$.

PROPOSITION $(X, \|\cdot\|)$ normed space. Then

1) The norm $\|\cdot\|$ is WEAKLY SEQUENTIALLY LOWER SEMICONTINOUS, i.e.,

$$x_n \rightharpoonup x_0 \Rightarrow \|x_0\| \leq \liminf_{n \to +\infty} \|x_n\|$$

2) The norm $\|\cdot\|_{X^*}$ is WEAKLY* SEQUENTIALLY LOWER SEMICONTINOUS, i.e.,

$$\varphi_n \overset{*}{\rightharpoonup} \varphi_0 \Rightarrow \|\varphi_0\|_{X^*} \leq \liminf_{n \to +\infty} \|\varphi_n\|_{X^*}$$

LESSON 2 - 10 MARCH 2021

HILBERT SPACES

HILBERT $\subseteq$ BANACH $\subseteq$ COMPLETE METRIC $\subseteq$ TOPOLOGICAL

INNER PRODUCT SPACES Let H be a real vector space. A function $\langle \cdot, \cdot \rangle : H \times H \to \mathbb{R}$ is an INNER PRODUCT if

- $\langle x, y \rangle = \langle y, x \rangle$, $\forall x, y \in H$ (Symmetric)
- $\langle \lambda x + \mu y, z \rangle = \lambda \langle x, z \rangle + \mu \langle y, z \rangle$, $\forall \lambda, \mu \in \mathbb{R}$, $x, y \in H$ (Bilinear)
- $\langle x, x \rangle \geq 0$, $\forall x$ and $\langle x, x \rangle = 0$ iff $x = 0$ (Positive definite)

The pair $(H, \langle \cdot, \cdot \rangle)$ is called an INNER PRODUCT SPACE

REMARK $(H, \langle \cdot, \cdot \rangle)$ inner product space. Then $\|x\| = \sqrt{\langle x, x \rangle}$ defines a norm on X.

CAUCHY - SCHWARZ INEQUALITY H inner prod space. Then $|\langle x, y \rangle| \leq \|x\| \, \|y\|$

HILBERT SPACE $(H, \langle \cdot, \cdot \rangle)$ inner product space. We say that H is a HILBERT SPACE if $(H, \|\cdot\|)$ is COMPLETE, with $\|x\| = \sqrt{\langle x, x \rangle}$

BASIS H Hilbert. A set of elements $\{e_n\}_{n \in \mathbb{N}} \subseteq H$ is called BASIS if

- $\langle e_i, e_j \rangle = 0$ for all $i \neq j$, $\langle e_i, e_i \rangle = 1$ $\forall i$
- $\operatorname{span}\{e_j\}$ is dense in H

where for a set $A \subseteq H$ we define $\operatorname{span} A = \left\{ \sum_{j=1}^{n} \lambda_j x_j \;\middle|\; \lambda_j \in \mathbb{R}, x_j \in A, n \in \mathbb{N} \right\}$

THEOREM If H is Hilbert separable then $\exists \{e_n\} \subseteq H$ basis

NOTATION Given a basis $\{e_n\}$ and $x \in H$ we define $x_k := \langle x, e_k \rangle$ k-th coordinate of x WRT $\{e_n\}$

PROPOSITION H Hilbert with basis $\{e_n\}$. Then

1) $\|v\|^2 = \sum_{j=1}^{+\infty} v_n^2$, $v_n := \langle v, e_n \rangle$

2) $\langle v, w \rangle = \sum_{j=1}^{+\infty} v_n w_n$

For a separable Hilbert space there is a natural correspondence between H and ℓ^2. Thus we can think of H as $\mathbb{R}^\infty$. To make this statement precise we need the following definition

(Recall: $\ell^2 = \{ (x_1, x_2, \dots) \mid \sum_{j=1}^{+\infty} |x_j|^2 < +\infty \}$. This is normed by $\|x\| = \left(\sum_{j=1}^{+\infty} |x_j|^2\right)^{1/2}$ and is Hilbert with inner product $\langle x, y \rangle = \sum_{j=1}^{+\infty} x_j y_j$)

DEF X normed space, $\{x_n\} \subseteq X$. We say that

$$\sum_{n=1}^{+\infty} x_n = x_0$$

if $S_n \to x_0$, where $S_n := \sum_{j=1}^{n} x_j$ partial sums.

THEOREM H Hilbert with basis $\{e_n\}$. Let $\{v_n\} \subseteq \mathbb{R}$. Then

$$\sum_{n=1}^{+\infty} v_n e_n \text{ converges in } H \iff \sum_{n=1}^{+\infty} v_n^2 \text{ converges in } \mathbb{R}$$

In particular $H \cong \ell^2$, with isomorphism $v \in H \mapsto (v_1, v_2, \dots, v_n, \dots) \in \ell^2$

Another nice aspect of Hilbert spaces is that $H = H^*$ dual space.

THEOREM (RIESZ) H Hilbert. Define the map $\Phi : H \to H^*$

$$\Phi(x)(z) := \langle x, z \rangle, \ \forall z \in H$$

Then Φ is invertible and $\|\Phi(x)\|_{H^*} = \|x\|_H$.

Thus $H \cong H^*$ isomorphic.

In particular H can be identified with H^*. Thus weak* and weak topologies coincide, and we can characterize weak convergence by

$$x_n \rightharpoonup x_0 \iff \langle x_n, z\rangle \to \langle x_0, z\rangle, \ \forall z \in H.$$

FURTHER PROPERTIES OF WEAK CONVERGENCE IN HILBERT

PROP H Hilbert with basis $\{e_n\}$. If $x_n \rightharpoonup x_0$ then

$$(x_n)_k \to (x_0)_k, \quad \forall k \in \mathbb{N}$$

WARNING We know that $x_n \rightharpoonup x_0$ does not imply $x_n \to x_0$. However it is also NOT true that $\|x_n\| \to \|x_0\|$ (i.e. the norm is not weakly continuous).

However, the following proposition relating strong convergence to weak convergence holds.

PROP H Hilbert. Then

$$x_n \to x_0 \iff x_n \rightharpoonup x_0 \text{ and } \|x_n\| \to \|x_0\|$$

NOTE: This is not saying that $\|x_n - x_0\| \to 0$

Another useful proposition is that weak convergence can be tested against a dense subset

PROP H Hilbert. Assume that $\{x_n\} \subseteq H$ is bounded, i.e. $\sup_n \|x_n\| < +\infty$. Suppose that $W \subseteq H$ is s.t $\overline{\text{span } W} = H$ and

$$\langle x_n, w\rangle \to \langle x_0, w\rangle, \ \forall w \in W.$$

Then $x_n \rightharpoonup x_0$.

COROLLARY

H Hilbert with basis $\{e_n\}$. Let $\{x_n\}$ be BOUNDED. Then if

$$(x_n)_k \to (x_0)_k \quad , \forall k \in \mathbb{N}$$

we have $x_n \to x_0$.

[Proof: take $W = \{e_n\}$]

EXAMPLE

$$X = C[a,b] \quad \text{with} \quad \|u\|_\infty := \max_{x \in [a,b]} |u(x)|$$

$$Y = C^1[a,b] \quad \text{with} \quad \|u\|_1 := \|u\|_\infty + \|u'\|_\infty$$

Then $(X, \|\cdot\|_\infty)$ and $(Y, \|\cdot\|_1)$ are Banach spaces, but not Hilbert spaces.

Hint to show this: in an inner product space the parallelogram law holds:

$$\|x+y\|^2 + \|y-x\|^2 = 2(\|x\|^2 + \|y\|^2)$$

$\forall x, y \in H$

1. CALCULUS IN NORMED SPACES

Reference : S. KESAVAN
"NON LINEAR FUNCTIONAL ANALYSIS, A FIRST COURSE"
HINDUSTAN BOOK AGENCY, 2004

Throughout this section X, Y are real NORMED SPACES, $U \subseteq X$ is OPEN and $F: U \to Y$ is a given function.

GOAL: Construct a theory of differentiation for maps $F: U \to Y$

DEFINITION 1.1 We say that F is FRÉCHET DIFFERENTIABLE at $u_0 \in U$ if $\exists\, A_{u_0} \in \mathcal{L}(X,Y)$ s.t.

$$(*) \qquad \lim_{\|v\|_X \to 0} \frac{\| F(u_0+v) - F(u_0) - A_{u_0}(v) \|_Y}{\|v\|_X} = 0 .$$

REMARK If F is diff. at u_0 then $A_{u_0} \in \mathcal{L}(X,Y)$ satisfying $(*)$ is UNIQUE (Check it by exercise)

NOTATION If F is diff. at $u_0 \in U$ we call A_{u_0} the Fréchet derivative (or just derivative) of F at u_0. We denote

$$F'(u_0) := A_{u_0} \in \mathcal{L}(X,Y)$$

NOTE This generalizes diff. for maps $F: \mathbb{R}^n \to \mathbb{R}^m$. In this case the differential is $F'(u)(v) = DF(u)\, v$, with $DF(u) \in \mathbb{R}^{m \times n}$ matrix of partial derivatives of F, i.e.

$$[DF(u)]_{ij} = \frac{\partial F_i}{\partial x_j}(u), \qquad F = (F_1, \dots, F_m)$$

DEFINITION 1.2 Assume that F is diff. $\forall u_0 \in U$. Then we can define the map

$$F' : U \to \mathcal{L}(X,Y)$$
$$u \mapsto F'(u)$$

If F' is continuous we say that $F \in C^1(U,Y)$

WRT norm on X and operator norm on $\mathcal{L}(X,Y)$

In words we say F is continuously diff.

EXAMPLES The most common examples throughout the course will be real valued functions, i.e., $Y = \mathbb{R}$

1) X normed, $U \subseteq X$ open, $F : U \to \mathbb{R}$. If F is diff at $u_0 \in U$ then $F'(u_0) \in \mathcal{L}(X,\mathbb{R}) = X^*$
Then if F diff in U, the derivative defines an application $F' : U \to X^*$ $\quad (u \mapsto F'(u) \in X^*)$

2) H Hilbert, $U \subseteq H$ open, $F : U \to \mathbb{R}$. If F is diff. at $u_0 \in U$ then $F'(u_0) \in \mathcal{L}(H,\mathbb{R})$. By Riesz's Thm $\exists!$ $z_0 \in H$ s.t.

$$F'(u_0)(w) = \langle z_0, w \rangle \quad , \forall w \in H$$

We denote $z_0 := \operatorname{grad} F(u_0)$ (gradient of F at u_0).

PROPOSITION 1.3 Assume that F is diff at $u_0 \in U$. Then F is continuous at u_0.

Proof Introduce the notation $o(\|v\|_X)$ for a quantity such that

$\frac{o(\|v\|_X)}{\|v\|_X} \to 0$ as $\|v\|_X \to 0$. Since U is open and $u_0 \in U$ then $\exists \varepsilon > 0$ such that $B_\varepsilon(u_0) \subseteq U$. Let $v \in B_\varepsilon(0)$ so that $u_0 + v \in B_\varepsilon(u_0) \subseteq U$: then

$$
\begin{aligned}
\| F(u_0+v) - F(u_0)\|_Y &\le \| F(u_0+v) - F(u_0) - A_{u_0}(v)\|_Y + \| A_{u_0}(v)\|_Y \\
&\le o(\|v\|_X) + \|A_{u_0}\|_{\mathcal{L}(X,Y)} \|v\|_X \quad (\text{since } A_{u_0} \in \mathcal{L}(X,Y)) \\
&= o(\|v\|_X) \to 0 \text{ as } \|v\|_X \to 0
\end{aligned}
$$
□

THEOREM 1.4 (CHAIN RULE) X, Y, Z normed, $U \subseteq X$, $V \subseteq Y$ open, $F: U \to V$, $G: V \to Z$. Assume F is diff at $u_0 \in U$ and G is diff. at $v_0 := F(u_0) \in V$. Then $G \circ F : U \to Z$ is diff. at u_0 with

$$(G \circ F)'(u_0) = G'(v_0) \circ F'(u_0) \in \mathcal{L}(X, Z)$$

composition of linear continuous operators in $\mathcal{L}(X,Y)$ and $\mathcal{L}(Y,Z)$

The proof is very simple, and we thus omit it. If you are interested you can find it in the book of KESAVAN, PROPOSITION 1.1.1 page 7

DEFINITION 1.5 We say that F is GÂTEAUX DIFFERENTIABLE at $u_0 \in U$ in the DIRECTION $v \in X$ if

$$F'_g(u_0)(v) := \lim_{t \to 0} \frac{F(u_0 + tv) - F(u_0)}{t} \in Y,$$

Here $t \in \mathbb{R}$

i.e., if the above limit exists.

NOTE The Gâteaux derivative (G. derivative) generalizes the directional derivative for maps $F: \mathbb{R}^n \to \mathbb{R}$.
In this case we have

$$F'_G(u)(v) = \frac{\partial F}{\partial v}(u) = \nabla F(u_0) \cdot v$$

If F diff at u

Scalar product in $\mathbb{R}^n$

WARNING $F'_G(u): X \to Y$ is always linear. But in general we don't have that $F'_G(u) \in \mathcal{L}(X, Y)$, as it happens for Fréchet derivatives.

PROPOSITION 1.6 If F diff. at $u_0 \in U$ then F is G. diff at u_0 in every direction v and

$$F'_G(u_0)(v) = F'(u_0)(v)$$

Gâteaux derivative can be computed by Frechet derivative, and viceversa

Proof

$$\frac{F(u_0+tv) - F(u_0)}{t} =$$

$$= \underbrace{\frac{F(u_0+tv) - F(u_0) - F'(u_0)(tv)}{t\|v\|_X}}_{\to 0 \text{ as } t \to 0,} \|v\|_X + \frac{F'(u_0)(tv)}{t}$$

since F is diff. at u_0 (note this is converging in Y to 0)

$$= o(t) + F'(u_0)(v) \to F'(u_0)(v) \text{ as } t \to 0$$

Here we used that $F'(u_0)$ is lin. operator, so $F'(u_0)(tv) = tF'(u_0)(v)$

$\square$

WARNING The converse of PROP 1.6 does not hold, i.e.

$$\text{Gâteaux Diff.} \not\Rightarrow \text{Fréchet Diff.}$$

For example take $F : \mathbb{R}^2 \to \mathbb{R}$

$$F(x,y) := \begin{cases} \dfrac{x^5}{(y-x^2)^2 + x^4} & \text{if } (x,y) \neq (0,0) \\ 0 & \text{if } (x,y) = (0,0) \end{cases}$$

It is easy to check that $F_G'(0)(v) = 0$, $\forall v \in \mathbb{R}$. So F is G-diff at $0 = (0,0)$ in every direction. Thus, if F was Fréchet diff we would have (by PROP 1.6) that $F'(0)(v) = 0$. Thus by def of derivative

$$(*) \qquad \lim_{\|v\|_{\mathbb{R}^2} \to 0} \frac{|F(v) - F(0)|}{\|v\|_{\mathbb{R}^2}} = 0$$

However if in the above limit we consider $v \in \{y = x^2\}$ we obtain

$$\lim_{\substack{\|v\|_{\mathbb{R}^2} \to 0 \\ v \in \{y = x^2\}}} \frac{|F(v) - F(0)|}{\|v\|_{\mathbb{R}^2}} = 1,$$

which contradicts $(*)$. Thus F is not Fréchet diff at 0.

REMARK PROPOSITION 1.6 is very useful to guess the Fréchet derivative of a function $F : U \to Y$, as the Gâteaux derivative can be computed via a formula

EXAMPLE $X = C[0,1]$, with norm $\|\cdot\|_\infty$. Define $F: X \to \mathbb{R}$ by

$$F(u) = \int_0^1 \sin(u(x))\, dx$$

What could be the derivative of F? Let us compute the Gâteaux derivative at u in the direction v:

$$\lim_{t \to 0} \frac{F(u+tv) - F(u)}{t} = \lim_{t\to 0} \int_0^1 \frac{\sin(u+tv) - \sin(u)}{t}\, dx$$

From the chain Rule (THEOREM 1.4) we have:

$$\lim_{t\to 0} \frac{\sin(u(x) + tv(x)) - \sin(u(x))}{t} = \cos(u(x))\, v(x), \quad \text{uniformly in } x \in [0,1]$$

Thus we can pass the limit under the integral and obtain

$$F_g'(u)(v) = \int_0^1 \cos(u(x))\, v(x)\, dx. \qquad (F_g'(u) \text{ is linear!})$$

If F is Fréchet diff. then by PROPOSITION 1.6 we must have $F'(u) = F_g'(u)$. So we guess that $F_g'(u)$ is the Fréchet derivative of F at u. Indeed notice that $F_g'(u) \in \mathcal{L}(X, \mathbb{R})$, since

$$|F_g'(u)(v)| \leq \|v\|_\infty \int_0^1 |\cos(u(x))|\, dx \leq \|v\|_\infty \Rightarrow \sup_{\|v\|_X \leq 1} |F_g(u)(v)| < +\infty.$$

Moreover it is easy to see that

$$\lim_{\|v\|_X \to 0} \frac{|F(u+v) - F(u) - F_g'(u)(v)|}{\|v\|_X} = 0$$

showing that F is Fréchet diff. at u with $F'(u)(v) = \int_0^1 \cos(u(x))\, v(x)\, dx$.

QUESTION Assume that $F: U \to Y$ is Gâteaux differentiable. Under which assumptions on F are we guaranteed Fréchet differentiability?

To answer the above question, we need the following theorem.

THEOREM 1.7 (MEAN VALUE)

Suppose F is G. diff. in U in every direction. Let $x_1, x_2 \in U$ be such that the segment

$$[x_1, x_2] := \{ x_1 + t(x_2 - x_1),\ t \in [0,1] \} \subseteq U$$

Assume also that $F'_g(u) : X \to Y$ is s.t. $F'_g(u) \in \mathcal{L}(X,Y)$, $\forall u \in U$.
Then

$$\| F(x_2) - F(x_1) \|_Y \leq \sup_{u \in [x_1, x_2]} \| F'_g(u) \|_{\mathcal{L}(X,Y)} \| x_2 - x_1 \|_X$$

<u>Proof</u> If $F(x_1) = F(x_2)$ then the thesis is trivial. Thus assume that $F(x_1) \neq F(x_2)$. We now employ the following:

<u>FACT</u> (COROLLARY OF HAHN-BANACH) Y normed space, $z \in Y$, $z \neq 0$. Then $\exists\, \Lambda \in Y^*$ s.t.

$$\| \Lambda \|_{Y^*} = 1 \quad \text{and} \quad \Lambda(z) = \| z \|_Y .$$

Thus let $\Lambda \in Y^*$ be such that $\Lambda(F(x_2) - F(x_1)) = \| F(x_2) - F(x_1) \|_Y$ and $\| \Lambda \|_{Y^*} = 1$.
Define the segment function $\alpha : [0,1] \to U$ by $\alpha(t) := x_1 + t(x_2 - x_1)$.
Consider the map $H : [0,1] \to \mathbb{R}$ defined by

$$H := \Lambda \circ F \circ \alpha .$$

The thesis is obtained by applying the classical Mean Value Theorem to H. Thus, all we need to show is that H is differentiable.

<u>WARNING:</u> It would be tempting to apply the Chain Rule of Theorem 1.4 to H. However F is only Gâteaux differentiable, and in general the Chain Rule does not apply in this case.

We will check by hand that H is differentiable. Thus let $t \in [0,1]$, and $\tau \neq 0$ be such that $(t+\tau) \in [0,1]$. We have

$$(*) \quad \frac{H(t+\tau) - H(t)}{\tau} = \Lambda\left[\frac{F(\alpha(t+\tau)) - F(\alpha(t))}{\tau}\right] \qquad \text{(being } \Lambda \text{ linear)}$$

Note that $F(\alpha(t+\tau)) = F(\alpha(t) + \tau(x_2 - x_1))$. Since F is gâteaux diff. at $\alpha(t)$ we then get

$$\frac{F(\alpha(t+\tau)) - F(\alpha(t))}{\tau} = \frac{F(\alpha(t) + \tau(x_2-x_1)) - F(\alpha(t))}{\tau} \longrightarrow F_g'(\alpha(t))(x_2-x_1)$$

as $\tau \to 0$. Note that by definition the above convergence is WRT the norm of Y. As $\Lambda \in Y^*$ is continuous, by taking the limit as $\tau \to 0$ in $(*)$ we get

$$H'(t) = \lim_{\tau \to 0} \frac{H(t+\tau) - H(t)}{\tau} = \Lambda\left[F_g'(\alpha(t))(x_2 - x_1)\right].$$

In particular H is diff. in $[0,1]$. Therefore we can apply the Mean Value Theorem to find $\xi \in (0,1)$ such that

$$(**) \qquad H(1) - H(0) = H'(\xi) \qquad \left(= H'(\xi)(1-0)\right)$$

Now

$$\begin{aligned} H(1) - H(0) &= \Lambda[F(\alpha(1))] - \Lambda[F(\alpha(0))] \\ &= \Lambda[F(x_2)] - \Lambda[F(x_1)] \\ &= \Lambda[F(x_2) - F(x_1)] \\ &= \| F(x_2) - F(x_1) \|_Y \end{aligned}$$

by the properties of Λ.

On the other hand, as we computed, we have

$$H'(\xi) = \Lambda\left[F_g'(\alpha(\xi))\,(x_2 - x_1) \right]$$

and so

$$|H'(\xi)| \leq \|\Lambda\|_{Y^*} \, \|F_g'(\alpha(\xi))\|_{\mathcal{L}(X,Y)} \, \|x_2 - x_1\|_X$$

(Using that Λ and $F_g'(\alpha(\xi))$ are linear and bounded)

$$\leq \sup_{u \in [x_1, x_2]} \|F_g'(u)\|_{\mathcal{L}(X,Y)} \, \|x_2 - x_1\|_X$$

(Using that $\|\Lambda\|_{Y^*} = 1$ and that $\alpha(\xi) \in [x_1, x_2]$)

From ** we then get

$$\|F(x_2) - F(x_1)\|_Y = H(1) - H(0) = H'(\xi)$$

$$\leq \sup_{u \in [x_1, x_2]} \|F_g'(u)\|_{\mathcal{L}(X,Y)} \, \|x_2 - x_1\|_X \qquad \square$$

COROLLARY 1.8 (of MEAN VALUE) Make the same assumptions of Theorem 1.7. Then

$$\|F(x_2) - F(x_1) - F_g'(x_1)(x_2 - x_1)\|_Y \leq \sup_{u \in [x_1, x_2]} \|F_g'(u) - F_g'(x_1)\|_{\mathcal{L}(X,Y)} \, \|x_2 - x_1\|_X$$

Proof Define $H : U \to Y$ by $H(u) := F(u) - F_g'(x_1)(u)$. Note that $F_g'(x_1) \in \mathcal{L}(X,Y)$ by assumption. In particular $F_g'(x_1) : X \to Y$ is Fréchet differentiable with derivative constantly equal to itself.

(Check it by exercise: X, Y normed spaces, $T \in \mathcal{L}(X,Y)$. Then T is Fréchet diff with $T'(u) = T$, $\forall u \in X$).

Therefore H is gâteaux diff. with $H_g'(u) = F_g'(u) - F_g'(x_1)$. Thus $H_g'(u) \in \mathcal{L}(X,Y)$ for all $u \in U$. Thus H satisfies assumptions of THEOREM 1.7. Applying THM 1.7 to H we conclude. $\square$

We are finally ready to answer our question:

"When does Gâteaux diff. imply Fréchet diff.?"

THEOREM 1.9 Assume that $F: U \to Y$ is Gâteaux diff. at every point of U and in every direction. Also suppose that $F'_g(u) \in \mathcal{L}(X,Y)$ for all $u \in U$. Define the map

$$F'_g : U \longrightarrow \mathcal{L}(X,Y)$$
$$u \longmapsto F'_g(u)$$

If F'_g is continuous at u_0 then F is Fréchet diff at u_0 and

continuity is WRT norm on X and operator norm on $\mathcal{L}(X,Y)$

$$F'(u_0)(v) = F'_g(u_0)(v), \quad \forall v \in X$$

Proof. Apply COROLLARY 1.8 to points $x_1 := u_0$, $x_2 := u_0 + v$. Since U is open, for v s.t. $\|v\|_X$ is sufficiently small we have $[x_1, x_2] \subseteq U$. By COROLLARY 1.8 we have

$$(*) \quad \| F(u_0+v) - F(u_0) - F'_g(u_0)(v) \|_Y \le \sup_{u \in [x_1,x_2]} \| F'_g(u) - F'_g(u_0) \|_{\mathcal{L}(X,Y)} \|v\|_X$$

Recall that $[x_1, x_2] = [u_0, u_0+v]$. As F'_g is continuous at u_0 we have

$$\sup_{u \in [x_1,x_2]} \| F'_g(u) - F'_g(u_0) \|_{\mathcal{L}(X,Y)} \longrightarrow 0 \quad \text{as } \|v\|_X \to 0$$

(in practice this implies $[x_1, x_2] \to \{u_0\}$)

Therefore, dividing $(*)$ by $\|v\|_X$ and taking the limit as $\|v\|_X \to 0$ concludes.

$\square$

LESSON 3 - 17 MARCH 2021

HIGHER ORDER DERIVATIVES

Let X, Y be normed spaces, $U \subseteq X$ open, $F: U \to Y$. Suppose that F is Fréchet diff. in U. Then we can define the map

$$F' : U \longrightarrow \mathcal{L}(X, Y)$$
$$u \longmapsto F'(u)$$

Now we can try and differentiate the above expression. This would yield a second derivative.

DEFINITION 1.10 Assume that F is diff. in U. We say that F is twice FRÉCHET diff. at $u_0 \in U$ if $F' : U \to \mathcal{L}(X, Y)$ is FRÉCHET diff. at u_0, i.e., if $\exists\, A_{u_0} \in \mathcal{L}(X, \mathcal{L}(X, Y))$ s.t.

$$\lim_{\|v\|_X \to 0} \frac{\| F'(u_0 + v) - F'(u_0) - A_{u_0}(v) \|_{\mathcal{L}(X,Y)}}{\|v\|_X} = 0$$

NOTATION The second Fréchet derivative is unique, if it exists. We denote it by $F''(u_0) := A_{u_0} \in \mathcal{L}(X, \mathcal{L}(X, Y))$

REMARK 1.11 Introduce the set

$$\mathcal{L}_2(X, Y) := \{ T : X \times X \to Y \mid T \text{ is bilinear, continuous} \}$$

where by continuous we mean $\exists\, M > 0$ s.t.

$$\| T(u, v) \|_Y \leq M \|u\|_X \|v\|_X, \quad \forall u, v \in X$$

Then one can show that

$$\mathcal{L}(X, \mathcal{L}(X, Y)) \cong \mathcal{L}_2(X, Y)$$

topologically and as sets (simple exercise)

Therefore $\mathcal{L}_2(X,Y)$ is naturally a normed space (and Banach if Y is Banach) with the norm

$$\|T\|_{\mathcal{L}_2(X,Y)} := \sup_{\|u\|_X, \|v\|_X \leq 1} \|T(u,v)\|_Y$$

DEFINITION 1.12 Let $F: U \to Y$ and assume that F is twice Fréchet diff. at each point of U. Thus we can define

$$F'': U \to \mathcal{L}_2(X,Y)$$
$$u \mapsto F''(u)$$

If F'' is continuous we say that $F \in C^2(U,Y)$

WRT norm on X and operator norm on $\mathcal{L}_2(X,Y)$

In words we say that F is twice continuously diff. in U

THEOREM 1.13

Assume $F: U \to Y$ is twice Fréchet diff at $u_0 \in U$. In particular we have $F''(u_0) \in \mathcal{L}_2(X,Y)$ bilinear and continuous. Then $F''(u_0)$ is also SYMMETRIC, i.e.,

$$F''(u_0)(v,w) = F''(u_0)(w,v) \quad , \forall\, v,w \in X$$

COMMENT ON THE PROOF The proof of THM 1.13 is quite long, and I decided to skip it. The interested reader can find it in the book of HENRI CARTAN – "CALCUL DIFFERENTIEL", 1967 (IN FRENCH) in Theorem 5.1.1 at page 65.

The main ideas of the proof are the following. Introduce the map

$$A(v,w) := \left[F(u_0 + v + w) - F(u_0 + v) - F(u_0 + w) + F(u_0) \right] \in Y$$

for $v, w \in X$ with $\|v\|_X, \|w\|_X$ sufficiently small, so that $u_0 + v + w \in U$ and A is well defined.

Notice that A is symmetric. One can show that

$$(*) \quad \| A(v,w) - F''(u_0)(v,w) \|_Y = o\left(\left(\|v\|_X + \|w\|_X \right)^2 \right)$$

(The above is not difficult to show, but it would require further analysis which is outside the scope of this course)

As A is symmetric we can swap v and w in $(*)$ to obtain

$$\| A(v,w) - F''(u_0)(w,v) \|_Y = o\left(\left(\|v\|_X + \|w\|_X \right)^2 \right)$$

Therefore by triangle ineq. and the above estimates we get

$$\| F''(u_0)(v,w) - F''(u_0)(w,v) \|_Y = o\left(\left(\|v\|_X + \|w\|_X \right)^2 \right)$$

Thus $\forall \varepsilon > 0$, $\exists \delta > 0$ s.t.

$$\| F''(u_0)(v,w) - F''(u_0)(w,v) \|_Y \leq \varepsilon \left(\|v\|_X + \|w\|_X \right)^2, \quad \text{if } \left(\|v\|_X + \|w\|_X \right) < \delta$$

If $v, w \in X$ are arbitrary, we can find $\lambda \neq 0$ s.t. $\|\lambda v\|_X + \|\lambda w\|_X < \delta$. Applying the above ineq. to λv, λw yields

$$\lambda^2 \| F''(u_0)(v,w) - F''(u_0)(w,v) \|_Y \leq \lambda^2 \varepsilon \left(\|v\|_X + \|w\|_X \right)^2, \quad \forall v, w \in X$$

As $\lambda \neq 0$ and ε is arbitrary, we conclude. $\square$

NOTE THEOREM 1.13 is a generalization of the classical SCHWARZ THEOREM on second derivatives of maps $F : \mathbb{R}^2 \to \mathbb{R}$. Indeed, if F is $C^2(\mathbb{R}^2)$ then

$$F'(u) = \nabla F(u) \in \mathcal{L}(\mathbb{R}^2, \mathbb{R})$$

where the application is given by

$$F'(u)(v) = \nabla F(u) \cdot v, \quad \nabla F(u) = \left(\frac{\partial F}{\partial x_1}(u), \frac{\partial F}{\partial x_2}(u) \right)$$

Thus $F' : \mathbb{R}^2 \to \mathcal{L}(\mathbb{R}^2, \mathbb{R})$. Then $F''(u) \in \mathcal{L}(\mathbb{R}^2, \mathcal{L}(\mathbb{R}^2, \mathbb{R})) = \mathcal{L}_2(\mathbb{R}^2, \mathbb{R})$ with application given by

$$F''(u)(v,w) = v^T \nabla^2 F(u)\, w, \quad \nabla^2 F(u) = \begin{pmatrix} \frac{\partial^2 F}{\partial x_1^2} & \frac{\partial^2 F}{\partial x_1 \partial x_2} \\ \frac{\partial^2 F}{\partial x_2 \partial x_1} & \frac{\partial^2 F}{\partial x_2^2} \end{pmatrix}$$

Therefore $F''(u)$ is symmetric $\iff \dfrac{\partial^2 F}{\partial x_1 \partial x_2} = \dfrac{\partial^2 F}{\partial x_2 \partial x_1}$, which is true by the classical Schwarz Theorem.

MORE DERIVATIVES! If $F: U \to Y$ is twice diff. in U then we can define

$$F'': U \longrightarrow \mathcal{L}(X, \mathcal{L}(X,Y))$$
$$u \longmapsto F''(u)$$

The function F'' can be in turn differentiated again, with

$$F'''(u) \in \mathcal{L}(X, \mathcal{L}(X, \mathcal{L}(X,Y)))$$

This procedure can be of course iterated, as F''' defines

$$F''': U \longrightarrow \mathcal{L}(X, \mathcal{L}(X, \mathcal{L}(X,Y)))$$

In general we have that

$$\underbrace{\mathcal{L}(X, \mathcal{L}(X, \mathcal{L}(X, \ldots\ldots \mathcal{L}(X,Y))))}_{n \text{ times}} \cong \mathcal{L}_n(X,Y)$$

where

$$\mathcal{L}_n(X,Y) := \{ T : \overbrace{X \times \cdots \times X}^{n \text{ times}} \to Y \mid T \text{ } n\text{-linear, bounded} \}$$

meaning that $\exists M > 0$ s.t. $\| T(u_1, \ldots, u_n) \|_Y \leq M \|u_1\|_X \cdots \|u_n\|_X$.
The space $\mathcal{L}_n(X,Y)$ is normed by

$$\|T\|_{\mathcal{L}_n(X,Y)} := \sup_{\|u_i\|_X \leq 1} \| T(u_1, \ldots, u_n) \|_Y$$

and $\mathcal{L}_n(X,Y)$ is Banach if Y is Banach. In particular the n-th Fréchet derivative is s.t.

$$F^{(n)}(u) \in \mathcal{L}_n(X,Y)$$

THEOREM 1.14 (TAYLOR FORMULA) (For a proof see book by CARTAN page 75)

X, Y Banach, $U \subseteq X$ open, $F: U \to Y$ $(n-1)$-times diff in U and n-times diff at $u_0 \in U$. Then

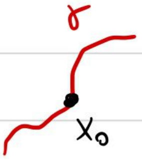
$$F(u_0 + v) = F(u_0) + F'(u_0)(v) + \frac{1}{2} F''(u_0)(v,v) + \cdots + \frac{1}{n!} F^{(n)}(u_0)\underbrace{(v, \ldots, v)}_{n\text{-times}} + o(\|v\|_X^n)$$

where $\dfrac{o(\|v\|_X^n)}{\|v\|_X^n} \to 0$ as $\|v\|_X \to 0$.

2. FIRST VARIATION (INDIRECT METHOD)

IDEA The FIRST VARIATION represents the DERIVATIVE of a functional $F : X \to \mathbb{R}$ with X = set

DEFINITION 2.1 Let $x_0 \in X$. A VARIATION at x_0 is a curve $\gamma : [-\delta, \delta] \to X$ s.t. $\gamma(0) = x_0$ (for some $\delta > 0$).



Note this is a scalar function of real variable

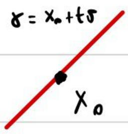
Consider the composition $\psi : [-\delta, \delta] \to \mathbb{R}$, $\psi(t) := F(\gamma(t))$.
If ψ is diff. at $t = 0$ we denote

$$\boxed{\delta F(x_0, \gamma) := \psi'(0)}$$

and call $\delta F(\gamma, x_0)$ the FIRST VARIATION of F at x_0 along γ.

NOTE We are not assuming any regularity on F or structure on X

EXAMPLE 2.2 1) $X = \mathbb{R}^d$, $x_0 \in \mathbb{R}^d$ fixed and $F : \mathbb{R}^d \to \mathbb{R}$ differentiable.
For $v \in \mathbb{R}^d$ consider the variation $\gamma(t) = x_0 + tv$, s.t. $\gamma(0) = x_0$.



Then

$$\delta F(x_0, \gamma) := \frac{d}{dt} F(x_0 + tv) \Big|_{t=0} = \nabla F(x_0) \cdot v$$

is the directional derivative of F at x_0 in direction v

2) X normed space, $x_0 \in X$, $F : X \to \mathbb{R}$ Gâteaux diff. at x_0 in every direction. Let $v \in X$ and consider the variation $\gamma(t) := x_0 + tv$.
Then

$$\delta F(x_0, \gamma) := (F \circ \gamma)'(0) = \lim_{t \to 0} \frac{F(\gamma(t)) - F(\gamma(0))}{t} =$$

$$= \lim_{t \to 0} \frac{F(x_0 + tv) - F(x_0)}{t} = F'_g(x_0)(v),$$

The first variation is just the Gâteaux derivative of F at x_0 in direction v.

PROPOSITION 2.3 $X =$ Set, $F : X \to \mathbb{R}$. Assume that $x_0 \in X$ is a minimizer for F, that is,

$$F(x) \geq F(x_0) \quad , \forall x \in X$$

Let $\gamma : [-\delta, \delta] \to X$ s.t. $\gamma(0) = x_0$. If $\psi = F \circ \gamma$ is differentiable at $t=0$ then

$$\boxed{\delta F(x_0, \gamma) = 0 .}$$

Proof By assumption ψ is diff at $t=0$. Therefore

$$\psi'(0) = \lim_{t \to 0} \frac{\psi(t) - \psi(0)}{t} = \lim_{t \to 0^+} \frac{\psi(\gamma(t)) - \psi(\gamma(0))}{t} = \lim_{t \to 0^+} \frac{F(\gamma(t)) - F(x_0)}{t} \geq 0$$

so that $\delta F(x_0, \gamma) \geq 0$. Similarly:

$$\psi'(0) = \lim_{t \to 0^-} \frac{\psi(t) - \psi(0)}{t} = \lim_{t \to 0^-} \frac{F(\gamma(t)) - F(x_0)}{t} \leq 0$$

So $\delta F(x_0, \gamma) \leq 0 \quad \Rightarrow \quad \delta F(x_0, \gamma) = 0 .$ $\square$

THE CASE OF AFFINE SPACES

REMINDER Let V real vector space. A set X is called an AFFINE SPACE with REFERENCE SPACE V if it is defined the addition operation

$$+ : X \times V \to X$$

with properties:

1) $x + 0 = x$, $\forall x \in X$, where $0 \in V$ is the zero in V
2) $x + (v + w) = (x + v) + w$, $\forall x \in X$, $\forall v, w \in V$
3) For every $x \in X$ the map $V \to X$, $v \mapsto x + v$ is a bijection

From the above reminder, the only important part for us is that:
If X is affine space with ref vector space V, then

$$\boxed{x \in X,\ v \in V \Rightarrow x + v \in X}$$

DEFINITION 2.4 X affine space with reference V. Let $x_0 \in X$. A VARIATION at x_0 in direction $v \in V$ is a curve $\gamma(t) := x_0 + tv \in X$

(Thus we restrict ourselves to the case of straight line variations)

REMARK 2.5 Notice that $\gamma : \mathbb{R} \to X$ by def. of affine space (i.e. γ is defined for all $t \in \mathbb{R}$). If $\psi := F \circ \gamma$ is diff at $t=0$ then

$$\psi'(0) = \lim_{t \to 0} \frac{\psi(t) - \psi(0)}{t} = \lim_{t \to 0} \frac{F(x_0 + tv) - F(x_0)}{t} .$$

Therefore the FIRST VARIATION reads

$$\boxed{\delta F(x_0, v) := \delta F(x_0, \gamma) = \lim_{t \to 0} \frac{F(x_0 + tv) - F(x_0)}{t}}$$

DEFINITION 2.6 X affine space over V, $F : X \to \mathbb{R}$. If $\delta F(v, x_0)$ exists then

$$\boxed{\delta F(x_0, v) = 0}$$

is called EULER–LAGRANGE EQUATION (ELE). If $x_0 \in X$ is such that (ELE) holds for all $v \in V$ then x_0 is called a CRITICAL POINT OF F (or STATIONARY POINT)

REMARK 2.7 X affine space over V, $F : X \to \mathbb{R}$ s.t. $\delta F(x_0, v)$ exists $\forall v \in V$. If x_0 minimizes F then x_0 is a CRITICAL POINT, i.e.

$$\delta F(x_0, v) = 0 , \quad \forall v \in V$$

Proof Apply PROPOSITION 2.3 to $\gamma(t) := x_0 + tv$. $\square$

THREE EXAMPLES

Let $a<b$, and $A<B$. Consider the set

$$X = \{ u \in C^1[a,b] \mid u(a)=A, \ u(b)=B \}$$

Then X is an affine space with reference vector space

$$V = \{ u \in C^1[a,b] \mid u(a)=0, \ u(b)=0 \}$$

We consider functionals in integral form:

$$u \in X \longmapsto \int_a^b L(x, u(x), u'(x))\, dx$$

for LAGRANGIANS $L : [a,b] \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}$. Specifically, consider $F, G, H : X \to \mathbb{R}$ defined by

$$F(u) := \int_a^b |\dot{u}(x)|^2 dx, \quad G(u) := \int_a^b |u(x)|\, dx, \quad H(u) = \int_a^b |\dot{u}(x)|^{1/2} dx$$

GOAL: We want to solve

$$\min_{u \in X} F(u), \quad \min_{u \in X} G(u), \quad \min_{u \in X} H(u)$$

and possibly characterize the solutions (if they exist!)

MINIMIZATION FOR F

By REMARK 2.7 we know that minimizers solve (ELE). Therefore, let us compute the first variation of F.

To do that, we compute the Fréchet derivative of F (This is actually not needed. The Gâteaux derivative of F would be sufficient, as seen in EXAMPLE 2.2 point (2). However we compute the Fréchet derivative as an exercise.)

PROPOSITION 2.8 Set $\tilde{X} := C^1[a,b]$ with norm $\|u\|_{C^1} := \|u\|_\infty + \|\dot{u}\|_\infty$. Extend F by

$$F(u) = \int_a^b |\dot{u}(x)|^2 \, dx , \qquad u \in \tilde{X}$$

Then F is Fréchet differentiable in $\tilde{X}$ with $F'(u) \in \mathcal{L}(\tilde{X}, \mathbb{R})$ given by

$$F'(u)(v) = 2 \int_a^b \dot{u}(x) \, \dot{v}(x) \, dx , \qquad \forall v \in \tilde{X}$$

Proof Let us start by computing the Gâteaux derivative of F at $u \in \tilde{X}$ in direction $v \in \tilde{X}$. We have

$$F(u+tv) = \int_a^b |\dot{u} + t\dot{v}|^2 dx = \int_a^b |\dot{u}|^2 dx + 2t \int_a^b \dot{u}\dot{v} \, dx + t^2 \int_a^b |\dot{v}|^2 dx$$

Therefore

$$F_g'(u)(v) = \lim_{t \to 0} \frac{F(u+tv) - F(u)}{t} =$$

$$= 2 \int_a^b \dot{u}\dot{v} \, dx + \lim_{t\to 0} t \int_a^b |\dot{v}|^2 dx$$

$$= 2 \int_a^b \dot{u}\dot{v} \, dx$$

Now notice that $F_g'(u) \in \mathcal{L}(\tilde{X}, \mathbb{R})$: In fact

$$|F_g'(u)(v)| \le 2 \int_a^b |\dot{u}| \, |\dot{v}| \, dx \le 2 \|v\|_{C^1} \int_a^b |\dot{u}| \, dx$$

so that

$$\sup_{\|v\|_{C^1} \le 1} |F_g'(u)(v)| \le 2 \int_a^b |\dot{u}| \, dx < +\infty$$

and so $F_g'(u) \in \mathcal{L}(\tilde{X}, \mathbb{R})$. Note that this is true for all $u \in \tilde{X}$. Consider

$$F_g' : \tilde{X} \longrightarrow \mathcal{L}(\tilde{X}, \mathbb{R})$$
$$u \longmapsto F_g'(u)$$

Then F_g' is continuous in $\tilde{X}$. Indeed, let $u, v \in \tilde{X}$:

$$\| F_g'(u) - F_g'(v) \|_{\mathcal{L}(\tilde{X}, \mathbb{R})} = \sup_{\|w\|_{C^1} \leq 1} | F_g'(u)(w) - F_g'(v)(w) |$$

$$\leq 2 \sup_{\|w\|_{C^1} \leq 1} \int_a^b |\dot{u} - \dot{v}| \, |\dot{w}| \, dx$$

$$\leq 2 \int_a^b |\dot{u} - \dot{v}| \, dx \leq 2(b-a) \|u - v\|_{C^1}$$

Showing continuity. We can then apply THEOREM 1.9 and conclude that F is Fréchet diff. in $\tilde{X}$ with $F'(u) = F_g'(u)$. □

Since X affine space, REMARK 2.5 tells us that for $u_0 \in X$ we have

$$\delta F(u_0, v) = \lim_{t \to 0} \frac{F(u_0 + tv) - F(u_0)}{t}, \quad \forall v \in V$$

if the limit exists. Note that $X, V \subseteq \tilde{X}$. Therefore, as we just computed the Fréchet derivative of F (PROPOSITION 2.8), we get

$$F'(u_0)(v) = F_g'(u_0)(v) = \lim_{t \to 0} \frac{F(u_0 + tv) - F(u_0)}{t}$$

and so in particular

$$\boxed{\delta F(u_0, v) = 2 \int_a^b \dot{u}(x) \, \dot{v}(x) \, dx, \quad \forall u_0 \in X, v \in V}$$

We now look for solutions to (ELE) in order to find STATIONARY POINTS.

Thus assume $u_0 \in X$ is a min. of F over X. By REMARK 2.7 we have

$$\boxed{\delta F(u_0, v) = 0 \,, \quad \forall v \in V} \qquad \text{(ELE)}$$

Assuming also that $u_0 \in C^2[a,b]$ we get

$$0 = \delta F(u_0, v) = 2 \int_a^b \dot{u}_0 \dot{v} \, dx$$

(Integrate by parts WRT to $(\dot{u} v)' = \ddot{u} v + \dot{u} \dot{v}$)

$$= 2 \left[\dot{u}_0 v \right]_{x=a}^{x=b} - 2 \int_a^b \ddot{u}_0 v \, dx$$

(as $v(a) = v(b) = 0$)

$$= -2 \int_a^b \ddot{u}_0 v \, dx$$

Thus

$$(*) \quad \int_a^b \ddot{u}_0 v \, dx = 0 \,, \quad \forall v \in V$$

It looks like $(*)$ can hold iff $\ddot{u}_0 \equiv 0$. Let's say this is true (it actually is true and we will show it soon). Therefore, as $u_0(a) = A$, $u_0(b) = B$, we have

$$\ddot{u}_0 \equiv 0 \implies u_0 \text{ is straight line connecting } (a, A) \text{ and } (b, B)$$

WARNING This does not prove that u_0 minimizes F over X. We just proved that:

"If $u_0 \in C^2[a,b]$ is a min. of F over X $\implies$ u_0 straight line"

PROPOSITION 2.9 Let $u_0 \in X$ be a straight line. Then u_0 is the unique solution to

$$\min_{u \in X} F(u) \,.$$

Recall: $X = \{ u \in C^1[a,b] \mid u(a) = A, u(b) = B \}$, $F(u) := \int_a^b |\dot{u}|^2 dx$, $A < B$

<u>Proof</u> Let $w \in X$ be arbitrary. We need to prove:

1) $F(u_0) \leq F(w), \ \forall w \in X$ (thus u_0 is a minimizer)

2) $F(u_0) = F(w) \iff u_0 = w$ (thus u_0 is unique minimizer)

Let us show (1): As $u_0, w \in X$ then $v := w - u_0 \in V$, since $v(a) = v(b) = 0$.

$$
\begin{aligned}
F(w) = F(u_0 + v) &= \int_a^b |\dot{u}_0 + \dot{v}|^2 \, dx \\
&= \int_a^b |\dot{u}_0|^2 \, dx + 2\int_a^b \dot{u}_0 \dot{v} \, dx + \int_a^b |\dot{v}|^2 \, dx \\
&= F(u_0) + 2\int_a^b \dot{u}_0 \dot{v} \, dx + F(v)
\end{aligned}
$$

Now u_0 is actually $C^2[a,b]$ (being a straight line). As $v(a) = v(b) = 0$, we can proceed as above (integrating by parts to obtain)

$$
\int_a^b \dot{u}_0 \dot{v} \, dx = \underbrace{\Big[\dot{u}_0 v\Big]_a^b}_{=0 \text{ as } v(a)=v(b)=0} - \underbrace{\int_a^b \ddot{u}_0 v \, dx}_{=0 \text{ as } \ddot{u}_0 = 0, \text{ since } u_0 \text{ is a line}} = 0
$$

Thus

$$
F(w) = F(u_0) + F(v) \geq F(u_0), \quad (\text{Since } F \geq 0 \text{ by definition})
$$

showing (1). Let us prove (2): We know that

$$
F(w) = F(u_0) + F(v)
$$

So $F(w) = F(u_0)$ iff $F(v) = 0$. By def of F this is true iff $\dot{v} \equiv 0$. Thus $v \equiv$ constant. Since $v(a) = v(b) = 0 \implies v \equiv 0$. Recalling that $v = w - u_0$, we infer $w \equiv u_0$, as claimed. $\square$

Recall: $X = \{u \in C^1[a,b] \mid u(a)=A, u(b)=B\}$

MINIMIZATION FOR G

$$G(u) := \int_a^b |\dot u| \, dx, \quad A<B$$

PROPOSITION 2.10 We have that

$$(*) \qquad \min_{u\in X} G(u) = B-A$$

and the minimum exists. Moreover the only solutions to $(*)$ are the monotonic functions, that is, $u_0 \in X$ with $\dot u_0 \geq 0$.

Proof Let $u \in X$ be arbitrary. Then

$$G(u) = \int_a^b |\dot u(x)| \, dx \geq \left| \int_a^b \dot u(x) \, dx \right| = |u(b)-u(a)| = B-A$$

Hence

$$(LB) \qquad G(u) \geq B-A, \qquad \forall u \in X$$

This lower bound is achieved by u_0 straight line between (a,A), (b,B). Indeed,

$$G(u_0) = \int_a^b |\dot u_0(x)| \, dx = \int_a^b \frac{B-A}{b-a} \, dx = B-A$$

Therefore u_0 solves $(*)$, as (LB) implies

$$(**) \qquad G(u) \geq G(u_0) = B-A, \quad \forall u \in X.$$

In particular $(*)$ holds.

Assume now that $w_0 \in X$ solves $(*)$. Thus $G(w_0) = B-A$. Since (as above)

$$G(w_0) = \int_a^b |\dot w_0(x)| \, dx \geq \left| \int_a^b \dot w_0(x) \, dx \right| = B-A,$$

we have $G(w_0) = B-A$ iff

$$\int_a^b |\dot w_0| \, dx = \left| \int_a^b \dot w_0(x) \, dx \right|.$$

But this is true iff $\operatorname{sign}(\dot w_0)$ is constant $\iff$ $\dot w_0 \geq 0$ (as $A<B$). $\square$

Recall: $X = \{ u \in C^1[a,b] \mid u(a) = A,\ u(b) = B \}$

$H(u) := \int_a^b \sqrt{|\dot u|}\, dx \quad , \quad A < B$

MINIMIZATION FOR H

PROPOSITION 2.11 We have that

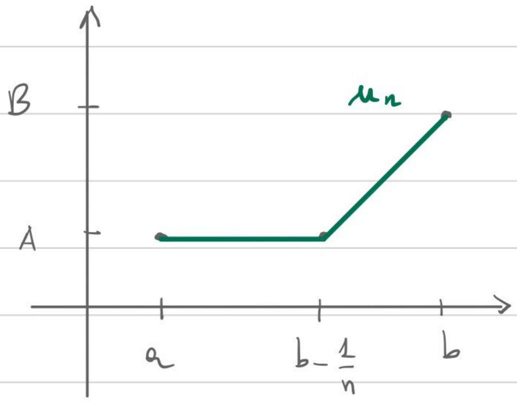
$$(*) \quad \min_{u \in X} H(u)$$

has no solutions, and

$$(**) \quad \inf_{u \in X} H(u) = 0 .$$

Proof Let us first show $(**)$. As $H \geq 0$, then $\inf_X H \geq 0$. Thus we need to find $\{u_n\} \subseteq X$ s.t. $H(u_n) \to 0$, so that $\inf H = 0$ follows.

Define $u_n : [a,b] \to \mathbb{R}$ as in the picture below



That is:
- $u_n \equiv A$ in $[a, b - \frac{1}{n}]$
- u_n straight line in $[b - \frac{1}{n}, B]$, so that $u_n(b - \frac{1}{n}) = A$ and $u_n(b) = B$

Then we have

$$H(u_n) = \int_a^b \sqrt{|\dot u_n|}\, dx = \int_{b-1/n}^{b} \sqrt{|\dot u_n|}\, dx$$

As $\dot u_n \equiv 0$ in $[a, b - \frac{1}{n})$ and $\dot u_n \equiv \frac{B-A}{b-(b-1/n)} = n(B-A)$ in $(b - 1/n, b]$

$$= \int_{b-1/n}^{b} \sqrt{n(B-A)}\, dx =$$

$$= \frac{1}{\sqrt{n}} \sqrt{B-A} \longrightarrow 0 \quad \text{as } n \to +\infty$$

PROBLEM: This almost shows (**). The only issue is that $u_n \notin C^1[a,b]$, as $\dot{u}_n$ has a jump at $x = b - 1/n$.

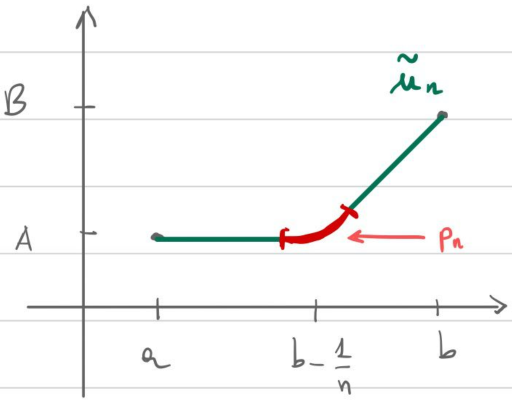
FIX: Smooth u_n around $x = b - 1/n$. For example one could define

$$\tilde{u}_n(x) := \begin{cases} u_n(x) & \text{if } x \in [a, b - \frac{2}{n}] \cup [b - \frac{1}{2n}, b] \\ p_n(x) & \text{if } x \in [b - \frac{2}{n}, b - \frac{1}{2n}] \end{cases}$$

where p_n is a polynomial such that

$$\begin{cases} p_n\left(b - \frac{2}{n}\right) = u_n\left(b - \frac{2}{n}\right), & \dot{p}_n(b - 2/n) = \dot{u}_n(b - 2/n) \\ p_n(b - 1/2n) = u_n(b - 1/2n), & \dot{p}_n(b - 1/2n) = \dot{u}_n(b - 1/2n) \end{cases}$$

so that $\tilde{u}_n \in C^1[a,b]$ and so $\tilde{u}_n \in X$ is admissible. This would look like



Since the region where we changed u_n is infinitesimal as $n \to +\infty$, we get

$$H(\tilde{u}_n) = H(u_n) + o(1) \to 0 \quad \text{as } n \to +\infty$$

(Showing these details will be left as an exercise in the Ex. course)

showing (**). We are left to show that (*) admits no minimizers. Assume by contradiction that the infimum is achieved by some $u_0 \in X$. As (**) holds we get $H(u_0) = 0$, which is possible iff $u_0 \equiv \text{constant}$. However, since $u_0(a) = A$ and $u_0(b) = B$, and since we are assuming $A \neq B$, we get a contradiction.

$\square$

Summary

We considered functionals on $X = \{ u \in C^1(a,b) \mid u(a)=A,\ u(b)=B \}$,

$$F(u) = \int_a^b \dot{u}^2\,dx \quad , \quad G(u) = \int_a^b |\dot{u}|\,dx \quad , \quad H(u) = \int_a^b \sqrt{|\dot{u}|}\,dx$$

For these the solutions were as follows:

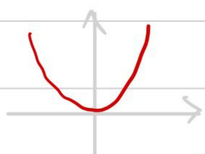
$\min\limits_{u \in X} F(u)$ → UNIQUE MINIMIZER: u_0 STRAIGHT LINE BETWEEN $(a,A),(b,B)$

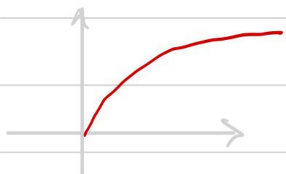
$\min\limits_{u \in X} G(u)$ → INFINITELY MANY MINIMIZERS: ALL THE MONOTONIC FUNCTIONS $\dot{u} \geq 0$

$\min\limits_{u \in X} H(u)$ → NO MINIMIZERS

To understand what is going on, let us considers the LAGRANGIANS associated to F, G, H

$$L_F(x,s,p) = p^2 \quad , \quad L_G(x,s,p) = |p| \quad , \quad L_H(x,s,p) = |p|^{1/2}$$





We observe:

- L_F is STRICTLY CONVEX in p $\rightsquigarrow$ $\exists!$ minimizer $u_0 \in C^\infty[a,b]$ (smooth)

- L_G is CONVEX in p, but NOT STRICTLY $\rightsquigarrow$ $\exists$ minimizer, but no uniqueness, no smooth

- L_H is NOT CONVEX in p $\rightsquigarrow$ Non Existence and no regularity

LESSON 4 - 24 MARCH 2021

3. FUNDAMENTAL LEMMAS

We now prove two fundamental Lemmas which will be ubiquitous throughout the course (we already used one of them after in the example of F, right before PROPOSITION 2.9).

DEFINITION 3.1 Let $u : (U \subseteq \mathbb{R}) \to \mathbb{R}$. The **SUPPORT** of u is the set

$$\operatorname{supp} u := \overline{\{x \in U \mid u(x) \neq 0\}}$$

We define the space of SMOOTH COMPACTLY SUPPORTED functions on (a,b) as

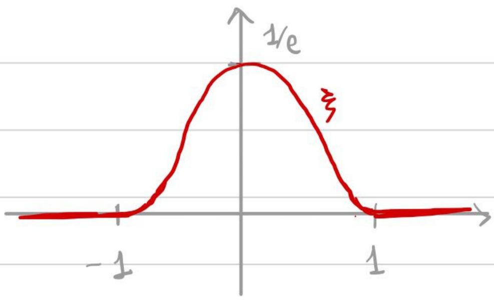
$$C^\infty_c(a,b) := \{ u \in C^\infty(a,b) \mid \operatorname{supp} u \text{ is COMPACT} \}$$

In other words, $u \in C^\infty_c(a,b)$ iff $\exists$ $[c,d] \subseteq (a,b)$ s.t. $\operatorname{supp} u \subseteq [c,d]$, i.e., $u \equiv 0$ $(a,b) \setminus [c,d]$.

REMARK 3.2 We can construct $u \in C^\infty_c(a,b)$ having PRESCRIBED support in some interval $[c,d] \subseteq (a,b)$, and having the same sign, i.e., either $u \geq 0$ or $u \leq 0$.

To do that, consider the **BUMP FUNCTION**

$$\xi(x) := \begin{cases} \exp\left(-\dfrac{1}{1-x^2}\right), & \text{if } |x| < 1 \\ 0, & \text{if } |x| \geq 1 \end{cases}$$



Then $\xi \in C^\infty_c(\mathbb{R})$, $\operatorname{supp} \xi \subseteq [-1,1]$ and $\xi > 0$ in $(-1,1)$.

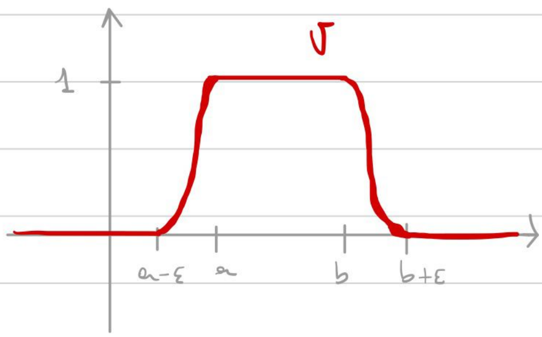
For $x_0 \in \mathbb{R}$, $r > 0$ fixed define

$$(*) \qquad u(x) := \xi\left(\frac{x - x_0}{r}\right)$$

Then $u \in C^\infty_c(\mathbb{R})$, $\operatorname{supp} u \subseteq [x_0 - r, x_0 + r]$ and $u > 0$ in $(x_0 - r, x_0 + r)$

(To get $u < 0$ just consider $-\xi$ in the definition $(*)$)

REMARK 3.3 Using the function ξ at REMARK 3.2 and CONVOLUTIONS, it is possible to construct $\varphi \in C_c^\infty(\mathbb{R})$ such that $0 \leq \varphi \leq 1$ and



$$\varphi(x) = \begin{cases} 1 & x \in [a,b] \\ 0 \quad \text{if} & x \notin [a-\varepsilon, b+\varepsilon] \end{cases}$$

where a, b and $\varepsilon > 0$ can be chosen arbitrarily. Such φ is called CUT-OFF function

(We omit the proof of this fact for the moment. It will be left as an exercise in the EXERCISES COURSE).

LEMMA 3.4 (FUNDAMENTAL LEMMA OF CALCULUS OF VARIATIONS) (FLCV)

Let $f: [a,b] \to \mathbb{R}$ be continuous. Assume that

$$\int_a^b f(x)\, \varphi(x)\, dx = 0, \qquad \forall\, \varphi \in C_c^\infty(a,b)$$

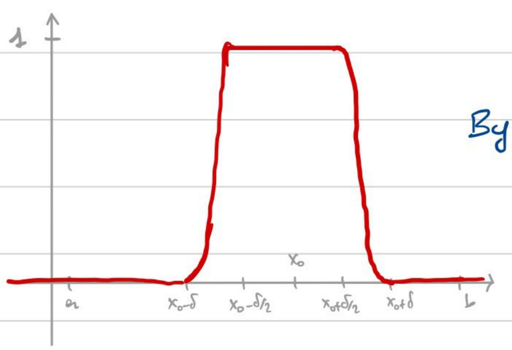
Then $f \equiv 0$.

We give 2 proofs of this Lemma, to show different and interesting techniques;

PROOF 1 OF LEMMA 3.4 (By contradiction)

Assume by contradiction that $f \not\equiv 0$. Then WLOG $\exists\, x_0 \in (a,b)$ such that $f(x_0) > 0$. By continuity also $\exists\, \delta > 0$ s.t.

$$f(x) \geq \frac{f(x_0)}{2}, \quad \forall\, x \in (x_0-\delta, x_0+\delta) \subseteq [a,b]$$



By REMARK 3.3 $\exists\, \varphi \in C_c^\infty(\mathbb{R})$ s.t. $0 \leq \varphi \leq 1$ and

$$\varphi(x) = \begin{cases} 1 & \text{for } x \in [x_0 - \delta/2,\, x_0 + \delta/2] \\ 0 & \text{for } x \notin [x_0-\delta, x_0+\delta] \end{cases}$$

Thus by assumption we have

$$\int_a^b f(x)\, v(x)\, dx = 0.$$

On the other hand,

$$\int_a^b f(x)\, v(x)\, dx = \int_{x_0-\delta}^{x_0+\delta} f(x)\, v(x)\, dx \geq \int_{x_0-\delta/2}^{x_0+\delta/2} f(x)\, v(x)\, dx \geq \frac{f(x_0)}{2}\, \delta > 0$$

As $v = 0$ outside of $[x_0-\delta, x_0+\delta]$

As $v \geq 0$ always, while $f \geq \frac{f(x_0)}{2} > 0$ in $[x_0-\delta, x_0+\delta]$

Since $v \equiv 1$ and $f(x) \geq f(x_0)/2$ here

which is a contradiction. $\square$

Before proceeding with the second proof of LEMMA 3.4, we make the following remark (a proof of which is left for the exercises course)

REMARK 3.5 Let $v : [a,b] \to \mathbb{R}$ continuous. There exists a sequence $\{v_n\} \subseteq C_c^\infty(a,b)$ s.t.

1) $\{v_n\}$ is uniformly bounded, i.e., $\exists M > 0$ s.t.

$$\sup_n \| v_n \|_\infty \leq M$$

2) For each $K \subseteq [a,b]$ compact we have that $v_n \to v$ uniformly on K, that is,

$$\lim_{n \to +\infty} \sup_{x \in K} | v_n(x) - v(x) | = 0.$$

PROOF 2 OF LEMMA 3.4 (By Density)

We claim the following

$$(*) \quad \int_a^b f(x)\, v(x) = 0, \ \forall v \in C_c^\infty(a,b) \implies \int_a^b f(x)\, v(x) = 0, \ \forall v \in C(a,b)$$

Notice that if (*) holds then the thesis of Lemma 3.4 follows: indeed, as we are assuming that f satisfies

$$\int_a^b f(x)\varphi(x)\,dx = 0 \quad , \forall \varphi \in C_c^\infty(a,b) \,,$$

then by (*) we get that

$$\int_a^b f(x)\varphi(x)\,dx = 0 \quad , \forall \varphi \in C(a,b) \,.$$

Thus we can choose $\varphi = f$ in the above (as f is continuous by assumption) and obtain

$$\int_a^b |f|^2\,dx = 0 \quad \Rightarrow f \equiv 0 \,,$$

which concludes the proof.

Thus, we are left to show (*). To this end, fix $\varphi \in C(a,b)$. By REMARK 3.5 $\exists \{\varphi_n\} \subseteq C_c^\infty(a,b)$ s.t. $\{\varphi_n\}$ is unif. bounded and $\varphi_n \to \varphi$ uniformly on each $K \subset (a,b)$ compact. As φ_n is smooth, by assumption we have

$$(**) \qquad \int_a^b f(x)\varphi_n(x)\,dx = 0 \quad , \forall \; n \in \mathbb{N} \,.$$

On the other hand, let $K \subset (a,b)$ be compact. Then

$$\left| \int_a^b f\varphi_n\,dx - \int_a^b f\varphi\,dx \right| \leq \|f\|_\infty \int_a^b |\varphi_n - \varphi|\,dx =$$

$$= \|f\|_\infty \left(\int_K |\varphi_n - \varphi|\,dx + \int_{K^c} |\varphi_n - \varphi|\,dx \right) \qquad \left(K^c := (a,b) \setminus K \right)$$

Now the first integral:

$$\int_K |v_n - v|\,dx \leq |K| \sup_{x\in K} |v_n(x) - v(x)| \to 0 \quad \text{as } n\to+\infty$$

by the properties of v_n. For the second integral we have:

$$\int_{K^c} |v_n - v|\,dx \leq |K^c| \left(\|v_n\|_\infty + \|v\|_\infty \right) \leq |K^c| \left(M + \|v\|_\infty \right)$$

In total,

$$\limsup_{n\to+\infty} \left| \int_a^b f v_n \,dx - \int_a^b f v \,dx \right| \leq \|f\|_\infty |K^c| \left(M + \|v\|_\infty \right).$$

Now, remember that $K \subset (a,b)$ is an arbitrary compact set. Thus $|K^c|$ is as small as we wish, from which we infer

$$\int_a^b f v_n \,dx \to \int_a^b f v\,dx \quad \text{as } n\to+\infty$$

Since (**) holds, we conclude that $\int_a^b f v \,dx = 0$, and the CLAIM is proven. $\square$

The second proof immediately suggests possible generalizations of LEMMA 3.4, which will allow us to test f against a smaller set of functions.

REMARK 3.6 Assume that $f \in C(a,b)$ satisfies

$$\int_a^b f(x) v(x)\,dx = 0, \quad \forall v \in V$$

where $V \subset C(a,b)$ is some set. Then

1) By linearity of the integral we have

$$\int_a^b f(x)\,v(x)\,dx = 0 \,, \quad \forall\, v \in \text{span}\, V$$

2) By a density argument similar to the one of PROOF 2 of LEMMA 3.4 we have

$$\int_a^b f(x)\,v(x)\,dx = 0 \,, \quad \forall\, v \in \overline{V} \,,$$

where the closure is taken WRT the uniform convergence of bounded sequences on compact sets $K \subset (a,b) \setminus \{x_1, .., x_N\}$ where the collection of points $\{x_1, .., x_N\}$ is FINITE.

As a consequence of REMARK 3.6, and following the arguments of PROOF 2 of LEMMA 3.4 we get:

LEMMA 3.7 (Generalized FLCV)

Let $f \in C(a,b)$, $V \subset C(a,b)$ such that $\overline{\text{span}\, V} = C(a,b)$, where the closure is as in REMARK 3.6 point (2), i.e.,

$$\overline{\text{span}\, V} := \Big\{ v \in C(a,b) \;\Big|\; \exists\, \{v_n\} \subseteq \text{span}\, V, \text{ with } \sup_n \|v_n\|_\infty < +\infty \text{ and } v_n \to v \text{ uniformly on each compact } K \subset (a,b) \setminus I \Big\}$$

with $I := \{x_1, .., x_N\}$ is a fixed finite collection of points. Then

$$\int_a^b f(x)\,v(x) = 0 \,, \quad \forall\, v \in V \;\Rightarrow\; f \equiv 0 \,.$$

We now state and prove a second "fundamental" lemma, which again will be very useful in the rest of the course.

LEMMA 3.8 (DU BOIS REYMOND) (DBR Lemma)

Let $f \in C(a,b)$ and assume that

Zero average function

$$(*) \quad \int_a^b f(x) v(x)\, dx = 0, \quad \forall v \in C_c^\infty(a,b) \text{ s.t. } \int_a^b v(x)\, dx = 0.$$

Then $f \equiv c$ for some $c \in \mathbb{R}$.

Proof The idea is to apply the FLCV (LEMMA 3.4). Thus let $v \in C_c^\infty(a,b)$. It would be nice if we could use

$$\tilde{v}(x) := v(x) - \frac{1}{b-a} \int_a^b v(y)\, dy$$

as a test function in $(*)$, seeing that $\int_a^b \tilde{v}(x)\, dx = 0$. However $\tilde{v}$ is not compactly supported.

To make this attempt rigorous, take $w \in C_c^\infty(a,b)$ s.t. $\int_a^b w(x)\, dx = 1$, and define

$$\varphi(x) := v(x) - w(x) \int_a^b v(y)\, dy$$

Then $\varphi \in C_c^\infty(a,b)$ and $\int_a^b \varphi(x)\, dx = 0$. By using φ as a test function in $(*)$ we get

$$0 = \int_a^b f(x)\varphi(x)\, dx = \int_a^b f(x) v(x)\, dx - \int_a^b f(x) w(x) \left(\int_a^b v(y)\, dy \right) dx$$

$$= \int_a^b f(x) v(x)\, dx - c \int_a^b v(x)\, dx,$$

where $c := \int_a^b f(x) w(x)\, dx$

Thus

$$0 = \int_a^b f(x)v(x)\,dx - c\int_a^b v(x)\,dx$$

$$= \int_a^b [f(x)-c]\,v(x)\,dx$$

Since this is true for all $v \in C_c^\infty(a,b)$, by FLCV LEMMA 3.4 we conclude $f-c \equiv 0 \Rightarrow f \equiv c$. □

A simple (but useful) equivalent formulation of the DBR Lemma is the following one.

LEMMA 3.9 (DBR - Second formulation)

Let $f \in C(a,b)$ and assume that

$$(*) \quad \int_a^b f(x)\dot{v}(x)\,dx = 0, \quad \forall v \in C_c^\infty(a,b)$$

Then $f \equiv c$ for some $c \in \mathbb{R}$.

Proof For $v \in C_c^\infty(a,b)$ we have

$$(**) \quad \int_a^b v(x)\,dx = 0 \iff \exists\, w \in C_c^\infty(a,b) \text{ s.t. } \dot{w} = v$$

Indeed, if $w \in C_c^\infty(a,b)$ is s.t. $\dot{w} = v$, then

$$\int_a^b v(x)\,dx = \int_a^b \dot{w}(x)\,dx = w(b) - w(a) = 0 \quad \left(\text{w is compactly supported}\right)$$

Conversely, assume $\int_a^b v(x)\,dx = 0$, and let $\varepsilon > 0$ be s.t

$\operatorname{supp} v \subset [a+\varepsilon, b-\varepsilon]$ (since v is compactly supported)

For $x \in [a,b]$ define

$$w(x) := \int_a^x v(y)\,dy$$

Then $\dot{w} = v$, and in particular $w \in C^\infty(a,b)$. Moreover

$$w(x) = \int_a^x v(y)\,dy = 0 \quad \text{if} \quad x \in [a, a+\varepsilon]$$

as $v \equiv 0$ in $[a, a+\varepsilon]$, while

We are assuming this

$$w(x) = \int_a^x v(y)\,dy = \int_a^b v(y)\,dy = 0$$

if $x \in [b-\varepsilon, b]$, as the whole support of v is in $[a, b-\varepsilon]$.

Thus $(**)$ is proven. Now assume that $(*)$ holds. Let $v \in C_c^\infty(a,b)$ be such that $\int_a^b v(x)\,dx = 0$. Then by $(**)$ $\exists\, w \in C_c^\infty(a,b)$ s.t. $\dot{w} = v$. Therefore, by $(*)$, we have $\int_a^b f(x)\,\dot{w}(x) = 0$. Then, as $\dot{w} = v$,

$$\int_a^b f(x)\,v(x) = \int_a^b f(x)\,\dot{w}(x) = 0$$

As v is arbitrary, then $f \equiv c$ by DBR LEMMA 3.8. □

As for the FLCV, also in the DBR lemma we can test f against a smaller set of functions, since the DBR can also be proven with a density argument (very similar to PROOF 2 of LEMMA 3.4). Such argument makes use of the following remark (Again, left for the exercise course)

REMARK 3.10 Let $v \in C(a,b)$ with $\int_a^b v(x)\,dx = 0$. Then $\exists$ $\{v_n\} \subseteq C_c^\infty(a,b)$ such that

1) $\sup_n \|v_n\|_\infty \le M$, for some $M>0$

2) $v_n \to v$ uniformly on compact sets $K \subset (a,b)$

3) $\int_a^b v_n(x)\,dx = 0$, $\forall n \in \mathbb{N}$.

We have the following alternative proof of the DBR LEMMA 3.8.

ALTERNATIVE PROOF OF LEMMA 3.8 (by density)

By proceeding exactly as in PROOF 2 of LEMMA 3.4 (using REMARK 3.10 in place of REMARK 3.5) we can show that

$$\int_a^b f(x)\,v(x)\,dx = 0, \quad \forall v \in C_c^\infty(a,b) \text{ with } \int_a^b v(x)\,dx = 0$$

$$\Downarrow \qquad (*)$$

$$\int_a^b f(x)\,v(x)\,dx = 0, \quad \forall v \in C(a,b) \text{ with } \int_a^b v(x)\,dx = 0$$

Now the thesis of LEMMA 3.8 follows immediately by $(*)$. Indeed, assume that $f \in C(a,b)$ is such that

$$\int_a^b f(x)v(x)\,dx = 0, \quad \forall v \in C_c^\infty(a,b) \text{ with } \int_a^b v(x)\,dx = 0.$$

As v has zero average, then also $f+c$ for any $c \in \mathbb{R}$ satisfies the above.

Thus, by (*),

(**) $\int_a^b [f(x)+c]\, v(x)\, dx = 0$, $\forall v \in C(a,b)$ with $\int_a^b v(x)\,dx = 0$

In particular, take $c = -\frac{1}{b-a}\int_a^b f(x)\,dx$, so that $\int_a^b f+c = 0$.

Thus, we can test (**) against $v := f+c$ to get $\int_a^b (f+c)^2 = 0$
$\Rightarrow$ $f \equiv -c$. $\square$

Following a similar reasoning to the one in REMARK 3.6, and arguments similar to the ones contained in the above proof, we can obtain a generalized version of the DBR Lemma (which we state without proof).

LEMMA 3.11 (Generalized DBR)

Consider the space

$$V = \{ v \in C(a,b) \mid \int_a^b v(x)\,dx = 0 \}$$

Assume that $F \subseteq V$ is such that $\overline{\text{span}\, F} = V$, where $\overline{\text{span}\, V}$ is

$$\overline{\text{span}\, V} := \{ v \in C(a,b) \mid \exists \{v_n\} \subseteq \text{span}\, V, \text{ with } \sup_n \|v_n\|_\infty < +\infty \text{ and } v_n \to v \text{ uniformly on each compact } K \subset (a,b) \setminus I \}$$

with $I := \{x_1, \dots, x_N\}$ is a fixed finite collection of points. Let $f \in C(a,b)$. If

$$\int_a^b f(x)\, v(x)\, dx = 0, \quad \forall v \in F$$

then $f \equiv c$ for some $c \in \mathbb{R}$.

BOUNDARY CONDITIONS (By Examples)

EXAMPLE 1 (DIRICHLET BOUNDARY CONDITIONS)

$$F(u) = \int_0^1 \dot{u}^2 + u^2 \, dx \quad \text{with} \quad u \in X,$$

$$X := \{ u \in C^1[0,1] \mid u(0) = \alpha, u(1) = \beta \}$$

We want to find solutions to

$$\min_{u \in X} F(u).$$

Let us start by computing the first variation. Thus let

$$V = \{ v \in C^1[0,1] \mid v(0) = v(1) = 0 \}$$

So that X is an affine space over V. For $u \in X$, $v \in V$ we get

$$F(u+tv) = \int_0^1 (\dot{u} + t\dot{v})^2 + (u+tv)^2 \, dx =$$

$$= \int_0^1 \dot{u}^2 + 2t \int_0^1 \dot{u}\dot{v} + t^2 \int_0^1 \dot{v}^2 +$$

$$\int_0^1 u^2 + 2t \int_0^1 uv + t^2 \int_0^1 v^2$$

$$= F(u) + t^2 F(v) + 2t \int_0^1 (uv + \dot{u}\dot{v}) \, dx$$

Therefore

$$\delta F(u,v) = \lim_{t\to 0} \frac{F(u+tv) - F(u)}{t} =$$

$$= \lim_{t\to 0} \; tF(v) + 2\int_0^1 (uv + \dot{u}\dot{v})\,dx$$

$$= 2\int_0^1 (uv + \dot{u}\dot{v})\,dx$$

Therefore the EULER-LAGRANGE EQUATION reads

$$(*) \qquad \boxed{\int_0^1 (uv + \dot{u}\dot{v})\,dx = 0, \quad \forall v \in V}$$

Assuming that $u \in C^2[0,1]$, we can integrate $(*)$ by parts to obtain

$$(**) \qquad \boxed{\int_0^1 (-\ddot{u} + u)\, v\,dx = 0, \quad \forall v \in V}$$

where we used $v(0) = v(1) = 0$.

NOTATION
- $(*)$ is called 1st INTEGRAL FORM OF (ELE)
- $(**)$ is called 2nd INTEGRAL FORM OF (ELE)

Thus, if u is minimum of F and $u \in C^2[0,1]$, then u solves $(**)$.
As $C_c^\infty(0,1) \subseteq V$, we can apply FLCV (LEMMA 3.4) to $(**)$ and obtain

$$-\ddot{u} + u \equiv 0$$

Recalling that u satisfies BC, we then need to solve the ORDINARY DIFFERENTIAL EQUATION (ODE)

$$\begin{cases} \ddot{u}(x) = u(x), & \forall x \in (0,1) \\ u(0) = \alpha \\ u(1) = \beta \end{cases}$$

DIRICHLET BOUNDARY CONDITIONS (DBC)

Now this is solved by

$$(*) \qquad u(x) = A\cosh(x) + B\sinh(x)$$

for appropriate A, B (as well known from basic analysis courses).

WARNING Recall that this just proves that if $u \in C^2[0,1]$ is a minimizer for F in X, then u is of the form $(*)$. Showing that u as in $(*)$ is actually a minimum requires a proof (energy estimates)

EXAMPLE 2 (DBC and NEUMANN BOUNDARY CONDITION (NBC))

Same functional F from the previous example, but defined on

$$X = \{ u \in C^1[0,1] \mid u(0) = \alpha \}$$

Note: we do not assign a condition for $u(1)$.

Let us compute the first variation. This time the reference vector space is

$$V = \{ v \in C^1[0,1] \mid v(0) = 0 \}.$$

Note that, as a consequence of the def. of X, we do not need to assign conditions on $v(1)$.

As before, the first variation at $u \in X$ along the direction $v \in V$ is

$$\delta F(u,v) = 2 \int_0^1 (uv + \dot{u}\dot{v})\, dx$$

Assuming $u \in C^2[0,1]$ and integrating by parts:

$$\delta F(u,v) = 2 \int_0^1 uv\, dx + \underbrace{2\, \dot{u} v \Big|_0^1}_{} - 2 \int_0^1 \ddot{u} v\, dx$$

This time this term is not zero, but it is equal to $2\, \dot{u}(1)\, v(1)$

Thus the 2nd integral form of (ELE) is

$$\text{(ELE)} \qquad \boxed{\int_0^1 (-\ddot{u} + u)\, v\, dx + \dot{u}(1)\, v(1) = 0 \,, \quad \forall v \in V}$$

Thus if $u \in C^2[0,1]$ and u minimizes F in X, then (ELE) holds.
How do we proceed? We cannot apply FLCV or DBR straightforwardly.
So we proceed in 2 steps:

- Step 1: Consider only test function $v \in V$ such that $v(1) = 0$.
  In this case (ELE) reads

$$\int_0^1 (-\ddot{u} + u)\, v\, dx = 0, \quad \forall v \in C^1[0,1] \text{ s.t. } v(0) = v(1) = 0$$

In particular (as in EXAMPLE 1) we can apply FLCV to get

$$-\ddot{u} + u \equiv 0$$

and hence the ODE

$$\begin{cases} \ddot{u}(x) = u(x) & , \forall x \in (0,1) \\ u(0) = \alpha \end{cases}$$

• Step 2: Now we know that $\ddot{u} \equiv u$. Therefore (ELE) becomes

$$\dot{u}(1)\, v(1) = 0 \quad , \quad \forall v \in V$$

Thus, by testing against $v \in V$ s.t. $v(1) \neq 0$ we get

$$\dot{u}(1) = 0$$

In total, we found that u solves

$$\begin{cases} \ddot{u}(x) = u(x) & , \forall x \in (0,1) \\ u(0) = \alpha & \text{(DIRICHLET BOUNDARY CONDITION)} \\ \dot{u}(1) = 0 & \text{(NEUMANN BOUNDRY CONDITION NBC)} \end{cases}$$

NOTICE: By not imposing a DIRICHLET BOUNDARY CONDITION on $u(1)$ for $u \in X$, we see that minimizers must satisfy a homogeneous condition on $\dot{u}(1)$.

This will be true in general. Also note that the NBC is of one less order than the highest derivative appearing in F.

EXAMPLE 3 (NEUMANN BOUNDARY CONDITIONS – NBC)

F as before but $X := C^1[0,1]$, with no additional conditions.

Note that in this case it is trivially true that $u \equiv 0$ minimizes F. However, for instructive purposes, let us ignore this fact and proceed with our usual method.

This time the ref. vector space is $V = C^1[0,1]$. The first variation is always the same,

$$\delta F(u,v) = 2\int_0^1 (uv + \dot{u}\dot{v})\,dx.$$

Assuming that $u \in C^2[0,1]$ minimizes F on X, and integrating by parts

$$\text{(ELE)} \quad \boxed{\int_0^1 (-\ddot{u} + u)\,v\,dx + \dot{u}(1)v(1) - \dot{u}(0)v(0) = 0, \quad \forall v \in V}$$

We now proceed in 2 steps:

- Step 1: Test (ELE) against $v \in C_c^\infty(0,1) \subseteq V$, so that

$$\int_0^1 (-\ddot{u} + u)\,v\,dx = 0, \quad \forall v \in C_c^\infty(0,1)$$

Thus FLCV implies

$$-\ddot{u} + u \equiv 0$$

• Step 2: Since $-\ddot{u}+u\equiv 0$, (ELE) becomes

$$(*) \quad \dot{u}(1)v(1)-\dot{u}(0)v(0)=0, \qquad \forall v\in V$$

Testing $(*)$ against $v\in V$ s.t. $v(0)\neq 0$, $v(1)=0$ yields

$$\dot{u}(0)=0$$

Testing $(*)$ against $v\in V$ s.t. $v(0)=0$, $v(1)\neq 0$ yields

$$\dot{u}(1)=0$$

In total, u solves

$$\begin{cases} \ddot{u}(x)=u(x) & , \quad x\in(0,1) \\ \dot{u}(0)=0 \\ \dot{u}(1)=0 \end{cases} \quad \text{NEUMANN BOUNDARY CONDITIONS (NBC)}$$

EXAMPLE 4 (PERIODIC BOUNDARY CONDITIONS – PBC)

F as before, but

$$X=\{u\in C^1[0,1] \mid u(0)=u(1)\}$$

(Also now the solution is trivially $u\equiv 0$. But let's ignore this).

Note X is vector space, so we can take $V=X$. The first variation δF is the same. Assuming $u\in C^2[0,1]$ minimizes F on X and integrating by parts:

(ELE) $$\boxed{\int_0^1 (-\ddot{u} + u)v \, dx + v(0)\{\dot{u}(1) - \dot{u}(0)\} = 0, \quad \forall v \in V}$$

where we used that $v(0) = v(1)$. We proceed in 2 steps:

- <u>Step 1</u> As usual, we can test against all $v \in C_c^\infty(0,1) \subseteq V$ and get

$$-\ddot{u} + u \equiv 0$$

- <u>Step 2</u>: We know that

$$v(0)\{\dot{u}(1) - \dot{u}(0)\} = 0, \quad \forall v \in V$$

Testing against $v \in V$ with $v(0) \neq 0$ (and $v(1) = v(0)$) we conclude

$$\dot{u}(0) = \dot{u}(1)$$

Recalling that $u(0) = u(1)$ as $u \in X$, we then get

$$\begin{cases} \ddot{u}(x) = u(x), & \forall x \in (0,1) \\ u(0) = u(1) \\ \dot{u}(0) = \dot{u}(1) \end{cases}$$

$\left.\begin{matrix} u(0) = u(1) \\ \dot{u}(0) = \dot{u}(1) \end{matrix}\right\}$ PERIODIC BOUNDARY CONDITIONS (PBC)

<u>EXAMPLES</u> F the same, $X = \{u \in C^1[0,1] \mid u(1) = u(0) + 2\}$

X is not a vector space. It is however affine space over

$$V = \{C^1[0,1] \mid v(0) = v(1)\}$$

By very similar calculations to the previous 4 examples, we get that if $u \in C^2[0,1]$ minimizes F over X, then

$$\begin{cases} \ddot{u}(x) = u(x), & \forall x \in (0,1) \\ u(1) = u(0) + 5 & \text{(This was enforced in } X\text{)} \\ \dot{u}(0) = \dot{u}(1) & \text{(NBC / PBC)} \end{cases}$$

EXAMPLE 6 (TOO MANY BOUNDARY CONDITIONS!)

F the same,

$$X = \{ u \in C^1[0,1] \mid u(\tfrac{1}{2}) = \alpha \}.$$

X is affine over $V = \{ v \in C^1[0,1] \mid v(\tfrac{1}{2}) = 0 \}$. If $u \in C^2[0,1]$ minimizes F over X, we integrate by parts to find

$$\text{(ELE)} \quad \boxed{\int_0^1 (-\ddot{u} + u) v \, dx + \dot{u}(1) v(1) - \dot{u}(0) v(0) = 0, \quad \forall v \in V}$$

- Step 1: Define

$$W := \{ v \in C^1[0,1] \mid v(0) = v(\tfrac{1}{2}) = v(1) = 0 \} \subseteq V$$

By (ELE) we have

$$(*) \quad \int_0^1 (-\ddot{u} + u) v \, dx = 0, \quad \forall v \in W$$

Now notice that $\overline{\operatorname{span} W} = C(0,1)$, where the closure is taken WRT the uniform convergence on compact subsets of $[0,1] \setminus \{\frac{1}{2}\}$. Then we can apply the GENERALIZED FLCV (LEMMA 3.7) to (*) and infer

$$\begin{cases} -\ddot{u} + u \equiv 0 \\ u(\frac{1}{2}) = \alpha \quad \text{(this is from } u \in X\text{)} \end{cases}$$

• Step 2: As $-\ddot{u} + u \equiv 0$, from (ELE) we get

$$\dot{u}(1)v(1) - \dot{u}(0)v(0) = 0 \quad , \quad \forall v \in V$$

Now just take $v \in V$ s.t. $v(1) = 0$, $v(0) \neq 0$ and $\tilde{v} \in V$ s.t. $\tilde{v}(1) \neq 0$, $\tilde{v}(0) = 0$ and obtain

$$\dot{u}(1) = \dot{u}(0) = 0 .$$

In total, u solves

$$\text{(ODE)} \begin{cases} \ddot{u}(x) = u(x) & , \ x \in (0,1) \\ u(1/2) = \alpha \\ \dot{u}(0) = \dot{u}(1) = 0 \end{cases}$$

As the ODE is of order 2 and we get 3 pointwise conditions, it is very unlikely that (ODE) admits a solution.

Notice that solvin (ODE) is equivalent to solving 2 seperate ODEs and then hoping that the solutions can be glued at 1/2 in a C^2 way where the two ODEs are

$$\text{(P1)} \begin{cases} \ddot{u} = u & \text{in } (0,1/2) \\ \dot{u}(0) = 0 \\ u(1/2) = \alpha \end{cases} \quad , \quad \text{(P2)} \begin{cases} \ddot{u} = u & \text{in } (1/2,1) \\ \dot{u}(1) = 0 \\ u(\frac{1}{2}) = \alpha \end{cases}$$

So there are two possibilities:

1) (ODE) admits a solution u $\Rightarrow$ with energy arguments we show that u minimizes F over X.

2) (ODE) does not admit a solution. Thus

$$\min_{u \in X} F(u)$$

admits no minimizer

$$\Downarrow$$

We solve (P1) and (P2), say with solutions $u_1 \in C^1[0,1/2]$, $u_2 \in C^1[1/2,1]$ respectively. Then

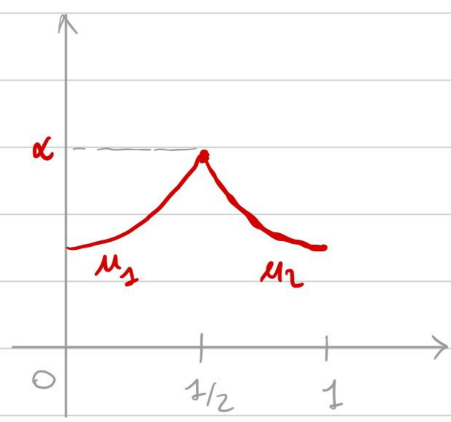
$$\hat{u}(x) := \begin{cases} u_1(x) & \text{if } x \in [0,1/2] \\ u_2(x) & \text{if } x \in [1/2,1] \end{cases}$$

DOES NOT BELONG to $C^1[0,1]$ (otherwise it would be a minimum).

One can show that

$$\inf_{u \in X} F(u) = F(\hat{u}) \leftarrow$$

(Note $F(\hat{u})$ is well defined by splitting the integral)

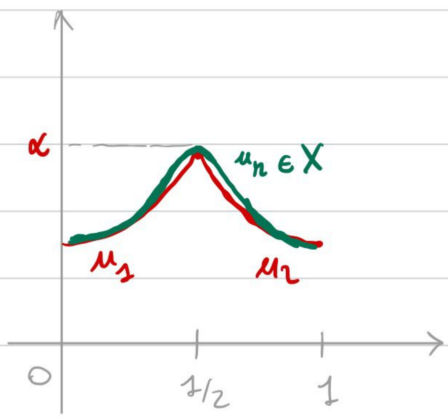


Idea of the proof:

① Show that $F(u) \geq F(\hat{u})$ for all $u \in X$, by the usual energy estimates

② Construct $\{u_n\} \subseteq X$ s.t. $u_n \to u$ uniformly on each $K \subseteq [0,1] \setminus \{1/2\}$ compact and

$$F(u_n) \to F(\hat{u})$$



This is done in the usual way: ROUNDING the corner of $\hat{u}$ at $x = 1/2$.

LESSON 5 - 14 APRIL 2021

4. THE EULER-LAGRANGE EQUATION

After the many examples seen so far, we look at the general theory for the minimization of integral functionals

$$F(u) := \int_a^b L(x, u(x), \dot{u}(x))\, dx$$

where $L : [a,b] \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}$, $L = L(x, s, p)$, is the LAGRANGIAN and $u : [a,b] \to \mathbb{R}$. We want to make sufficient assumptions on L so that F admits the first variation δF in some appropriate domain of definition. Specifically, we have:

THEOREM 4.1 Suppose that $L : [a,b] \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ is continuous, and continuously partially differentiable WRT to the variables s, p. Let $X \subseteq C^1([a,b])$ be an affine space, with reference vector space $V \subseteq C^1[a,b]$.
Define $F : X \to \mathbb{R}$ by setting

$$F(u) := \int_a^b L(x, u(x), \dot{u}(x))\, dx$$

Then F is Gâteaux differentiable at all points $u \in X$ and all directions $v \in V$, with

$$F_g'(u)(v) = \int_a^b L_s(x, u, \dot{u})\, v + L_p(x, u, \dot{u})\, \dot{v}\, dx$$

with $L_s := \partial_s L$, $L_p := \partial_p L$. In particular $\delta F(u, v)$ exists, with

$$\boxed{\delta F(u, v) = \int_a^b L_s(x, u, \dot{u})\, v + L_p(x, u, \dot{u})\, \dot{v}\, dx}$$

NOTE: Here $C^1[a,b]$ is equipped with the norm $\|u\| := \|u\|_\infty + \|u'\|_\infty$.

Proof Let $u \in X$, $v \in V$. As X is affine space over V, then $u + tv \in X$, $\forall t \in \mathbb{R}$. Then

$$(*) \quad \frac{F(u+tv) - F(u)}{t} = \int_a^b \underbrace{\frac{L(x, u+tv, \dot{u}+t\dot{v}) - L(x,u,\dot{u})}{t}}_{=: \Lambda(t,x)} dx$$

Now suppose $|t| \le \varepsilon$. Then

$$\Lambda(t,x) = \frac{1}{t} \int_0^t \left\{ \frac{d}{d\tau} L(x, u(x) + \tau v(x), \dot{u}(x) + \tau \dot{v}(x)) \right\} d\tau$$

As L diff. in s, p $\rightarrow$
$$= \frac{1}{t} \int_0^t \left\{ L_s(x, u+\tau v + \dot{u} + \tau\dot{v}) v + L_p(x, u+\tau v, \dot{u}+\tau\dot{v}) \dot{v} \right\} d\tau$$

Adding and subtracting $\rightarrow$
$$= L_s(x, u, \dot{u}) v + L_p(x, u, \dot{u}) \dot{v}$$
$$+ \frac{1}{t} \int_0^t \left\{ L_s(x, u+\tau v, \dot{u}+\tau\dot{v}) v - L_s(x,u,\dot{u}) v \right\} d\tau \quad (=: R_1(t,x))$$
$$+ \frac{1}{t} \int_0^t \left\{ L_p(x, u+\tau v, \dot{u}+\tau\dot{v}) \dot{v} - L_p(x,u,\dot{u}) \dot{v} \right\} d\tau \quad (=: R_2(t,x))$$

Thus, by $(*)$,

$$\frac{F(u+tv) - F(u)}{t} = \int_a^b \Lambda(t,x)\, dx$$
$$= \int_a^b L_s(x,u,\dot{u}) v + L_p(x,u,\dot{u}) \dot{v}\, dx + \int_a^b R_1(t,x)\, dx + \int_a^b R_2(t,x)\, dx$$

To see that F is Gâteaux diff it is sufficient to show that

$$\lim_{t \to 0} \int_a^b R_j(t,x)\, dx = 0, \quad \text{for } j = 1,2.$$

To this end, notice that, as $u, v \in C^1[a,b]$, then

$$K := \{ (x, u(x) + \tau v(x), \dot{u}(x) + \tau \dot{v}(x)) \mid x \in [a,b], |\tau| \leq \varepsilon/2 \}$$

is compact in $[a,b] \times \mathbb{R} \times \mathbb{R}$. As L_s is continuous on $[a,b] \times \mathbb{R} \times \mathbb{R}$, then in particular it is uniformly continuous on K (continuous on compact $\Rightarrow$ U.C.)
Then $\forall\ \tilde{\varepsilon} > 0$, $\exists\ \delta > 0$ s.t.

$$| L_s(x, u(x) + \tau v(x), \dot{u}(x) + \tau \dot{v}(x)) - L_s(x, u(x), \dot{u}(x)) | < \tilde{\varepsilon} \quad (*)$$

for all $x \in [a,b]$ and $|\tau| \leq \frac{\varepsilon}{2}$, such that $|\tau| (|v(x)| + |\dot{v}(x)|) < \delta$.
The last condition is fullfilled for τ s.t.

$$|\tau| < \min \left\{ \frac{\varepsilon}{2}, \frac{\delta}{\|v\|} \right\} \quad (B)$$

Therefore let $\hat{\varepsilon} > 0$ be arbitrary and fix $\tilde{\varepsilon} > 0$ s.t.

$$\tilde{\varepsilon} < \frac{\hat{\varepsilon}}{\|v\|} \quad (**)$$

Let also $\tilde{\delta} := \min \{ \varepsilon/2, \frac{\delta}{\|v\|} \}$. Then for $|t| < \delta$ we have

$$|R_1(t,x)| \leq \frac{1}{|t|} \int_0^t | L_s(x, u + \tau v, \dot{u} + \tau \dot{v}) - L_s(x, u, \dot{u}) | d\tau \, |v(x)|$$

$$\left(\begin{array}{c} \text{by } (*), \text{ as} \\ |\tau| \leq |t| < \delta, \text{ so} \\ \text{that (B) holds} \end{array} \right) \leq \frac{\|v\|}{|t|} \int_0^t \tilde{\varepsilon} \, d\tau = \|v\| \tilde{\varepsilon} < \hat{\varepsilon} \quad (\text{by } (**))$$

As $\hat{\varepsilon}$ is arbitrary, and $\tilde{\delta}$ does not depend on x, we conclude

$$\lim_{t \to 0} \sup_{x \in [a,b]} |R_1(t,x)| = 0.$$

By similar arguments also $\lim_{t \to 0} \sup_{x \in [a,b]} |R_2(t,x)| = 0$.

Then $\int_a^b R_i(t,x)\,dx \to 0$ as $t \to 0$. Taking the limit in

$$\frac{F(u+tv) - F(u)}{t} = \int_a^b L_s(x,u,\dot{u})v + L_p(x,u,\dot{u})\dot{v}\,dx + \int_a^b R_1(t,x)\,dx + \int_a^b R_2(t,x)\,dx$$

yields that

$$F_g'(u)(v) = \int_a^b L_s(x,u,\dot{u})v + L_p(x,u,\dot{u})\dot{v}\,dx,$$

as claimed. Now just recall that for affine spaces which are also normed,

$$\delta F(u,v) := \lim_{t \to 0} \frac{F(u+tv) - F(u)}{t}$$

(see REMARK 2.5). This concludes. □

DEFINITION 4.2 In the setting of THEOREM 4.1, we call

$$(*) \quad \boxed{\delta F(u,v) = \int_a^b L_s(x,u,\dot{u})v + L_p(x,u,\dot{u})\dot{v}\,dx}$$

the FIRST INTEGRAL FORM of the FIRST VARIATION.

CASE OF DIRICHLET BOUNDARY CONDITIONS

Assume $L : [a,b] \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ continuous, and continuously diff. in s, p. Let

$$X = \{ u \in C^1[a,b] \mid u(a) = \alpha, \ u(b) = \beta \}$$

which is an affine space over

$$V = \{ v \in C^1[a,b] \mid v(a) = v(b) = 0 \}.$$

Then we can apply THEOREM 4.1, and $\delta F(u,v)$ is given by $(*)$.

Assume also that

$$L \in C^2([a,b]\times\mathbb{R}\times\mathbb{R}), \qquad u \in C^2[a,b] \cap X$$

Then the second term in ✱ can be integrated by parts:

$$\int_a^b L_p(x,u,\dot{u})\,\dot{v}\,dx = L_p(x,u,\dot{u})\,v\Big|_a^b - \int_a^b (L_p(x,u,\dot{u}))'\,v(x)\,dx$$

$$= -\int_a^b (L_p(x,u,\dot{u}))'\,v(x)\,dx \qquad (\text{as } v(a)=v(b)=0)$$

Therefore ✱ reads

✱✱ $$\boxed{\delta F(u,v) = \int_a^b \left\{ L_s(x,\dot{u},\dot{u}) - (L_p(x,u,\dot{u}))' \right\} v(x)\,dx}$$

Note that, in the above assumptions, we can explicitly compute

$$(L_p(x,u,\dot{u}))' = L_{px}(x,u,\dot{u}) + L_{ps}(x,u,\dot{u})\,\dot{u} + L_{pp}(x,u,\dot{u})\,\ddot{u}$$

DEFINITION 4.3 ✱✱ is called the SECOND INTEGRAL FORM of the FIRST VARIATION

Assume in addition that u minimizes F over X.

Then by REMARK 3.7 we know that $\delta F(u,v)=0$, $\forall v \in V$. Note that $C_c^\infty(a,b) \subseteq V$. Hence we can apply the FLCV (LEMMA 3.4) to ✱✱ (equated to zero), and obtain

✱✱✱ $$\boxed{\left[L_p(x,u,\dot{u}) \right]' = L_s(x,u,\dot{u})}$$

Note that, in addition to ✱✱✱, u satisfies also the DIRICHLET BC imposed in X, that is,

$$u(a) = \alpha, \quad u(b) = \beta$$

DEFINITION 4.4 (***) is called **EULER-LAGRANGE EQUATION in DIFFERENTIAL FORM.**

We therefore have proven the following theorem.

THEOREM 4.5 Let $L: [a,b] \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ be continuous and continuously partially differentiable in s, p.
Define

$$X := \{ u \in C^1[a,b] \mid u(a) = \alpha, u(b) = \beta \}$$

$$V := \{ v \in C^1[a,b] \mid v(a) = v(b) = 0 \}$$

Define the functional $F: X \to \mathbb{R}$ s.t.

$$F(u) := \int_a^b L(x, u, \dot{u}) \, dx$$

(1) If $u \in X$ minimizes F over X, then u solves the **ELE in INTEGRAL FORM**:

$$\int_a^b L_s(x, u, \dot{u}) \, v + L_p(x, u, \dot{u}) \, \dot{v} \, dx = 0$$

for all $v \in V$.

(2) Assume in addition $L \in C^2([a,b] \times \mathbb{R} \times \mathbb{R})$.
If $u \in X \cap C^2[a,b]$ minimizes F over X, then u solves the **ELE in DIFFERENTIAL FORM**:

$$\begin{cases} \dfrac{d}{dx}\left[L_p(x, u(x), \dot{u}(x)) \right] = L_s(x, u(x), \dot{u}(x)), & \forall x \in (a,b) \\ u(a) = \alpha, \quad u(b) = \beta \end{cases}$$

THE CASE OF NEUMANN BOUNDARY CONDITIONS

Again, suppose $L : [a,b] \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ is continuous and continuously diff. in s, p. Define

$$X = \{ u \in C^1[a,b] \mid u(a) = \alpha \}$$

which is affine over

$$V = \{ v \in C^1[a,b] \mid v(a) = 0 \}.$$

We can then apply THEOREM 4.1 to obtain the FIRST INTEGRAL FORM of the FIRST VARIATION:

(*) $$\boxed{\delta F(u,v) = \int_a^b L_s(x,u,\dot{u})\, v + L_p(x,u,\dot{u})\, \dot{v} \; dx}$$

Assume in addition that

$$L \in C^2([a,b] \times \mathbb{R} \times \mathbb{R}), \qquad u \in C^2[a,b] \cap X.$$

Then the second term in (*) can be integrated by parts

$$\int_a^b L_p(x,u,\dot{u})\, \dot{v}\, dx = L_p(x,u,\dot{u})\, v \Big|_a^b - \int_a^b (L_p(x,u,\dot{u}))'\, v(x)\, dx$$

$$= L_p(b, u(b), \dot{u}(b))\, v(b) - \int_a^b \left[L_p(x,u,\dot{u}) \right]' v\, dx$$

obtaining the SECOND INTEGRAL FORM of the FIRST VARIATION:

(**) $$\boxed{\delta F(u,v) = \int_a^b \left\{ L_s(x,u,\dot{u}) - \left[L_p(x,u,\dot{u}) \right]' \right\} v\, dx + L_p(b, u(b), \dot{u}(b))\, v(b)}$$

Assume now that u is also a minimizer. Then by REMARK 3.7 we have $\delta F(u,v)=0$, $\forall v\in V$. In particular we can test for

$$v \in V \quad \text{s.t.} \quad v(b)=0$$

to obtain

$$\int_a^b \left\{ L_s(x,u,\dot{u}) - \left[L_p(x,u,\dot{u})\right]' \right\} v \, dx = 0, \quad \forall v \in C^1[a,b] \text{ s.t. } v(a)=v(b)=0.$$

Then by FLCV we obtain the EULER-LAGRANGE EQUATION in DIFF. FORM:

$$(***) \qquad \boxed{\left[L_p(x,u,\dot{u})\right]' = L_s(x,u,\dot{u})}$$

Now, the first boundary condition to pair to $(***)$ is already given in X:

$$u(a)=\alpha$$

For the second BC, just test $(**)$ against $v\in V$ s.t. $v(b)\neq 0$, and recall $(***)$, to get

$$L_p(b,u(b),\dot{u}(b))\,v(b)=0, \ \forall v\in V \xrightarrow[\text{s.t. } v(b)\neq 0]{\text{Taking } v\in V} L_p(b,u(b),\dot{u}(b))=0$$

which is a NEUMANN BOUNDARY CONDITION.

NOTE If we took $X=V=C^1[a,b]$ in the above example, we would have obtained that minimizers $u\in C^2[a,b]\cap X$ satisfy $(***)$ with two NEUMANN BC

$$L_p(a,u(a),\dot{u}(a)) = L_p(b,u(b),\dot{u}(b))=0$$

To summarize, we have proven the following Theorem:

THEOREM 4.6 Let $L : [a,b] \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ be continuous and continuously partially differentiable WRT s, p.
Define sets

$$X := \{ u \in C^1[a,b] \mid u(a) = \alpha \}$$

$$V := \{ v \in C^1[a,b] \mid v(a) = 0 \}$$

Define $F : X \to \mathbb{R}$ by

$$F(u) := \int_a^b L(x, u, \dot{u}) \, dx$$

(1) Suppose u minimizes F over X. Then u solves the ELE in INTEGRAL FORM:

$$\int_a^b L_s(x, u, \dot{u}) v + L_p(x, u, \dot{u}) \dot{v} \, dx = 0 , \quad \forall v \in V$$

(2) Suppose in addition $L \in C^2([a,b] \times \mathbb{R} \times \mathbb{R})$, and that $u \in X \cap C^2[a,b]$ minimizes F over X. Then u solves ELE in DIFFERENTIAL FORM:

$$\begin{cases} \dfrac{d}{dx}\left[L_p(x, u(x), \dot{u}(x)) \right] = L_s(x, u(x), \dot{u}(x)) , & \forall x \in (a,b) \\ u(a) = \alpha , \quad L_p(b, u(b), \dot{u}(b)) = 0 \end{cases}$$

ELE IN ERDMANN FORM

Consider the special case of Lagrangians not depending on x, i.e.,

$$F(u) = \int_a^b L(u, \dot{u})\, dx \quad , \quad L: \mathbb{R} \times \mathbb{R} \to \mathbb{R}$$

with $F: X \to \mathbb{R}$, $X \subseteq C^1[a,b]$ affine space over $V \subseteq C^1[a,b]$.

As done previously, if $L \in C^2(\mathbb{R} \times \mathbb{R})$ and $u \in C^2[a,b] \cap X$ minimizes F over X, the ELE reads

$$(*) \quad \left[L_p(u, \dot{u}) \right]' = L_s(u, \dot{u}) \quad , \quad \forall x \in (a,b)$$

Multiplying by $\dot{u}$ yields

$$(**) \quad \left[L_p(u, \dot{u}) \right]' \dot{u} = L_s(u, \dot{u})\, \dot{u}$$

Now the LHS is

$$\left[L_p(u, \dot{u}) \right]' \dot{u} = \left[L_p(u, \dot{u})\, \dot{u} \right]' - L_p(u, \dot{u})\, \ddot{u}$$

so that, from $(**)$

$$\left[L_p(u, \dot{u})\, \dot{u} \right]' \overset{\text{by } (**)}{=} L_s(u, \dot{u})\, \dot{u} + L_p(u, \dot{u})\, \ddot{u} \overset{\text{by direct calculation}}{=} \left[L(u, \dot{u}) \right]'$$

Therefore

$$\boxed{L_p(u, \dot{u})\, \dot{u} = L(u, \dot{u}) + \text{constant}}$$

ELE in ERDMANN FORM

REMARK 4.5 ELE and ELE-ERDMANN are <u>not</u> equivalent. It holds:

1. If u satisfies ELE $\Rightarrow$ u satisfies ELE-ERDMANN

   (we just proved this)

2. If u satisfies ELE-ERDMANN $\Rightarrow$ u satisfies ELE in the points $x \in [a,b]$ s.t. $\dot{u}(x) \neq 0$

   (To show this, just go backwards in the above calculation)

ELE FOR GENERAL LAGRANGIANS

HIGHER ORDER

$X \subseteq C^k[a,b]$ affine space over $V \subseteq C^k[a,b]$, $L : [a,b] \times \mathbb{R} \times \mathbb{R}^k \to \mathbb{R}$

$L = L(x, s, p_1, \dots, p_k)$, $F : X \to \mathbb{R}$ defined by:

$$F(u) := \int_a^b L(x, u, \dot{u}, \ddot{u}, \dots, u^{(k)})\, dx$$

Assume L is continuous and continuously differentiable WRT $s, p_1, \dots, p_k$.

Analogously to THEOREM 4.1, one can compute the Gâteaux derivative of F and obtain the FIRST INTEGRAL FORM of the FIRST VARIATION

$$(*) \qquad \boxed{\delta F(u,v) = \int_a^b L_s(x, u, \dots, u^{(k)})\, v + \sum_{i=1}^{k} L_{p_i}(x, u, \dots, u^{(k)})\, v^{(i)}\, dx}$$

Assume now that $L \in C^2([a,b] \times \mathbb{R} \times \mathbb{R}^k)$, $u \in C^{k+1}[a,b]$, and $v \in V$ is s.t. $v^{(i)}(a) = v^{(i)}(b) = 0$ for all $i = 0, .., k-1$. Integrating ⊛ by parts we get the SECOND INTEGRAL FORM of the FIRST VARIATION

$$\boxed{\delta F(u,v) = \int_a^b \left\{ L_s(x,u,..,u^{(k)}) + \sum_{i=1}^{k} (-1)^i \frac{d^i}{dx^i} L_{p_i}(x,u,..,u^{(k)}) \right\} v \, dx}$$

Finally, if in addition u is a minimizer, then $\delta F(u,v) = 0$, and by the FLCV we get the ELE in DIFFERENTIAL FORM

$$\boxed{\sum_{i=1}^{k} (-1)^{i+1} \frac{d^i}{dx^i} L_{p_i}(x,u,..,u^{(k)}) = L_s(x,u,..,u^{(k)})} \quad , \forall x \in (a,b)$$

MORE UNKNOWNS

$X \subseteq C^1[a,b]$ affine space over $V \subseteq C^1[a,b]$, $L : [a,b] \times \mathbb{R}^k \times \mathbb{R}^k \to \mathbb{R}$,

$L = L(x, s_1, .., s_k, p_1, .., p_k)$, $F : \underbrace{X \times \dots \times X}_{k \text{ times}} \to \mathbb{R}$ defined by

$$F(u_1, .., u_k) := \int_a^b L(x, u_1, .., u_k, \dot{u}_1, .., \dot{u}_k) \, dx$$

Assume L is continuous and continuously differentiable in $s_1, .., s_k, p_1, .., p_k$.

Analogously to THEOREM 4.1, one can compute the Gâteaux derivative of F and obtain the FIRST INTEGRAL FORM of the FIRST VARIATION

$$\boxed{\delta F(u,v) = \int_a^b \sum_{i=1}^{k} \left[L_{s_i}(x,u,\dot{u})\, v_i + L_{p_i}(x,u,\dot{u})\, \dot{v}_i \right] dx} \quad (*)$$

where $u = (u_1, \dots, u_k) \in X^k$, $v = (v_1, \dots, v_k) \in X^k$.

Suppose in addition that $L \in C^2([a,b] \times \mathbb{R}^k \times \mathbb{R}^k)$, $u_i \in C^2[a,b] \cap X$ and that $v_i \in V$ are s.t. $v_i(a) = v_i(b) = 0$, for all $i = 1, \dots, k$. Then we can integrate $(*)$ by parts to get the SECOND INTEGRAL FORM of the FIRST VARIATION

$$\boxed{\delta F(u,v) = \int_a^b \sum_{i=1}^{k} \left[L_{s_i}(x,u,\dot{u}) - L_{p_i}(x,u,\dot{u})' \right] v_i \, dx}$$

Finally, taking $u \in X^k$ minimum of F and $v_1 \in C_c^\infty(a,b)$, $v_2 = v_3 = \dots = v_k = 0$ and applying FLCV, we get

$$L_{p_1}(x,u,\dot{u})' = L_{s_1}(x,u,\dot{u})$$

Similarly, by taking the other components of v to be zero, except for one, we obtain the ELE in DIFFERENTIAL FORM

$$\boxed{L_{p_i}(x,u,\dot{u})' = L_{s_i}(x,u,\dot{u}), \quad i = 1, \dots, k} \quad , \ \forall x \in (a,b)$$

which in this case is a SYSTEM of k ODEs of ORDER 2.