University of Verona

Master's Program in Mathematics

Lecture notes of Optimization

Antonio Marigonda

Academic Year 2018 – 2019

Contents

CHAPTER 1

First part

1. Lecture of 1 october 2018: Introduction (3h)

"... nihil omnino in mundo contingit, in quo non maximi minime ratio quaepiam eluceat [1]*..."*
Leonhard Euler, 1744

We will speak of *optimization problem*, when we have a rational and coherent agent and a set of possible (mutually exclusive) alternatives X, among which the agent must choose. The agent must choose one of the alternatives basing on a choice criterion, characterized by a preference relation, that allows the comparison of pairs of alternatives $x, y \in X$.

The rationality and coherence of the agent impose that the preference relation must be a total order relation[2], and the simplest way is to assume the existence of a function $F : X \to [-\infty, +\infty]$ (called *cost function*) modeling the preference relation in the following way: the agent will prefer $x \in X$ to $y \in X$ if and only if $F(x) \le F(y)$.

Thus, the basic form of an optimization problem we are going to deal with is the following one.

Let X be a set, $F : X \to [-\infty, +\infty]$ be a function:

i. determine $\inf\limits_{x\in X} F(x)$, in particular establish if $\inf\limits_{x\in X} F(x) > -\infty$;
ii. establish if there exist points $\bar{x} \in X$ such that $F(\bar{x}) = \inf\limits_{x\in X} F(x)$. In such case, the set
$$\arg\min_{x\in X} F(x) = \left\{ \bar{x} \in X : \, F(\bar{x}) = \inf_{x\in X} F(x) \right\}$$
is called *the set of minimum points* of F on X, and $\inf\limits_{x\in X} F(x) = \min\limits_{x\in X} F(x)$;
iii. characterize the points of $\arg\min\limits_{x\in X} F(x)$.

We recall that $\sup\limits_{x\in X} F(x) = -\inf\limits_{x\in X}(-F(x))$, so also the problem of the maximization could be reformulated as minimization problems.

Maximization/minimizartion problems, apart from their mathematical interest, are at the basis of almost every field of knowledge, e.g. physics, biology models, economics, social sciences, but also industry, design, programming, resource management, transport...

Some examples:

- the *distance function* of a point $x \in \mathbb{R}^n$ from a nonempty set $S \subset \mathbb{R}^n$ is defined as
$$d_S(x) := \inf_{y\in S}\{\|y - x\|\},$$
in this case x is fixed and $F(y) = \|y - x\|$;

- *Fermat's Rule* in geometrical optics states that the path followed by the ligth to go from a point A to a point B is the path which minimizes the time needed to travel from A to B among all possible paths joining A and B;

[1] nothing in all the world will occur in which no maximum or minimum rule is somehow shining forth...

[2] In some framework, in particular in economics, there are also other possibility, modeling the fact that the agent can have multiple contrasting criteria of choice. This lead to drop the total order assumption and assuming that the preference relation is a partial order. We will not deal with this problems in this course

- the *least action principle* for a mechanical system subject to conservative forces and smooth constraints establish that if T is the kinetic energy of the systems and V is the potential energy, introducing the Lagrangian $L = T - V$, the trajectories followed by the system from an initial state q_0 at an initial time a to a final state q_1 at a final time b are the minimizers of the action
$$F(q(\cdot)) := \int_a^b L(t, q(t), \dot{q}(t))\, dt,$$
among all the trajectories $q : [a, b] \to \mathbb{R}^n$ satisfying $q(a) = q_0$, $q(b) = q_1$ and the constraints;

- usually students wants to pass exams minimizing the study time and maximizing the final mark.

We recall the following definitions.

DEFINITION 1.1. Let X be a set, $f : X \to \mathbb{R} \cup \{\pm\infty\}$ a function.
- a point $a \in X$ is an *absolute minimizer* of f if $f(y) \geq f(a)$ for all $y \in X$. The minimizer is strict if $f(y) > f(a)$ for all $y \in X \setminus \{a\}$.
- a point $a \in X$ is an *absolute maximizer* of f if $f(y) \leq f(a)$ for all $y \in X$. The maximizer is strict if $f(y) < f(a)$ for all $y \in X \setminus \{a\}$.

If X is endowed with a topology, we can give a *local* version of the above definitions.

DEFINITION 1.2. Let X be a topological space, $f : X \to \mathbb{R} \cup \{\pm\infty\}$ be a function.
- a point $a \in X$ is a *local minimizer* of f if there exists a neighborhood U of a such that $f(y) \geq f(a)$ for all $y \in U$. The local minimizer is strict if $f(y) > f(a)$ for all $y \in U \setminus \{a\}$.
- a point $a \in X$ is a *local maximizer* of f if there exists a neighborhood U of a such that $f(y) \leq f(a)$ for all $y \in U$. The local maximizer is strict if $f(y) < f(a)$ for all $y \in U \setminus \{a\}$.

Local maxima and minima are called *local extremals*.

EXAMPLE 1.3. Let $X = \mathbb{R}$. Consider $F(x) = x^4 - 12x + 1$. We want to minimize F over $\mathbb{R}$. As well known from the previous courses in Mathematical Analysis, we compute $F'(x) = 4x^3 - 12$. We have $F'(x) = 0$ if and only if $x = \sqrt[3]{3}$. Noticing that $\lim_{x\to\pm\infty} F(x) = +\infty$, and since $F''(x) = 12x^2$ and $F''(\sqrt[3]{3}) > 0$, we have that $\bar{x} = \sqrt[3]{3}$ is the unique minimum point for F and the value of the minimum is $F(\sqrt[3]{3}) = 1 - 9\sqrt[3]{3}$.

REMARK 1.4. Let us review carefully step by step the above example:

(1) **existence**: notice that $F \in C^0(\mathbb{R}, \mathbb{R})$, moreover $\lim_{|x|\to+\infty} F(x) = +\infty$. Since F is not identically $+\infty$, we have that $\inf_{x\in\mathbb{R}} F(x) < +\infty$. Thus let $\{x_n\}_{n\in\mathbb{N}}$ be a sequence in $\mathbb{R}$ such that $F(x_n) \to \inf_{x\in\mathbb{R}} F(x)$ (such kind of sequences are called *minimizing sequences*). If $|x_n|$ was unbounded, it would be possible to find a subsequence $\{x_{n_k}\}_{k\in\mathbb{N}}$ such that $|x_{n_k}| \to +\infty$, but in this case we would have
$$\inf_{x\in\mathbb{R}} F(x) = \lim_{n\to\infty} F(x_n) = \lim_{k\to\infty} F(x_{n_k}) = \lim_{|x|\to+\infty} F(x) = +\infty,$$
which is a contradiction. Thus there exists $M > 0$ such that $|x_n| \leq M$. But this implies that
$$\inf_{x\in\mathbb{R}} F(x) = \inf_{x\in[-M,M]} F(x),$$
and since F is continuous and the interval $[-M, M]$ is compact, we can apply *Weierstrass' Theorem*: every real-valued continuous function defined on a compact set admits absolute maxima and minima. So there **exists** at least one minimum point of F.

(2) **necessary condition**: let $\bar{x}$ be a minimum point. Since F is of class C^1, necessarily we must have $F'(\bar{x}) = 0$, this lead to isolate the candidate point $\bar{x} = \sqrt[3]{3}$.

(3) **sufficient condition**: Since $F \in C^2$, $F'(\bar{x}) = 0$ e $F''(\bar{x}) > 0$ then $\bar{x} = \sqrt[3]{3}$ is a minimum point for F on $\mathbb{R}$ and it is the unique minimum point.

REMARK 1.5. **Attention**: using necessary conditions *without an existence theorem* can be extremely misleading.

EXAMPLE 1.6 (Perron's Paradox). Consider the problem $\sup_{x \in \mathbb{R}^+} x$. Let $\bar{x} > 0$. We consider two possibilities: if $\bar{x} > 1$ then $\bar{x}^2 > \bar{x}$, thus $\bar{x}$ cannot be a maximum point, and if $0 < \bar{x} < 1$ then $\sqrt{\bar{x}} > \bar{x}$, so again $\bar{x}$ cannot be a maximum point. So these two necessary conditions lead us to isolate the point $\bar{x} = 1$. However $\sup_{x \in \mathbb{R}^+} x = +\infty$, there are no maximum points, and the necessary conditions are useless.

The core of the existence part of the proof is Weierstrass Theorem, which can be summarized by saying:

$$\textbf{continuity} + \textbf{compactness} = \textbf{existence of maxima and minima}.$$

Our aim is now to preserve *existence* weakening the other two assumptions. We preliminary notice that Weierstrass Theorem yields existence of *both* maxima and minima, but we are intrested *only* in the minimum points.

Thus it is natural to search for a *weakened* notion of continuity in some sense *still respecting the notion of minimum*.

Assume that $\bar{x} \in X$ is a local minimum point, in this case we have that there exists a neighborhood U of x such that

$$F(y) \geq F(\bar{x}), \text{ for all } y \in U \setminus \{\bar{x}\}.$$

In particular, we have

$$\inf_{y \in U \setminus \{\bar{x}\}} F(y) \geq F(\bar{x}),$$

which leads to

$$\sup_{\substack{V \text{ open} \\ \bar{x} \in V}} \inf_{y \in V \setminus \{\bar{x}\}} F(y) \geq F(\bar{x}).$$

By definition, the right hand side is the lim inf for y tending to $\bar{x}$, thus we obtain :

$$\liminf_{y \to \bar{x}} F(y) \geq F(\bar{x}).$$

If F is continuous and X is a topological Hausdorff space, we have that for every $x \in X$ it holds

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

in particular

$$\liminf_{y \to \bar{x}} F(y) = \lim_{y \to \bar{x}} F(y) = F(\bar{x});$$

thus a natural weakening of the cointinuity which repects the notion of minimum would be to require that for every $x \in X$ it holds

$$\liminf_{y \to x} F(y) \geq F(x).$$

DEFINITION 1.7 (Limsup and Liminf in topological spaces). Let X be a topological Hausdorff space, $x \in X$, $F : X \to [-\infty, +\infty]$. We define

$$\liminf_{y \to x} F(y) := \sup_{\substack{V \text{ open} \\ x \in V}} \inf_{y \in V \setminus \{x\}} F(y),$$
$$\limsup_{y \to x} F(y) := \inf_{\substack{V \text{ open} \\ x \in V}} \sup_{y \in V \setminus \{x\}} F(y),$$

LEMMA 1.8. *Let X be a topological Hausdorff space, $x \in X$, $F : X \to [-\infty, +\infty]$. We have always*

$$\liminf_{y\to x} F(y) \leq \limsup_{y\to x} F(y).$$

Moreover, $\lim\limits_{y\to x} F(y)$ exists if and only if

$$\liminf_{y\to x} F(y) = \limsup_{y\to x} F(y),$$

and in this case we have

$$\liminf_{y\to x} F(y) = \limsup_{y\to x} F(y) = \lim_{y\to x} F(y).$$

PROOF. Let $a > \limsup\limits_{y\to x} F(y)$. In particular, for any $\varepsilon > 0$ there exists an open neighborhood V_ε of x such that

$$\sup_{y\in V_\varepsilon\setminus\{x\}} F(y) - \varepsilon \leq \limsup_{y\to x} F(y),$$

and recalling the choice of a we obtain

$$\sup_{y\in V_\varepsilon\setminus\{x\}} F(y) \leq a + \varepsilon,$$

Given any open neighborhood V of x, we have

$$\inf_{y\in V\setminus\{x\}} F(y) \leq \inf_{y\in V\cap V_\varepsilon\setminus\{x\}} F(y) \leq \sup_{y\in V_\varepsilon\setminus\{x\}} F(y) \leq a + \varepsilon,$$

hence by taking the supremum on V we obtain

$$\liminf_{y\to x} F(y) \leq a + \varepsilon.$$

By letting $\varepsilon \to 0^+$ and $a \to \left[\limsup\limits_{y\to x} F(y)\right]^+$, we have

$$\liminf_{y\to x} F(y) \leq \limsup_{y\to x} F(y).$$

Indeed, assume that $\lim\limits_{y\to x} F(y) = \ell$ exists, and take any neighborhood W of ℓ. Then there exists a neighborhood V of x such that if $y \in V$ then $F(y) \in W$, thus

$$\inf W \leq \inf_{y\in V\setminus\{x\}} F(y),$$
$$\sup W \geq \sup_{y\in V\setminus\{x\}} F(y)$$

and so for all neighborhood W of ℓ we have

$$\inf W \leq \liminf_{y\to x} F(y) \leq \limsup_{y\to x} F(y) \leq \sup W.$$

If $\ell \in \mathbb{R}$ take $W = B(\ell, 1/i)$, if $\ell = -\infty$ take $W =]-\infty, -i]$, if $\ell = +\infty$ take $W =]i, +\infty[$. In all cases, by letting $i \to +\infty$ we have

$$\ell = \lim_{y\to x} F(y) = \liminf_{y\to x} F(y) = \limsup_{y\to x} F(y).$$

Assume now that

$$\ell = \liminf_{y\to x} F(y) = \limsup_{y\to x} F(y).$$

For any $\varepsilon > 0$ there exist open sets V_ε and V^ε such that $x \in V(\varepsilon) := V_\varepsilon \cap V^\varepsilon$ and

$$\sup_{y\in V_\varepsilon\setminus\{x\}} F(y) - \varepsilon \leq \ell \leq \inf_{y\in V^\varepsilon\setminus\{x\}} F(y) + \varepsilon,$$

hence

$$\sup_{y\in V(\varepsilon)\setminus\{x\}} F(y) - \varepsilon \leq \ell \leq \inf_{y\in V(\varepsilon)\setminus\{x\}} F(y) + \varepsilon,$$

So for every $\varepsilon > 0$, if $y \in V(\varepsilon) \setminus \{\bar{x}\}$ we have $F(y) \in B(\ell, \varepsilon)$, hence we have that the limit exists and

$$\lim_{y \to x} F(y) = \ell.$$

□

The following simple remark will be often used.

LEMMA 1.9. *Let X, Y, Z be nonempty sets, and let $f : X \times Y \times Z \to [-\infty, +\infty]$ be a map. Then*

$$\sup_{y \in Y}\left[\sup_{z \in Z} f(x,y,z)\right] = \sup_{z \in Z}\left[\sup_{y \in Y} f(x,y,z)\right],$$

$$\inf_{y \in Y}\left[\inf_{z \in Z} f(x,y,z)\right] = \inf_{z \in Z}\left[\inf_{y \in Y} f(x,y,z)\right],$$

$$\sup_{y \in Y}\left[\inf_{z \in Z} f(x,y,z)\right] \leq \inf_{z \in Z}\left[\sup_{y \in Y} f(x,y,z)\right].$$

PROOF. Assume that $M(x) \geq \sup_{z \in Z} f(x,y,z)$ for all $y \in Y$, then we have $M(x) \geq f(x,y,z)$ for all $y \in Y$, $z \in Z$, and so $M(x) \geq \sup_{y \in Y} f(x,y,z)$ for all $z \in Z$. Conversely, by reversing the role of y and z we have that if $M(x) \geq \sup_{y \in Y} f(x,y,z)$ for all $z \in Z$ then $M(x) \geq \sup_{z \in Z} f(x,y,z)$ for all $y \in Y$. This implies that $M(x) \geq \sup_{y \in Y} \sup_{z \in Z} f(x,y,z)$ if and only if $M(x) \geq \sup_{z \in Z} \sup_{y \in Y} f(x,y,z)$, and so equality hold, proving the first relation. The second relation can be obtained similarly, by replacing sup with inf and $\geq$ with $\leq$.

For the third relation, notice that for any $(x, \bar{y}, \bar{z}) \in X \times Y \times Z$ we have

$$\inf_{z \in Z} f(x,\bar{y},z) \leq f(x,\bar{y},\bar{z}) \leq \left[\sup_{y \in Y} f(x,y,\bar{z})\right],$$

in particular, for every $\bar{y} \in Y$ we have

$$\inf_{z \in Z} f(x,\bar{y},z) \leq \left[\sup_{y \in Y} f(x,y,\bar{z})\right], \text{ for all } \bar{z} \in Z,$$

and so for every $\bar{y} \in Y$ we have

$$\inf_{z \in Z} f(x,\bar{y},z) \leq \inf_{z \in Z}\left[\sup_{y \in Y} f(x,y,\bar{z})\right],$$

thus

$$\sup_{y \in Y} \inf_{z \in Z} f(x,y,z) \leq \inf_{z \in Z}\left[\sup_{y \in Y} f(x,y,\bar{z})\right].$$

□

REMARK 1.10. In the case in which X is a metric space, we have that the above definition of lim inf reduces to the usual one. More precisely, we can characterize the topological lim inf and lim sup by mean of sequences recalling that in every metric space a base for the topology is given by balls (in particular, every point has a countable base of neighborhoods).

LEMMA 1.11 (Liminf and limsup in metric spaces). *Let X be a metric space, $x \in X$, $F : X \to \mathbb{R} \cup \{\pm\infty\}$. Then the following are equivalent:*

(I_1) $\liminf_{y \to x} F(y) = \ell$;

(I_2) *for all sequences $\{y_j\}_{j \in \mathbb{N}} \subseteq X \setminus \{x\}$ such that $y_j \to x$ and $\{F(y_j)\}_{j \in \mathbb{N}}$ has a limit, we have $\lim_{j \to \infty} F(y_j) \geq \ell$, and there exists a sequence $\bar{y}_i \to \bar{x}$ such that equality holds.*

Symmetrically, the following are equivalent:

(S_1) $\limsup_{y\to x} F(y) = \ell$;

(S_1) *for all sequences* $\{y_j\}_{j\in\mathbb{N}} \subseteq X\setminus\{x\}$ *such that* $y_j \to x$ *and* $\{F(y_j)\}_{j\in\mathbb{N}}$ *has a limit, we have* $\lim_{j\to\infty} F(y_j) \leq \ell$, *and there exists a sequence* $\bar{y}_i \to \bar{x}$ *such that equality holds.*

PROOF. We prove only the equivalence of (I_1) and (I_2). The corresponding results for (S_1) and (S_2) can be deduced by applying the results for (I_1) and (I_2) to the function $G = -F$, since

$$\liminf_{y\to\bar{x}} G(y) := \sup_{\substack{V \text{ open}\\ \bar{x}\in V}} \inf_{y\in V\setminus\{\bar{x}\}} G(y) = -\inf_{\substack{V \text{ open}\\ \bar{x}\in V}} \sup_{y\in V\setminus\{\bar{x}\}} F(y) = -\limsup_{y\to x} F(y),$$

$(I_1 \implies I_2)$ For any $i\in\mathbb{N}$ there exists an open neighborhood U_i of x such that

$$\sup_{\substack{V \text{ open}\\ x\in V}} \inf_{y\in V\setminus\{x\}} F(y) \leq \inf_{y\in U_i\setminus\{x\}} F(y) + \frac{1}{2^i},$$

moreover, there exists $r_i > 0$ such that $B(x,r_i)\subseteq U_i$. Without loss of generality, we may assume also that if $i>0$ we have $0 < r_i < \min\{1/2^i, r_{i-1}\}$. We have

$$\sup_{\substack{V \text{ open}\\ x\in V}} \inf_{y\in V\setminus\{x\}} F(y) \leq \inf_{y\in U_i\setminus\{x\}} F(y) + \frac{1}{2^i} \leq \inf_{y\in B(x,r_i)\setminus\{x\}} F(y) + \frac{1}{2^i}.$$

Let $\{y_j\}_{j\in\mathbb{N}} \subseteq X\setminus\{x\}$ be such that $y_j\to x$ and $\{F(y_j)\}_{j\in\mathbb{N}}$ converges. For every $i\in\mathbb{N}$ we have that there exists $j_i\in\mathbb{N}$ such that $y_j \in B(x,r_i)$ for all $i \geq j_u$ since $y_j \to x$. This implies that

$$\ell = \sup_{\substack{V \text{ open}\\ x\in V}} \inf_{y\in V\setminus\{x\}} F(y) \leq F(y_j) + \frac{1}{2^i}, \text{ for all } j \geq j_i.$$

By letting $j\to\infty$ and then $i\to+\infty$ we obtain the first part of (2).

Choose $\bar{y}_i \in B(x,r_i)\setminus\{\bar{x}\}$ such that

$$\inf_{y\in B(x,r_i)\setminus\{\bar{x}\}} F(y) \leq F(\bar{y}_i) \leq \inf_{y\in B(x,r_i)\setminus\{\bar{x}\}} F(y) + \frac{1}{2^i}.$$

Thus we have

$$\begin{aligned} F(\bar{y}_i) - \frac{1}{2^i} &\leq \inf_{y\in B(\bar{x},r_i)\setminus\{\bar{x}\}} F(y) \leq \sup_{\substack{V \text{ open}\\ \bar{x}\in V}} \inf_{y\in V\setminus\{\bar{x}\}} F(y) = \ell \\ &\leq \inf_{y\in B(x,r_i)\setminus\{x\}} F(y) + \frac{1}{2^i} \leq F(\bar{y}_i) + \frac{1}{2^i}. \end{aligned}$$

We conclude that:

(a) since $0 < r_i < 1/2^i$ for $i > 0$, we have $r_i \to 0^+$ and so $\bar{y}_i \to x$;

(b) since $r_i < r_{i-1}$ for $i>0$, we have that $\left\{ \inf_{y\in B(\bar{x},r_i)\setminus\{x\}} F(y) \right\}_{i\in\mathbb{N}\setminus\{0\}}$ is an increasing sequence (the infimum is made on shrinking sets), thus it admits a limit that we denote by ℓ';

(c) we have that $F(y_i)\to\ell'$ as $y\to\infty$ by the choice of y_i;

(d) we have $\ell' = \ell$ for the above chain of inequalities.

$(I_2 \implies I_1)$ If $\{y_i\}$ is any sequence such that $y_i\to x$ and $\{F(y_i)\}_{i\in\mathbb{N}}$ has a limit, we have that for every open neighborhood V of x there exists $j_V\in\mathbb{N}$ such that $y_i\in V$ for all $i > j_V$, hence

$$\inf_{y\in V\setminus\{x\}} F(y) \leq F(y_i), \qquad i > j_V.$$

By letting $i \to +\infty$, we obtain

$$\inf_{y \in V \setminus \{x\}} F(y) \leq \lim_{i \to \infty} F(y_i),$$

and taking the sup on V we have

$$\liminf_{y \to x} F(y) \leq \lim_{i \to \infty} F(y_i),$$

for every sequence $\{y_i\}_{i \in \mathbb{R}}$ such that $y_i \to x$ and $\{F(y_i)\}_{i \in \mathbb{N}}$ has a limit. In particular, we have that

$$\liminf_{y \to x} F(y) \leq \lim_{i \to \infty} F(\bar{y}_i) = \ell,$$

Let $\{r_i\}_{i \in \mathbb{N}} \subseteq [0, +\infty[$ a monotone decreasing sequence. Then $\left\{ \inf_{y \in B(\bar{x}, r_i) \setminus \{x\}} F(y) \right\}_{i \in \mathbb{N}}$ is an increasing sequence, thus it admits a limit ℓ'. Choose $\bar{y}_i \in B(x, r_i) \setminus \{x\}$ such that

$$\inf_{y \in B(\bar{x}, r_i) \setminus \{x\}} F(y) \leq F(\bar{y}_i) \leq \inf_{y \in B(x, r_i) \setminus \{x\}} F(y) + \frac{1}{2^i},$$

and notice that $\bar{y}_i \to x$ and $F(\bar{y}_i) \to \ell'$, thus by assumption we must have $\ell \leq \ell'$. On the other hand, we must have

$$\sup_{\substack{V \text{ open} \\ x \in V}} \inf_{y \in V \setminus \{x\}} F(y) \geq \inf_{y \in B(x, r_i) \setminus \{x\}} F(y),$$

and letting $i \to \infty$ we obtain

$$\ell \geq \sup_{\substack{V \text{ open} \\ x \in V}} \inf_{y \in V \setminus \{x\}} F(y) \geq \ell',$$

so $\ell = \ell'$ and the proof is completed.

$\square$

DEFINITION 1.12 (Semicontinuity). Let X be a topological Hausdorff space, $F : X \to [-\infty, +\infty]$. We say that F is *lower semicontinuous* (shortly l.s.c.) if for every $x \in X$ it holds

$$\liminf_{y \to x} F(y) \geq F(x),$$

where

$$\liminf_{y \to x} F(y) = \sup_{\substack{V \text{ open} \\ x \in V}} \inf_{y \in V \setminus \{x\}} F(y).$$

Symmetrically, we say that F is *upper semicontinuous* (shortly u.s.c.) if for every $x \in X$ it holds

$$\limsup_{y \to x} F(y) \leq F(x),$$

where

$$\limsup_{y \to x} F(y) = \inf_{\substack{V \text{ open} \\ x \in V}} \sup_{y \in V \setminus \{x\}} F(y).$$

We will give now another characterization of semicontinuous functions.

DEFINITION 1.13. Let X be a topological Hausdorff space, $F : X \to [-\infty, +\infty]$. Define

$$\begin{aligned} \operatorname{dom} F &:= \{x \in X : F(x) \in \mathbb{R}\}, \text{the } \textit{domain} \text{ of } F; \\ \operatorname{epi} F &:= \{(x, \alpha) \in X \times \mathbb{R} : \alpha \geq F(x)\}, \text{the } \textit{epigraph} \text{ of } F; \\ \operatorname{hypo} F &:= \{(x, \beta) \in X \times \mathbb{R} : \beta \leq F(x)\}, \text{the } \textit{hypograph} \text{ of } F. \end{aligned}$$

LEMMA 1.14. *Let X be a topological vector space, $F : X \to [-\infty, +\infty]$. Then*

(1) *F is l.s.c. if and only* epi F *is closed in $X \times \mathbb{R}$ (endowed with product topology);*
(2) *F is u.s.c. if and only* hypo F *is closed in $X \times \mathbb{R}$ (endowed with product topology);*
(3) *F is continuous if and only if it is both l.s.c. and u.s.c.*

PROOF. Assume that F is l.s.c. and want to prove that $\operatorname{epi} F$ is closed. Let $(x,\alpha) \notin \operatorname{epi} F$, in particular we have

$$\alpha < F(x) \leq \liminf_{y\to x} F(y) = \sup_{\substack{V \text{ open} \\ x\in V}} \inf_{y\in V\setminus\{x\}} F(y).$$

Take $\varepsilon > 0$ such that $\alpha + \varepsilon/2 < F(x)$ and let V_ε be open such that $x \in V_\varepsilon$ and

$$\sup_{\substack{V \text{ open} \\ x\in V}} \inf_{y\in V\setminus\{x\}} F(y) \leq \inf_{y\in V_\varepsilon\setminus\{x\}} F(y) + \frac{\varepsilon}{4},$$

hence

$$\alpha + \frac{\varepsilon}{2} < F(x) \leq \inf_{y\in V_\varepsilon\setminus\{x\}} F(y) + \frac{\varepsilon}{4},$$

so $\alpha + \varepsilon/4 < F(y)$ for all $y \in V_\varepsilon$, hence $V_\varepsilon \times]-\infty, \alpha+\varepsilon/4[$ is an open neighborhood of (x,α) that has empty intersection with $\operatorname{epi} F$, hence the complement set of $\operatorname{epi} F$ is open, thus $\operatorname{epi} F$ is closed.

Assume now that $\operatorname{epi} F$ is closed in $X \times \mathbb{R}$ and want to prove that F is l.s.c. Given $(x,\alpha) \notin \operatorname{epi} F$ we have that there exists an open neighborhood V of x and $\varepsilon > 0$ such that $V \times B(\alpha,\varepsilon)$ has empty intersection with $\operatorname{epi} F$, since the complement of has $\operatorname{epi} F$ is open. In particular, we have that $F(y) > \alpha + \varepsilon$ for every $y \in V$, hence

$$\inf_{y\in V\setminus\{x\}} F(y) \geq \alpha + \varepsilon,$$

and so by passing to the sup on V

$$\liminf_{y\to x} F(y) \geq \alpha + \varepsilon.$$

This holds for every $\alpha < F(x)$ and for every $\varepsilon > 0$ sufficiently small, hence by letting $\varepsilon > 0$ and $\alpha \to F(x)^-$ we have

$$\liminf_{y\to x} F(y) \geq F(x),$$

thus F is l.s.c.

The statement on upper semicontinuity can be proved with an analogous argument, and it is left to the reader.

Assume that F is continuous, then is trivially l.s.c. and u.s.c. Conversely, if we assume that F is both l.s.c. and u.s.c. we have

$$\limsup_{y\to x} F(y) \leq F(x) \leq \liminf_{y\to x} F(y),$$

which implies

$$F(x) = \limsup_{y\to x} F(y) = \liminf_{y\to x} F(y) = \lim_{y\to x} F(y),$$

hence F is continuous. □

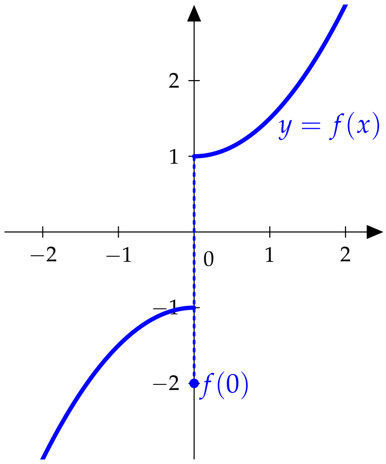
EXAMPLE 1.15. A lower semicontinuous function. Since

$$\lim_{x\to 0^\pm} f(x) = \pm 1,$$

we have that

$$\liminf_{x\to 0} f(x) = -1 > -2 = f(0).$$

All the points $(0, y)$ with $y \geq -2$ belong to $\operatorname{epi} f$, which then is closed in $\mathbb{R} \times \mathbb{R}$.



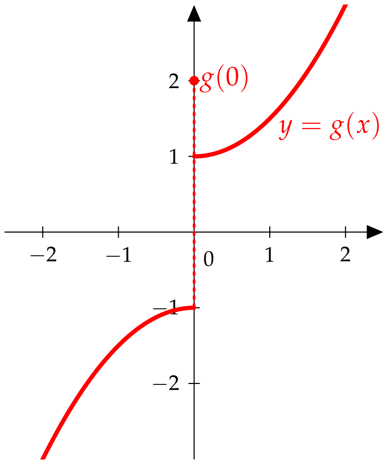
EXAMPLE 1.16. An upper semicontinuous function. Since

$$\lim_{x\to 0^\pm} f(x) = \pm 1,$$

we have

$$\limsup_{x\to 0} f(x) = 1 < 2 = g(0).$$

All the points $(0, y)$ with $y \leq 2$ belong to $\operatorname{hypo} g$ which then is closed in $\mathbb{R} \times \mathbb{R}$.



After having discussed the weakened notion of continuity, we pass to examine compactness assumption. Even in the first example (minimization on $\mathbb{R}$) the space X was not compact a priori, however the fact that $\lim\limits_{|x|\to\infty} F(x) = +\infty$, toghether with the fact that F was not identically $+\infty$ leaded us to say that every minimizing sequence indeed belongs to a suitable compact set. Thus the following definition is natural:

DEFINITION 1.17 (Coercivity). Let X be a metric space, $F : X \to [-\infty, +\infty]$. We say that F is *coercive* if for every $t \in \mathbb{R}$ there exists $K(t) \subset X$, with $K(t)$ compact subset of X, such that

$$\{x \in X : F(x) \leq t\} \subseteq K(t).$$

We notice that if X is compact, coercivity property is trivial, since we can choose $K(t) = X$ for every $t \in \mathbb{R}$.

We are ready now to state the weakened version of Weierstrass Theorem.

THEOREM 1.18 (Tonelli-Weierstrass). *Let X be a metric space, $F : X \to [-\infty, +\infty]$. Assume that F is lower semicontinuous and coercive. Then F admits at least a minimum point in X.*

PROOF. If $F(x) = +\infty$ for every $x \in X$ then there is nothing to prove, since every $x \in X$ is a minimum point. So we can assume that $\inf\limits_{x\in X} F(x) = m < +\infty$. Let $t > m$ and by assumption we can consider the compact set $K(t)$ containing $\{x \in X : F(x) \leq t\}$. We have

$$\inf_{x\in X} F(x) = \inf_{x\in K(t)} F(x).$$

Let now $\{x_n\}_{n\in\mathbb{N}}$ be a minimizing sequence. For sufficiently large n we have $F(x_n) < t$ thus $x_n \in K(t)$ which is compact. So there exists a subsequence $\{x_{n_k}\}_{k\in\mathbb{N}}$ converging to $\bar{x} \in K(t)$. By lower semicontinuity:

$$F(\bar{x}) \leq \liminf_{k\to\infty} F(x_{n_k}) = \lim_{n\to\infty} F(x_n) = m,$$

and this implies $F(\bar{x}) = m$ so $\bar{x}$ is a minimum point for F in $K(t)$ and in X. $\square$

LEMMA 1.19. *Let* $X = \mathbb{R}^n$, $F : X \to [-\infty, +\infty]$. *Then* F *is coercive if and only if* $\lim_{\|x\|\to\infty} F(x) = +\infty$.

PROOF. Fix $t \in \mathbb{R}$. By assumption, there exists $M = M(t) > 0$ such that if $\|x\| > M(t)$ then $F(x) > t$. But this implies

$$\{x \in X : F(x) \le t\} \subseteq \{x \in X : \|x\| \le M(t)\}$$

Since we are in $\mathbb{R}^n$, the set $K(t) := \overline{B(0, M(t))}$ is compact, so F is coercive.

Conversely, assume that F is coercive. In that case, if $\{x_n\}_{n\in\mathbb{N}}$ is a sequence with $\|x_n\| \to \infty$ and such that $\lim_{n\to\infty} F(x_n)$ exists, we must necessarily have $\lim_{n\to\infty} F(x_n) = +\infty$. In fact, if by contradiction $\lim_{n\to\infty} F(x_n) \le \ell \in \mathbb{R}$, we should have $x_n \in K(\ell + 1)$, for n sufficiently large. But in this case, since $K(\ell + 1)$ is compact thus bounded, we could not have $\|x_n\| \to \infty$. So for every sequence such that $\|x_n\| \to \infty$ and $\lim_{n\to\infty} F(x_n)$ exists, necessarily we have $\lim_{n\to\infty} F(x_n) = +\infty$, thus

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

but this implies

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

$\square$

Summarizing, we have obtained the following fact:

lower semicontinuity + **coercivity** = **existence of points of minimum**.

EXERCISE 1.20. We will prove Tonelli-Weierstrass Theorem without assuming that X is a metric space.

(1) Prove the following topological version of Weierstrass theorem: let $f : Z \to Y$ be a continuous map between two topological spaces Z and Y. If Z is compact then $f(Z)$ is compact in Y.
(2) Define $\mathcal{T} := \{]a, +\infty] : a \in \mathbb{R} \cup \{-\infty\}\} \cup \{[-\infty, +\infty], \emptyset\}$. Prove that $\mathcal{T}$ is a topology on $\mathbb{R} \cup \{\pm\infty\}$.
(3) Let X be a topological space, $f : X \to \mathbb{R} \cup \{\pm\infty\}$ be a map, and $x_0 \in X$. Prove that f is continuous at x_0 when we endow $\mathbb{R} \cup \{\pm\infty\}$ with the topology $\mathcal{T}$ if and only if f is lower semicontinuous at x_0 (w.r.t. the usual topology on $\mathbb{R} \cup \{\pm\infty\}$)
(4) Prove that $K \subseteq \mathbb{R} \cup \{\pm\infty\}$ is compact for the topology $\mathcal{T}$ if and only if $\inf K \in K$.
(5) Conclude the proof of the theorem.

SOLUTION.

(1) Let $\{V_\alpha\}_{\alpha\in I}$ be an open covering of $f(Z)$. Set
$$U_\alpha = f^{-1}(V_\alpha) = \{x \in Z : f(z) \in V_\alpha\}.$$
Since f is continuous, U_α is open. Given $x \in Z$, there exists $\alpha_x \in I$ such that $f(x) \in V_{\alpha_x}$ since
$$f(x) \in f(Z) \subseteq \bigcup_{\alpha\in I} V_\alpha,$$
thus $x \in U_{\alpha_x}$. In particular, we have that $\{U_\alpha\}_{\alpha\in I}$ is an open covering of the compact set Z. So there exist $N \in \mathbb{N} \setminus \{0\}$, and $\alpha_1, \dots, \alpha_N \in I$ such that
$$Z = \bigcup_{i=1}^{N} U_{\alpha_i},$$
and so
$$f(Z) = f\left(\bigcup_{i=1}^{N} U_{\alpha_i}\right) = \bigcup_{i=1}^{N} f(U_{\alpha_i}) = \bigcup_{i=1}^{N} V_{\alpha_i},$$
thus from the open covering $\{V_\alpha\}_{\alpha\in I}$ of $f(Z)$ we have extracted a finite subcovering, and so $f(Z)$ is compact.

(2) We have $\mathbb{R}\cup\{\pm\infty\}, \emptyset \in \mathcal{T}$. Given a family of elements of $\mathcal{T}$ if at least one of them coincides with $[-\infty,+\infty]$, then the union is $[-\infty,+\infty] \in \mathcal{T}$. If all the elements of the family are $\emptyset$, then the union is $\emptyset \in \mathcal{T}$. In the remaining cases, we have that the union of all the elements of the family coincides with $\bigcup_{\alpha\in I}]r_\alpha,+\infty]$ for a family $\{r_\alpha\}_{\alpha\in I} \subseteq \mathbb{R}\cup\{-\infty\}$. If $r \in \bigcup_{\alpha\in I}]r_\alpha,+\infty]$ we have that there exists $\hat{\alpha} \in I$ such that $r > r_{\hat{\alpha}}$, and so $r > \inf_{\alpha\in I} r_\alpha$. Conversely, let $r > \inf_{\alpha\in I} r_\alpha$. Then there exists $\hat{\alpha} \in I$ such that $r > r_{\hat{\alpha}}$, in particular

$$\bigcup_{\alpha\in I}]r_\alpha,+\infty] = \left]\inf_{\alpha\in I} r_\alpha,+\infty\right] \in \mathcal{T},$$

since $\inf_{\alpha\in I} r_\alpha \in \mathbb{R}\cup\{-\infty\}$. It is trivial to verify that a finite intersection of element of $\mathcal{T}$ belongs again to $\mathcal{T}$.

(3) Suppose that f is l.s.c. at x_0. We prove that the inverse image of every open neighborhood of $f(x_0)$ w.r.t. the topology $\mathcal{T}$ is a neighborhood of x_0 in X. An open neighborhood of $f(x_0)$ is of the form $W_r =]f(x_0)-r,+\infty]$ for a certain $r \in \mathbb{R}$. Since

$$\sup_{\substack{V \text{ open}\\ x_0\in V}} \inf_{y\in V\setminus\{x\}} f(y) \geq f(x),$$

there exists an open neighborhood V_r of x_0 such that

$$\inf_{y\in V_r\setminus\{x_0\}} f(y) + \frac{r}{2} \geq \sup_{\substack{V \text{ open}\\ x_0\in V}} \inf_{y\in V\setminus\{x_0\}} f(y) \geq f(x),$$

and so for all $y \in V_r$ we obtain

$$f(y) \geq f(x) - \frac{r}{2} > f(x) - r,$$

thus for any W_r there exists a neighborhood V_r of x_0 such that $f(y) \in W_r$ for all $y \in V_r$.

Conversely, assume that the inverse image of every open neighborhood of $f(x_0)$ w.r.t. the topology $\mathcal{T}$ is a neighborhood of x_0 in X. In particular, for any $r > 0$ the set $\{y : f(y) > f(x_0) - r\} = f^{-1}(]f(x_0)-r,+\infty])$ is a neighborhood of x_0 in X. In particular, there exists an open neighborhood $\hat{V}$ of x_0 such that $f(y) > f(x_0) - r$ for all $y \in \hat{V}$. So for all $r > 0$,

$$\sup_{\substack{V \text{ open}\\ x_0\in V}} \inf_{y\in V\setminus\{x_0\}} f(y) \geq \inf_{y\in \hat{V}\setminus\{x_0\}} f(y) \geq f(x_0) - r$$

which, by letting $r \to 0^+$, implies

$$\liminf_{y\to x} f(y) = \sup_{\substack{V \text{ open}\\ x_0\in V}} \inf_{y\in V\setminus\{x_0\}} f(y) \geq f(x_0),$$

i.e., f is l.s.c. at x_0.

(4) Let K satisfying $\inf K \in K$ and take an open covering of K. If $[-\infty,+\infty]$ belongs to the covering, then trivially we can extract the finite subcovering consisting only on $[-\infty,+\infty]$ to cover K. Otherwise, any union of sets of the form $]r_\alpha,+\infty]$, $\alpha \in I$, cannot cover K unless there exists $\hat{\alpha} \in I$ such that $r_{\hat{\alpha}} < \inf K$. In this case, we have that $K \subseteq]r_{\hat{\alpha}},+\infty]$, thus we have extract a finite subcovering of K. Conversely, assume that $\inf K \notin K$, and take a sequence $\{x_n\}_{n\in\mathbb{N}}$, $x_n \to \inf F$ with $x_n > k$. We have

$$\bigcup_{n\in\mathbb{N}}]x_n,+\infty] \supseteq K,$$

but we cannot extract a finite subcovering of K from $]x_n,+\infty]$, $x \in \mathbb{N}$.

(5) Let $f : X \to \mathbb{R} \cup \{\pm\infty\}$ be l.s.c. and coercive. If $\inf F(x) = +\infty$, then the minimum is $+\infty$ and it is achieved at every $x \in X$. Otherwise, given $m \in \mathbb{R}$, $m > \inf F(x)$, there exists a compact $K \subseteq X$ such that $F_{|K} \leq m$ and $\inf_{x\in K} F(x) = \inf_{x\in X} F(x)$. Endowing $\mathbb{R} \cup \{\pm\infty\}$ with the topology $\mathcal{T}$, we have that F is continuous, and so $F(K)$ is compact. In particular, $\inf F(K) \in F(K)$, thus there exists $\bar{x} \in K$ such that $F(\bar{x}) = \inf F(K)$, and so $\bar{x}$ is a point of minimum for F on X.

Summary of Lecture 1

- Given a nonempty set X and a map $F : X \to \mathbb{R} \cup \{\pm\infty\}$ we consider the *basic minimization problem* $\inf_{x\in X} F(x)$.
- The main features of the problem are:
  - existence of solutions, i.e. existence of minimizers $\bar{x}$ for F with $F(\bar{x}) \in \mathbb{R}$,
  - necessary conditions, i.e. properties that a minimizer satisfies,
  - sufficient conditions, i.e. properties that, if satisfied, give a minimizer.
- Perron's paradox taught us that using necessary conditions without an existence result can be quite dangerous.
- We notice that *a priori* there is no topology given on X.
- Without a topology on X, we speak only about *absolute* minimizers, which are *independent from* the topology. *Local* minimizers are instead related to a topology.
- We are free to choose *any* topology on X in order to study the problem, in particular to have existence.
- A topological criterion yielding existence is the *Weierstrass theorem*: if X is compact and F is continuous then F admits maximum and minimum.
- This overexceeds our needs: we want to *relax* the assumptions of Weierstrass theorem, but still *retaining* the existence of the minimum.
- To this aim we introduce
  - the concept of *lower semicontinuity*, which is a weakened form of continuity. A function is l.s.c. iff it has closed graph.
  - the concept of *coercivity*, i.e. for each sublevel there is compact set containing it.
- We ended up with the *Tonelli-Weierstrass theorem*: if F is l.s.c. and coercive, then F admits minimum on X.

2. Lecture of 5 october 2018: Weak topologies, convex sets and convex functions (3h)

During the previous lecture, we discussed the problem of the **existence** of minima, obtaining a more general version of Weierstrass Theorem suitable for our purpouses:

lower semicontinuity + coercivity = existence of points of minimum.

When we study the general minimization problem $\inf_{x\in X} F(x)$, is *not* given a priori any topology on X. To determine a suitable topology on X to have existence of solutions is a part of the problem.

L.s.c. functions enjoy intresting stability properties:

Lemma 2.1. *Let $\{F_i\}_{i\in I}$ with $F_i : X \to [-\infty, +\infty]$ for every $I \in I$ be a family of l.s.c. functions, let $F_1, F_2 : X \to [-\infty, +\infty]$ be l.s.c. functions, and let $\lambda \geq 0$. Then:*

(1) *$\lambda F_1 + F_2$ is l.s.c.;*
(2) *Defined $F(x) := \sup_{i\in I} F_i(x)$, we have that $F : X \to [-\infty, +\infty]$ and F is l.s.c.*

Proof. The first statement follows from the properties of the limits:

$$\lambda F_1(x) + F_2(x) \leq \lambda \liminf_{y\to x} F_1(y) + \liminf_{y\to x} F_2(y) \leq \liminf_{y\to x}(\lambda F_1(y) + F_2(y)),$$

For the second, we have

$$\operatorname{epi} F = \bigcap_{i\in I} \operatorname{epi} F_i,$$

since by assumption epi F_i is closed for all $i \in I$, we have that epi F is also closed, thus F is l.s.c. □

We notice that:

(1) a very strong topology on X make easier to obtain lower semicontinuity, but more difficult to have compactness and then coercivity;
(2) a very weak topology on X make more difficult to have lower semicontinuity, but easier to have compactness and then coercivity.

EXAMPLE 2.2. Let $g \in X := L^2(0,1)$ be fixed. We are intrested in the following minimization problem:

$$\inf_{u \in X} F_g(u),$$

where $F_g : L^2(0,1) \to \mathbb{R}$ is defined by

$$F_g(u) := \int_0^1 |u(t)|^2\,dt - 2\int_0^1 g(t)u(t)\,dt.$$

We notice that, indeed, the problem is trivial: in fact we have

$$\begin{aligned}
F_g(u) &= \int_0^1 |u(t)|^2\,dt - 2\int_0^1 g(t)u(t)\,dt + \int_0^1 |g(t)|^2\,dt - \int_0^1 |g(t)|^2\,dt \\
&= \int_0^1 \left(|u(t)|^2 - 2g(t)u(t) + |g(t)|^2\right)\,dt - \int_0^1 |g(t)|^2\,dt \\
&= \int_0^1 (u(t)-g(t))^2\,dt - \int_0^1 |g(t)|^2\,dt \geq -\int_0^1 |g(t)|^2\,dt
\end{aligned}$$

and, on the other hand,

$$F_g(g) = -\int_0^1 |g(t)|^2\,dt,$$

and so we conclude that g is a minimum point in X for F_g and the value at the minimum is $F_g(g)$.

We are now intrested in testing the assumptions of the generalized Weierstrass Theorem in the above simple situation. Suppose to endow $X = L^2(0,1)$ with the strong topology of the norm $\|\cdot\|_{L^2}$ coming from the scalar product of L^2. Then we have

$$F_g(u) := \|u\|_{L^2}^2 - 2\langle g, u\rangle_{L^2}.$$

Let us check the assumption of the Tonelli-Weierstrass Theorem:

(1) The functional is continuous, indeed, given a sequence $\{u_n\}_{n\in\mathbb{N}}$ in $L^2(0,1)$ converging to $u \in L^2(0,1)$ according to the norm of $L^2(0,1)$, by Schwarz's inequality we have:

$$\begin{aligned}
|F_g(u_n) - F_g(u)| &= \left| \|u_n\|_{L^2}^2 - \|u\|_{L^2}^2 - 2\langle g, u_n - u\rangle_{L^2} \right| \\
&\leq \left| \|u_n\|_{L^2}^2 - \|u\|_{L^2}^2 \right| + 2\,|\langle g, u_n - u\rangle_{L^2}| \\
&\leq |\langle u_n - u, u_n + u\rangle_{L^2}| + \|g\|_{L^2}\|u_n - u\|_{L^2} \\
&\leq \|u_n - u\|_{L^2}\|u_n + u\|_{L^2}| + \|g\|_{L^2}\|u_n - u\|_{L^2} \\
&\leq \|u_n - u\|_{L^2}\left(\|u_n + u\|_{L^2} + \|g\|_{L^2}\right) \to 0, \text{ per } n \to \infty,
\end{aligned}$$

since the term in brackets is bounded. Being continuous, in particular the functional is lower semicontinuous.

(2) We check now the coercivity property. We preliminary observe that

$$F_g(u) \leq \|u\|_{L^2}^2 + 2\|u\|_{L^2}\|g\|_{L^2}.$$

Let us consider the sequence of functions $\{u_k(t) := \sin(2\pi k t)\}_{k\in\mathbb{N}}$. We have

$$\|u_k\|_{L^2}^2 = \int_0^1 |\sin(2\pi k t)|^2\,dt = \frac{1}{2\pi}\int_0^{2\pi} \sin^2(kx)\,dx = \frac{1}{2},$$

thus

$$F_g(u_k) \leq \frac{1}{2} + \sqrt{2}\|g\|_{L^2} =: M.$$

If F_g was coercive, it should exist a compact $K = K(M)$ such that $u_k \in K(M)$ for every $k \in \mathbb{N}$. But in this case it would be possible to extract from $\{u_k\}_{k\in\mathbb{N}}$ a subsequence $\{u_{k_j}\}_{j\in\mathbb{N}}$ converging in L^2 norm to a function $\bar{u} \in K(M)$. According to Riemann-Lebesgue's Lemma, we have for every $\varphi \in L^2(0,1)$

$$\lim_{j\to\infty} \langle u_{k_j}, \varphi \rangle = \lim_{j\to\infty} \int_0^1 \varphi(x) \sin(2\pi k_j x)\, dx = 0,$$

and in particular:

$$\lim_{j\to\infty} \langle u_{k_j}, \bar{u} \rangle = 0,$$

but we should have also

$$\lim_{j\to\infty} \langle u_{k_j}, \bar{u} \rangle = \langle \bar{u}, \bar{u} \rangle_{L^2} = \|\bar{u}\|^2_{L^2},$$

thus $\bar{u} = 0$. But this leads to the following contradiction

$$\frac{1}{2} = \|u_{k_j}\|_{L^2} \to \|\bar{u}\|_{L^2} = 0.$$

Thus the functional is not coercive.

We recall now from the previous courses in calculus the following definitions:

DEFINITION 2.3 (Topological dual). Let E be a Banach space, we denote by E' its *topological dual*, i.e.,

$$E' := \{f : E \to \mathbb{R} : f \text{ is linear and continuous}\}$$

The space E' is a Banach space endowed with the following norm:

$$\|f\|_{E'} := \sup_{x\in E\setminus\{0\}} \frac{|f(x)|}{\|x\|_E}.$$

We recall that if $\dim E < +\infty$ then *every* linear map is also automatically *continuous*. This fact is no longer true when E has infinite dimension.

DEFINITION 2.4 (weak topology). The *weak topology* $\sigma(E,E')$ on E is the weaker topology on E (i.e. with minimal number of open sets) such that allthe elements of E' are continuous. Given a sequence $\{x_n\}_{n\in\mathbb{N}}$ in E and $x \in E$ we say that x_n *weakly converges* to x and write $x_n \rightharpoonup x$ or $x_n \to x$ in $\sigma(E,E')$ if

$$\langle f, x_n \rangle_{E',E} \to \langle f, x \rangle \text{ in } \mathbb{R}, \qquad \text{for all } f \in E',$$

where we denote with $\langle f, x_n \rangle_{E',E} = f(x_n) \in \mathbb{R}$ the evaluation of f at x_n.

THEOREM 2.5 (properties of the weak topology). *The weak topology enjoys the following properties:*

1. *The weak topology is Hausdorff (equivalently, if the weak limit exists, then it is unique).*
2. *Given $x_0 \in E$, a basis of neighborhoods of x_0 for $\sigma(E,E')$ is given by finite intersection of sets of the form*
$$V := \{x \in E : |\langle f, x - x_0 \rangle_{E',E}| < \varepsilon\}$$
where $\varepsilon > 0$, $f \in E'$.
3. *If $x_n \to x$ strongly (i.e., according to the norm of E, equivalently if $\|x_n - x\|_E \to 0$) then $x_n \rightharpoonup x$, the converse in general does not hold.*
4. *If $x_n \rightharpoonup x$ then $\|x_n\|_E$ is bounded and $\|x\|_E \le \liminf_{n\to\infty} \|x_n\|_E$.*
5. *If $x_n \rightharpoonup x$ and $\|f_n - f\|_{E'} \to 0$ then $f_n(x_n) \to f(x)$.*
6. *If E is finite-dimensional, the weak and the strong topology coincide on E, otherwise the weak topology is strictly weaker.*

PROOF. See Propositions III.3, III.4, III.5, III.6 at pp.52–54 of [**3**]. □

DEFINITION 2.6 (Convex sets). Let X be a vector space on $\mathbb{R}$ or $\mathbb{C}$, $F \subseteq X$ a nonempty set. We say that F is *convex* if for every $\lambda \in [0,1]$, $x, y \in F$ we have $\lambda x + (1-\lambda)x \in F$.

It is immediate to prove that a nonempty arbitrary intersection of convex sets is convex, thus given a nonempty set S, is possible to consider the intersection of all the subsets of X containing it. This intersection, called $\mathrm{co}(S)$, is the smaller convex set (in the sense of inclusion) containing S and is named the *convexification* or *convex hull* of S. If holds:

$$\mathrm{co}(S) := \left\{ \sum_{j=0}^{n} \lambda_j x_j : n \in \mathbb{N},\, \lambda_j \in [0,1], x_j \in S \text{ for all } 0 \leq j \leq n,\, \sum_{j=0}^{n} \lambda_j = 1 \right\},$$

set of all *finite* convex combination of elements of S.

Since the intersection of an arbitrary family of closed sets is closed, is possibile to dermine also the smallest closed convex set containing S, which is the intersection of all the closed convex sets containing S and coincides with $\overline{\mathrm{co}}(S)$, i.e., the closure of the convex hull.

Attention: the closure of the convex hull in general *does not coincide* with the convex hull of the closure!

EXAMPLE 2.7. Consider a separable Hilbert space with orthonormal basis $S = \{e_i\}_{i\in\mathbb{N}}$. Given two elements $e_i, e_j \in S$ with $i \neq j$ we have $\|e_i - e_j\| = \sqrt{2}$, so $S = \bar{S}$ is closed. Thus $\mathrm{co}(\bar{S})$ is the set of *finite* convex combinations of elements of S. We show now that $0 \notin \mathrm{co}(\bar{S})$. Indeed, since all the elements of $S = \bar{S}$ are linearly independent, we have

$$0 = \sum_{i=1}^{N} \lambda_i e_i$$

if and only if $\lambda_i = 0$ for all i, but in this case $0 \notin \mathrm{co}(\bar{S})$ because we must have $\lambda_i \geq 0$ and $\sum \lambda_i = 1$ so at least one of the λ_i must be nonzero.

However, it is well known that $e_i \rightharpoonup 0$ for $i \to \infty$, in particular $0 \in \overline{\mathrm{co}}S$ (since e_i can be viewed as a convex combination made of a single element) because strong and weak topology coincides on convex sets.

EXAMPLE 2.8. Set $X := L^2(-\pi, \pi)$ with the strong topology. Consider

$$S := \left\{ \frac{1}{\sqrt{2\pi}}, \frac{\cos(kx)}{\sqrt{\pi}}, \frac{\sin(kx)}{\sqrt{\pi}}, k \in \mathbb{N} \setminus \{0\} \right\}$$

Given $f_1, f_2 \in S$, $f_1 \neq f_2$, we have $\|f_1 - f_2\| > \sqrt{2}$, so $S = \bar{S}$ and it is closed. Thus $\mathrm{co}(\bar{S})$ is the set of *finite* convex combinations of elements of S, in particular $\mathrm{co}(\bar{S}) \subseteq C^\infty(]-\pi, \pi[)$. On the other hand, for every $N \in \mathbb{N}$, consider the function $f_N : [-\pi, \pi] \to \mathbb{R}$ defined by

$$f_N(x) := \frac{\displaystyle\sum_{n=1}^{N} \frac{\sin(nx)}{\sqrt{\pi} n^2}}{\displaystyle\sum_{n=1}^{N} \frac{1}{n^2}}.$$

Such function belongs to $\mathrm{co}(S)$ for every $n \in \mathbb{N}$ and converges in L^2 to

$$f_\infty(x) := \frac{6}{\pi^{3/2}} \sum_{n=1}^{\infty} \frac{\sin(nx)}{n^2}.$$

The sequence of the first derivatives converges in L^2 to

$$f'_\infty(x) := \frac{6}{\pi^{3/2}} \sum_{n=1}^{\infty} \frac{\cos(nx)}{n}.$$

but such function is not continuous in $x = 0$, thus $f_\infty \in \overline{\mathrm{co}}(S) \setminus \mathrm{co}(\bar{S})$.

DEFINITION 2.9. Let E be a vector space, A, B nonempty subsets of E, $f : E \to \mathbb{R}$ be a linear function not identically 0, and $\alpha \in \mathbb{R}$. We say that

(1) H is the hyperplane of equation $f = \alpha$ if

$$H := \{x \in E : f(x) = \alpha\};$$

(2) H separates A and B in the weak sense if $f(a) \leq \alpha$ and $f(b) \geq \alpha$ for every $a \in A$, $b \in B$;
(3) H separates A and B in the strict sense if there exists $\varepsilon > 0$ with $f(a) \leq \alpha - \varepsilon$ and $f(b) \geq \alpha + \varepsilon$ for every $a \in A$, $b \in B$.

If E is a normed space, then H is closed in E if and only if f is continuous.

We recall now the geometric form of Hahn-Banach's Theorem:

THEOREM 2.10 (Hahn-Banach, geometric form). *Let E be a normed vector space, $A, B \subseteq E$ convex, nonempty, disjoint.*

(1) *If A is open, then there exists a closed hyperplane separating A and B in the weak sense;*
(2) *If A is closed and B is compact, then there exists a closed hyperplane separating A e B in the strict sense.*

PROOF. Omitted. See Theorem I.6, I.7, pp. 5–10 in **[3]**. □

EXERCISE 2.11. Let X be a topological vector space (i.e. a vector space endowed with an Hausdorff topology such that the sum and the multiplication by scalars are continuous), $C \subset X$ be a nonempty convex set.

(1) if $\operatorname{int} C \neq \emptyset$ then $\operatorname{int} C$ is convex;
(2) $\overline{C}$ is convex;
(3) if $\operatorname{int} C \neq \emptyset$ then $\overline{C} = \overline{\operatorname{int} C}$.

SOLUTION. Let $y \in X$ be fixed. The function $f :]0,1[\times X \times X \to X$ defined by

$$f(\lambda, x, y) := \lambda y + (1 - \lambda)x$$

is continuous. Moreover, for every fixed $\lambda \in]0,1[$ and $y \in X$, the function $x \mapsto f_{y,\lambda}(x) := f(\lambda, x, y)$ is invertible with continuous inverse given by:

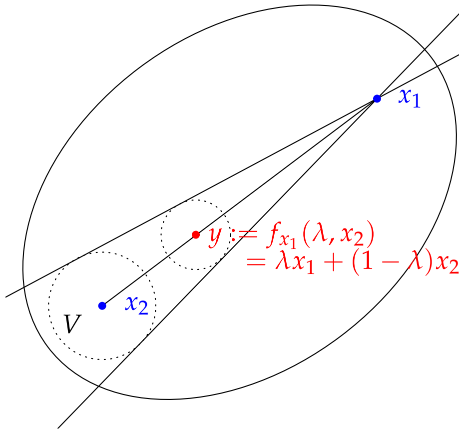
$$g_{y,\lambda}(z) := \frac{z - \lambda y}{1 - \lambda}.$$

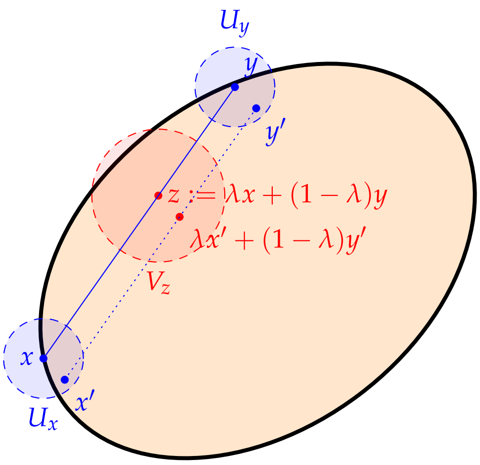
So for every fixed $\lambda \in]0,1[$, $y \in X$, the function $x \mapsto f_{y,\lambda}(x)$ maps open sets on open sets. By convexity, we have also that if $y \in C$ and $\lambda \in]0,1[$ then $f_{y,\lambda}(x) \in C$ for every $x \in C$.

Let C be a convex set with nonempty interior, $x_1, x_2 \in \operatorname{int} C$. We must prove that

$$y := \lambda x_1 + (1-\lambda)x_2 \in \operatorname{int} C,$$

for every $\lambda \in]0,1[$. Fix $\lambda \in]0,1[$, by assumption there exists an open neighborhood V of x_2 contained in C, and so $f_{x_1,\lambda}(V) \subseteq C$ by convexity. Since $y \in f_{x_1,\lambda}(V)$ and $f_{x_1,\lambda}(V)$ is open (since $f_{x_1,\lambda}(\cdot)$ maps open sets to open sets), we conclude that $y \in \operatorname{int} C$.





Suppose now that $x, y \in \overline{C}$ and take $\lambda \in]0,1[$. Set $z = \lambda x + (1-\lambda)y$, and we want to prove that $z \in \overline{C}$. Since the map $(x', y') \mapsto \lambda x' + (1-\lambda)y'$ is continuous, for any neighborhood V_z of z there are neighborhoods U_x and U_y of x and y respectively such that if $x' \in U_x$ and $y' \in U_y$ then $\lambda x' + (1-\lambda)y' \in V_z$. By assumption, we have $U_x \cap C \neq \emptyset$ and $U_y \cap C \neq \emptyset$, since $x, y \in \overline{C}$. By convexity, $z' = \lambda x' + (1-\lambda)y' \in C \cap V_z$ for all $x' \in U_x \cap C$ and $y' \in U_y \cap C$. In particular, for any neighborhood V_z of z we have that $z' \in V_z \cap C \neq \emptyset$ and so $z \in \overline{C}$, thus $\overline{C}$ is convex.

We prove that $\overline{C} = \overline{\operatorname{int}(C)}$. Trivially, $\overline{C} \supseteq \overline{\operatorname{int}(C)}$. We prove the converse inclusion. Let now $z \in \overline{C}$. We have to prove that for every open neighborhood U_z of z we have $U_z \cap \operatorname{int} C \neq \emptyset$. By construction, we have $U_z \cap C \neq \emptyset$, thus let $x \in U_z \cap C$. Choose $y \in \operatorname{int}(C)$, $\lambda \in]0,1[$, and consider the open map $f_{x,\lambda}(\cdot)$. Since by assumption there exists $V_y \subseteq C$, V_y open, we have that $f_{x,\lambda}(V_y) \subseteq C$ for all $\lambda \in]0,1[$, and it is open. Thus $f_{x,\lambda}(y) = \lambda x + (1-\lambda)y \in f_{x,\lambda}(V_y) \subseteq \operatorname{int} C$ for all $\lambda \in]0,1[$. By continuity of the map $\lambda \mapsto f(\lambda, x, y)$ on $[0,1]$ and the fact that $x = f(1, x, y)$, the inverse image of the open set U_z, which is also a neighborhood of x, is an open subset of $[0,1]$, in particular, there exists $\hat{\lambda} \in]0,1[$ such that $f(\hat{\lambda}, x, y) \in U_z$. Hence $f(\hat{\lambda}, x, y) \in U_z \cap \operatorname{int} C$, which is nonempty.

EXERCISE 2.12. Let X, Y be topological vector spaces Let $L : X \to Y$ and $\ell : X \to \mathbb{R}$ be linear, $b \in Y$, $c \in \mathbb{R}$. Define $P : X \times]0, +\infty[\to X$ by $P(x,t) = x/t$, $Q : X \to Y$ by $Q(x) = Lx + b$, $R : X \to \mathbb{R}$ by $R(x) = \ell(x) + c$, $S(x) = \dfrac{Q(x)}{R(x)}$. Prove the following assertions

- for any family of convex sets $\{C_i\}_{i \in I}$, if $C := \bigcap_{i \in I} C_i \neq \emptyset$, then it is convex;
- for all $C \subseteq X$ convex, the set $Q(C')$ is convex;
- for all $C' \subseteq Q(X) \subseteq Y$ convex, the set $Q^{-1}(C')$ is convex;
- for all $C \subseteq X$ convex, we have that $P(C,t)$ is convex for all $t > 0$;
- for all $C' \subseteq P(X) \subseteq X$ convex, we have that $P^{-1}(C')\}$ is convex;
- for all $C \subseteq X$ convex such that $R(C) \subseteq]0, +\infty[$, we have that $S(C)$ is convex;
- for all $C' \subseteq S(X) \cap R^{-1}(]0, +\infty[) \subseteq X$ convex, we have that $S^{-1}(C')$ is convex.

DEFINITION 2.13. Let E be a normed space, H be a closed hyperplane of equation $f = \alpha$, $A \subset E$, $a \in A$. We say that H is a *supporting hyperplane* to A at the *supporting point* a if $a \in H$ and at least one of the following conditions holds true:

(1) either $f(x) \leq \alpha$ for all $x \in A$;
(2) or $f(x) \geq \alpha$ for all $x \in A$.

Geometrically, we have that A is *entirely contained* in at least one of the closed half-spaces:

$$H^+ := \{x \in E : f(x) \geq \alpha\}, \qquad H^- := \{x \in E : f(x) \leq \alpha\}.$$

We say that a convex set is *strictly convex* if every supporting hyperplane intersects it in a unique point. 22

Hahn-Banach's Theorem now yields:

COROLLARY 2.14. *Let E be a normed space, $A \subseteq E$ be convex.*

(1) *if* $\mathrm{int}\,(A) \neq \emptyset$*, then every point of the boundary of A is a supporting point;*
(2) *every closed convex set is the intersection of all the half-spaces containing it;*
(3) *a convex set is weakly closed if and only if it is strongly closed.*

We notice that if $E = \mathbb{R}^d$, every nonempty closed subset A admits supporting hyperplanes at every point of the boundary without additional assumptions.

THEOREM 2.15 (Mazur). *Let X be a normed space, E be a totally bounded set. Then $H := \mathrm{co}(E)$ is totally bounded.*

PROOF. Let $\varepsilon > 0$, $U = B(0, \varepsilon)$. Set $V = B(0, \varepsilon/2)$. There exists a finite set E_1 such that $E \subseteq E_1 + V$. Let $H_1 = \mathrm{co}(E_1)$. Denoted $E_1 = \{e_1, ..., e_m\}$, we consider the set

$$S := \left\{ (t_1, ..., t_m) \in \mathbb{R}^m : t_i \geq 0 \text{ for all } i = 1...m, \sum_{i=1}^{m} t_i = 1 \right\},$$

(m-dimensional simplex) and the map $\sigma : S \to H_1$, $\sigma(t_1, ..., t_m) = \sum_{i=1}^m t_i e_i$. The map σ is continuous and surjective, thus since S is compact also H_1 is compact. Given $x \in H$, we have that $x = \alpha_1 x_1 + ... + \alpha_n x_n$ where $\alpha_i \in [0, 1]$, $x_i \in E$ for all $i = 1...n$ and $\alpha_1 + ... + \alpha_n = 1$. By definition of E_1, there exist $y_i \in E_1$ such that $x_i - y_i \in V$ for $i = 1, ..., n$. We decompose x in the sum $x = x' + x''$ with $x' = \sum \alpha_i y_i \in H_1$ and $x'' = \sum \alpha_i (x_i - y_i)$. By convexity we have that $x'' \in V$, so $E \subseteq H_1 + V$. By compactness, there exists a finite set F such that $H_1 \subseteq F + V$, so $E \subseteq F + V + V \subseteq F + U$. By arbitrariness of ε and hence of U the proof is concluded. $\square$

COROLLARY 2.16. *If X is a normed space and $\{x_n\}_{n \in \mathbb{N}} \subseteq X$ weakly converges to $\bar{u}$ then there exists a sequence of convex combinations $\{v_n\}_{n \in \mathbb{N}}$ of $\{x_n\}_{n \in \mathbb{N}} \subseteq X$ such that v_n converges to $\bar{u}$ and*

$$v_n = \sum_{k=1}^{n} \lambda_k x_k, \qquad \sum_{k=1}^{n} \lambda_k = 1, \qquad \lambda_k \geq 0, \qquad 1 \leq k \leq n.$$

We have already noticed that if F is a l.s.c. function with convex epigraph w.r.t. strong topology, it remains l.s.c. also if we equip the space with the weak topology. So if we pass from the strong topology to the weak topology, from one side we are *not threatening the lower semicontinuity* of the functions with convex epigraph, and on the other side in general, may help to *prove coecivity and level set compactness*.

It turns out to be natural, in this framework, to study the minimization problems of lower semicontinuous functionals with convex epigraph.

DEFINITION 2.17. Let X be a vector space, $F : X \to [-\infty, +\infty]$. We say that F is *convex* if

$$F(\lambda x + (1 - \lambda) y) \leq \lambda F(x) + (1 - \lambda) F(y),$$

for every $\lambda \in [0, 1]$ and for every $x, y \in X$ such that we have not $F(x) = -F(y) = \pm\infty$, i.e. the right hand side is not $+\infty - \infty$ or $-\infty + \infty$.

By induction, if $x_1, \dots, x_n \in X$ and $\lambda_1, \dots, \lambda_n \in [0,1]$ with $\sum_{i=1}^{n} \lambda_i = 1$ then

$$F\left(\sum_{i=1}^{n} \lambda_i x_i\right) \le \sum_{i=1}^{n} \lambda_i F(x_i),$$

if the right hand side is well-defined (i.e. it has no expressions $+\infty - \infty$).

EXERCISE 2.18. If F is convex, then $\{x : F(x) \le u\}$ and $\{x : F(x) < u\}$ are either convex or empty for every $u \in [-\infty, +\infty]$. The converse in general fails: there are nonconvex functions F such that $\{x : F(x) \le u\}$ and $\{x : F(x) < u\}$ are either empty or convex for every $u \in [-\infty, +\infty]$. In particular, if F is convex then $\operatorname{dom} F$ is convex.

We motivate now the choice to allow F to take the value $\pm\infty$. Consider the following *constrained* minimization problem: let $V \subseteq X$ and $F : V \to [-\infty, +\infty]$. We are interested in

$$\inf_{x \in V} F(x).$$

If is natural to redefine F by setting $\tilde{F} : X \to [-\infty, \infty]$

$$\tilde{F}(x) := \begin{cases} F(x) & \text{if } x \in V \\ +\infty, & \text{if } x \notin V, \end{cases}$$

and study

$$\inf_{x \in X} \tilde{F}(x).$$

This problem has the same properties of the original one, but now we can take advantage of the fact that $\tilde{F}$ is defined on the whole space.
So we will *always* assume that our function are defined on the whole space, unless explicitely stated. To add constraints, it turns out to be very useful the following

DEFINITION 2.19 (Indicator function). Let X be a set, V a nonempty subset of X. We define the *indicator function* $I_V : X \to [0, +\infty[$ setting $I_V(x) = 0$ if $x \in V$ and $I_V(x) = +\infty$ if $x \notin V$. So

$$\inf_{x \in V} F(x)$$

becomes

$$\inf_{x \in X} (F(x) + I_V(x)).$$

Notice that I_V is a convex function if and only if V is convex, and it is l.s.c. if and only if V is closed.

We will treat later the case $-\infty$.

DEFINITION 2.20. Let X be a vector space, $F : X \to [-\infty, +\infty]$ be convex. We say that F is *proper* if $F(x) > -\infty$ for every $x \in X$ and there exists at least one $y \in X$ for which $F(y) \in \mathbb{R}$ (equivalently, $F > -\infty$ and $\operatorname{dom} F \neq \emptyset$).

We now clarify the link between convex functions and function with convex epigraph.

LEMMA 2.21. *Let X be a vector space, $F : X \to [-\infty, +\infty]$. Then F is convex if and only if* epi F *is convex. Hence, if $F, G, \{F_i\}_{i \in I}$ are proper convex functions and $\lambda > 0$ we have that:*

(1) *$\lambda F + G$ is convex;*
(2) *$H(x) := \sup_{i \in I} F(x)$ is convex.*

PROOF. Left as an exercise. □

DEFINITION 2.22. We say that F is *strictly* convex if epi F is stricly convex.

An immediate consequence of the definition is the following.

LEMMA 2.23. *Let X be a Banach space. $F : X \to [-\infty, +\infty]$ convex and strongly l.s.c. Then it is weakly l.s.c.*

Summary of Lecture 2

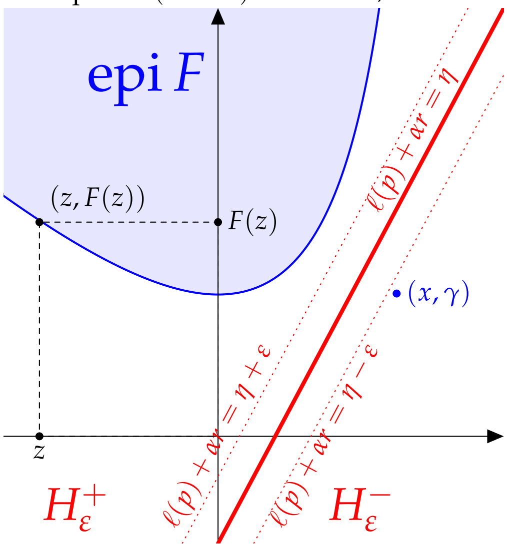
- In the previous lecture we ended sup with Tonelli-Weierstrass theorem, a powerful *topological* criterion to establish the existence of the minimizer of the basic optimization problem $\inf_{x\in X} F(x)$ where $F: X \to \mathbb{R} \cup \{\pm\infty\}$.
- More precisely, we asked F to be l.s.c. and coecive to obtain existence of a minimizer.
- Thus our aim now will be to *identify* a good topology on X in order to apply the above theorem.
- It easy to see that a too weak topology on X will give us compactness, and hence coercivity, for free, but it will make almost impossible to have l.s.c. property. On the other hand, a too strong topology on X will give us continuity, and hence l.s.c., for free, but it will make hard to have the coercivity property.
- We have seen an example where $X = L^2([0,1])$ in which the natural topology on X, i.e., the topology of the norm, will prevent us to apply Tonelli-Weierstrass criterion *even for a very simple* F.
- Thus, even if we restrict to the cases when X is a normed space, we have to move to the *weak topology*. According to this topology, a sequence $\{x_n\}_{n\in\mathbb{N}} \subseteq X$ converges to $x \in X$ if and only if we have the following convergence in $\mathbb{R}$
$$\lim_{n\to+\infty} \langle f, x_n\rangle_{X',X} = \langle f, x\rangle_{X',X},$$
for all $f: X \to \mathbb{R}$ linear and continuous. We use the notation $\langle f, x\rangle_{X',X}$ in place of $f(x)$.
- We made a quick review of the properties of the weak topology. The weak topology is weaker than the strong topology of the norm. Thus it will be easier to have compactness and hence coercivity. However to switch from strong to weak topology may destroy the l.s.c. property.
- The geometric characterization of the l.s.c. told us that l.s.c. is equivalent to the closedness of the epigraph. Thus the first step to preserve this property is to identify a class of sets for which the *strong* closure and the *weak* closure coincides.
- To this aim we introduced and studied some properties of the *convex sets*. Convexity is a property which does not depend on the topology, but just on the linear structure of the space.
- If we add a topological structure *compatible* with the linear one, we have seen that convexity is preserved in passing to interior or closure.
- The core result of all the convex analysis is Hahn-Banach theorem, whose geometric forms we revised carefully.
- Among the countless important consequences of Hahn-Banach theorem, we mention the fact that a convex set is *strongly* closed if and only if it is *weakly* closed. By applying this fact on the epigraph of F, we conclude that a convex function is l.s.c. for the strong topology if and only if it is l.s.c. for the weak topology.
- In particular, for convex functions, to pass from strong to weak topology does not destroy the l.s.c. property.

3. Lecture of 8 october 2018: Continuity and affine minorants of convex functions (3h)

We prove now a particular case of Hahn-Banach Theorem applied to convex functions. Before it, we recall that we can identify $(X\times\mathbb{R})'$ and $X'\times\mathbb{R}$ in the following way: given $f \in (X\times\mathbb{R})'$, the couple $(\ell,\alpha) \in X'\times\mathbb{R}$ defined as $\ell(x) = f(x,0)$ and $\alpha = f(0,1)$ satisfies $f(x,r) = \ell(x) + \alpha r$. Conversely, given $(\ell,\alpha) \in X'\times\mathbb{R}$, we can define $f \in (X\times\mathbb{R})'$ by setting $f(x,r) = \ell(x) + \alpha r$.

LEMMA 3.1 (affine minorants of convex l.s.c. functions). *Let X be a Banach space. $F: X \to]-\infty,+\infty]$ be convex and l.s.c. Let $(x,\gamma) \notin \operatorname{epi} F$. Then there exists a continuous linear functional $\ell: X\to\mathbb{R}$, $\ell \neq 0$ and $\alpha, \varepsilon > 0$ such that*
$$\ell(x) + \alpha\gamma < \ell(y) + \alpha\beta - \varepsilon$$
for all $(y,\beta) \in \operatorname{epi} F$.

PROOF. If $F \equiv +\infty$ then $\operatorname{epi} F = \emptyset$ and there is nothing to prove. So we assume the existence of $z \in X$ with $F(z) < +\infty$. By assumption the closed set $\operatorname{epi} F$ and the compact set $\{(x,\gamma)\}$ are disjoint, so by Hahn-Banach Theorem in its second geometric form, it is possible to find a continuous linear functional $f : X \times \mathbb{R} \to \mathbb{R}$ and $\eta \in \mathbb{R}$ such that they are stricly separated by the closed hyperplane (closed in $X \times \mathbb{R}$) whose equation is $f(p,r) = \eta$. Recalling the isomorphism $(X \times \mathbb{R})' \simeq X' \times \mathbb{R}$, we can identify f with a couple $(\ell, \alpha) \in X' \times \mathbb{R}$.



To avoid triviality assume $\operatorname{epi} F \neq \emptyset$ or, equivalently, $\operatorname{dom} F \neq \emptyset$.

Choose $z \in \operatorname{dom} F$, thus $F(z) < +\infty$.

Separate the closed $\operatorname{epi} F$ and the compact $\{(x,\gamma)\}$ with an hyperplane whose equation is $\ell(p) + \alpha r = \eta$ in such a way that

$(x,\gamma) \in H_\varepsilon^- := \{(p,r) : \ell(p) + \alpha r \leq \eta - \varepsilon\}$,

$\operatorname{epi} F \subseteq H_\varepsilon^+ := \{(p,r) : \ell(p) + \alpha r \geq \eta + \varepsilon\}$.

The separation is strict, thus $\varepsilon > 0$.

Since $\{z\} \times [F(z), +\infty[\subseteq H_\varepsilon^+$, we must have $\alpha \geq 0$.

If $x \in \operatorname{dom} F$ then $\alpha > 0$: otherwise, if $\alpha = 0$ the conditions $(x,\gamma) \in H_\varepsilon^-$ and $(x, F(z)) \in \operatorname{epi} F \subseteq H_\varepsilon^+$ lead to a contradiction.

Assume $x \notin \operatorname{dom} F$ and $\alpha = 0$, thus $\ell(x) < \eta < \ell(y)$ for all $y \in \operatorname{dom} F$, or equivalently, $\ell(y) - \ell(x) \geq \eta - \ell(x) > 0$ for all $y \in \operatorname{dom} F$. Since $z \in \operatorname{dom} F$, we take $u < F(z)$ and we can separate $(z,u) \notin \operatorname{epi} F$ from $\operatorname{epi} F$ by an hyperplane of equation $\ell'(p) + \alpha'(r) = \eta'$ in such a way that $\ell'(z) + \alpha' u < \ell(y) + \alpha' \beta$ for all $(y,\beta) \in \operatorname{epi} F$ and $\alpha' > 0$. Set $\bar{l} = k\ell + \ell'$, $\bar{\alpha} = \alpha' > 0$ and choose k such that $\bar{\ell}(x) + \bar{\alpha}\gamma + \varepsilon < \bar{\ell}(y) + \bar{\alpha}\beta$ for all $(y,\beta) \in \operatorname{epi} F$. Substituting, we obtain

$$k > \frac{\ell'(x) + \alpha'\gamma - (\ell'(y) + \alpha'\beta) + \varepsilon}{\ell(y) - \ell(x)}, \text{ for all } (y,\beta) \in \operatorname{epi} F.$$

The denominator of the right hand side is larger than $\eta - \ell(x)$ and the numerator is less than $\ell'(x) + \alpha'(x) - (\ell'(z) + \alpha' u) + \varepsilon$, thus the inequality is fulfilled by choosing

$$k > \frac{\ell'(x) + \alpha'\gamma - (\ell'(z) + \alpha' u) + \varepsilon}{\eta - \ell(x)}.$$

□

EXAMPLE 3.2. Assume that $X = \mathbb{R}$, $F : X \to]-\infty, +\infty]$ be convex and l.s.c. Let $(x_0, y_0) \notin \operatorname{epi} F$. In this case, $X' = \mathbb{R}$, thus according to the previous theorem, we have that there exists $m' \in \mathbb{R}$, $\alpha > 0$, and η such that

$$m' x_0 + \alpha y_0 < \eta < m' x_1 + \alpha y_1 - \varepsilon$$

for all $(x_1, y_1) \in \operatorname{epi} F$. Since $\alpha > 0$, we can divide all by α and define $m = m'/\alpha$, $q = \eta/\alpha$ obtaining that the line $y = mx + q$ stricly separates (x_0, y_0) and $\operatorname{epi} F$. So since $\alpha > 0$ we have find that the separating line is not vertical. More precisely, if we can separate (x_0, y_0) from $\operatorname{epi} F$ with a vertical line, we can also make the separation with a nonvertical line.

We discuss now the case of convex l.s.c. functions which take the value $-\infty$.

LEMMA 3.3. *Let X be a Banach space, $F : X \to [-\infty, +\infty]$ be convex and l.s.c. If there exists $u_0 \in X$ such that $F(u_0) = -\infty$ then* $\operatorname{dom} F = \emptyset$.

PROOF. By contradiction, suppose $\operatorname{dom} F \neq \emptyset$, in particular there exist $\bar{u} \in X$ and $\bar{a} \in \mathbb{R}$ such that $\bar{a} < F(\bar{u}) \in \mathbb{R}$. By lemma 3.1 there exist $\ell \in X'$ and $\alpha > 0$ such that

$$\ell(\bar{u}) + \alpha\bar{a} < \ell(y) + \alpha\beta$$

for all $(y, \beta) \in \operatorname{epi} F$. Taking $y = u_0$, since $(u_0, \beta) \in \operatorname{epi} F$ for all $\beta \in \mathbb{R}$, and dividing by $\alpha > 0$ we have

$$\frac{1}{\alpha}\ell(\bar{u} - u_0) + \bar{a} < \beta.$$

By letting $\beta \to -\infty$ we obtain a contradiction, since the left hand side is finite. □

We are going now to state some results about continuity of convex functions. We will prove that a convex function cannot be have as discontinuity jumps of finite length (in other words, we cannot have that lim sup and lim inf are different and both finite). Convexity gives to the function a *rigid* structure, which we are going to examine.
The *rigidity* of the convex functions allows us in a certain sense to *transfer* information from one point of the domain to another one.

PROPOSITION 3.4. *Let X be a topological vector space. Let $F : X \to [-\infty, +\infty]$ be convex. The following are equivalent:*

(i.) *there exists $u \in X$ with $F(u) > -\infty$, $a \in \mathbb{R}$ and a neighborhood W of u, such that $F(v) \le a$ for all $v \in W$;*
(ii.) *there exists $u \in \operatorname{dom} F$ such that F is continuous at u;*
(iii.) *there exists a nonempty open set $\mathcal{O}$ and $a \in \mathbb{R}$ such that $F_{|\mathcal{O}}$ is not identically $-\infty$ and $F(v) \le a$ for all $v \in \mathcal{O}$;*
(iv.) *F is proper and is continuous in the interior of its domain, which is nonempty;*
(v.) *if X is a normed space, then the domain of F has nonempty interior and F is locally Lipschitz continuous on the interior of $\operatorname{dom} F$.*

REMARK 3.5. Notice that if we have just information on a subset of the domain of F (*local* information), we are still able to derive *global* regularity properties for F.

PROOF.

$(i.) \iff (ii.)$. Trivially $(ii.) \implies (i.)$. Let us prove the converse implication. Up to translation, without loss of generality, we can assume $u = 0$ and $F(u) = 0$ and that $\beta W \subseteq W$ for all $|\beta| \le 1$ (if X is a Banach space, we take $W = B(0, \delta)$ for a suitable $\delta > 0$). By assumption, we have $\operatorname{co}\{W \times \{a\}, (0,0)\} \subseteq \operatorname{epi} F$, thus for every $0 < \varepsilon < 1$ if $x \in \varepsilon W$, hence $x/\varepsilon \in W$, we must have

$$F(x) = F\left(\varepsilon \cdot \frac{x}{\varepsilon} + (1-\varepsilon)\cdot 0\right) \le \varepsilon F\left(\frac{x}{\varepsilon}\right) + (1-\varepsilon)F(0) \le \varepsilon a.$$

Moreover, since also $-x \in \varepsilon W$, we have $F(-x) \le \varepsilon a$, thus

$$0 = F(0) = F\left(\frac{1}{2}\cdot x + \frac{1}{2}(-x)\right) \le \frac{1}{2}F(x) + \frac{1}{2}F(-x) \le \frac{1}{2}F(x) + \frac{1}{2}\cdot \varepsilon a,$$

which implies $F(x) \ge -\varepsilon a$. Thus we have $|F(x)| \le \varepsilon a$ for all $x \in \varepsilon W$, which yields continuity of F at u.

$(ii.) \iff (iii.)$ By continuity, we have $(ii.) \implies (iii.)$. Conversely, $(iii.)$ implies that there exists at least one point in $\mathcal{O}$ fulfilling $(i.)$, hence F is continuous at that point, thus $(ii.)$ is fulfilled.

$(iii.) \iff (iv.)$ Trivially, $(iv.) \implies (iii.)$. Conversely, assume $(iii.)$. Since F is continuous in at least one point of $\mathcal{O}$, in particular is bounded around that point, by possibly shrinking $\mathcal{O}$ we can assume that $\mathcal{O} \subseteq \operatorname{int} \operatorname{dom} F$. and $F_{|\mathcal{O}}$ is bounded, hence there exists $a \in \mathbb{R}$ such that $F(x) \le a$ for all $x \in \mathcal{O}$. In particular, $\operatorname{int} \operatorname{dom} F \neq \emptyset$. Let $u \in \mathcal{O}$, $v \in \operatorname{int} \operatorname{dom} F$, $v \neq u$. For $\mu > 0$ sufficiently small $z_\mu := u + (1+\mu)(v-u)$ belongs to $\operatorname{int} \operatorname{dom} F$ since $z_\mu \to v$ for $\mu \to 0^+$. Fix $\mu > 0$ such that $z_\mu \in \operatorname{int} \operatorname{dom} F$, and let $\beta > F(z_\mu)$. For any fixed $\lambda \in]0,1[$ consider the continuous map $f_\lambda(x) = \lambda x + (1-\lambda)z_\mu$. As already observed, this map is invertible and its inverse is continuous. Moreover,

$$f_\lambda(u) = \lambda u + (1-\lambda)\left(u + (1+\mu)(v-u)\right) = (\lambda - \mu + \lambda\mu)u + (1 - (\lambda - \mu + \lambda\mu))v,$$

thus choosing $\lambda = \dfrac{\mu}{1+\mu} \in]0,1[$, we have $f_\lambda(u) = v$. Moreover, $f_\lambda(\mathcal{O})$ is an open neighborhood of v and given $y \in f_\lambda(\mathcal{O})$, we have $y = f_\lambda(x)$ for a certain $x \in \mathcal{O}$

$$F(y) = F(f_\lambda(x)) = F(\lambda x + (1-\lambda)z_\mu) \le \lambda F(x) + (1-\lambda)F(z_\mu) \le \lambda a + \beta.$$

Since F is bounded from above in a neighborhood of v, we have that it is continuous at v. By the arbitrariness of $v \in \operatorname{int}\operatorname{dom} F$, we have that F is continuous in $\operatorname{int}\operatorname{dom} F$.

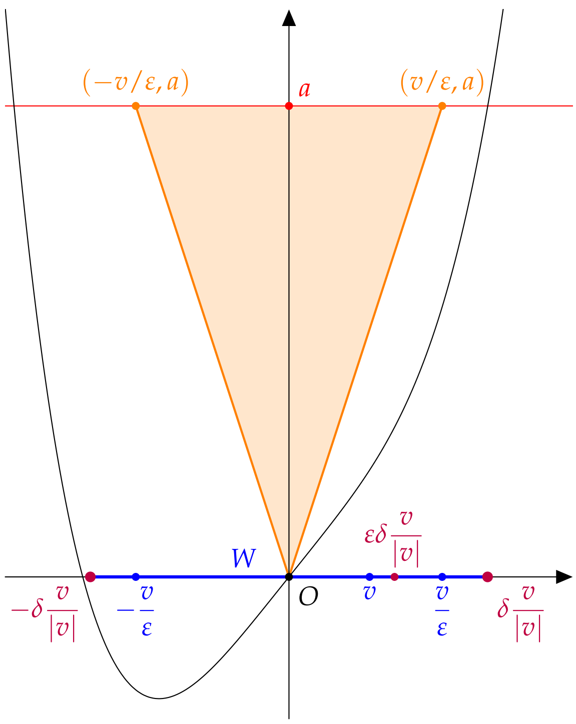
$(i.) \iff (v.)$ Trivially, $(v.) \implies (i.)$. Conversely, assume $(i.)$. Without loss of generality, we can take $u = 0$, $F(u) = 0$ and $W = B(0,\delta)$. Take $v_1, v_2 \in B(0,\delta/2)$, $v_1 \neq v_2$ and set $\mu = \dfrac{\delta}{2\|v_1 - v_2\|} > 1$. Notice that $z_\mu := v_2 + \mu(v_2 - v_1) \in B(0,\delta)$. As before, we consider $f_\lambda(x) = x + \lambda(z_\mu - x)$, noticing that for $\lambda = \dfrac{1}{1+\mu}$ we have $f_\lambda(v_1) = v_2$, moreover,

$$F(v_2) = F(f_{\frac{1}{1+\mu}}(v_1)) \le F(v_1) + \frac{1}{1+\mu}[F(z_\mu) - F(v_1)] \le F(v_1) + \frac{2a}{1+\mu},$$

thus $F(v_2) - F(v_1) \le \dfrac{2a}{1+\mu}$. Reversing the roles of v_1, v_2 we have

$$|F(v_2) - F(v_1)| \le \frac{2a}{1+\mu} = \frac{2a}{1 + \frac{\delta}{2\|v_1 - v_2\|}} = \frac{4a\|v_1 - v_2\|}{\delta + 2\|v_1 - v_2\|} \le \frac{4a}{\delta}\|v_1 - v_2\|.$$

□



Without loss of generality, $(u, F(u)) = (0,0)$, and $W = B(0,\delta)$.

We have $\varepsilon W \subseteq W$ for all $|\varepsilon| \le 1$.

Given $v \in \varepsilon W$, we have $v/\varepsilon \in W$ and $\pm v \in \varepsilon W$, thus by convexity of $\operatorname{epi} F$, the triangle $\operatorname{co}\{(0,0), (\pm v/\varepsilon, a)\}$ is contained in $\operatorname{epi} F$, hence the graph of F restricted to the segment joining $(\pm v/\varepsilon, a)$ must be below that triangle.

We write $\pm v$ as convex combination of 0 and $\pm v/\varepsilon$ with coefficients $\frac{1}{1+\varepsilon}$ and ε, obtaining $F(\pm v) \le \varepsilon F(\pm v/\varepsilon) = \varepsilon a$.

On the other hand, since 0 is the midpoint of that segment, by convexity the value of the function at 0 (i.e. 0) must be below the midpoint of the values of the function at $\pm v$.

This implies that the sum of the values of the function at $\pm v$ must be positive, and since $F(-v) \le \varepsilon a$, we have $F(v) \ge -\varepsilon a$.

Thus $|F(v)| \le \varepsilon a$, which yields continuity by letting $\varepsilon \to 0^+$.

REMARK 3.6. Taking $F(x) = 1$ for $x > 0$ and $F(x) = -1$ for $x \le 0$, we see that these properties are far from being true if F is not convex!

EXERCISE 3.7. For every $\bar{x} \in H_0^1(\Omega)$, $\delta, M > 0$ and $x^* \in H^{-1}(\Omega) \setminus L^2(\Omega)$ there exists $x_M \in H_0^1(\Omega)$ with $\|\bar{x} - x_M\|_{L^2} < \delta$ and $\langle x^*, x_M\rangle_{H^{-1}, H_0^1} > M$.

PROOF. Since $x^* \notin L^2$, it cannot be continuous for the L^2-norm, otherwise we have $x^* \in L^2$, thus there exists a sequence $\{x_n\}_{n\in\mathbb{N}} \subseteq C_c^\infty(\Omega)$ such that $x_n \to 0$ in L^2 and $\langle x^*, x_n\rangle \to +\infty$. The general case can be obtained by translation. □

We have already seen that a closed convex set coincides with the intersection of all the closed half-spaces containing it. If the closed convex set is the epigraph of a convex l.s.c function $G : X \to [-\infty, +\infty]$, this can be equivalently stated by saying that $\operatorname{epi} G$ coincides with the epigraph of the function defined as pointwise supremum of all the affine continuous function whose graphs are below the graph of G, since their epigraphs form a family of closed half-space containing $\operatorname{epi} G$, and the pointwise supremum yields the intersection of the family. We will precise now that notion.

DEFINITION 3.8. Let X be a normed space, $F : X \to [-\infty, +\infty]$ be a function (not necessarily convex). Define

$$\bar{F}(x) = \sup\{h(x) : \ h : X \to [-\infty, +\infty] \text{ is l.s.c. and } h(y) \le F(y) \text{ for all } y \in X\}.$$

Since the pointwise supremum of l.s.c. functions is l.s.c., we have that $\bar{F}$ is *the largest* l.s.c. function everywhere less or equal to F, we will call $\overline{F}$ the *l.s.c. regularization* of F.

LEMMA 3.9 (closure and convexification of epigraphs). *Let X be a normed space, $F : X \to [-\infty, +\infty]$ be a function (not necessarily convex). Then:*

(1) *there exists $G : X \to [-\infty, +\infty]$ such that $\operatorname{epi} G = \overline{\operatorname{epi} F}$, moreover we have*

$$G(x) = \lim_{r\to 0^+} \inf_{y\in B(x,r)} F(y).$$

(2) *there exists $\tilde{G} : X \to [-\infty, +\infty]$ such that $\operatorname{epi} \tilde{G} = \overline{\operatorname{co}}(\operatorname{epi} F)$;*
(3) $\operatorname{epi} \overline{F} = \overline{\operatorname{epi} F}$.

PROOF. We define:

$$\begin{aligned} G(x) &= \inf\{\beta \in \mathbb{R} : \ (x, \beta) \in \overline{\operatorname{epi} F}\}, \\ \tilde{G}(x) &= \inf\{\beta \in \mathbb{R} : \ (x, \beta) \in \overline{\operatorname{co}}(\operatorname{epi}) F\}, \end{aligned}$$

where we set $\inf\{\emptyset\} = +\infty$.

(1) If $(\bar{x}, \bar{\beta}) \in \operatorname{epi} F$ then $(\bar{x}, \bar{\beta}) \in \overline{\operatorname{epi} F}$ thus $G(\bar{x}) \le \bar{\beta}$ and so $(\bar{x}, \bar{\beta}) \in \operatorname{epi} G$. Hence $\operatorname{epi} F \subseteq \operatorname{epi} G$. Conversely, assume that $(x, \bar{\beta}) \in \operatorname{epi} G$. By definition, there exists $\{\xi_n\}_{n\in\mathbb{N}} \subseteq \mathbb{R}$ such that $\{(x, \xi_n)\}_{n\in\mathbb{N}} \subseteq \overline{\operatorname{epi} F}$ and $\bar{\beta} \le \xi_n \le \bar{\beta} + 1/n$ for all $n \in \mathbb{N}$. Since $(x, \xi_n) \to (x, \bar{\beta})$ in $X \times \mathbb{R}$, by the closedness of $\overline{\operatorname{epi} F}$ we obtain $(x, \bar{\beta}) \in \overline{\operatorname{epi} F}$.

So we have $\operatorname{epi} F \subseteq \operatorname{epi} G \subseteq \overline{\operatorname{epi} F}$, and by taking the closure we have $\operatorname{epi} G = \overline{\operatorname{epi} F}$. Define now $G'(x) = \lim\limits_{r\to 0^+} \inf\limits_{y\in B(x,r)} F(y)$, we want to prove $\operatorname{epi} G' = \overline{\operatorname{epi} F}$, thus $G = G'$.

Suppose to have $(x, \beta) \in \overline{\operatorname{epi} G'}$, then there are sequences $(x_n, \beta_n) \in \operatorname{epi} G'$ such that $x_n \to x$, $\beta_n \in \beta$, and

$$\beta_n \ge G'(x_n) = \lim_{r\to 0^+} \inf_{y\in B(x_n,r)} F(y) \ge \lim_{r\to 0^+} \inf_{y\in B(x,r+|x-x_n|)} F(y) = G'(x),$$

and so $\beta \ge G'(x)$, thus $\operatorname{epi} G'$ is closed. Moreover, suppose to have $(x, \beta) \in \overline{\operatorname{epi} F}$, then there are sequences $(x_n, \beta_n) \in \operatorname{epi} F$ such that $x_n \to x$, $\beta_n \to \beta$, and

$$\beta_n \ge F(x_n) \ge \lim_{r\to 0^+} \inf_{y\in B(x_n,r)} F(y) \ge \lim_{r\to 0^+} \inf_{y\in B(x,r+|x-x_n|)} F(y) = G'(x),$$

and so $\beta \ge G'(x)$, thus $\operatorname{epi} G' \supseteq \overline{\operatorname{epi} F}$.

Take now a sequence $r_n \to 0^+$, then $\left\{ \inf\limits_{y\in B(x,r_n)} F(y) \right\}_{n\in\mathbb{N}}$ is an increasing sequence in $\mathbb{R}$, and so

$$G'(x) = \sup_{r>0} \inf_{y\in B(x,r)} F(y) = \min\{\liminf_{y\to x} F(y), F(x)\}.$$

In particular, we have that $G'(x) \le F(x)$ and equality holds if and only if $F(x) \le \liminf\limits_{y\to x} F(y)$.

We have two possibilities: either $(x, G'(x)) = (x, F(x))$ or there exists a sequence $x_n \to x$ such that $(x_n, F(x_n)) \to (x, G'(x))$ and $(x_n, F(x_n)) \in \operatorname{epi} F$. In any case, we have $(x, G'(x)) \in \overline{\operatorname{epi} F}$, since even if $(x, G'(x)) = (x, F(x))$ we can see $(x, G'(x))$ as the limit along the constant sequence $x_n \equiv x$ of $(x_n, F(x_n)) \in \operatorname{epi} F$. We conclude that $\overline{\operatorname{epi} F} = \operatorname{epi} G'$

(2) Repeat first part of the above argument replacing G and $\overline{\operatorname{epi}} F$ with $\tilde{G}$ and $\overline{\operatorname{co}}(\operatorname{epi} F)$, respectively.
(3) Since $\overline{F}$ is l.s.c. and is everywhere less or equal than F, we have that $\operatorname{epi} \overline{F}$ is closed and contains $\operatorname{epi} F$, thus by taking the closure we have $\operatorname{epi} \overline{F} \supseteq \overline{\operatorname{epi} F}$. Moreover, by definition we have that G is everywhere less than $\overline{F}$, since G is a l.s.c. function below F and $\overline{F}$ is the pointwise supremum of all such functions, thus $\operatorname{epi} G \supseteq \operatorname{epi} \overline{F}$.

Recalling that $\operatorname{epi} G = \overline{\operatorname{epi} F}$, we have

$$\overline{\operatorname{epi} F} = \operatorname{epi} G \supseteq \operatorname{epi} \overline{F} \supseteq \overline{\operatorname{epi} F},$$

so equality holds.

□

LEMMA 3.10. *A function $f : X \to [-\infty, +\infty]$ is l.s.c. at $x_0 \in X$ if and only if $f(x_0) = \overline{f}(x_0)$.*

PROOF. We recall that $\overline{f}(y) \leq f(y)$ for all $y \in X$. Assume that $\overline{f}(x_0) = f(x_0)$, then

$$f(x_0) = \overline{f}(x_0) \leq \liminf_{x \to x_0} \overline{f}(x) \leq \liminf_{x \to x_0} f(x).$$

Conversely, let f be l.s.c. at x_0. If $\overline{f}(x_0) = +\infty$ the proof is concluded, otherwise let $(x_0, t) \in \operatorname{epi} \overline{f} = \overline{\operatorname{epi} f}$. We have a sequence $(x_n, t_n) \in \operatorname{epi} f$ with $x_n \to x_0$, $t_n \to t$. Then

$$f(x_0) \leq \liminf_{n \to \infty} f(x_n) \leq \liminf_{n \to \infty} t_n = t$$

so $f(x_0) \leq t$ for all $t \geq \overline{f}(x_0)$. But then $f(x_0) \leq \overline{f}(x_0)$ and so equality holds. □

Among the l.s.c. minorants, a distinguished role is played by the continuous affine functions, i.e. functions f of the form $f(v) = \ell(v) + \alpha$ where $\ell \in X'$ (so linear and continuous from X to $\mathbb{R}$) and $\alpha \in \mathbb{R}$. Thus the following definition is quite natural.

DEFINITION 3.11. Let X be a topological vector space where every points has a basis of neighborhoods made by convex sets (such spaces are called *locally convex* l.c.s.), we define

$$\Gamma(X) := \left\{ F : X \to [-\infty, +\infty] : \ F(x) = \sup_{i \in I} \{f_i(x)\} \text{ where } f_i \text{ is continuous affine} \right\},$$

We define its subset $\Gamma_0(X)$ by removing from $\Gamma(X)$ the constant functions taking everywhere vaues $+\infty$ or $-\infty$.

REMARK 3.12. The pointwise supremum of elements of $\Gamma(X)$ is still an element of $\Gamma(X)$. Indeed, let $G(x) = \sup_{i \in I} k_i(x)$ with $k_i \in \Gamma(X)$ for all $i \in I$, and let $\bar{x}$ be fixed. Take any sequence $\varepsilon_i \to 0^+$, and notice by definition that there exists $i \in I$ such that $k_i(x) \leq G(x)$ for all $x \in X$ and

$$G(\bar{x}) - \varepsilon_i \leq k_i(\bar{x}) \leq G(\bar{x}).$$

Since $k_i \in \Gamma(X)$, there exists a continuous affine function f_i such that $f_i(x) \leq k_i(x) \leq G(x)$ for all $x \in X$ and satisfying $k_i(\bar{x}) - \varepsilon_i \leq f_i(\bar{x})$. So we have

$$G(\bar{x}) - 2\varepsilon_i \leq k_i(\bar{x}) - \varepsilon_i \leq f_i(\bar{x}) \leq G(\bar{x}).$$

So $G(\bar{x}) = \sup_{i \in I} f_i(\bar{x})$ hence $G \in \Gamma(X)$.

PROPOSITION 3.13. *We have that $F \in \Gamma(X)$ if and only if $F : X \to [-\infty, +\infty]$ is a convex l.s.c. function such that if there exists at least a point where it take the value $-\infty$, then it is identically $-\infty$.*

PROOF. Assume that $F \in \Gamma(X)$. Since it is pointwise supremum for convex l.s.c. functions, it is a convex l.s.c. function. Moreover, if $I = \emptyset$ then F is identically $-\infty$, otherwise, F cannot take the value $-\infty$ at any point.

Conversely, let $F : X \to [-\infty, +\infty]$ be a convex l.s.c. function such that $F(x) > -\infty$ for every $x \in X$. If F is identically $+\infty$ then F is the pointwise supremum of *all* continuous affine functions from X to $\mathbb{R}$, if $F \equiv -\infty$, it is the supremum of the empty family of affine continuous function, thus it remains the case $\operatorname{epi} F \neq \emptyset$ and $\operatorname{epi} F \neq X \times R$.
We have already seen in Lemma 3.1 that if $P := (x, u) \notin \operatorname{epi} F$, there exist $\ell = \ell_P \in X'$, $\alpha = \alpha_P > 0$, $\varepsilon = \varepsilon_P$ such that $\ell_P(x) + \alpha_P u < \ell_P(y) + \alpha_P \beta - \varepsilon_P$ for all $(y, \beta) \in \operatorname{epi} F$, or, equivalently,

$$f_P(y) := \frac{1}{\alpha_P} [-\ell_P(y) + \ell_P(x) + \varepsilon_P] + u < \beta$$

for all $(y, \beta) \in \operatorname{epi} F$. Notice that the map $y \mapsto f_P(y)$ is continuous and affine (since $P = (x, u)$ is fixed) and we have $\operatorname{epi} F \subseteq \operatorname{epi} f_P$. Moreover, we have $f_P(x) = \dfrac{\varepsilon}{\alpha} + u > u$, thus $P \notin \operatorname{epi} f_P$.
Define $G(x) = \sup\{f_P(x) : P \notin \operatorname{epi} F\}$. Clearly $G \in \Gamma$. Moreover,

$$\operatorname{epi} G = \bigcap_{P \notin \operatorname{epi} F} \operatorname{epi} f_P \supseteq \operatorname{epi} F,$$

since $\operatorname{epi} f_P \supseteq \operatorname{epi} F$ for all $P \notin \operatorname{epi} F$. On the other hand, given $Q \notin \operatorname{epi} F$, we have $Q \notin \operatorname{epi} f_Q$ thus $Q \notin \operatorname{epi} G$. Hence $\operatorname{epi} F = \operatorname{epi} G$ thus $F = G$. □

DEFINITION 3.14. Let $F : X \to [-\infty, +\infty]$ be a function (not necessarily convex). We say that G is the Γ-regularization of F if it is the largest function of $\Gamma(X)$ such that $G(x) \leq F(x)$ for every $x \in X$, i.e. if it is the pointwise supremum of all function $g \in \Gamma(x)$ such that $g(x) \leq F(x)$ for all $x \in X$.

PROPOSITION 3.15. *Let $F, G : X \to [-\infty, +\infty]$.*
(1) *G is the Γ-regularization of F if and only if G is the pointwise supremum of all continuous affine minorants of F;*
(2) *if $F \in \Gamma(X)$ then F coincides with its Γ-regularization;*
(3) *if F admits at least one continuous affine minorant, and G is the Γ-regularization of F, then $\operatorname{epi} G = \overline{\operatorname{co}} \operatorname{epi} F$;*
(4) *if G is the Γ-regularization of F then $G \leq \bar{F} \leq F$;*
(5) *if F is convex and admits at least one continuous affine minorant, and G is the Γ-regularization of F, then $G = \bar{F}$.*

PROOF.
(1) We use the convention that $\inf \emptyset = +\infty$, and define

$$G_1(x) := \sup\{h(x) : h(y) \leq F(y) \text{ for all } y \in X,\ h \text{ continuous and affine}\},$$
$$G_2(x) := \sup\{k(x) : k(y) \leq F(y) \text{ for all } y \in X,\ k \in \Gamma(X)\}.$$

We have $G_1, G_2 \in \Gamma(X)$ since the pointwise supremum of elements of $\Gamma(X)$ is in $\Gamma(X)$.

Since all the continuous affine functions belong to $\Gamma(X)$, being sup of a family made by a single element of continuous affine functions), the following holds:
(a) $G_1(x) \leq G_2(x)$ for all $x \in X$ by definition (the sup in the definition of G_2 is on a larger set than in the definition of G_1), so all the continuous affine minorants of G_1 are continuous affine minorants of G_2.
(b) since $G_2(x) \leq F(x)$ for all $x \in X$, every continuous affine minorant of G_2 is a function of $\Gamma(X)$ which is also a minorant of F. So by definition of G_1, every continuous affine minorant of G_2 must be also a continuous affine minorant of G_1.

So G_1 and G_2 are elements of $\Gamma(X)$ which have the same set of continuous affine minorants, thus $G_1 = G_2$.
(2) Trivial.

(3) Let f be a continuous affine minorant of F. Take a l.s.c. convex function $\tilde{G}$ such that $\operatorname{epi}\tilde{G} = \overline{\operatorname{co}}\operatorname{epi} F$. Since

$$\operatorname{epi} F \subseteq \overline{\operatorname{co}}\operatorname{epi} F = \operatorname{epi}\tilde{G} \subseteq \operatorname{epi} f,$$

we have $f \leq \tilde{G} \leq F$, and by passing to the pointwise supremum on the minorants f, we have that $G \leq \tilde{G}$. Since $\tilde{G} \in \Gamma(X)$ we have by definition $\tilde{G} \leq G$, so we obtain $\tilde{G} = G'$.

(4) Trivial by the previous items.

(5) Trivial by the previous items.

□

Summary of Lecture 3

- We have seen that a good framework for our minimization problem $\inf_{x\in X} F(x)$, with $F : X \to \mathbb{R} \cup \{\pm\infty\}$ is to take as X a normed space, and as $F(\cdot)$ a l.s.c. convex function. In this case, endowing X with the weak topology will not compromise the l.s.c., since for convex functions strong l.s.c. and weak l.s.c. coincides. We recall that compactness, and hence coercivity, has much more possibility to be proved under weak topology.
- Hahn-Banach Theorem tells us that every convex and closed subset C of $X \times \mathbb{R}$ is the intersection of the closed half-spaces containing it. This follows from the separation of each point $P \notin C$ by a closed hyperplane.
- When F is convex l.s.c., its epigraph is a closed convex subset of $X \times \mathbb{R}$. Among the closed hyperplanes of $X \times \mathbb{R}$, a distinguished role is played by the *graphs* of continuous affine functions. Every graph (resp. epigraph) of a continuous affine functions defines uniquely a closed hyperplane in $X \times \mathbb{R}$ (resp. a closed half space of $X \times \mathbb{R}$). Unfortunately, it is not possible in general to represent *every* closed hyperplane in $X \times \mathbb{R}$ by a graph of a function defined on X (e.g. the line $x = 0$ is a closed hyperplane in $\mathbb{R} \times \mathbb{R}$ which cannot be represented by the graph of a function defined on $\mathbb{R}$).
- Thus we proved a version of Hahn-Banach Theorem in order to separate every $P \notin \operatorname{epi} F$ from $\operatorname{epi} F$ by mean of the graph of a continuous affine function. A consequence is that the epigraph of every convex l.s.c. function is *contained* in the interserction of all the epigraphs of continuous affine function below it (called continuous affine *minorants*). They actually *coincides* if the set of continuous affine minorants is nonempty.
- Moreover, since the intersection of epigraph is the epigraph of the pointwise supremum, the previous representation allows us to construct functions whose epigraph is the closed convex hull of the epigraph of a given function. This procedure can be viewed as a *regularization* of the original function.
- The rigid structure of convex functions allows us to prove strong continuity property (and in some case local Lipschitz continuity) for a convex function $F : X \to \mathbb{R} \cup \{+\infty\}$ provided that the interior of the domain is nonempty. In particular, we have seen that if we have at least a point u where F does not assume the value $-\infty$ and is uniformly bounded from above, then F is locally Lipschitz continuous on the whole of the interior of the domain of F, which is notempty since u belongs to it.

4. Lecture of 12 october 2018: Conjugate of convex functions (3h)

The importance of the study of continuous affine minorants of convex functions can be motivated by the following simple remark. If $\operatorname{epi} F$ admits at $(\bar{x}, \bar{a})$ the supporting hyperplane of equation $f = \alpha\bar{a}$, where $f(x, a) = \ell(x) + \alpha a$ and $\alpha > 0$, $\ell(x) = 0$, necessarily for every given $(x, a) \in \operatorname{epi} F$ we have $a \geq \bar{a}$ for every $x \in X$. Thus, since $F(\bar{x}) = \bar{a}$ and we can always take $a = F(x)$, we obtain that $\bar{x}$ is a minimum point for F and the value of the minimum is $\bar{a}$.

Let $F : X \to [-\infty, +\infty]$, $x^* \in X'$, $\alpha \in \mathbb{R}$. The continuous affine function $f(x) := \langle x^*, x\rangle_{X',X} - \alpha$ is everywhere less than F if and only if for every $x \in X$ we have.

$$\alpha \geq \langle x^*, x\rangle_{X',X} - F(x).$$

So the following definition is quite natural:

DEFINITION 4.1 (Convex conjugate). Let $F : X \to [-\infty, +\infty]$, $x^* \in X'$, $\alpha \in \mathbb{R}$. Define the *convex conjugate* $F^* : X' \to [-\infty, +\infty]$ (also called *polar* or Legendre-Fenchel transformed) of F by setting for all $x^* \in X'$

$$F^*(x^*) = \sup_{x\in X} \left\{ \langle x^*, x\rangle_{X',X} - F(x) \right\}.$$

Notice that $F^*(x^*) \leq \alpha$ if and only if the continuous affine function $f(x) := \langle x^*, x\rangle_{X',X} - \alpha$ is everywhere less than F. Equivalently, $(x^*, \alpha) \in \operatorname{epi} F^*$ if and only if $\operatorname{epi} F \subseteq \operatorname{epi} f_{(x^*,\alpha)}$, where $f_{(x^*,\alpha)}(x) = \langle x^*, x\rangle - \alpha$.

Now our problem is: in order to have good regularity properties on F^*, which topology is better to endow X' with? F^* is a pointwise supremum of affine functions $x^* \mapsto \langle x^*, x\rangle_{X',X} - F(x)$ (here x is fixed). A minimum natural requirement is to ask all these functions to be continuous. This amounts to endow X' with the weaker topology such that the maps $x^* \mapsto \langle x^*, x\rangle$ are continuous for every $x \in X$.

From the course in Functional Analysis, we know that this topology is the *weak***-topology* $\sigma(X', X)$. We will review now some basic notions about that topology.

DEFINITION 4.2. Let X be a Banach space, X' its dual and $X'' := (X')'$ its *bidual*, i.e., the dual of X' endowed with his norm as dual space. Then there exists an isometric embedding $J : X \to X''$ defined as follows

$$\langle Jx, x^*\rangle_{X'',X'} := \langle x^*, x\rangle_{X',X}$$

for every $x \in X$ and $x^* \in X'$. J is trivially linear, and

$$\|Jx\|_{X''} = \|x\|_X.$$

If J is surjective, the space X is called *reflexive*.

DEFINITION 4.3. The *weak***-topology* $\sigma(X', X)$ on X' is the smallest topology which make all the functions of the set $J_X := \{Jx : x \in X\}$ be continuous. Since $J_X \subseteq X''$, this topology is weaker than the weak topology $\sigma(X', X'')$ which make continuous all the functions of X''.

PROPOSITION 4.4. *The following properties hold:*

(1) *the weak** *topology is Hausdorff;*
(2) *given* $f \in X'$, *a basis for the set of neighborhoods of* f *in the weak** *topology is given by*

$$V = \{f \in X' : |\langle f - f_0, x_i\rangle_{X',X}| < \varepsilon, \text{ for all } i \in I\},$$

where I *is finite,* $x_i \in X$ *for all* $i \in I$ *and* $\varepsilon > 0$;

Let $\{f_n\}_{n\in\mathbb{N}}$ *be a sequence in* X', $f \in X'$, $\{x_n\}_{n\in\mathbb{N}}$ *be a sequence,* $x \in X$. *Then:*

(1) $f_n \rightharpoonup^* f$ *(i.e.,* f_n *weakly** *converges to* f*) if and only if* $\langle f_n, x\rangle_{X',X} \to \langle f, x\rangle_{X',X}$ *for all* $x \in X$;
(2) *If* $f_n \to f$ *strongly, then* $f_n \rightharpoonup f$ *weakly in* $\sigma(X', X'')$, *and if* $f_n \rightharpoonup f$ *weakly in* $\sigma(X', X'')$ *then* $f_n \rightharpoonup^* f$ *(i.e., weakly***, or in* $\sigma(X', X)$*);*
(3) *If* $f_n \rightharpoonup^* f$, *then* $\|f_n\|$ *is bounded and* $\|f\| \leq \liminf \|f_n\|$;
(4) *If* $f_n \rightharpoonup^* f$ *and* $x_n \to x$ *strongly in* X, *then* $\langle f_n, x_n\rangle \to \langle f, x\rangle$.

PROOF. See **[3]**, Section III.4, in particular Propositions III.10, III.11, III.12 pp. 59–61. □

We recall now the following fundamental theorems:

THEOREM 4.5 (Banach-Alaoglu-Bourbaki). *The closed unit ball* $\mathbb{B}' := \{x^* \in X' : \|x^*\|_{X'} \leq 1\}$ *of* X' *is weakly***-compact.*

PROOF. See Theorem III.15 in **[3]** pp. 64–66. □

THEOREM 4.6 (Kakutani). *The closed unit ball* $\mathbb{B} := \{x \in X : \|x\|_X \leq 1\}$ *of* X *is weakly compact if and only if* X *is reflexive.*

PROOF. See Theorem III.16 in [3] p. 66. □

If we endow X' with the weak* topology, we have the following result.

PROPOSITION 4.7. *Let $F : X \to [-\infty, +\infty]$. The convex conjugate $F^* : X' \to [-\infty, +\infty]$ enjoys the following properties:*

(1) *F^* is always convex and w^*-lower semicontinuous (i.e., l.s.c. with respect to weak*-topology);*
(2) *if F is proper (not necessarily convex), then $F^*(x^*) > -\infty$ for all x^*;*
(3) *if F is convex, l.s.c. and proper, then F^* is proper;*
(4) *if F is proper, for all $x^* \in X'$ and $x \in X$ the* Young's inequality *holds:*

$$\langle x^*, x\rangle_{X',X} \leq F(x) + F^*(x^*).$$

PROOF.

(1) F^* is a pointwise supremum of continuous affine functions, so it is a pointwise supremum of convex l.s.c. functions, so is convex l.s.c.
(2) If F is proper, then $\operatorname{dom} F \neq \emptyset$, so there exists $x \in X$ with $-\infty < F(x) < +\infty$. But then $F^*(x^*) \geq \langle x^*, x\rangle_{X',X} - F(x) > -\infty$.
(3) Since F is proper, convex and l.s.c., we have $F \in \Gamma(X)$, thus there exists a continuous affine minorant $\langle x^*, x\rangle - \alpha$ dove $x^* \in X'$ and $\alpha \in \mathbb{R}$. But then , by using the definition of F^*, we have $F^*(x^*) \leq \alpha$, so F^* is not identically $+\infty$.
(4) Trivial since $F^*(x^*) \geq \langle x^*, x\rangle_{X',X} - F(x)$ for all $x \in X$, $x^* \in X'$, and the sum $F(x) + F(x^*)$ is always well-defined (cannot be $+\infty - \infty$).

□

Immediately from the definition of F^* we have the following properties:

PROPOSITION 4.8. *Let $F, G, F_i : X \to [-\infty, +\infty]$, $i \in I$. Then:*

(1) *$F^*(0) = -\inf_{x\in X} F(x)$;*
(2) *if $F \leq G$ then $F^* \geq G^*$;*
(3) $\left(\inf\limits_{i\in I} F_i\right)^* = \left(\sup\limits_{i\in I} F_i^*\right)$ *and* $\left(\sup\limits_{i\in I} F_i\right)^* \leq \left(\inf\limits_{i\in I} F_i^*\right)$*;*
(4) *$(\lambda F)^*(x^*) = \lambda F^*(x^*/\lambda)$, for all $\lambda > 0$, $x^* \in X'$;*
(5) *defined $(F + \alpha)(x) = F(x) + \alpha$, we have $(F + \alpha)^* = F^* - \alpha$, for all $\alpha \in \mathbb{R}$*
(6) *fix $w \in X$ and set $F_w(x) = F(x - w)$, then $F_w^*(x^*) = F^*(x^*) + \langle x^*, w\rangle_{X',X}$.*

PROOF. Trivial from the definition. For item (3), use Lemma 1.9. □

DEFINITION 4.9 (Bipolar). We can iterate the construction of convex conjugate considering the *bipolar* of $F : X \to [-\infty, +\infty]$, i.e., the function $F^{**} : X'' \to \mathbb{R}$ defined as $F^{**} = (F^*)^*$. We have in particular

$$F^*(x^*) \geq \langle x^*, x\rangle_{X',X} - F(x)$$

which implies

$$F(x) \geq \langle x^*, x\rangle_{X',X} - F^*(x^*) = \langle Jx, x^*\rangle_{X'',X} - F^*(x^*)$$

and by taking the sup on $x^* \in X'$ we have

$$F(x) \geq \sup_{x^*\in X'} \langle Jx, x^*\rangle_{X'',X} - F^*(x^*) = F^{**}(Jx).$$

The following theorem gives a full characterization of the cases in which equality holds.

THEOREM 4.10 (Fenchel-Moreau). *Let X be a normed space and $F : X \to]-\infty, +\infty]$ be not identically $+\infty$. Then $F = F^{**} \circ J$ if and only if F is convex and l.s.c.*

PROOF. Suppose $F = F^{**} \circ J$. We already know that F^{**} is convex. Since J is linear, for $x, y \in X$, $\lambda \in [0, 1]$ we have

$$\begin{aligned} F(\lambda x + (1-\lambda)y) &= F^{**}(J(\lambda x + (1-\lambda)y)) = F^{**}(\lambda Jx + (1-\lambda)Jy) \\ &\leq \lambda F^{**}(Jx) + (1-\lambda)F^{**}(Jy) = \lambda F(x) + (1-\lambda)F(y), \end{aligned}$$

so F is convex. Moreover F^{**} is l.s.c. for weak* topology on X^{**}, so $F = F^{**} \circ J$ is convex and l.s.c. for both the weak and the strong topology on X.

Conversely, assume that F is convex and l.s.c. Since $F^{**} \circ J \le F$, the thesis is trivial if $F^{**}(x) = +\infty$. Since F is proper, F^* is not identically $+\infty$, and so $F^{**} > -\infty$. So we have to prove that if $F^{**}(Jx) \in \mathbb{R}$ we have $F(x) = F^{**}(Jx)$. By contradiction, suppose that there exists $x_0 \in X$ such that $F^{**}(Jx_0) < F(x_0)$ so $(x_0, F^{**}(Jx_0)) \notin \operatorname{epi} F$. We can separate the closed epi F (recalling that F is l.s.c.) and the compact $(x_0, F^{**}(Jx_0))$ in the same way as we did in the characterization of $\Gamma(X)$, obtaining $\ell \in X'$, $\alpha > 0$, $\beta \in \mathbb{R}$ such that

$$\begin{aligned}\ell(x_0) + \alpha F^{**}(Jx_0) < \beta := & \inf_{(x,t)\in \operatorname{epi} F} (\ell(x) + \alpha t) \\ & = \inf_{x \in \operatorname{dom} F} (\ell(x) + \alpha F(x)) = -\alpha \sup_{x\in\operatorname{dom} F} \left(-\frac{\langle \ell, x\rangle_{X',X}}{\alpha} - F(x)\right) \\ & = -\alpha F^*(-\ell/\alpha).\end{aligned}$$

We obtain

$$\ell(x_0) + \alpha F^{**}(Jx_0) < -\alpha F^*(-\ell/\alpha)$$

and so

$$\langle Jx_0, -\frac{\ell}{\alpha}\rangle_{X'',X'} > F^{**}(Jx_0) + F^*(-\ell/\alpha),$$

however by Young's inequality, being $F^{**} > -\infty$, we should have

$$F^{**}(Jx_0) + F^*(-\ell/\alpha) \ge \langle Jx_0, -\frac{\ell}{\alpha}\rangle_{X'',X'},$$

which lead to a contradiction. □

COROLLARY 4.11. *Let X be a normed, $F : X \to [-\infty, +\infty]$ convex but not necessarily l.s.c. Then $F^{**} \circ J = \bar{F}$. In particular, $F^{**} \circ J(x) = F(x)$ if and only if F is l.s.c. at x.*

PROOF. $F^{**} \circ J$ is a l.s.c. minorant of F, so $F^{**} \circ J \le \bar{F}$. On the other hand, if G is a convex l.s.c. functions with $G \le F$ we have $G^* \ge F^*$ e $G^{**} \le F^{**}$ so by Fenchel-Moreau Theorem, we have $G = G^{**} \circ J \le F^{**} \circ J$. By the arbitrariness of G, we have $\bar{F} \le F^{**} \circ J$. □

REMARK 4.12. A consequence of the previous lemma, we have $F^{***} = F^*$, in fact F^* is convex and l.s.c., so $(F^*)^{**}$ coincides with F^* (we identify X' with its image in X''').

We give now some examples of convex conjugate:

EXERCISE 4.13. Let $X = \mathbb{R}$. Prove that

(1) the convex conjugate of $F(x) = |x|$ is $F^*(y) = I_{[-1,1]}(y)$;
(2) the convex conjugate of $G(x) = e^x$ is defined by $G^*(y) = +\infty$ if $y < 0$, $G^*(0) = 0$, and $G^*(y) = y(\log y - 1)$ if $y > 0$;
(3) the convex conjugate of $H(x) = |x|^p/p$ where $1 < p < +\infty$ is $H^*(y) = |y|^q/q$ with $1/p + 1/q = 1$;

SOLUTION.

(1) We have

$$\begin{aligned}F^*(y) = \sup_{x\in\mathbb{R}} \Big\{ yx - |x| \Big\} &= \max \left\{ \sup_{x\ge 0}\{yx - |x|\}, \sup_{x\le 0}\{yx-|x|\} \right\} \\ &= \max\left\{ \sup_{x\ge 0}\{x(y-1)\}, \sup_{x\le 0}\{x(y+1)\}\right\}.\end{aligned}$$

Notice that if $y > 1$ then $\sup\limits_{x\ge 0}\{x(y-1)\} = +\infty$, so $F^*(y) = +\infty$, similarly, if $y < -1$ we have $|y+1| = -(y+1)$, so

$$\sup_{x\le 0}\{x(y+1)\} = \sup_{x\le 0}\{-|x|(y+1)\} = \sup_{x\ge 0}\{|x| \cdot |y+1|\} = +\infty,$$

so $F^*(y) = +\infty$ even in this case. If $-1 \le y \le 1$ we have $x(y-1) \le 0$ for all $x \ge 0$ and the supremum on $x \ge 0$ is attained at $x = 0$ and its value is 0, and similarly $x(y+1) \le 0$ for all $x \le 0$, and the supremum on $x \le 0$ is attained at $x = 0$ and its value is 0. Thus $F^*(y) = 0$ for $|y| \le 1$.

(2) Fix $y \in \mathbb{R}$ and consider $g_y(x) = yx - e^x$, thus $G^*(y) = \sup_{x\in\mathbb{R}} g_y(x)$. We have $\dfrac{d}{dx} g_y(x) = y - e^x$ and $\dfrac{d^2}{dx^2} g_y(x) = -e^x < 0$. This allows to conclude that $g_y(x)$ has at most one critical point, and if $g_y(x)$ has a critical point, then it is a maximum. We have that g_y admits a critical point $\bar{x} \in \mathbb{R}$ if and only if $y = e^{\bar{x}}$, i.e., $\bar{x} = \log y$ and this implies $y > 0$. Since

$$\lim_{x\to+\infty} g_y(x) = -\infty, \qquad \lim_{x\to-\infty} g_y(x) = \begin{cases} +\infty, & \text{if } y < 0, \\ 0, & \text{if } y = 0, , \\ -\infty, & \text{if } y > 0, \end{cases}$$

we can make the following conclusions:
- if $y \le 0$ we have that g_y has no critical points, and it is strictly decreasing since $\dfrac{d}{dx} g_y(x) < 0$. So $G^*(y) = \sup_{x\in\mathbb{R}} g_y(x) = \lim_{x\to-\infty} g_y(x)$, thus $G^*(0) = 0$ and $G^*(y) = +\infty$ if $y < 0$;
- if $y > 0$ we have that g_y attains its unique maximum at $\bar{x} = \log y$, thus $G^*(y) = g_y(\bar{x}) = y \log y - e^{\log y} = y(\log y - 1)$.

(3) Fix $y \in R$ and consider $h_y(x) = yx - \dfrac{1}{p}|x|^p$, thus $H^*(y) = \sup_{x\in\mathbb{R}} h_y(x)$. If $y = 0$ we have $H^*(0) = \sup_{x\in\mathbb{R}} \left\{-\dfrac{1}{p}|x|^p\right\} = 0$, attained at $x = 0$. Assume $y \ne 0$ and notice that given $x \in \mathbb{R}$ with $|x| = \lambda$, $\lambda \ge 0$, we have always

$$h_y\left(\lambda \frac{y}{|y|}\right) = \lambda|y| - \frac{1}{p}\lambda^p = |x||y| - \frac{1}{p}|x|^p \ge h_y(x),$$

and so for $y \ne 0$ we have

$$H^*(y) = \sup_{\lambda\ge 0} \left\{\lambda|y| - \frac{1}{p}\lambda^p\right\}.$$

A quick study of the map $\hat{h}_y(\lambda) = \lambda|y| - \dfrac{1}{p}\lambda^p$ for $\lambda \ge 0$ gives

$$\frac{d}{d\lambda}\hat{h}_y(\lambda) = |y| - \lambda^{p-1},$$

hence $h_y(\lambda)$ attains its unique maximum in $[0, +\infty[$ at $\bar{\lambda} \ge 0$ satisfying $\bar{\lambda} = |y|^{\frac{1}{p-1}}$, so

$$H^*(y) = \hat{h}_y(\bar{\lambda}) = |y|^{1+\frac{1}{p-1}} - \frac{1}{p}|y|^{\frac{p}{p-1}} = \left(1 - \frac{1}{p}\right)|y|^{\frac{p}{p-1}} = \frac{1}{q}|y|^q.$$

REMARK 4.14. The very same idea of (3) in the above exercise can be used to prove that if $X = \mathbb{R}^n$ and we take $F(x) = \frac{1}{p}\|x\|^p$ then $F^*(y) = \frac{1}{q}\|y\|^q$.

EXERCISE 4.15.
(1) Let $X = \mathbb{R}^2$ and $f(x, y) = ax + by + c$, where $a, b, c \in \mathbb{R}$. Find f^*;
(2) Let $X = \mathbb{R}^2$, $f_1(x, y) = a_1x + b_1y + c_1$, $f_2(x, y) = a_2x + b_2y + c_2$, $g(x, y) = \min\{f_1(x, y), f_2(x, y)\}$. Find g^*.

PROOF.

(1) We have to compute
$$\begin{aligned} f^*(x^*,y^*) &= \sup_{(x,y)\in\mathbb{R}^2} \{\langle (x^*,y^*),(x,y)\rangle - (ax+by+c)\} \\ &= \sup_{(x,y)\in\mathbb{R}^2} \{\langle (x^*-a,y^*-b),(x,y)\rangle\} - c \end{aligned}$$
If $x^*=a$ and $y^*=b$ then $f^*(a,b)=-c$. Otherwise, by taking $x=\lambda(x^*-a)$ and $y=\lambda(y^*-b)$, we have
$$\begin{aligned} f^*(x^*,y^*) &\ge \sup_{\lambda\in\mathbb{R}} \left\{\lambda\left[(x^*-a)^2+(y^*-b)^2\right]\right\} - c \\ &= \left[(x^*-a)^2+(y^*-b)^2\right]\sup_{\lambda\in\mathbb{R}}\{\lambda\} - c = +\infty. \end{aligned}$$
Thus $f^*(x^*,y^*) = I_{\{(a,b)\}}(x^*,y^*) - c$.

(2) From the previous item, we have $f_i^*(x^*,y^*) = I_{\{(a_i,b_i)\}}(x^*,y^*) - c_i$, $i=1,2$. Thus
$$\begin{aligned} g^*(x^*,y^*) &= (\min\{f_1,f_2\})^*(x^*,y^*) = \max\{f_1^*(x^*,y^*), f_2^*(x^*,y^*)\} \\ &= \max\{I_{\{(a_1,b_1)\}}(x^*,y^*) - c_1, I_{\{(a_2,b_2)\}}(x^*,y^*) - c_2\}, \end{aligned}$$
and so
$$g^*(x^*,y^*) = \begin{cases} +\infty, & \text{if } (a_1,b_1)\neq(a_2,b_2), \\ \\ I_{\{(a_1,b_1)\}}(x^*,y^*) - \min\{c_1,c_2\}, & \text{if } a_1=a_2 \text{ and } b_1=b_2. \end{cases}$$

□

LEMMA 4.16 (Marginals). *Let X,Y be normed spaces, and $\Phi : X\times Y \to]-\infty,+\infty]$ be a convex function. We define the functions $\Phi_X : Y \to]-\infty,+\infty]$ and $\Phi_Y \to]-\infty,+\infty]$ by setting $\Phi_X(y) = \inf_{x\in X}\Phi(x,y)$ and $\Phi_Y(x) = \inf_{y\in Y}\Phi(x,y)$. Then Φ_X and Φ_Y are convex functions, moreover $\Phi_X^*(y^*) = \Phi^*(0,y^*)$ and $\Phi_Y^*(x^*) = \Phi^*(x^*,0)$. The functions Φ_X and Φ_Y are called* first *and* second *marginal of Φ.*

PROOF. We know that the pointwise supremum of convex l.s.c. functions is a convex l.s.c. function, hence since we deal with a pointwise infimum a proof is required. We will prove only the statements for Φ_X, being the others completely symmetric. Given $\lambda\in[0,1]$, $y_1,y_2\in Y$, $a_1>\Phi_X(y_1)$ and $a_2>\Phi_X(y_2)$, there exists $x_1,x_2\in X$ such that $\Phi_X(y_i)\le\Phi(x_i,y_i)<a_i$, $i=1,2$.
$$\begin{aligned} \Phi_X(\lambda y_1+(1-\lambda)y_2) &\le \Phi(\lambda x_1+(1-\lambda)x_2),\lambda y_1+(1-\lambda)y_2) = \Phi(\lambda(x_1,y_1)+(1-\lambda)(x_2,y_2)) \\ &\le \lambda\Phi(x_1,y_1)+(1-\lambda)\Phi(x_2,y_2) \le \lambda a_1+(1-\lambda)a_2. \end{aligned}$$
By the arbitrariness of a_1 and a_2, we can let $a_i\to\Phi_X(y_i)^+$, $i=1,2$, obtaining convexity. We have
$$\begin{aligned} \Phi_X^*(y^*) &= \sup_{y\in Y}\left\{\langle y^*,y\rangle_{Y',Y} - \Phi_X(y)\right\} = \sup_{y\in Y}\left\{\langle y^*,y\rangle_{Y',Y} - \inf_{x\in X}\Phi(x,y)\right\} \\ &= \sup_{y\in Y}\sup_{x\in X}\left\{\langle y^*,y\rangle_{Y',Y} - \Phi(x,y)\right\} = \sup_{y\in Y}\sup_{x\in X}\left\{\langle (0,y^*),(x,y)\rangle_{X'\times Y',X\times Y} - \Phi(x,y)\right\} \\ &= \Phi^*(0,y^*). \end{aligned}$$

□

DEFINITION 4.17 (Support function). Let $U\subseteq X$ be nonempty, we consider its indicator function I_U. Its convex conjugate is called *support function to U*:
$$\sigma_U(p) := \sup_{u\in U}\langle p,u\rangle_{X',X}.$$
On the other hand, the minimal (w.r.t. inclusion) closed convex set containing U is $\overline{\text{co}}\,U$, so for the previous result we have
$$I_U^{**}\circ J = I_{\overline{\text{co}}\,U}.$$
By passing to the conjugate again, $\sigma_U(p) = \sigma_{\overline{\text{co}}\,U}(p)$.

Summary of Lecture 4

- We have seen that all the information about a proper, l.s.c., convex function is enclosed in the set of its continuous affine minorants.
- More generally, we can define the convex conjugate $F^* : X' \to [-\infty, +\infty]$ of *any* function $F : X \to [-\infty, +\infty]$ by setting $F^*(x^*) = \sup_{x^* \in X'} \{\langle x^*, x\rangle_{X',X} - F(x)\}$. Its interpretation is as follows: $F^*(x^*) \leq \alpha$ if and only if $(x^*, \alpha) \in \operatorname{epi} F^*$ if and only if the continuous affine function $f_{x^*,\alpha} : X \to \mathbb{R}$, defined by $f_{x^*,\alpha}(x) = \langle x^*, x\rangle_{X',X} - \alpha$, is a minorant of F, i.e., $f_{x^*,\alpha}(x) \leq F(x)$ for all $x \in X$.
- The more convenient choice for the topology on X' is to endow X' with the weak* topology. In this case F^* turns out to be always w^*-l.s.c. and convex, *regardless of the choice of F*.
- The class of proper, l.s.c., convex functions behaves well under the operation of conjugation, since the convex conjugate of a proper, l.s.c., convex function is again a proper, l.s.c., convex function.
- We can iterate the conjugation, defining the *biconjugate* $F'' : X'' \to [-\infty, +\infty]$ of F. Recalling that there exists a canonical injection, which is a linear isometry, $J : X \to X''$, it is natural to compare the values of F and $F^{**} \circ J$ on the elements of X. It turns out that $F \geq F^{**} \circ J$, and equality holds if and only if F is convex and l.s.c. (Fenchel-Moreau theorem).

5. Lecture of 15 october 2018: Normal cone and subdifferential of convex analysis (3h)

Consider now a function $f : \mathbb{R}^d \to \mathbb{R}$. If f is differentiable and if x is a critical point for f, then the affine hyperplane in $\mathbb{R}^{d+1}$ defined by the equation $x_{d+1} = f(x)$ is tangent at $(x, f(x))$ to the graph of f (the hyperplanes defined by equations of the type $x_{d+1} =$ cost. will be often called *horizontal hyperplanes*, this terminology being imprecise but suggestive).

If f is a convex differentiable function, and x is a critical point for f, then epi f is all contained in the half space $\{(y, \beta) : y \in \Omega, \beta \geq f(x)\}$ thus a posteriori *for convex differentiable function, the notion of critical point and point of minimum coincide*.

If f is convex, not necessarily differentiable at x, but there exists an horizontal supporting hyperplane to epi f at $(x, f(x))$, then epi f is all contained in the half space $\{(y, \beta) : y \in \Omega, \beta \geq f(x)\}$, and so x is a point of minimum.

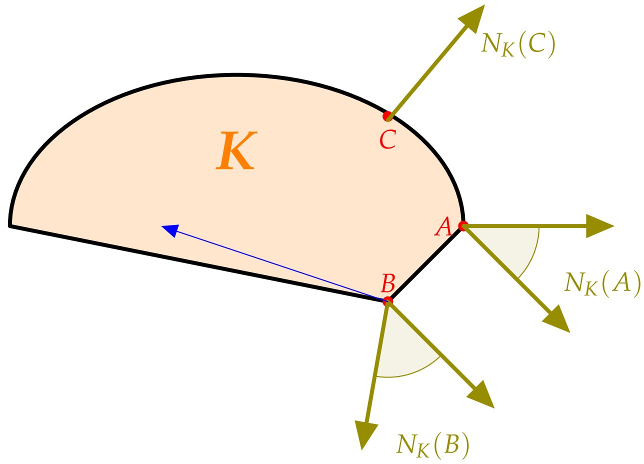
It is thus natural to associate at every $(x, f(x))$ the set of supporting hyperplanes to epi f at $(x, f(x))$. Each of such hyperplane is completely determined by the direction of its normal, whose orientation is choosen in order to point towards the half space not containing epi f, thus we associate to $(x, f(x))$ the set of all the normals of all the supporting hyperplanes to epi f passing by $(x, f(x))$ and pointing towards the half space not containing epi f.

More generally, given a closed convex set C, we can associate to each $x \in C$ the set of the normals to the supporting hyperplanes to C passing through x and pointing towards the half space not containing C.

DEFINITION 5.1 (Normal cone in the sense of convex analysis). Let X be a normed space, C be a closed convex nonempty subset of X, $x \in C$. We define the *normal cone in the sense of convex analysis* to C at x by setting:

$$N_C(x) := \{v \in X' : \langle v, y - x\rangle_{X',X} \leq 0 \text{ for all } y \in C\} = \{v \in X' : \sigma_C(v) = \langle v, x\rangle_{X',X}\}.$$

The normal cone $N_C(x)$ is trivially a w^*-closed, convex and nonempty subset of X' and if $\lambda > 0$, $v \in N_C(x)$ then $\lambda v \in N_C(x)$.



Geometrical interpretation of the normal cone. The scalar product between every element of the cone at x and every segment joining x to any other point of K must be nonpositive.

REMARK 5.2.

(1) To recover the geometrical interpretation, let $v \in N_C(x)$, $v \neq 0$. Set $\alpha = \langle v, x\rangle_{X',X}$ and consider the hyperplane H of equation $\langle v, y\rangle = \alpha$. Clearly, we have $x \in H$, moreover by definition
$$\langle v, y - x\rangle_{X',X} = \langle v, y\rangle_{X',X} - \alpha \leq 0, \text{ for all } y \in C,$$
hence $C \subseteq \{y \in X : \langle v, y\rangle \leq \alpha\}$.

(2) If X is an Hilbert space and C is a closed convex nonempty subset of H, then by the projection theorem we have always $z - \pi_C(z) \in N_C(\pi_C(z))$.

(3) Notice that the action of the vector $v \in N_C(x)$ (which should be imagined as applied at x) on all the vectors $y - x$ (that also must be imagined as applied at x) yields a real nonpositive number for all $y \in C$.

REMARK 5.3. If $C = \operatorname{epi} F$, with $F : X \to]-\infty, +\infty]$ convex and l.s.c., given $(x, \beta) \in \operatorname{epi} F$ we have
$$N_C(x, \beta) := \{v \in X' \times \mathbb{R} : \langle v, (y, \alpha) - (x, \beta)\rangle_{X'\times\mathbb{R}, X\times\mathbb{R}} \leq 0 \text{ for all } (y, \alpha) \in \operatorname{epi} F\}.$$
In particular, if we choose $y = x$, $\alpha \geq f(x) = \beta$, we have that if $v = (v_x, \xi) \in N_C(x, \beta)$ where $v_x \in X'$ and $\xi \in \mathbb{R}$ then necessarily $\xi \leq 0$. Thus if $\xi \neq 0$ we have $\left(\dfrac{v_x}{|\xi|}, -1\right) \in N_C(x, \beta)$.

DEFINITION 5.4 (Subdifferential of convex analysis). Let $F : X \to]-\infty, +\infty]$ a convex l.s.c. function, $x \in \operatorname{dom} F$. We define the *subdifferential in the sense of convex analysis* by setting
$$\partial F(x) := \{v_x \in X' : (v_x, -1) \in N_{\operatorname{epi} F}(x, F(x))\}.$$
If $x \notin \operatorname{dom} F$, we will set $\partial F(x) = \emptyset$. Equivalently, $v_x \in \partial F(x)$ if and only if $x \in \operatorname{dom} F$ and
$$F(y) - F(x) \geq \langle v_x, y - x\rangle_{X',X}, \text{ for all } y \in X.$$
We will set $\operatorname{dom} \partial F := \{x \in X : \partial F(x) \neq \emptyset\}$.

REMARK 5.5. If f is classically (Fréchet) differentiable at x, then $\partial f(x) = \{f'(x)\}$. Indeed, if f is differentiable at x then it is continuous at x, thus subdifferentiable at x. Moreover, given $\lambda \in [0, 1]$ sufficiently small, we have for all $y \neq x$
$$f(x + \lambda(y - x)) \leq f(x) + \lambda[f(y) - f(x)],$$
leading to
$$\frac{f(x + \lambda(y - x)) - f(x) - \langle f'(x), \lambda(y - x)\rangle}{\lambda|y - x|} \leq \frac{f(y) - f(x) - \langle f'(x), y - x\rangle}{|y - x|}.$$
By letting $\lambda \to 0^+$, the left hand side vanishes by definition of differential. Multiplying by $|y - x|$, we obtain
$$f(y) - f(x) \geq \langle f'(x), y - x\rangle,$$

hence $f'(x) \in \partial f(x)$. Conversely, given $\xi \in \partial f(x)$, we have for all $t \in \mathbb{R}$ with $|t|$ sufficiently small that

$$f(x + t(y - x)) - f(x) - t\langle \xi, y - x\rangle \geq 0.$$

Assuming $t > 0$, dividing by $t|y - x| > 0$ and letting $t \to 0^+$, setting $z = x + t(y - x)$ we have

$$\liminf_{z\to 0} \frac{f(z) - f(x) - \langle \xi, z - x\rangle}{|z - x|} \geq 0.$$

Assuming $t < 0$ and performing the same computations, we have

$$\limsup_{z\to 0} \frac{f(z) - f(x) - \langle \xi, z - x\rangle}{|z - x|} \leq 0.$$

Hence

$$\lim_{z\to 0} \frac{f(z) - f(x) - \langle \xi, z - x\rangle}{|z - x|} = 0,$$

and by the uniqueness of the differential, we have $\xi = f'(x)$.

EXAMPLE 5.6 (Subdifferential of the norm). Let X be a normed real space, and consider $p : X \to \mathbb{R}$, $p(x) = \|x\|_X$. Then $\partial p(0) = \overline{B_{X'}(0, 1)}$. Indeed, we have that $\varphi \in X'$ belongs to $\partial p(0)$ if and only if $p(x) \geq p(0) + \langle \varphi, x\rangle$ for all $x \in X$, hence $\|x\|_X \geq \langle \varphi, x\rangle$. But we have also $p(-x) \geq p(0) + \langle \varphi, -x\rangle$ for all $x \in X$, thus $\|x\|_X \geq -\langle \varphi, x\rangle$. So $\varphi \in \partial p(0)$ if and only if $|\langle \varphi, x\rangle| \leq \|x\|_X$, therefore if and only if $\varphi \in \overline{B_{X'}(0, 1)}$.

PROPOSITION 5.7 (Properties of the subdifferential). *Let X be a normed space, $f : X \to \mathbb{R} \cup \{+\infty\}$ be a proper convex function, $x_0 \in \operatorname{dom} f$.*

(1) *if $\partial f(x_0) \neq \emptyset$, then $\partial f(x_0)$ is convex and w^*-closed in X';*
(2) *assume f l.s.c. and let $\{(x_n, \xi_n)\}_{n\in\mathbb{N}} \subseteq X \times X'$ be such that $x_n \to x$ strongly in X, $\xi_n \rightharpoonup^* \xi$ weakly* in X', and $\xi_n \in \partial f(x_n)$ for all $n \in \mathbb{N}$, then $\xi \in \partial f(x)$;*
(3) *if f is continuous at x_0 then $\partial f(x_0)$ is bounded, thus w^*-compact;*
(4) *$f(x_0) = \min_{x\in X} f(x)$ if and only if $0 \in \partial f(x_0)$;*
(5) *if $\partial f(x_0) \neq \emptyset$ then $f(x_0) = (f^{**} \circ J)(x_0)$;*
(6) *if $f(x_0) = f^{**}(J_{x_0})$ then $\partial f(x_0) = \partial(f^{**} \circ J)(x_0)$;*
(7) *we have $\varphi_0 \in \partial f(x_0)$ if and only if $f(x_0) + f^*(\varphi_0) = \langle \varphi_0, x_0\rangle_{X',X}$. In this case $\varphi_0 \in \operatorname{dom} f^*$;*
(8) *we have $\varphi_0 \in \partial f(x_0)$ if and only if $Jx_0 \in \partial f^*(\varphi_0)$ and $f(x_0) = f^{**} \circ J(x_0)$.*
(9) *if f is proper, convex, l.s.c. and X is reflexive, then $p^* \in \partial f(x)$ if and only if $x \in \partial f^*(p^*)$.*

PROOF.

(1) Given $\xi_1, \xi_2 \in \partial f(x_0)$, $\lambda \in [0, 1]$ we have

$$\begin{aligned} f(x) - f(x_0) &= \lambda(f(x) - f(x_0)) + (1 - \lambda)(f(x) - f(x_0)) \\ &\geq \lambda\langle \xi_1, x - x_0\rangle + (1 - \lambda)\langle \xi_2, x - x_0\rangle = \langle \lambda\xi_1 + (1 - \lambda)\xi_2, x - x_0\rangle, \end{aligned}$$

thus $\lambda\xi_1 + (1 - \lambda)\xi_2 \in \partial f(x_0)$, which therefore turns out to be convex. Given $\{\xi_n\}_{n\in\mathbb{N}} \subseteq \partial f(x_0)$ w^*-converging to $\xi \in X'$, we have $\langle \xi_n, x - x_0\rangle \to \langle \xi, x - x_0\rangle$ for all $x \in X$, therefore $f(x) - f(x_0) \geq \langle \xi_n, x - x_0\rangle \to \langle \xi, x - x_0\rangle$, which yields the w^*-closure.

(2) For all $y \in X$ we have

$$\begin{aligned} f(y) - f(x_n) &\geq \langle \xi_n, y - x_n\rangle = \langle \xi_n - \xi, y - x_n\rangle + \langle \xi, y - x_n\rangle \\ &= \langle \xi_n - \xi, y - x\rangle + \langle \xi_n - \xi, x - x_n\rangle + \langle \xi, y - x_n\rangle \\ &\geq \langle \xi_n - \xi, y - x\rangle - \|\xi_n - \xi\|_{X'} \cdot \|x - x_n\|_X + \langle \xi, y - x_n\rangle \end{aligned}$$

The first term of the last line vanishes as $n \to +\infty$ by definition of weak* convergence. The second one vanishes since, by the properties of w^*-convergence, we have that $\|\xi_n - \xi\|_{X'}$ remains bounded (and $\|x_n \to x\|_X \to 0$ by assumption). The third term converges to $\langle \xi, y - x\rangle$ since $\xi \in X'$ thus it is continuous. So we have for all $y \in X$

$$f(y) - f(x) \geq f(y) - \liminf_{n\to+\infty} f(x_n) = \limsup_{n\to\infty} f(y) - f(x_n) \geq \langle \xi, y - x\rangle,$$

therefore $\xi \in \partial f(x)$.

(3) We have already proved the w^*-closure of the subdifferential. According to Banach-Alaoglu theorem, to prove its w^* compactness it remains to prove its boundedness. By continuity of f at x_0, fixed $\varepsilon = 1$ there exists $\delta > 0$ such that if $\|x - x_0\|_X \leq \delta$ then $|f(x) - f(x_0)| \leq 1$. Let $\varphi \in \partial f(x_0)$. Then for all $x \in \overline{B(x_0, \delta)}$ we have
$$\langle \varphi, x - x_0 \rangle \leq f(x) - f(x_0) \leq |f(x) - f(x_0)| \leq 1$$
and so $\sup\limits_{\|u\| \leq \delta} |\langle \varphi, u \rangle| \leq 1$, therefore $\|\varphi\|_{X'} \leq 1/\delta$.

(4) Trivial.

(5) Let $\xi \in \partial f(x_0)$ and set $\psi(x) = f(x_0) + \langle \xi, x - x_0 \rangle$. We have that ψ is a continuous affine minorant of f such that $\psi(x) \leq f(x)$ at all $x \in X$ and $\psi(x_0) = f(x_0)$. Since $f^{**} \circ J = \overline{f}$ is the largest lower semicontinuous minorant of f, we must have $\psi(x) \leq f^{**} \circ J(x) \leq f(x)$, so at $x = x_0$ we have equality.

(6) Let $\varphi \in \partial(f^{**} \circ J)(x_0)$. Then
$$f(x) \geq f^{**} \circ J(x) \geq f^{**} \circ J(x_0) + \langle \varphi, x - x_0 \rangle = f(x_0) + \langle \varphi, x - x_0 \rangle,$$
hence $\varphi \in \partial f(x_0)$, therefore $\partial(f^{**} \circ J)(x_0) \subseteq \partial f(x_0)$. Conversely, let $\varphi \in \partial f(x_0)$. This implies $f(x_0) \in \mathbb{R}$. Then
$$f(x) \geq f(x_0) + \langle \varphi, x - x_0 \rangle = f^{**} \circ J(x_0) + \langle \varphi, x - x_0 \rangle =: \psi(x),$$
Since ψ is a continuous affine minorant of f, we have $f^{**} \circ J(x) \geq \psi(x)$ thus $\varphi \in \partial(f^{**} \circ J)(x_0)$, hence $\partial(f^{**} \circ J)(x_0) \supseteq \partial f(x_0)$. Thus equality holds.

(7) By definition, $\varphi_0 \in \partial f(x_0)$ if and only if for all $x \in X$ we have $f(x) \geq f(x_0) + \langle \varphi_0, x - x_0 \rangle$, which is equivalent to
$$\langle \varphi_0, x_0 \rangle - f(x_0) \geq \langle \varphi_0, x \rangle - f(x).$$
By taking the sup on $x \in X$ we have
$$\langle \varphi_0, x_0 \rangle - f(x_0) \geq f^*(\varphi_0).$$
On the other hand, by definition we have $\langle \varphi_0, x_0 \rangle - f(x_0) \leq f^*(\varphi_0)$, thus equality holds.

(8) Suppose $Jx_0 \in \partial f^*(\varphi_0)$ and $f(x_0) = f^{**} \circ J(x_0)$. Applying (7) to f^* we have
$$f^*(\varphi_0) + f^{**}(Jx_0) = \langle Jx_0, \varphi_0 \rangle$$
and so, recalling the assumptions,
$$f^*(\varphi_0) + f(x_0) = \langle \varphi_0, x_0 \rangle,$$
thus by (7) applied to f we have $\varphi_0 \in \partial f(x_0)$. Conversely, let $\varphi_0 \in \partial f(x_0)$. By (7) we have
$$f^*(\varphi_0) + f(x_0) = \langle \varphi_0, x_0 \rangle,$$
and recalling (4) we have $f(x_0) = f^{**} \circ J(x_0)$ thus
$$f^*(\varphi_0) + f^{**}(Jx_0) = \langle Jx_0, \varphi_0 \rangle,$$
and we conclude by (7) that $Jx_0 \in \partial f^*(\varphi_0)$.

(9) Trivial by (8).

□

PROPOSITION 5.8 (Subdifferentiability criterion). *Let X be a normed space and $f : X \to \mathbb{R} \cup \{+\infty\}$ be a convex function. If there exists $x_0 \in X$ such that f is continuous at x_0, then $\partial f(x) \neq \emptyset$ for all $x \in \operatorname{int}\operatorname{dom} f$ and in particular $\partial f(x_0) \neq \emptyset$.*

PROOF. By continuity, f is upper bounded in a neighborhood of x_0, hence $x_0 \in \operatorname{int}\operatorname{dom} f$. Moreover, $\operatorname{epi} f$ is a convex set with nonempty interior, thus if admits a supporting hyperplane at every point of the boundary. So given $x \in \operatorname{int}\operatorname{dom} f$, we have that there exists $(0, 0) \neq (v_x, \alpha) \in X' \times [0, +\infty[$ such that
$$\langle (v_x, \alpha), (y, f(y)) - (x, f(x)) \rangle \leq 0, \text{ for all } y \in \operatorname{dom} f.$$

If $\alpha = 0$ we would have $\langle v_x, y - x\rangle \leq 0$ for all $y \in \operatorname{dom} f$ and since $x \in \operatorname{intdom} f$ we can choose $\delta > 0$ such that $y - x \in B(0, \delta)$, thus obtaining $v_x = 0$, a contradiction. Hence $\alpha > 0$ and so $v_x/|\alpha| \in \partial f(x)$. □

Summary of Lecture 5

- In the last lectures, from the hyperplanes defining half-spaces containing a given closed convex set, we moved to the study of continuous affine minorants to convex l.s.c. functions, defining the convex conjugate. In this lecture, we concentrate on a special subclass of such hyperplanes, namely, the supporting hyperplanes. Our aim is to find the corresponding objects for functions.
- A sufficient condition ensuring the existence of supporting hyperplanes at every point of the boundary of a convex set is the nonemptiness of its interior.
- The information about the supporting hyperplane at a point of a convex set is enclosed in the *normal cone in the sense of convex analysis*, while the corresponding object for functions is the *subdifferential in the sense of convex analysis*.
- The subdifferential of the convex analysis $\partial F(x)$, roughly speaking, is the set of the slopes of affine continuous minorants of F passing through $(x, F(x))$, and if F is classically differentiable, it reduced to the usual differential.
- We proved several properties of the subdifferential, which follow directly from the definition. Among them, we recall
  - the strong-weak* closure of its graph,
  - its boundedness, hence w^*-compactness, at every continuity point,
  - $0 \in \partial F(x)$ is a *necessary and sufficient* condition for having a *global* minimum of f at x,
  - the *inversion formula* $p^* \in \partial f(x)$ if and only if $x \in \partial f^*(p^*)$, for every proper, convex, l.s.c. function f on a *reflexive* space
- In analogy with the criterion of existence of supporting hyperplane, we can give a sufficient condition for the subdifferentiability of a function F in the interior of its domain, by asking that $\operatorname{epi} F$ has nonempty interior. For instance, it is enough to ask the continuity of F at least at a point of the domain.

6. Lecture of 19 october 2018: Subdifferential calculus and minimization problems (3h)

PROPOSITION 6.1 (Subdifferential calculus). *Let X be a real normed space, $f, g : X \to]-\infty, +\infty]$ be proper convex functions. Then*

(1) *if $\lambda \geq 0$ we have $\partial(\lambda f)(x_0) = \lambda \partial f(x_0)$;*
(2) *$\partial(f + g)(x) \supseteq \partial f(x) + \partial g(x)$, the inclusion may be strict;*
(3) *if there exists $x_0 \in \operatorname{dom}(f) \cap \operatorname{dom}(g)$ such that f is continuous at x_0 then $\partial(f + g)(x) = \partial f(x) + \partial g(x)$ for all $x \in \operatorname{dom}(f) \cap \operatorname{dom}(g)$.*
(4) *let Y be a normed space, $\bar{y} \in Y$, $\Lambda : Y \to X$ be linear and continuous, g be continuous and finite at $\Lambda(\bar{y})$. Then $\partial(g \circ \Lambda)(y) = \Lambda^* \partial g(\Lambda y)$ for all $y \in Y$.*

PROOF.
(1) Given $\lambda \geq 0$, $\xi \in X'$, we have $f(x) \geq f(x_0) + \langle \xi, x - x_0\rangle$ if and only if $\lambda f(x) \geq \lambda f(x_0) + \langle \lambda\xi, x - x_0\rangle$, thus $\xi \in \partial f(x_0)$ if and only if $\lambda\xi \in \partial(\lambda f)(x_0)$.
(2) Let $\xi_f \in \partial f(x_0)$, e $\xi_g \in \partial g(x_0)$. Then
$$\begin{aligned}(f+g)(x) - (f+g)(x_0) &= (f(x) - f(x_0)) + (g(x) - g(x_0)) \\ &\geq \langle \xi_f, x - x_0\rangle + \langle \xi_g, x - x_0\rangle = \langle \xi_f + \xi_g, x - x_0\rangle.\end{aligned}$$
Therefore $\xi_f + \xi_g \in \partial(f + g)(x_0)$. However the inclusion $\partial(f + g)(x) \supseteq \partial f(x) + \partial g(x)$ may be strict, as shown in the example below.
(3) According to (2), we have to prove that given $\varphi \in \partial(f + g)(x)$ there exist $\varphi_1 \in \partial f(x)$ and $\varphi_2 \in \partial g(x)$ such that $\varphi = \varphi_1 + \varphi_2$. By assumption, for all $y \in X$ we have
$$f(y) + g(y) - f(x) - g(x) \geq \langle \varphi, y - x\rangle.$$

Define

$$A := \{(y,a) \in X \times \mathbb{R} : f(y) - f(x) - \langle \varphi, y-x \rangle \le a\}$$
$$B := \{(z,b) \in X \times \mathbb{R} : b \le g(x) - g(z)\},$$

and notice that they are nonempty, for instance we have $(x_0, 0) \in A \cap B$. Since the map $y \mapsto \varphi(y) := f(y) - f(x) - \langle \varphi, y - x\rangle$ is convex and $A = \operatorname{epi}\varphi$, we have that A is convex, moreover, since φ is continuous at x_0, we have $\operatorname{int} A \neq \emptyset$. Since f must be upper bounded in a neighborhood V of x_0 by continuity, we have that also $\varphi(\cdot)$ is upper bounded in a neighborhood of x_0, thus there is a constant a_{x_0} such that $(y, a_{x_0}) \in \operatorname{int} A$ for all $y \in V$. In particular, $(x_0, a_{x_0}) \in \operatorname{int} A$. In the same way, by convexity $z \mapsto g(z) - g(x)$, we have that B is convex.

We prove now that $\operatorname{int} A \cap B = \emptyset$. Assume $(y,a) \in \operatorname{int} A \cap B$. Then we have that there exists $\varepsilon > 0$ such that $(y, a-\varepsilon) \in A$, and so we have $f(y) - f(x) - \langle \varphi, y-x\rangle \le a - \varepsilon$ and $a \le g(x) - g(y)$, yielding $f(y) + g(y) - f(x) - g(x) \le \langle \varphi, y - x\rangle - \varepsilon$, contradicting $\varphi \in \partial(f+g)(x)$.

By Hahn-Banach Theorem, we can separate $\operatorname{int} A$ and B by a closed affine hyperplane, thus there exist $\psi \in X'$, $t \in \mathbb{R}$, $\beta \in \mathbb{R}$ such that $(\psi, t) \neq (0,0)$ and

$$\psi(z) + tb \le \beta \le \psi(y) + ta, \text{ for all } (z,b) \in B \text{ and } (y,a) \in \operatorname{int} A.$$

Since A is an epigraph, we can send $a \to +\infty$, thus we must have $t \ge 0$. We prove that $t > 0$ by contradiction. Assume $t = 0$, hence necessarily $\psi \neq 0$ and $\psi(z) \le \beta \le \psi(y)$ for all $(z,b) \in B$ and $(y,a) \in \operatorname{int} A$. We choose $z = x_0$, $b = g(x_0) - g(x)$, $y = x_0$, $a = a_{x_0}$, recalling that $(x_0, a_{x_0}) \in \operatorname{int} A$, obtaining $\beta = \psi(x_0)$, and so we must then have $\psi(x_0) \le \psi(y)$, i.e. $0 \le \psi(y - x_0)$. Since $(y, a_{x_0}) \in \operatorname{int} A$ for all y in a neighborhood of x_0 we have that $0 \le \psi(\eta)$ for all η in a neighborhood of 0. But this implies $\psi = 0$, because if $\psi(\eta) > 0$ for some $\eta \in B(0,\delta)$ then $\psi(-\eta) < 0$ and $-\eta \in B(0,\delta)$, leading to a contradiction, hence $t > 0$.

Since $(x, 0) \in A \cap B$, by applying the separation inequality with $z = y = x$ and $a = b = 0$ we obtain $\beta = \psi(x)$. If we take now the separation inequality with $b = g(z) - g(x)$ and $a = f(y) - f(x) - \langle \varphi, y - x\rangle$, we obtain

$$g(z) - g(x) \ge \frac{1}{t}[\psi(x) - \psi(z)] = \langle \frac{\psi}{t}, z - x\rangle$$
$$f(y) - f(x) \ge \frac{1}{t}\frac{\psi(x) - \psi(y)}{t} + \langle \varphi, y - x\rangle = \langle \varphi - \frac{\psi}{t}, y - x\rangle.$$

Setting $\xi_f = \varphi - \dfrac{\psi}{t} \in \partial f(x)$, $\xi_g = \dfrac{\psi}{t} \in \partial g(x)$, we have that $\varphi = \xi_f + \xi_g \in \partial f(x) + \partial g(x)$.

(4) For all $z, y \in Y$, $\xi \in \partial g(\Lambda y)$ we have

$$g \circ \Lambda(z) - g \circ \Lambda(y) = g(\Lambda z) - g(\Lambda y) \ge \langle \xi, \Lambda z - \Lambda y\rangle = \langle \xi, \Lambda(z - y)\rangle = \langle \Lambda^*\xi, z - y\rangle,$$

hence $\Lambda^*\xi \in \partial(g \circ \Lambda)(y)$, so $\Lambda^*\partial g(\Lambda y) \subseteq \partial(g\circ\Lambda)(y)$.

Conversely, let $\xi \in \partial(g \circ \Lambda)(y)$, hence

$$g \circ \Lambda(z) \ge g \circ \Lambda(y) + \langle \xi, z - y\rangle, \text{ for all } z \in Y.$$

Define:

$$A := \{(x,a) \in X \times \mathbb{R} : g(x) \le a\} := \operatorname{epi} g,$$
$$B := \{(\Lambda z, b) \in X \times \mathbb{R} : \langle \xi, z - y\rangle + g \circ \Lambda(y) \ge b, \text{ for all } z \in Y\}.$$

As in (3), we have that A, B are closed, convex and nonempty, moreover $\operatorname{int} A \neq \emptyset$ and $\operatorname{int} A \cap B = \emptyset$, By Hahn-Banach Theorem, we can separate $\operatorname{int} A$ and B by a closed affine hyperplane, thus there exist $\psi \in X'$, $t \in \mathbb{R}$, $\beta \in \mathbb{R}$ such that $(\psi, t) \neq (0,0)$ and

$$\langle \psi, \Lambda z\rangle_{X',X} + tb \le \beta \le \langle \psi, x\rangle_{X',X} + ta, \text{ for all } (\Lambda z, b) \in B \text{ and } (x,a) \in A.$$

Exactly as before, we can prove that $t > 0$.

If we take now the separation inequality with $b = \langle \xi, z - y \rangle + g \circ \Lambda(y)$, and recall that $\langle \psi, \Lambda z \rangle_{X',X} = \langle \Lambda^* \psi, z \rangle_{Y',Y}$ we have

$$\langle \Lambda^* \psi, z \rangle_{Y',Y} + t(\langle \xi, z - y \rangle_{Y',Y} + g \circ \Lambda(y)) \leq \beta$$

thus for all $z \in Y$ we must have

$$\langle \frac{\Lambda^* \psi}{t} + \xi, z \rangle_{Y',Y} \leq \frac{\beta}{t} - g \circ \Lambda(y) + \langle \xi, y \rangle_{Y',Y},$$

and this implies $\Lambda^* \frac{\psi}{t} + \xi = 0$, otherwise we can send the left hand side to $+\infty$ by choosing a suitable sequence $\{z_n\}_{n \in \mathbb{N}} \subseteq Y$ with $|z_n| \to +\infty$, while the right hand side is bounded. We have then

$$0 \leq \frac{\beta}{t} - g \circ \Lambda(y) + \langle -\Lambda^* \frac{\psi}{t}, y \rangle_{Y',Y},$$

thus the least $\beta \in \mathbb{R}$ that we can take is

$$\beta = tg \circ \Lambda(y) - \langle -\frac{\psi}{t}, \Lambda y \rangle_{X',X}.$$

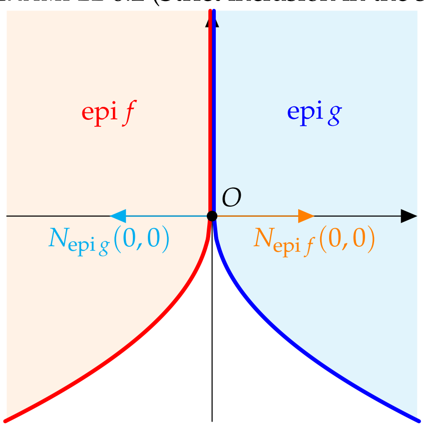
The other separation inequality with $a = g(x)$ yields $\psi(x) + tg(x) \geq \beta$, thus

$$g(x) - g(\Lambda y) \geq \langle -\frac{\psi}{t}, x - \Lambda y \rangle_{X',X}, \text{ for all } x \in X,$$

and so $\frac{-\psi}{t} \in \partial g(\Lambda y)$. We conclude that $\xi = \Lambda^* \left(\frac{-\psi}{t} \right) \in \Lambda^* \partial g(\Lambda y)$, thus $\Lambda^* \partial g(\Lambda y) \supseteq \partial (g \circ \Lambda)(y)$.

□

EXAMPLE 6.2 (Strict inclusion in the subdifferential of the sum).



Let $f(x) = -\sqrt{|x|}(1 + I_{]-\infty,0]})(x)$ and $g(x) = f(-x)$.
We have $0 \in \operatorname{dom} f \cap \operatorname{dom} g$ and $\partial f(0) = \partial g(0) = \emptyset$, thus $\partial f(0) + \partial g(0) = \emptyset$.
However $f(x) + g(x) = I_0(x)$, in particular $f + g$ has a minimum at 0, hence $0 \in \partial(f + g)(0)$.
Therefore $\partial(f + g)(x) \neq \emptyset = \partial f(x) + \partial g(x)$.

EXAMPLE 6.3 (Normal cone as subdifferential). Let X be a Banach space, $K \subseteq X$ closed and convex. Then $\partial I_K(x) = N_K(x)$ if $x \in K$, otherwise $\partial I_K(x) = \emptyset$. Indeed, if $x \neq K$ we $I_K(x) = +\infty$, thus $\partial I_K(x) = \emptyset$. Let $x \in K$, hence $I_K(x) = 0$. We have $\varphi \in \partial I_K(x)$ if and only if

$$I_K(y) - I_K(x) = I_K(y) \geq \langle \varphi, y - x \rangle, \text{ for all } y \in X.$$

If $y \notin K$, the inequality holds for every φ, but if $y \in K$ we must have $\langle \varphi, y - x \rangle \leq 0$, hence $\varphi \in N_K(x)$.

EXAMPLE 6.4 (Subdifferential of the norm squared). Given an Hilbert space Z, $v \in Z$, and defined $w_v(x) = \frac{1}{2}\|x - v\|_H^2$, we want to compute $\partial w_v(x)$. We have that $\xi \in \partial w_v(x)$ if and only if for all $y \in H$

$$\frac{1}{2}\|y - v\|_H^2 \geq \frac{1}{2}\|x - v\|_H^2 + \langle \xi, y - x \rangle_H,$$

equivalently,

$$\frac{1}{2}\|y - v\|_H^2 \geq \frac{1}{2}\|x - v\|_H^2 + \langle \xi, (y - v) - (x - v) \rangle_H,$$

hence
$$\frac{1}{2}\|y-v\|_H^2 - \langle \xi, y-v\rangle_H + \frac{1}{2}\|\xi\|_H^2 \geq \frac{1}{2}\|x-v\|_H^2 - \langle \xi, x-v\rangle_H + \frac{1}{2}\|\xi\|_H^2,$$
thus we must have $\|y-v-\xi\|_H^2 \geq \|x-v-\xi\|_H^2$ for all $y \in H$. We notice that this relation is satisfied if $\xi = x-v$, while if we take $\xi \neq x-v$ and $y = \xi + v$ it fails. We conclude that $\partial w_v(x) = x - v$.

LEMMA 6.5 (Jensen's inequality). *Let $(\Omega, \mathscr{M}, \mu)$ be a measure space such that $\mu(\Omega) = 1$. Given $g \in L^1_\mu(\Omega;\mathbb{R})$ and a convex function $\Phi : \mathbb{R} \to \mathbb{R}$ we have*
$$\Phi\left(\int_\Omega g(x)\,d\mu(x)\right) \leq \int_\Omega \Phi \circ g(x)\,d\mu(x).$$

PROOF. Since Φ is defined on the whole of $\mathbb{R}$, in particular it is continuous at $z_0 := \int_\Omega g(x)\,d\mu(x)$, so it is subdifferentiable at z_0 and we have $a \in \mathbb{R}$ such that $\Phi(z) \geq \Phi(z_0) + a(z - z_0)$ for all $z \in \mathbb{R}$, in particular by taking $z = g(x)$ we have
$$\Phi \circ g(x) \geq \Phi(z_0) + a(g(x) - z_0) = \Phi\left(\int_\Omega g(x)\,d\mu(x)\right) + a\left(g(x) - \int_\Omega g(x)\,d\mu(x)\right).$$
Integrating in x w.r.t. μ, and recalling that $\mu(\Omega) = 1$, the thesis follows. □

We will face the problem of the *existence* of solutions of $\inf_{u \in X} F(u)$ where X is a normed space, $F : X \to]-\infty, +\infty]$ a proper, convex, l.s.c. function.
We notice that in Banach reflexive space the following result holds.

PROPOSITION 6.6 (Coercivity in reflexive spaces). *Let X be a Banach reflexive space, $K \subset X$.*

(1) *If K is bounded closed and convex, then it is weak compact (i.e., compact in $\sigma(X, X')$)*
(2) *If K is closed convex (and nonempty) and F is a proper l.s.c. convex functions such that:*
   (a) *if K is unbounded it holds* $\lim_{\substack{\|x\|\to\infty \\ x \in K}} F(x) = +\infty$,
   (b) *otherwise if K is bounded no additional assumptions,*

   then F admits a minimum point in K.

PROOF.

(1) We recall that the ball is weak compact in reflexive spaces by Kakutani Theorem. Since K is convex, weak and strong closure coincides, so K is weakly closed. Being bounded, it is contained in a closed ball, so it is a closed subset of a weak compact set, and hence it is compact.
(2) If F is identically $+\infty$ on K, the result is trivial, otherwise there exists $x \in K$, $\lambda \in \mathbb{R}$ such that $F(x) = \lambda$. But then
$$\inf_{x \in K} F(x) = \inf_{x \in C} F(x),$$
where $C := K \cap F^{-1}(]-\infty, \lambda])$. C is convex and weakly closed, since F is l.s.c. and convex, so its sublevels are convex and closed w.r.t. both strong and weak topology, and K is closed and convex, so weakly closed. If K is bounded, we have that C is convex and closed w.r.t. both strong and weak topology, so weakly compact and so we can apply Tonelli-Weierstrass Theorem to conclude. Otherwise, we notice that by assumption there exists $R > 0$ such that if $\|x\| > R$ we have $F(x) > \lambda$, so we consider $C \cap \overline{B(0, R)}$ concluding the proof in the same way of bounded K.

□

EXERCISE 6.7. For every $x = (x_1, \dots, x_d) \in \mathbb{R}^d$, $1 \leq p < +\infty$ set
$$H_p(x) = \|x\|_{\ell^p} = \left(\sum_{i=1}^d |x_i|^p\right)^{1/p}, \qquad H_\infty(x) = \|x\|_{\ell^\infty} = \max_{i=1,\dots,d} |x_i|.$$
Given $1 \leq p \leq +\infty$, $x \in \mathbb{R}^d$, $r \geq 0$ denote the *open p-ball* of radius r by
$$B_p(x, r) := \{y \in \mathbb{R}^d : \|y - x\|_{\ell^p} \leq r\},$$

and set $\overline{B_p} = \overline{B_p(0,1)}$. Assume $1 < p < +\infty$ and let $1 < q < +\infty$ be such that $\dfrac{1}{p} + \dfrac{1}{q} = 1$. Prove that we have the following relations

$$\begin{aligned} H_p^*(x^*) &= I_{\overline{B_q}}(x^*), & \|x\|_{\ell^p} &= \sigma_{\overline{B_q}}(x), \\ H_1^*(x^*) &= I_{\overline{B_\infty}}(x^*), & \|x\|_{\ell^\infty} &= \sigma_{\overline{B_1}}(x), \\ H_\infty^*(x^*) &= I_{\overline{B_1}}(x^*), & \|x\|_{\ell^1} &= \sigma_{\overline{B_\infty}}(x), \\ \partial H_p(0) &= \overline{B_q}. \end{aligned}$$

DEFINITION 6.8 (Primal and dual problems). Let X be a normed space, $F : X \to]-\infty, +\infty]$ a proper, convex, l.s.c. function. Given another normed space Y and a proper, convex, l.s.c. function $\phi : X \times Y \to]-\infty, +\infty]$ with $\phi(x, 0) = F(x)$, we consider the first marginal

$$h(y) := \inf_{x\in X} \phi(x, y),$$

i.e., we embed the problem $\inf\limits_{x\in X} F(x)$ in a family of minimization problems (*perturbed problem*).
We will call:

- *primal problem* the minimization of F over X;
- *dual problem* the maximization of $-\phi^*(0, \cdot)$ over Y'.

According to Lemma 4.16, h is convex and $h^*(\varphi) = \phi^*(0, \varphi)$, moreover we have

$$\sup_{y^*\in Y'} -\phi^*(0, y^*) = \sup_{y^*\in Y'} -h^*(y^*) = h^{**} \circ J(0) = \overline{h}(0) \le h(0) = \inf_{x\in X} \phi(x, 0) = \inf_{x\in X} F(x).$$

This quantity is less than $+\infty$ since F is proper, but it can assume the value $-\infty$.

We say that the primal problem is

- *normal* if $h(0) \in \mathbb{R}$ and h is l.s.c. at 0.
- *stable* if $\partial h(0) \neq \emptyset$.

DEFINITION 6.9 (Lagrangian function). We introduce the *Lagrangian function* $L : X \times Y' \to [-\infty, +\infty]$ of the primal problem by setting

$$L(x, y^*) = \inf_{y\in Y} \{\phi(x, y) - \langle y^*, y\rangle_{Y',Y}\},$$

i.e., $y^* \mapsto L(x, y^*)$ is the opposite of the conjugate of $y^* \mapsto \phi(x, y^*)$ for every fixed $x \in X$. For this reason, we have that $y^* \mapsto L(x, y^*)$ is u.s.c. and concave for all $x \in X$, while by Lemma 4.16 we have that $x \mapsto L(x, y^*)$ is convex for all $y^* \in Y'$.

DEFINITION 6.10 (Saddle point). We say that $(\hat{x}, \hat{y}^*) \in X \times Y'$ is a *saddle point* for L if for every $x \in X$, $y^* \in Y'$ it holds

$$L(\hat{x}, y^*) \le L(\hat{x}, \hat{y}^*) \le L(x, \hat{y}^*).$$

Equivalently, $(\hat{x}, \hat{y}^*) \in X \times Y'$ is a saddle point for L, if and only if

$$\sup_{y^*\in Y'} L(x, \hat{y}^*) = L(\hat{x}, \hat{y}^*) = \inf_{x\in X} L(x, \hat{y}^*).$$

Below we collected some consequences of the above definitions.

PROPOSITION 6.11. *Same notation of Definition 6.8.*

(1) *the primal problem can be written as* $\inf\limits_{x\in X} \sup\limits_{y^*\in Y'} \{L(x, y^*)\}$;

(2) *the dual problem can be written as* $\sup\limits_{y^*\in Y'} \inf\limits_{x\in X} \{L(x, y^*)\}$;

(3) *In general we have*

$$\sup_{y^*\in Y'} \inf_{x\in X} \{L(x, y^*)\} \le \inf_{x\in X} \sup_{y^*\in Y'} \{L(x, y^*)\}.$$

(4) *the primal problem is normal if and only if*

$$\sup_{y^*\in Y'} -\phi^*(0, y^*) = \inf_{x\in X} \phi(x, 0).$$

(5) *the primal problem is stable if and only if is normal and the dual problem has a solution;*
(6) *the set of solutions of the dual problem is given by $\partial h^{**} \circ J(0)$;*
(7) *if the primal problem is stable, then the set of solutions of the dual problem if $\partial h(0)$;*

PROOF.
(1) We have (recalling Lemma 1.9)
$$-\phi^*(0,y^*) = -\sup_{\substack{x\in X\\ y\in Y}} \{\langle y^*, y\rangle_{Y',Y} - \phi(x,y)\} = \inf_{x\in X} -[\phi(x,\cdot)]^*(y^*) = \inf_{x\in X} L(x,y^*).$$
(2) Similarly, since ϕ is convex l.s.c.,
$$\phi(x,0) = [\phi(x,\cdot)]^{**}\circ J(0) = \sup_{y^*\in Y'} \{-[\phi(x,\cdot)]^*(y^*)\} = \sup_{y^*\in Y'} L(x,y^*).$$
(3) See Lemma 1.9.
(4) Suppose that the primal problem is normal, then $h^{**}\circ J(0) = \bar{h}(0) = h(0) \in \mathbb{R}$. Conversely, if $h^{**}\circ J(0) = \bar{h}(0) = h(0) \in \mathbb{R}$ then the problem is normal.
(5) Assume that the primal problem is stable, in particular $h(0)\in\mathbb{R}$. Then there exists $y_0^* \in Y'$ such that $h(y) \geq h(0) + \langle y_0^*, y\rangle_{Y',Y}$ for all $y \in Y$. By taking the liminf for $\|y\| \to 0$ we have
$$\liminf_{\|y\|_Y\to 0} h(y) \geq h(0),$$
i.e., the l.s.c. of h at 0, so the primal problem is normal. For all $y^* \in Y'$, $y \in Y$, recalling that $J0 = 0$ in Y'', we have
$$-\phi^*(0,y^*) = -h^*(y^*) = \langle J0, \varphi\rangle_{Y',Y} - h^*(y^*) \leq h^{**}\circ J(0) = h(0) \leq h(y) - \langle y_0^*, y\rangle_{Y',Y}.$$
By taking the inf on $y\in Y$ we have for all $y^* \in Y'$
$$-\phi^*(0,y^*) \leq \inf_{y\in Y}\{h(y) - \langle y_0^*, y\rangle_{Y',Y}\} = -\sup_{y\in Y}\{\langle y_0^*, y\rangle_{Y',Y} - h(y)\} = -h^*(y_0^*) = -\phi^*(0,y_0^*),$$
hence y_0^* solves the dual problem.

Conversely, let the primal problem be normal and let y_0^* be a solution of the dual problem. Then $-h^*(y_0^*) \geq -h^*(y^*)$ for all $y^* \in Y'$. We have then
$$-h^*(y_0^*) = \sup_{y\in Y'} -h^*(y) = h^{**}\circ J(0) = h(0).$$
So $h(0) \leq -\langle y_0^*, y\rangle_{Y',Y} + h(y)$ for all $y \in Y$, hence the primal problem is stable.
(6) maximize $-\phi^*(0,\cdot) = -h^*(\cdot)$ is the same of minimizing the convex l.s.c. function $h^* : Y' \to [-\infty, +\infty]$, hence the set of solution is given by $\partial h^{**}\circ J(0)$.
(7) by the properties of the subdifferential, if the problem is normal then $h(0) = h^{**}\circ J(0)$ thus $\partial h(0) = \partial h^{**}\circ J(0)$.

□

We state now an important stability criterion.

LEMMA 6.12 (Stability criterion). *Same notation of Definition 6.8. Assume $h(0) > -\infty$ and that there exists $x_0 \in X$ such that $y \mapsto \phi(x_0, y)$ is continuous at 0 and $\phi(x_0, 0) \in \mathbb{R}$. Then the primal problem is stable.*

PROOF. The function $y \mapsto \phi(x_0,y)$ is continuous at 0, hence bounded above in an open neighborhood of V of 0 by a constant K. Given $y \in V$ we have
$$h(y) \leq \inf_{x\in X} \phi(x,y) \leq \phi(x_0,y) \leq K.$$
Since h is convex and bounded from above in V, it is also continuous and finite at 0, so epi h has nonempty interior and there exists a supporting hyperplane to epi h at $(0,h(0))$, in particular there are $y^* \in Y'$, $t > 0$ such that for every $(y,\beta) \in \text{epi}\, h$ we have
$$\langle y^*, 0\rangle_{Y',Y} + th(0) \leq \langle y^*, y\rangle_{Y,Y'} + t\beta.$$

Dividing by $t > 0$ and taking $y = h(y)$ we have for all $y \in Y$ that

$$h(0) \leq \langle \frac{y^*}{t}, y \rangle_{Y,Y'} + h(y),$$

which concludes the proof since it implies $-y^*/t \in \partial h(0)$. □

The most important necessary condition for the convex optimization problem that we are studying is the following.

THEOREM 6.13 (Extremality conditions). *Same notation of Definition 6.8. Assume that the problem is stable.*

(1) *let $\hat{y}^* \in Y'$ a solution of the dual problem. Given any solution $\hat{x} \in X$ of the primal problem, the following* extremality condition *holds*

$$\phi(\hat{x}, 0) + \phi^*(0, \hat{y}^*) = 0,$$

or, equivalently,

$$(0, \hat{y}^*) \in \partial\phi(\hat{x}, 0).$$

(2) *Conversely, given $\hat{x} \in X$ and $\hat{y}^* \in Y'$ such that the extremality condition holds, then*

$$\phi(\hat{x}, 0) = \min_{x \in X} \phi(x, 0) = \max_{y^* \in Y'} -\phi^*(0, y^*) = -\phi^*(0, \hat{y}^*),$$

i.e. $\hat{x} \in X$ solves the primal problem and $\hat{y}^* \in Y'$ solves the dual problem.

PROOF. By stability of the primal problem we have

$$\sup_{y^* \in Y'} -\phi^*(0, y^*) = \inf_{x \in X} \phi(x, 0).$$

Recalling that $\hat{x} \in X$ solves the primal problem and $\hat{y}^* \in Y'$ solves the dual problem, the extremality condition holds.
Conversely, assume that the extremality condition holds for given $\hat{x} \in X$ and $\hat{y}^* \in Y'$. By stability, the dual problem admits a solution, hence

$$\phi(\hat{x}, 0) \geq \inf_{x \in X} \phi(x, 0) = \max_{y^* \in Y'} -\phi^*(0, y^*) \geq -\phi^*(0, \hat{y}^*) = \phi(\hat{x}, 0),$$

so equality holds. □

COROLLARY 6.14. *If the problem is stable, $(\hat{u}, \hat{\varphi}) \in X \times Y'$ is a saddle point for L if and only if $\hat{u} \in X$ is a minimizer of the primal problem, $\hat{\varphi}$ is a maximizer of the dual problem, both the values of the minimum and maximum are finite, and the extremality relation*

$$\phi(\hat{u}, 0) + \phi^*(0, \hat{\varphi}) = 0,$$

holds. Moreover, if the problem is stable, $\hat{u} \in X$ is a minimizer of the primal problem if and only if there exists $\hat{\varphi} \in Y'$ such that $(\hat{u}, \hat{\varphi}) \in X \times Y'$ is a saddle point for L.

PROOF. Suppose that $(\hat{u}, \hat{\varphi})$ is a saddle point, then

$$-\phi^*(0, \hat{\varphi}) = \inf_{u \in X} \{L(u, \hat{\varphi})\} = L(\hat{u}, \hat{\varphi}) = \max_{\varphi \in Y'} L(\hat{u}, \varphi) = \phi(\hat{u}, 0),$$

and so extremality relation follows. Conversely, assuming the extremality relation, since

$$-\phi^*(0, \hat{\varphi}) = \inf_{u \in X} L(u, \hat{\varphi}) \leq L(\hat{u}, \hat{\varphi}) \leq \sup_{\varphi \in Y'} L(\hat{u}, \varphi) = \phi(\hat{u}, 0),$$

by extremality relation equality holds and so $(\hat{u}, \hat{\varphi})$ is a saddle point for L.

Since the problem is stable, there always exists a solution $\hat{\varphi}$ of the dual problem. Since $\hat{u} \in X$ is a minimizer if and only if extremality relation between $\hat{u}$ and $\hat{\varphi}$ holds, and by the previous part of the proof we have that $(\hat{u}, \hat{\varphi})$ is a saddle point if and only if extremality relations holds, the proof is completed. □

REMARK 6.15. We notice that the first statement in the extremality condition is just a *necessary condition* for the solution of stable primal problem. Indeed, the *existence* of $\hat{x}$ is *assumed* (the existence of $\hat{y}^*$ follows from stability itself). The second statement ensures its *sufficiency*: if we have a pair $(\hat{x}, \hat{y}^*)$ linked by the extremality conditions, then they solves the primal and the dual problem.

To have a useful statement, we have to grant a solution of the primal problem, by mean, for example, of Tonelli-Weierstrass Theorem.

COROLLARY 6.16. *Assume that X is a reflexive Banach space, that the primal problem is stable, and*

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

Then the primal problem admits a solution (a minimizer), the dual problem admits a solution (a maximizer), and the extremality relation holds.

PROOF. In reflexive space, the growth condition is equivalent to coercivity, so by Tonelli-Weierstrass Theorem, the primal problem admits a solution. By stability, also the dual problem has a solution. Thus extremality condition holds. □

Summary of Lecture 6

- We proved in this lecture the main calculus rules for the subdifferential in the sense of convex analysis, namely the *sum rule* and the *chain rule w.r.t. linear operators*. Both these sum rules, in order to hold, require a weak separation argument, thus it must be ensured the nonemptiness of the interior of the epigraph. This is granted by the continuity assumption.
- We introduced the concept of *perturbed problem*, i.e., we embed our problem $\inf_{x\in X} F(x)$, in a *family* of problems indexed by *variations* belonging to another normed space Y. It is important to stress that the choice of the family of perturbations is not *a priori* given. The *coupling* between the original problem and the variation is enclosed in a l.s.c. convex proper function $\phi : X \times Y \to [-\infty, +\infty]$, while h denotes the value of the solution of the minimization problem related to the variation y, i.e., $h(y) = \inf_{x\in X} \phi(x, y)$. We require to recover for $y = 0$ our original problem, i.e., $\phi(x, 0) = F(x)$.
- We introduce the *dual problem*, which a maximization problem in Y' and the *Lagrangian function*.
- When the primal problem is normal, primal and dual problem have the same value. If the problem is stable, then we have a precise description of the solutions of the dual problem in terms of the subdifferential of h at 0.
- Since stability is related to subdifferentiability, a stability criterion must involve the existence of supporting hyperplanes to $\operatorname{epi} h$. In particular it can be ensured asking some continuity property of ϕ, as usual (thus $\operatorname{epi} h$ will have nonempty interior).
- The extremality relations summarize the necessary and sufficient conditions for optimality in this class of problems.

7. Lecture of 22 october 2018: Special case of convex functionals (3h)

If the functional F has a particular form, the choice of the family of perturbations ϕ can be done in a quite standard way.

PROPOSITION 7.1. *Let X be a reflexive Banach space, Y be a normed space. Let $f : X \to]-\infty, +\infty]$ and $g : Y \to]-\infty, +\infty]$ be proper, convex, l.s.c., let $\Lambda : X \to Y$ be a linear and continuous operator. Assume that:*

(1) $\displaystyle\lim_{\|u\|_X\to+\infty} f(u) + g(\Lambda u) = +\infty$;

(2) *there exists $u_0 \in X$ such that $f(u_0) < +\infty$, $g(\Lambda u_0) < +\infty$ and g is continuous at Λu_0.*

Then if we set

$$F(u) := f(u) + g(\Lambda u), \qquad \phi(u,y) = f(u) + g(\Lambda u - y),$$

the primal problem and the dual problems admit solutions $\hat{u} \in X$ and $\hat{\varphi} \in Y'$, respectively, and

$$\begin{cases} f(\hat{u}) + f^*(\Lambda^*\hat{\varphi}) & = \langle \Lambda^*\hat{\varphi}, \hat{u}\rangle_{X',X}, \\ -g(\Lambda\hat{u}) - g^*(-\hat{\varphi}) & = \langle \hat{\varphi}, \Lambda\hat{u}\rangle_{Y',Y}, \end{cases},$$

or, equivalently

$$\begin{cases} -\hat{\varphi} \in \partial g(\Lambda\hat{u}), \\ \Lambda^*\hat{\varphi} \in \partial f(\hat{u}), \end{cases}$$

where $\Lambda^* : Y' \to X'$ is the adjoint operator of Λ.

PROOF. Under the above assumptions there are solutions both of the primal and of the dual problem. We have (set $p = \Lambda u - y$)

$$\begin{aligned} \phi^*(0,\varphi) &= \sup_{\substack{u\in X\\ y\in Y}} \langle \varphi, y\rangle_{Y',Y} - f(u) - g(\Lambda u - y) \\ &= \sup_{\substack{u\in X\\ p\in Y}} \langle \varphi, \Lambda u - p\rangle_{Y',Y} - f(u) - g(p) \\ &= \sup_{\substack{u\in X\\ p\in Y}} \langle \varphi, \Lambda u\rangle_{Y',Y} - \langle \varphi, p\rangle_{Y',Y} - f(u) - g(p) \\ &= \sup_{\substack{u\in X\\ p\in Y}} \langle \Lambda^*\varphi, u\rangle_{Y',Y} - f(u) + \langle -\varphi, p\rangle_{Y',Y} - g(p) \\ &= \sup_{u\in X} [\langle \Lambda^*\varphi, u\rangle_{Y',Y} - f(u)] + \sup_{p\in Y} [\langle -\varphi, p\rangle_{Y',Y} - g(p)] \\ &= f^*(\Lambda^*\varphi) + g^*(-\varphi). \end{aligned}$$

Extremality relation is

$$f(\hat{u}) + g(\Lambda\hat{u}) + f^*(\Lambda^*\hat{\varphi}) + g^*(-\hat{\varphi}) = 0,$$

thus,

$$f(\hat{u}) + f^*(\Lambda^*\hat{\varphi}) = -(g^*(-\varphi) + g(\Lambda\hat{u})).$$

According to Young's inequality, the right hand side is always greater or equal of $\langle \hat{\varphi}, \Lambda\hat{u}\rangle_{Y',Y}$, while the left hand side is always less or equal than $\langle \Lambda^*\hat{\varphi}, \hat{u}\rangle_{X',X} = \langle \hat{\varphi}, \Lambda\hat{u}\rangle_{Y',Y}$ (recall the definition of adjoint operator). So both the left and the right hand side must be equal to $\langle \hat{\varphi}, \Lambda\hat{u}\rangle_{Y',Y} = \langle \Lambda^*\hat{\varphi}, \hat{u}\rangle_{X',X}$.

Extremality condition may be also obtained directly by subdifferential calculus rules. Since $g \circ \Lambda$ is continuous at u_0, we have $\partial F(u) = \partial f(u) + \partial(g \circ \Lambda)(u)$ at all $x \in \operatorname{dom} f \cap \operatorname{dom} g \circ \Lambda$, moreover, since g is continuous at Λu_0 we have $\partial(g \circ \Lambda)(u) = \Lambda^* \partial g(\Lambda u)$. Thus $\hat{u} \in X$ is a minimizer if and only if $0 \in \partial F(\hat{u}) = \partial f(\hat{u}) + \Lambda^*\partial g(\Lambda u)$. In particular, there exists $-\hat{\varphi} \in \partial g(\Lambda u) \subseteq Y'$ such that $\Lambda^*\hat{\varphi} \in \partial f(\hat{u})$. □

EXERCISE 7.2. Let Ω be a bounded open subset of $\mathbb{R}^d$ and let $r, q \in L^2(\Omega;\mathbb{R})$ be fixed. Define $F : H_0^1(\Omega;\mathbb{R}) \to \mathbb{R}$ by setting

$$F(u) := \frac{1}{2}\int_\Omega \|\nabla u(x)\|^2\,dx + \frac{1}{2}\int_\Omega |r(x) - u(x)|^2\,dx - \int_\Omega q(x)u(x)\,dx.$$

Study the problem of minimization of F on $H_0^1(\Omega)$.

SOLUTION. Set $X = H_0^1(\Omega;\mathbb{R})$, $X' = H^{-1}(\Omega;\mathbb{R})$, $Y = L^2(\Omega;\mathbb{R}^d)$, $Y' = Y$. The operator $\Lambda = \nabla : X \to Y$ is linear and continuous. The functions $f : X \to \mathbb{R}$ and $g : Y \to \mathbb{R}$ defined as

$$\begin{aligned} f(u) &:= \frac{1}{2}\|r - u\|_{L^2}^2 - \langle q, u\rangle_{L^2}, \\ g(p) &:= \frac{1}{2}\|p\|_{L^2}^2, \end{aligned}$$

allow to write F in the form $F(u) = f(u) + g(\Lambda u)$.

We verify now the requirements of the previous result:

(1) We prove strictly convexity of g. Indeed, by triangular inequality, for every $\lambda \in [0,1]$ we have
$$\|\lambda p_1 + (1-\lambda)p_2\|_Y \le \lambda \|p_1\|_Y + (1-\lambda)\|p_2\|_Y$$
and since $r \to r^2$ is convex and strictly increasing on nonnegative reals
$$(\|\lambda p_1 + (1-\lambda)p_2\|_Y)^2 \le (\lambda\|p_1\|_Y + (1-\lambda)\|p_2\|_Y)^2 \le \lambda\|p_1\|_Y^2 + (1-\lambda)\|p_2\|_Y^2,$$
and, by the strict increasing property, equality holds if and only if $p_1 = p_2$. We notice that $\operatorname{dom} g = Y$ and g is bounded from above in a neighborhood of every $\bar p \in Y$ More precisely, given $\bar p \in Y$ we have that g is bounded from above in $B(\bar p, \delta)$ by $\frac{1}{2}\|\bar p\|_Y^2 + \delta\|\bar p\|_Y + \delta^2$. Thus g is continuous, and locally Lipschitz, on the whole of Y.

(2) f is the sum between $u \mapsto -\langle q, u\rangle_{L^2}$, which is a linear continuous function in X, hence convex, l.s.c. and proper, and the composition between the map $u \mapsto r - u$ and $s \mapsto \frac{1}{2}\|s\|_{L^2}^2$, both of which are convex and continuous (see the proof of the convexity of g), hence f is proper convex, and continuous.

(3) We prove that if $\|u\|_X \to +\infty$ then $F(u) \to +\infty$: indeed
$$F(u) \ge \frac{1}{2}\|\nabla u\|_{L^2}^2 - \|q\|_{L_2} \cdot \|u\|_{L^2} \ge \frac{1}{2}\|\nabla u\|_{L^2}^2 - \|q\|_{L_2} \cdot \|u\|_X.$$
Since Ω is bounded, according to Poincaré's inequality we have $\|u\|_X \le C\|\nabla u\|_{L^2}$ where $C > 0$ is a costant depending only on Ω, so
$$F(u) \ge C\|u\|_X \left(\frac{C}{2}\|u\|_X - \|q\|_{L_2}\right) \to +\infty \text{ if } \|u\|_X \to +\infty.$$

(4) It is trivial to prove that there exists $u_0 \in X$ such that $f(u_0) < +\infty$, $g(\Lambda u_0) < +\infty$ and g is continuous at Λu_0: we can take $u_0 = 0$.

We compute now the conjugate functions $f^* : X' \to [-\infty, +\infty]$ and $g^* : Y' = Y \to \mathbb{R}$.
We compute $g^* : Y' = Y \to \mathbb{R}$ as follows
$$\begin{aligned}
g^*(p^*) &= \sup_{p \in Y} \left\{ \langle p^*, p\rangle_{Y',Y} - \frac{1}{2}\int_\Omega |p(x)|^2\,dx \right\} \\
&= \frac{1}{2}\sup_{p\in Y}\left\{ \int_\Omega \left(2p^*(x)p(x) - |p(x)|^2\right)\,dx \right\} \\
&= \frac{1}{2}\sup_{p\in Y}\left\{ \int_\Omega \left(|p^*(x)|^2 - |p^*(x)|^2 + 2p^*(x)p(x) - |p(x)|^2\right)\,dx \right\} \\
&= \frac{1}{2}\int_\Omega |p^*(x)|^2\,dx - \frac{1}{2}\inf_{p\in Y}\left\{ \int_\Omega \left(|p^*(x)|^2 - 2p^*(x)p(x) + |p(x)|^2\right)\,dx \right\} \\
&= \frac{1}{2}\int_\Omega |p^*(x)|^2\,dx - \frac{1}{2}\inf_{p\in Y}\left\{ \int_\Omega \left(|p^*(x) - p(x)|^2\right)\,dx \right\} \\
&= \frac{1}{2}\int_\Omega |p^*(x)|^2\,dx,
\end{aligned}$$
thus the sup was attained for $p = p^*$.

We recall that $X \subseteq L^2(\Omega;\mathbb{R}) \subseteq X'$, thus
$$\begin{aligned}
f^*(u^*) &= \sup_{u\in X}\left\{ \langle u^*, u\rangle_{X',X} - \frac{1}{2}\|r-u\|_{L^2}^2 + \langle q, u\rangle_{L^2} \right\} \\
&= \sup_{u\in X}\left\{ \langle u^*, u\rangle_{X',X} + \langle q, u\rangle_{X',X} - \frac{1}{2}\|r-u\|_{L^2}^2 \right\} \\
&= \sup_{u\in X}\left\{ \langle u^* + q, u\rangle_{X',X} + \langle r, u\rangle_{L^2} - \frac{1}{2}\|r\|_{L^2}^2 - \frac{1}{2}\|u\|_{L^2}^2 \right\}
\end{aligned}$$

$$= \sup_{u\in X}\left\{\langle u^*+q+r,u\rangle_{X',X}-\frac{1}{2}\|u\|^2_{L^2}\right\}-\frac{1}{2}\|r\|^2_{L^2}.$$

We notice that if $u^*\in L^2$, we have

$$\begin{aligned}
f^*(u^*) &= \sup_{u\in X}\left\{-\frac{1}{2}\|u^*+q+r\|^2_{L^2}+\langle u^*+q+r,u\rangle_{L^2}-\frac{1}{2}\|u\|^2_{L^2}\right\}-\frac{1}{2}\|r\|^2_{L^2}+\frac{1}{2}\|u^*+q+r\|^2_{L^2}\\
&= \sup_{u\in X}\left\{-\frac{1}{2}\|u^*+q+r-u\|^2_{L^2}\right\}-\frac{1}{2}\|r\|^2_{L^2}+\frac{1}{2}\|u^*+q+r\|^2_{L^2}\\
&= -\frac{1}{2}\|r\|^2_{L^2}+\frac{1}{2}\|u^*+q+r\|^2_{L^2},
\end{aligned}$$

since X is dense in L^2, thus the sup is attained at $u=u^*+q+r$. On the other hand, if $u^*\in H^{-1}\setminus L^2$, we have that there exists a sequence $\{u_n\}_{H^1_0}$ such that $\|u_n\|_{L^2}\to 0$ and $\langle u^*,u_n\rangle\to+\infty$, hence $f^*(u^*)+\infty$. Hence

$$f^*(u^*)=\begin{cases}-\frac{1}{2}\|r\|^2_{L^2}+\frac{1}{2}\|u^*+q+r\|^2_{L^2}, & \text{if } u^*\in L^2;\\[2ex] +\infty, & \text{otherwise.}\end{cases}$$

According to Green's formulas, for every $u\in X$, $v\in Y'$ sufficiently smooth it holds

$$\langle v,\Lambda u\rangle_{Y',Y}=\int_\Omega v(x)\cdot\nabla u(x)\,dx=-\int_\Omega \operatorname{div} v(x)\cdot u(x)\,dx,$$

and so $\Lambda^*:Y'\to X'$ is $\Lambda^*=-\operatorname{div}$, where the divergence must be taken in the distributional sense.

According to the previous result, both the primal and the dual problems have solutions $\hat{u}\in X$ and $\varphi\in Y'$, respectively. Moreover, since $g=g^*$ is strictly convex these solutions are unique. We have the extremality condition

$$\begin{cases} f(\hat{u})+f^*(\Lambda^*\hat{\varphi}) &= \langle\Lambda^*\hat{\varphi},\hat{u}\rangle_{X',X},\\[2ex] -g(\Lambda\hat{u})-g^*(-\hat{\varphi}) &= \langle\hat{\varphi},\Lambda\hat{u}\rangle_{Y',Y}.\end{cases}$$

In our case, the first relation implies that f^* must be finite at $\Lambda^*\hat{\varphi}$, thus we have $\Lambda^*\hat{\varphi}\in L^2$. The second relation is

$$-g(\nabla\hat{u})-g^*(-\hat{\varphi})=\langle\hat{\varphi},\nabla\hat{u}\rangle_{L^2},$$

which corresponds to

$$g^*(-\hat{\varphi})=\langle-\hat{\varphi},\nabla\hat{u}\rangle_{L^2}-g(\nabla\hat{u}),$$

so since the sup in the computation of g^* was obtained for $p=p^*$, we obtain $-\hat{\varphi}=\nabla\hat{u}$ in $L^2(\Omega)$. We can arrive at the same result even directly: the second extremality relation amounts to say

$$-\frac{1}{2}\int_\Omega|\nabla\hat{u}(x)|^2\,dx-\frac{1}{2}\int_\Omega|-\hat{\varphi}|^2\,dx=\int_\Omega\hat{\varphi}(x)\cdot\nabla\hat{u}(x)\,dx,$$

thus

$$\int_\Omega(|\nabla\hat{u}(x)|^2+|\hat{\varphi}(x)|^2+2\hat{\varphi}(x)\cdot\nabla\hat{u}(x))\,dx=0,$$

hence $\hat{\varphi}+\nabla\hat{u}=0$ in $L^2(\Omega)$.
In the first relation, we have that $f^*(\Lambda^*\hat{\varphi})=\langle\Lambda^*\hat{\varphi},\hat{u}\rangle_{X',X}-f(\hat{u})$, and since the sup in the computation of f^* was obtained for $u=u^*+q+r$, we have $\hat{u}=\Lambda^*\hat{\varphi}+q+r$, hence we obtain

$$\begin{cases}-\Lambda^*\hat{\varphi}+\hat{u}=q+r\\ -\hat{\varphi}=\nabla u\end{cases}$$

Finally, we have that $\hat{u}$ solves (weakly in H^1):

$$\begin{cases}-\triangle u+\hat{u}=q+r, & \text{in } \Omega;\\ u_{|\partial\Omega}=0,\end{cases}$$

recalling that $\operatorname{div}\nabla u = \triangle u$.

REMARK 7.3. By following the same argument in the opposite sense, we can prove the following result given an open bounded $\Omega \subseteq \mathbb{R}^d$, for every $r, q \in L^2(\Omega)$ the problem

$$\begin{cases} -\triangle u + u = q + r, & \text{in } \Omega; \\ u_{|\partial\Omega} = 0. \end{cases}$$

admits only one solution $\hat{u} \in H_0^1(\Omega)$, which is characterized by being the minimizer on $H_0^1(\Omega)$ of the following functional

$$F(u) := \frac{1}{2}\int_\Omega \|\nabla u(x)\|^2\,dx + \frac{1}{2}\int_\Omega |r(x) - u(x)|^2\,dx - \int_\Omega q(x)u(x)\,dx.$$

REMARK 7.4. We want to solve the same problem with subdifferential calculus. Recall that, given an Hilbert space Z, $v \in Z$, and defined $w_v(x) = \frac{1}{2}\|x - v\|_H^2$, we have $\partial w_v(x) = x - v$. The functional F defined as

$$F(u) := \frac{1}{2}\|\Lambda u\|_{L^2}^2\,dx + \frac{1}{2}\|r - u\|_{L^2}^2 - \langle q, u\rangle_{L^2},$$

where $\Lambda = \nabla$, satisfies all the properties to apply subdifferential calculus, hence the minimizer are characterized by

$$\Lambda^*\Lambda\hat{u} + \hat{u} - r - q = 0,$$

hence $-\triangle u + \hat{u} = q + r$.

Summary of Lecture 7

- Although the class of perturbations and the coupling is not a priori given by the problem, in certain cases there is a *standard* choice. Namely, if we consider $F : X \to]-\infty, +\infty]$ which can be written as $F(x) = f(x) + g(\Lambda x)$, where $f : X \to]-\infty, +\infty]$ is convex and l.s.c., $g : Y \to]-\infty, +\infty]$ convex and l.s.c., $\Lambda : X \to Y$ linear and continuous. In this relevant case, we choose $\phi : X \times Y \to]-\infty, +\infty]$ as $\phi(x, y) = f(x) + g(\Lambda x - y)$. We notice that ϕ is convex, l.s.c. and the stability criterion is fulfilled if there exists $x_0 \in X$ such that $f(x_0) < +\infty$ and g is continuous at Λx_0. Extremality relations for this problem *decouples* in a relation involving f, f^*, Λ^* and another involving g, g^*, Λ.
- It is **extremely important** to familiarize with the **example** in this section, since it is a *model* for many concrete cases.

8. Lecture of 26 october 2018: Complements to the first part (3h)

8.1. Useful tools to conjugate function computation. We will present now some tools and arguments which turn useful in order to compute in practice the conjugate functions. For some background remarks on Sobolev spaces, we refer the reader to the Appendix.

Let $g : \mathbb{R}^n \to \mathbb{R}$ be a C^1 function, and suppose to have to compute g^*. According to the definition, we have

$$g^*(y) = \sup_{\xi \in \mathbb{R}^n} \{\xi \cdot y - g(\xi)\}.$$

So we must compute for *fixed* y the supremum of $\xi \mapsto q(\xi) := \xi \cdot y - g(\xi)$. Recalling basic Calculus results, $q \in C^1(\mathbb{R}^n; \mathbb{R})$, and so to maximize it we must study the limits for $\|\xi\| \to \infty$ and then study the critical points, i.e., the points $x \in \mathbb{R}^n$ where $\nabla q(x) = 0$.

If we assume moreover that g has *superlinear growth*, i.e.,

$$\lim_{\|\xi\| \to +\infty} \frac{g(\xi)}{\|\xi\|} = +\infty,$$

we have also

$$\lim_{\|\xi\| \to +\infty} q(\xi) = \lim_{\|\xi\| \to +\infty} \|\xi\| \left(\frac{\xi}{\|\xi\|} \cdot y - \frac{g(\xi)}{\|\xi\|}\right) \leq \lim_{\|\xi\| \to +\infty} \|\xi\| \left(\|y\| - \frac{g(\xi)}{\|\xi\|}\right) = -\infty,$$

so the supremum is actually a maximum attained at some critical points belonging to $\mathbb{R}^n$ (we recall that g cannot take value $-\infty$).

The critical point condition $\nabla q(x) = 0$ can be written as $y = \nabla g(x)$. If g is convex then q is concave, thus $\nabla q(x) = 0$ if and only if x is a maximum, hence

$$g^*(y) + g(x) = x \cdot y,$$

for every $y = \nabla g(x)$. Hence $\partial g(x) = \{\nabla g(x)\}$.

Assume now that the relation $y = \nabla g(x)$ is invertible for every y, i.e., for every y the maximum of $q(\cdot)$ is attained at a unique point $x = [\nabla g]^{-1}(y)$. This holds if and only if q is strictly concave (i.e., g is strictly convex). In this case

$$g^*(y) = \sup_{\xi \in \mathbb{R}^n} \{\xi \cdot y - g(\xi)\} = \bar{\xi} \cdot y - g(\bar{\xi}) \text{ if and only if } y = \nabla g(\bar{\xi}).$$

Let us assume that also g^* is of class C^1 with superlinear growth. Iterating the argument we have

$$g^{**}(x) = \sup_{\eta \in \mathbb{R}^n} \{\eta \cdot x - g^*(\eta)\} = \bar{\eta} \cdot y - g^*(\eta) \text{ if and only if } x = \nabla g^*(\bar{\eta}).$$

By the regularity of g we obtain $g = g^{**}$, thus

(1) if $y = \nabla g(x)$ then $g^*(y) = x \cdot y - g(x)$, hence $g^{**}(x) = g(x) = x \cdot y - g^*(y)$, and so $x = \nabla g^*(y)$;
(2) if $x = \nabla g^*(y)$ then $g^{**}(x) = g(x) = x \cdot y - g^*(y)$, hence $g^*(y) = x \cdot y - g(x)$, and so $y = \nabla g(y)$.

So in this particular case we obtain

$$\nabla g^*(y) = x \Longleftrightarrow \nabla g(x) = y,$$

which is a smooth versione of $x \in \partial g(y) \Longleftrightarrow y \in \partial g(x)$.

Recalling that $(\nabla g(x), -1)$ is the normal to $\operatorname{epi} g$ at $(x, g(x))$, and $(\nabla g^*(x), -1)$ is the normal to $\operatorname{epi} g^*$ at $(y, g^*(y))$, from a geometric point of view the above relation yields a relation between the normals to $\operatorname{epi} g$ and $\operatorname{epi} g^*$.

EXAMPLE 8.1. Consider the case $n = 1$, $g : \mathbb{R} \to \mathbb{R}$, $g(x) = e^x$. We have

$$\lim_{x \to +\infty} \frac{g(x)}{x} = +\infty, \qquad \lim_{x \to -\infty} g(x) = 0.$$

Set

$$g^*(y) = \sup_{x \in \mathbb{R}} \{x \cdot y - e^x\},$$

we have $g^*(0) = -\inf\{e^x : x \in \mathbb{R}\} = 0$. Since

$$\lim_{x \to +\infty} \frac{g(x)}{x} = +\infty,$$

the supremum in the definition of $g^*(y)$ cannot be attended by the limit $\xi \to +\infty$. In the 1-dimensional case, ∇ denotes the ordinary derivative, thus

$$\nabla g^*(y) = x \Longleftrightarrow \nabla g(x) = y,$$

reads as

$$\frac{d}{dy} g^*(y) = x \Longleftrightarrow \frac{d}{dx} g(x) = y.$$

In particular, $\frac{d}{dx} g(x) = y$ means $e^x = y$, which can holds only if $y > 0$ and $x = \log y$. So for $y > 0$ we have $\frac{d}{dy} g^*(y) = \log y$, hence

$$g^*(y) = \int_0^y \log s \, ds = y \log y - y,$$

recalling that $g^*(0) = 0$ (be careful: it is an improper integral). In the case $y < 0$ the map $\xi \mapsto \xi \cdot y - e^{\xi}$ admits no critical point $\bar{x} \in \mathbb{R}$, since the equation $y = e^x$ has no solution. So the supremum of $\xi \mapsto \xi \cdot y - e^{\xi}$ is attained either for $\xi \to +\infty$ or for $\xi \to -\infty$. Since the case $\xi \to +\infty$ has already been excluded, we have immediately that the supremum is achieved for $\xi \to -\infty$ and it holds $+\infty$. Finally, we have $g^*(y) = +\infty$ if $y < 0$, $g^*(0) = 0$, and $g^*(y) = y(\log y - 1)$ if

$y > 0$. We notice that since g was strictly increasing, we cannot have normals to epi g whose first component is negative. This impossibility reflects on the fact that g^* cannot be finite for $y < 0$.

In many concrete problems are involved integral functionals, thus it is common the need to compute the convex conjugate $F^* : X' \to \mathbb{R}$ of functionals

$$F(u) = \int_\Omega f(x, u(x))\, dx,$$

where $\Omega \subseteq \mathbb{R}^d$ and $f : \Omega \times \mathbb{R}^m \to \mathbb{R} \cup \{+\infty\}$ are measurable, and $u \in X$ (X is some normed space contained in the set of measurable functions from Ω to $\mathbb{R}^m$). By definition,

$$F^*(u^*) = \sup_{u \in X} \left\{ \langle u^*, u \rangle_{X',X} + \int_\Omega f(x, u(x))\, dx \right\}.$$

Assume for simplicity that the action $\langle u^*, u \rangle_{X',X}$ can also be written in integral form (for instance, this is true if $X = L^p(\Omega)$ hence $X' = L^q(\Omega)$ where $1 \leq p < \infty$, $1 < q \leq +\infty$ and $1/p + 1/q = 1$, with the convention $1/+\infty = 0$)

$$\langle u^*, u \rangle_{X',X} = \int_X u^*(x) u(x)\, dx.$$

In this case we have to compute

$$F^*(u^*) = \sup_{u \in X} \int_\Omega [u^*(x)u(x) - f(x, u(x))]\, dx.$$

On the other hand, we know that for every $p \in \mathbb{R}^m$, $x \in \Omega$ it holds

$$u^*(x) \cdot p - f(x, p) \leq f^*(x, u^*(x)),$$

where we denote by

$$f^*(x, p) = \sup_{q \in \mathbb{R}^m} p \cdot q - f(x, q)$$

the conjugate of f only with respect to the second variable, $f^* : \Omega \times \mathbb{R}^m \to \mathbb{R} \cup \{+\infty\}$. In particular, we can take $p = u(x)$, and so

$$u^*(x)u(x) - f(x, u(x)) \leq f^*(x, u^*(x)),$$

Integrating and taking the supremum on $u \in X$, we have

$$F^*(u^*) \leq \int_\Omega f^*(x, u^*(x))\, dx.$$

It is clear the importance of providing sufficient conditions yielding

$$F(u) = \int_\Omega f(x, u(x))\, dx \text{ implies } F^*(u^*) = \int_\Omega f^*(x, u^*(x))\, dx,$$

since in this case we can compute a conjugate function of the integral functional by computing the finite-dimension conjugate of $f : \Omega \times \mathbb{R}^m \to \mathbb{R}$, i.e., the problem is reduced to the finite-dimensional case.

This amounts to *switch the* sup *and the integral operator*. However, in general, this operation may be not allowed, since the *selections*

$$p(x) \in \arg \sup_{p \in \mathbb{R}^m} \{u^*(x) \cdot p - f(x, p)\},$$

which can be built for every fixed x, could lead to a function $x \mapsto p(x)$ not belonging to X, or not integrabile, and even not measurable. Nevertheless, if the integrand function f is sufficiently smooth, this case does not occurr.

DEFINITION 8.2. Let Ω be an open subset of $\mathbb{R}^n$ and $B \subseteq \mathbb{R}^p$ be a Borel set. We say that a function $f : \Omega \times B \to [-\infty, +\infty]$ is a *normal integrand* on $\Omega \times B$ if

(1) for a.e. $x \in \Omega$ the map $a \mapsto f(x, a)$ is l.s.c. on B
(2) there exists a Borel map $\tilde{f} : \Omega \times B \to [-\infty, +\infty]$ such that $f(x, \cdot) = \tilde{f}(x, \cdot)$ for a.e. $x \in \Omega$.

If $f, g, \{f_n\}_{n\in\mathbb{N}}$ are normal integrand, then also $\lambda f + g$ is normal for all $\lambda > 0$, and also $\inf\{f, g\}$, $\sup_n \tilde{f}_n$ are normal.

An important class of normal integrand (see Proposition VIII.1.1 in **[5]**) is given by *Carathéodory functions*, i.e., functions $f : \Omega \times B \to [-\infty, +\infty]$ satisfying

(1) for a.e. $x \in \Omega$ the map $a \mapsto f(x, a)$ is continuous on B;
(2) for a.e. $a \in B$ the map $x \mapsto f(x, a)$ is measurable in Ω.

We state now the measurable selection theorem:

THEOREM 8.3. *Let Ω be an open subset of $\mathbb{R}^n$, B be a compact subset of $\mathbb{R}^m$, and g be a normal integrand on $\Omega \times B$. Then there exists a measurable function $\bar{u} : \Omega \to B$ such that for every $x \in \Omega$ we have*

$$g(x, \bar{u}(x)) = \min_{a\in B} g(x, a).$$

PROOF. Omitted. See Theorem VIII.1.2 in **[5]**. □

The following result formalizes the computation of the conjugate of integral functionals

PROPOSITION 8.4. *Let Ω be an open bounded subset of $\mathbb{R}^n$. Let $1 \le \alpha \le +\infty$, $f : \Omega \times \mathbb{R}^m \to [0, +\infty[$ be a normal integrand where $v \mapsto f(x, v)$ is convex for all $x \in \Omega$. Define $F : L^\alpha(\Omega, \mathbb{R}^m) \to [0, +\infty]$ by setting*

$$F(u) = \int_\Omega f(x, u(x))\, dx.$$

Assume that there exists $u_0 \in L^\infty(\Omega, \mathbb{R}^m)$ such that $F(u_0) < +\infty$. Then if α' is the conjugate exponent of α, i.e., $1/\alpha + 1/\alpha' = 1$, we have

$$F^*(u^*) = \int_\Omega f^*(x, u^*(x))\, dx,$$

for all $u^* \in L^{\alpha'}(\Omega, \mathbb{R}^m)$.

PROOF. Fix $u^* \in L^{\alpha'}(\Omega, \mathbb{R}^m)$. Define

$$\Phi(x) := \sup_{\xi\in\mathbb{R}^m} [u^*(x)\cdot\xi - f(x, \xi)],$$
$$\Phi_n(x) := \max_{\substack{\xi\in\mathbb{R}^m\\|\xi|\le n}} [u^*(x)\cdot\xi - f(x, \xi)].$$

Clearly $\{\Phi_n\}_{n\in\mathbb{N}}$ is an increasing sequence of functions pointwise convergent to Φ in Ω. Moreover, for every $n \ge \|u_0\|_{L^\infty}$ we have

$$\Phi_n(x) \ge [u^*(x)\cdot u_0(x) - f(x, u_0(x))] =: \tilde{\Phi}(x),$$

since $|u_0(x)| \le \|u_0\|_{L^\infty}$, and the function $\tilde{\Phi}(\cdot)$ on the right hand side is integrable because $u_0 \in L^\infty(\Omega; \mathbb{R}^m) \subseteq L^\alpha(\Omega; \mathbb{R}^m)$. According to the measurable selection theorem, for every $n \in \mathbb{N}$ there exists a measurable function $\bar{u}_n : \Omega \to \mathbb{R}^m$ such that $\|\bar{u}_n\|_{L^\infty} \le n$ and

$$\Phi_n(x) = u^*(x)\cdot\bar{u}_n(x) - f(x, \bar{u}_n(x)).$$

This implies that $\{\Phi_n\}_{n\in\mathbb{N}}$ is a (increasing) sequence of measurable functions pointwise converging to Φ, and then Φ is also measurable. According to Beppo Levi's Monotone Convergence Theorem (applied to the nonnegative increasing sequence of measurable functions $\{\Phi_n - \tilde{\Phi}\}_{n\in\mathbb{N}}$ pointwise convergent to $\Phi - \tilde{\Phi}$), we have

$$\int_\Omega \Phi(x)\, dx = \sup_{n\in\mathbb{N}} \int_\Omega \Phi_n(x)\, dx.$$

Thus

$$\begin{aligned}\int_\Omega \Phi(x)\, dx &= \sup_{n\in\mathbb{N}} \int_\Omega \Phi_n(x)\, dx\\ &= \sup_{n\in\mathbb{N}} \int_\Omega u^*(x)\cdot\bar{u}_n(x) - f(x, \bar{u}_n(x))\, dx\\ &\le \sup_{u\in L^\infty\subseteq L^\alpha} \int_\Omega u^*(x)\cdot u(x) - f(x, u(x))\, dx\end{aligned}$$

$$\leq \sup_{u\in L^{\alpha}(\Omega;\mathbb{R}^m)} \int_\Omega u^*(x)\cdot u(x) - f(x,u(x))\,dx = F^*(u^*).$$

Conversely, since

$$\sup_{\xi\in\mathbb{R}^m}[u^*(x)\cdot\xi - f(x,\xi)] = \Phi(x),$$

for all $u\in L^{\alpha}(\Omega;\mathbb{R}^m)$ we have

$$u^*(x)\cdot u(x) - f(x,u(x)) \leq \Phi(x).$$

Integrating the above relation and taking the sup on $u\in L^{\alpha}(\Omega;\mathbb{R}^m)$ we obtain

$$F^*(u^*) = \sup_{u\in L^{\alpha}(\Omega;\mathbb{R}^m)} \int_\Omega [u^*(x)\cdot u(x) - f(x,u(x))]\,dx \leq \int_\Omega \Phi(x)\,dx.$$

So we have

$$F^*(u^*) = \int_\Omega \Phi(x)\,dx = \int_\Omega f^*(x,u^*(x))\,dx.$$

□

REMARK 8.5 (Conjugate in H^1). We will discuss now a frequent case occurring in the exercises. Let Ω be a bounded open subset of $\mathbb{R}^d$, and assume to have $F: H^1(\Omega)\to]-\infty,+\infty]$ given by

$$F(u) := \int_\Omega f(x,u(x))\,dx,$$

where f is a normal integrand, and $v\mapsto f(x,v)$ is a convex function. Assume moreover that there exists $u_0\in L^\infty(\Omega)$ such that $F(u_0)<+\infty$. Under these assumption we would know how to conjugate F if F was defined on L^2: in fact, we would have $F^*: L^2(\Omega)\to]-\infty,+\infty]$ defined by

$$F^*(u^*) := \int_\Omega f^*(x,u^*(x))\,dx.$$

However, since F is defined on $H^1(\Omega)$, its conjugate is defined on $F^*: H^{-1}(\Omega)\to[-\infty,+\infty]$.

We recall that given $u^*\in L^2(\Omega)\subset H^{-1}(\Omega)$ and $v\in H^1(\Omega)\subseteq L^2(\Omega)$, the action of u^* on v is given by

$$\langle u^*,v\rangle_{H^{-1},H^1} = \int_\Omega u^*(x)v(x)\,dx = \langle u^*,v\rangle_{L^2}.$$

For every $u\in H^1(\Omega)$ and $u^*\in H^{-1}$ we have that

$$F^*(u^*) + F(w) \geq \langle u^*,w\rangle_{H^{-1},H^1}.$$

Assume that F is the restriction on $H^1(\Omega)$ of a continuous functional on $L^2(\Omega)$, still denoted by F

We distinguish now two cases:

(1) For every $u^*\in H^{-1}(\Omega)\setminus L^2(\Omega)$ there exists a sequence $\{v_n\}_{n\in\mathbb{N}}\subseteq H^1(\Omega)$ and $v\in L^2(\Omega)$ with $v_n\to v$ in $L^2(\Omega)$ such that $\langle u^*,v_n\rangle_{H^{-1},H^1}\to+\infty$. In particular, we have

$$F^*(u^*) + F(v_n) \geq \langle u^*,v_n\rangle_{H^{-1},H^1},$$

and by taking the limit for $n\to+\infty$ and recalling that $F(v_n)\to F(v)<+\infty$ we have we obtain $F^*(u^*)=+\infty$.

(2) For every $u^*\in L^2(\Omega)$ we have

$$F^*(u^*) = \sup_{u\in H^1(\Omega)} \langle u^*,u\rangle_{H^{-1},H^1} - F(u) = \sup_{u\in H^1(\Omega)} \langle u^*,u\rangle_{L^2} - F(u).$$

Given $v\in L^2(\Omega)$ and $\varepsilon>0$ there exists $u\in H^1(\Omega)$ such that $\|u-v\|_{L^2}\leq\varepsilon$ and $|F(u)-F(v)|\leq\varepsilon$, so

$$\begin{aligned}\langle u^*,v\rangle_{L^2} - F(v) &\leq \langle u^*,u\rangle + \|u^*\|_{L^2}\|u-v\|_{L^2} - F(u) + |F(u)-F(v)|\\ &\leq \langle u^*,u\rangle - F(u) + \varepsilon(1+\|u^*\|_{L^2}).\end{aligned}$$

We obtain for every $\varepsilon>0$

$$\langle u^*,v\rangle_{L^2} - F(v) \leq \sup_{u\in H^1(\Omega)} \langle u^*,u\rangle - F(u) + \varepsilon(1+\|u^*\|_{L^2}),$$

thus
$$\langle u^*, v\rangle_{L^2} - F(v) \le \sup_{u\in H^1(\Omega)} \langle u^*, u\rangle - F(u),$$
and
$$\sup_{v\in L^2(\Omega)} \langle u^*, v\rangle_{L^2} - F(v) \le \sup_{u\in H^1(\Omega)} \langle u^*, u\rangle - F(u),$$
since the opposite inequality trivially holds, we have equality, thus
$$F^*(u^*) = \sup_{u\in L^2(\Omega)} \langle u^*, u\rangle_{L^2} - F(u),$$
and so
$$F^*(u^*) = \int_\Omega f^*(x, u^*(x))\, dx.$$
We conclude that in this case if $u^* \notin L^2$ automatically $F^*(u^*) = +\infty$ otherwise we compute F^* as in the case of $F : L^2(\Omega) \to L^2(\Omega)$.

EXERCISE 8.6. We consider the *Mossolov's problem*. Let $\alpha, \beta > 0$, $\Omega \subseteq \mathbb{R}^d$ be open and bounded, $q \in L^2(\Omega)$. Define $F : H^1_0(\Omega;\mathbb{R}) \to]-\infty, +\infty]$ by setting
$$F(u) := \frac{\alpha}{2}\int_\Omega |\nabla u(x)|^2\, dx + \beta\int_\Omega |\nabla u(x)|\, dx - \int_\Omega q(x)u(x)\, dx,$$
and study $\inf\limits_{u\in H^1_0} F(u)$.

SOLUTION. Set $X = H^1_0(\Omega;\mathbb{R})$, $X' = H^{-1}(\Omega;\mathbb{R})$, $Y = Y' = L^2(\Omega;\mathbb{R}^d)$, $\Lambda = \nabla : X \to Y$ and define $f : X \to \mathbb{R}$ and $g : Y \to \mathbb{R}$ by
$$f(u) = -\int_\Omega q(x)u(x)\, dx,$$
$$g(p) = \frac{\alpha}{2}\int_\Omega |p(x)|^2\, dx + \beta\int_\Omega |p(x)|\, dx,$$
thus we have
$$F(u) = f(u) + g(\Lambda u).$$
It is easy to see that f and g are continuous and convex functions, that F is coercive and strictly convex, and $\Lambda^* = -\mathrm{div} : Y' \to X'$. Moreover $f^*(u^*) = 0$ if and only if $u^* = -q$, otherwise $f^*(u^*) = +\infty$.

If $q = 0$ the the unique solution to the problem is $\hat{u} = 0$. We assume $q \neq 0$.

To compute the conjugate of g, we notice that
$$g(p) = \frac{\alpha}{2}\int_\Omega |p(x)|^2\, dx + \beta\int_\Omega |p(x)|\, dx = \int_\Omega r(x, p(x))\, dx,$$
where
$$r(x, a) = r(a) = \frac{\alpha}{2}|a|^2 + \beta|a|$$
for all $a \in B := \mathbb{R}^d$. We notice that $r(\cdot)$ is a normal integrand on $\Omega \times B$, moreover $r(\cdot) \ge 0$ and it is convex and superlinear, then
$$g^*(p^*) = \int_\Omega r^*(p^*(x))\, dx.$$
We will compute now the conjugate of $r(\cdot)$. We notice that $r^*(0) = \inf\limits_{a\in\mathbb{R}} r = 0$, thus we consider now $a^* \neq 0$. In particular, we write $a^* = \mu\eta$, where $\mu = \|a^*\| \ge 0$ and $\|\eta\| = 1$.
$$r^*(\mu\eta) = \sup_{a\in\mathbb{R}^d} \left\{\mu\langle\eta, a\rangle - \frac{\alpha}{2}|a|^2 - \beta|a|\right\}.$$
Since $\langle\eta, a\rangle \le |a|$ with equality if and only if $a = |a|\eta$, setting $\lambda = |a|$, we have
$$r^*(\mu\eta) = \sup_{a\in\mathbb{R}^d} \left\{(\mu - \beta)|a| - \frac{\alpha}{2}|a|^2\right\} = \sup_{\lambda\ge 0} \left\{(\mu - \beta)\lambda - \frac{\alpha}{2}\lambda^2\right\}.$$

The graph of the map $\psi(\lambda) := (\mu-\beta)\lambda - \frac{\alpha}{2}\lambda^2$ is a parabola passing through the origin with downward concavity, thus this map is strictly increasing for $\lambda \le \frac{\mu-\beta}{\alpha}$ and strictly decreasing for $\lambda \ge \frac{\mu-\beta}{\alpha}$.
In particular, if $\mu-\beta \le 0$, the map ψ is strictly decreasing on $[0,+\infty[$, thus its supremum on $[0,+\infty[$ is achieved at 0 and its value is $\psi(0)=0$. If $\mu-\beta>0$, the map ψ achieves its unique maximum on $\mathbb{R}$ at $\frac{\mu-\beta}{\alpha}>0$, its supremum on $[0,+\infty[$ is achieved at $\frac{\mu-\beta}{\alpha}$ and its value is $\psi\left(\frac{\mu-\beta}{\alpha}\right)=\frac{(\mu-\beta)^2}{2\alpha}$.
Summarizing, we have:

(1) if $\|a^*\| \le \beta$, then $r^*(a^*)=0$ and $r^*(a^*)=\langle a^*,a\rangle - r(a)$ if and only if $a=0$.

(2) if $\|a^*\|>\beta$, then $r^*(a^*)=\frac{(\|a^*\|-\beta)^2}{2\alpha}$ and $r^*(a^*)=\langle a^*,a\rangle - r(a)$ if and only if $a=\frac{\|a^*\|-\beta}{\alpha}\eta=\frac{\|a^*\|-\beta}{\|a^*\|\alpha}a^*$.

Thus

$$r^*(a^*)=\begin{cases}0, & \text{if } \|a^*\|\le\beta;\\[2ex] \frac{(\|a^*\|-\beta)^2}{2\alpha}, & \text{if } \|a^*\|\ge\beta.\end{cases}$$

So the dual problem is

$$\sup_{\substack{p^*\in L^2\\ \operatorname{div} p^*=q}} -\int_\Omega r^*(p^*(x))\,dx.$$

The duality theorem holds, and we obtain the extremality relations

$$\begin{cases} f(\hat{u})+f^*(\Lambda^*\hat{\varphi})=\langle\hat{\varphi},\Lambda\hat{u}\rangle_{Y',Y},\\[2ex] -g(\Lambda\hat{u})-g^*(-\hat{\varphi})=\langle\Lambda^*\hat{\varphi},\hat{u}\rangle_{Y',Y}.\end{cases}$$

In our case, from the first relation, to have a finite value of f^* (recalling that f^* assumes only the values 0 or $+\infty$), necessarily $\Lambda^*\hat{\varphi}=-q$, thus $\operatorname{div}\hat{\varphi}=q$. From the second we have,

$$-g(\nabla\hat{u})-g^*(-\hat{\varphi})=\langle\hat{\varphi},\nabla\hat{u}\rangle_{L^2},$$

which becomes

$$\int_\Omega \left(r(\nabla\hat{u})+r^*(-\hat{\varphi}(x))+\hat{\varphi}\cdot\nabla\hat{u}(x)\right)\,dx=0$$

By Young's inequality, the integrand is always nonnegative, thus for a.e. $x\in\Omega$:

$$r^*(-\hat{\varphi}(x))=\langle\nabla\hat{u}(x),-\hat{\varphi}\rangle - r(\nabla\hat{u}(x)).$$

In particular, this implies $-\hat{\varphi}(x)=\gamma\nabla\hat{u}(x)$, $\gamma\in\mathbb{R}$, since both r and r^* depends only by the modulus of their arguments.

Recalling that $r^*(a^*)=a^*\cdot a - r(a)$ if and only if $a=\bar{\lambda}a^*=\frac{\|a^*\|-\beta}{\|a^*\|\alpha}a^*$ for $\|a^*\|>\beta$, and $a=0$ for $\|a^*\|\le\beta$, we obtain

$$\nabla\hat{u}(x)=\begin{cases}0, & \text{if } \|\hat{\varphi}(x)\|\le\beta\\[2ex] -\hat{\varphi}(x)\frac{\|\hat{\varphi}(x)\|-\beta}{\|\hat{\varphi}(x)\|\alpha} & \text{if } \|\hat{\varphi}(x)\|\ge\beta.\end{cases}$$

where $\Lambda^*\hat{\varphi}=-q$, so $\operatorname{div}\hat{\varphi}=q$ and $\hat{\varphi}$ is the solution of the dual problem.

8.2. Relaxation and convexification. It may occurr that we are dealing with a problem $\inf_{x\in X} F(x)$ where

$$F(x) := \int_\Omega f(x, u(x), \nabla u(x))\, dx$$

fails to be convex and lower semicontinuous. In this case, even in presence of coercivity of the functional, the existence of a solution cannot be taken as granted, since Tonelli-Weierstrass theorem cannot be applied, moreover, the lack of convexity prevents to use any of the necessary condition we stated. Nevertheless, in many cases the problem is faced by introducing a new problem, called *relaxed problem* which exhibits good regularity properties and whose solutions are connected with the original problem.

We present first a generalization of Tonelli-Weierstrass to non-l.s.c. functionals.

PROPOSITION 8.7. *Let X be a reflexive Banach space, and let $F : X \to \mathbb{R} \cup \{+\infty\}$ be coercive. Consider the l.s.c. regularization $\overline{F}$ of F then*

(1) *$\bar{F} : X \to \mathbb{R} \cup \{+\infty\}$ is coercive and l.s.c., thus it admits point of minimum;*
(2) *every cluster point of a minimizing sequence of F is a minimum of $\overline{F}$;*
(3) *every minimum point of $\overline{F}$ is the limit of a minimizing sequence of F*

PROOF.

(1) By coercivity of F in a reflexive space, for any $M > 0$ there exists $N > 0$ such that if $\|y\|_X \geq N$ then $\dfrac{F(y)}{\|y\|_X} \geq M$ Thus, if we take $x \in X$ with $\|x\| > 2N$, we have

$$M \leq \liminf_{y\to x} \frac{F(y)}{\|y\|_X} = \frac{\bar{F}(y)}{\|x\|_X},$$

yielding coercivity. By Tonelli-Weierstrass, we have that $\bar{F}$ has points of minimum.

(2) let $\{x_n\}_{n\in\mathbb{N}} \subseteq X$ be a minimizing sequence of F, and assume that $x_n \to x$. Thus, $(x_n, F(x_n))$ is a sequence in $\operatorname{epi} \overline{F} = \overline{\operatorname{epi} F}$ wich converges to $(x, \inf F)$. By closedness of $\operatorname{epi} \overline{F}$, we have that $(x, \inf F) \in \operatorname{epi} \overline{F}$, thus $\overline{F}(x) \leq \inf F$ On the other hand, suppose by contradiction that there exists $y \in X$ such that $\overline{F}(y) < \inf F$, and let $\overline{F}(y) < a < \inf F$. By definition, we have $a > \overline{F}(y) = \liminf_{z\to y} F(y) \geq \inf F > a$, which is a contradiction, hence $\overline{F}(y) \geq \inf F$ for all $y \in X$ and thus $\overline{F}(x) = \inf F$, so x is a point of minimum for $\overline{F}$.

(3) Since $\overline{F} \leq F$, if x a minimum point of $\overline{F}$ we must have $\overline{F}(x) \leq F(y)$ for all $y \in X$, thus $\overline{F}(x) \leq \inf F$. On the other hand, we have already proved that $\overline{F}(y) \geq \inf F$ for all $y \in X$, hence if x is a minimum point for $\overline{F}$, then $\overline{F}(x) = \inf F$. Moreover, we have

$$\overline{F}(x) = \inf F = \liminf_{y\to x} F(y),$$

thus there exists a sequence $x_n \to x$ such that $F(x_n) \to \inf F$.

□

We present here, without proof, a general result on relaxation of integral functionals.

THEOREM 8.8 (Relaxation). *Let $\Phi : [0, +\infty[\to [0, +\infty[$ be a nonnegative, increasing, convex and l.s.c. function such that $\lim_{t\to+\infty} \dfrac{\Phi(t)}{t} = +\infty$. Let $g : \Omega \times \mathbb{R}^m \to \mathbb{R}$ be a normal integrand satisfying $\Phi(|\xi|) \leq g(x, \xi)$ for all $(x, \xi) \in \Omega \times \mathbb{R}^m$. Let $1 \leq \beta \leq +\infty$, and $f : \Omega \times \mathbb{R}^\ell \times \mathbb{R}^m \to \mathbb{R}$ be a normal integrand satisfying*

(1) *if $1 \leq \beta < +\infty$, there exists $a_1, a_2 \in L^1(\Omega)$, $b \geq 0$, $c \geq 1$ such that*

$$g(x, \xi) + a_2(x) \leq f(x, s, \xi) \leq cg(x, \xi) + b|s|^\beta + a_1(x).$$

(2) *if $\beta = +\infty$, there exists $a_2 \in L^1(\Omega)$ and, for all $k > 0$ there exists $c \geq 1$ and $a_1 \in L^2(\Omega)$ such that*

$$g(x, \xi) + a_2(x) \leq f(x, s, \xi) \leq cg(x, \xi) + a_1(x), \text{ for } |s| \leq k.$$

(3) *for a.e. $x \in \Omega$, the restriction of $f(x, \cdot, \cdot)$ to $\mathbb{R}^\ell \times \operatorname{dom} g(x, \cdot)$ is continuous.*

Let $\mathscr{G} : L^1 \to L^\beta$ be a map such that if $\{p_n\}_{n\in\mathbb{N}}$ converges weakly to $\bar{p}$ in $L^1(\Omega;\mathbb{R}^m)$, and if $\sup_{n\in\mathbb{N}} \int_\Omega \Phi(|p_n(x)|)\,dx < +\infty$, then $\{\mathscr{G}p_n\}_{n\in\mathbb{N}}$ converges to $\mathscr{G}\bar{p}$ in $L^\beta(\Omega;\mathbb{R}^m)$. We introduce the following problems:

$$(\mathscr{P}) \qquad \inf_{\substack{p\in L^\Phi(\Omega;\mathbb{R}^m)\\ u=\mathscr{G}p}} \int_\Omega f(x,u(x),p(x))\,dx,$$

$$(\mathscr{R}\mathscr{P}) \qquad \inf_{\substack{p\in L^\Phi(\Omega;\mathbb{R}^m)\\ u=\mathscr{G}p}} \int_\Omega f^{**}(x,u(x),p(x))\,dx,$$

where

$$L^\Phi(\Omega) := \left\{ p : \int_\Omega \Phi(|p(x)|)\,dx < +\infty \right\}.$$

Then

(1) *the problem $(\mathscr{R}\mathscr{P})$ has a solution,*
(2) *the minimum of $(\mathscr{R}\mathscr{P})$ equals the infimum of $(\mathscr{P})$,*
(3) *if $(\bar{u},\bar{p})$ with $\bar{u} = \mathscr{G}\bar{p}$ solves $(\mathscr{R}\mathscr{P})$, then there exists a minimizing sequence $\{(u_n,p_n)\}_{n\in\mathbb{N}}$ for $(\mathscr{P})$ such that $u_n = \mathscr{G}p_n$, $u_n \to \bar{u}$ in L^β, and $p_n \to \bar{p}$ weakly in L^1,*
(4) *if $\{(u_n,p_n)\}_{n\in\mathbb{N}}$ is a minimizing sequence for $(\mathscr{P})$, there exists $(\bar{u},\bar{p})$ with $\bar{u} = \mathscr{G}\bar{p}$ which solves $(\mathscr{R}\mathscr{P})$, and a subsequence $\{(u_{n_k},p_{n_k})\}_{k\in\mathbb{N}}$ such that $u_n \to \bar{u}$ in L^β, and $p_n \to \bar{p}$ weakly in L^1.*

PROOF. See Theorem 4.1 in Chapter IX of **[5]** at p.287. □

We consider now a concept of convergence of functionals introduced by De Giorgi in the '70s. Our problem is as follows: a sequence of functionals $\{F_h\}_{h\in\mathbb{N}}$ is given. Supposing that $\bar{x}_n$ is a minimum of F_n, we want to give conditions in order to have convergence of $\bar{x}_n$ to a point x_0 that is characterized as the minimum of a suitable *limit* functional F. The main reference for this part is **[2]**.

DEFINITION 8.9 (Γ-limit). Let X be a separable Banach space endowed with a topology $\mathscr{T}$, and let $\{F_h : X \to [-\infty,+\infty]\}_{h\in\mathbb{N}}$ be a sequence of functionals. We say that the sequence F_h Γ-converges to $F : X \to [-\infty,+\infty]$ (or $F = \Gamma - \lim F_h$) if

(1) For every $x \in X$ and every sequence $\{x_n\}_{n\in\mathbb{N}} \subseteq X$ converging to x for the topology $\mathscr{T}$ we have
$$F(x) \le \liminf_{h\to\infty} F_h(x_h).$$
(2) For every $x \in X$ there exists a sequence $\{x_n\}_{n\in\mathbb{N}} \subseteq X$ converging to x for the topology $\mathscr{T}$ we have
$$F(x) \ge \limsup_{h\to\infty} F_h(x_h).$$

In order to ensure convergence of the sequence of the minima of F_h, the following definition is quite natural.

DEFINITION 8.10 (Equicoercivity). Let X be a separable Banach space endowed with a topology $\mathscr{T}$, and let $\{F_h : X \to [-\infty,+\infty]\}_{h\in\mathbb{N}}$ be a sequence of functionals. We say that $\{F_h\}_{h\in\mathbb{N}}$ is *equicoercive* if for every $t \in \mathbb{R}$ there exists a compact set K_t in the topology $\mathscr{T}$ such that $\{x \in X :\in F_h(x) \le t\} \subseteq K_t$ for all $h \in \mathbb{N}$.

THEOREM 8.11 (Γ-convergence). *Let X be a separable Banach space endowed with a topology $\mathscr{T}$, and let $\{F_h : X \to [-\infty,+\infty]\}_{h\in\mathbb{N}}$ be a sequence of equicoercive functionals. Then*

(1) *if the Γ-limit of $\{F_h\}_{h\in\mathbb{N}}$ exists, then it is unique and l.s.c.;*
(2) *there exists a subsequence $\{F_{h_k}\}_{h\in\mathbb{N}}$ and F such that $F = \Gamma - \lim F_{h_k}$;*
(3) *if $F = \Gamma - \lim F_h$, then $F + G = \Gamma - \lim F_h + G$ for all continuous $G : X \to [-\infty,+\infty]$; there exists a subsequence $\{F_{h_k}\}_{h\in\mathbb{N}}$ and F such that $F = \Gamma - \lim F_{h_k}$;*

(4) *let $F = \Gamma - \lim F_h$ and assume that F admits x_0 as unique minimum point. Let $\{x_h\}_{h\in\mathbb{N}} \subseteq X$ and $\{\varepsilon_h\}_{h\in\mathbb{N}} \subseteq]0,+\infty[$ be such that $\varepsilon_h \to 0^+$ and $|F_h(x_h) - \inf F_h| \le \varepsilon_h$. Then $x_h \to x_0$ in $\mathscr{T}$ and $F_h(x_h) \to F(x_0)$.*

PROOF. Omitted, see [**2**]. □

PROPOSITION 8.12. *Let X be a separable Banach space endowed with a topology $\mathscr{T}$, and let $\{F_h : X \to [-\infty,+\infty]\}_{h\in\mathbb{N}}$ be a sequence of functionals, $F : X \to [-\infty,+\infty]$. Then*

(1) *If $\{F_h\}_{h\in\mathbb{N}}$ converges to F uniformly, then F_h Γ-converges to F;*
(2) *If $\{F_h\}_{h\in\mathbb{N}}$ is a decreasing sequence converging to F pointwise, then F_h Γ-converges to $\overline{F}$.*

PROOF. Omitted, see [**2**]. □

9. Some exercises in preparation to the first partial test

EXERCISE 9.1. Let Ω be an open bounded subset of $\mathbb{R}^2$. Consider the problem:

$$\inf_{u\in H^1_0(\Omega)} \int_\Omega \Big(2\,|\nabla u\,(x_1,x_2)|^2 + 4\partial_{x_2} u\,(x_1,x_2)\,\partial_{x_1} u\,(x_1,x_2) + \Big(\Big(x_2^2+2\Big)\,u\,(x_1,x_2) - 3x_1\Big)^2 +$$
$$+[\partial_{x_1} u(x_1,x_2)]^2 + 2[\partial_{x_2} u(x_1,x_2)]^2\Big)\,dx_1\,dx_2.$$

(1) Prove that the problem admits a unique solution.
(2) Formulate the problem in the form $\mathscr{F}(u) = F(u) + G\circ\Lambda(u)$ where $F : X \to]-\infty,+\infty]$, $G : Y \to]-\infty,+\infty]$, and $\Lambda : X \to Y$, carefully precising the functional spaces X, Y and discussing the regularity of F, G, Λ.
(3) Write the dual problem and the extremality relations. Establish if the dual problem admits an unique solution.
(4) Using the previous results, write down a partial differential equation satisfied by the minimum.

SOLUTION. Set $X = H^1_0(\Omega;\mathbb{R})$, $X' = H^{-1}(\Omega;\mathbb{R})$, $Y = Y' = L^2(\Omega;\mathbb{R}^2)$, $\Lambda = \nabla : X \to Y$, $\Lambda^* = -\mathrm{div} : Y' \to X'$. Denoted by $x = (x_1,x_2)$, $p = (p_1,p_2)$, we define

$$\begin{array}{llll}
r : \mathbb{R}^2 \to \mathbb{R}, & r(x) := & x_2^2 + 2, \\
q : \mathbb{R}^2 \to \mathbb{R}, & q(x) := & 3x_1, \\
A \in \mathrm{Mat}_{2\times 2}(\mathbb{R}), & A := & \begin{pmatrix} 3 & 2 \\ 2 & 4 \end{pmatrix}, \\
f : \mathbb{R}^2\times\mathbb{R} \to \mathbb{R}, & f(x,u) := & (r(x)u - q(x))^2, \\
g : \mathbb{R}^2 \to \mathbb{R}, & g(x,p) := & 3p_1^2 + 4p_2p_1(x) + 4p_2^2 = \langle Ap, p\rangle,
\end{array}$$

With this choices, we set

$$F(u) := \int_\Omega f(x,u(x))\,dx = \int_\Omega (r(x)u(x) - q(x))^2\,dx = \|ru - q\|^2_{L^2},$$
$$G(p) := \int_\Omega g(x,p(x))\,dx = \int_\Omega \Big(3p_1^2(x) + 4p_2(x)p_1(x) + 4p_2^2(x)\Big)\,dx = \langle Ap, p\rangle_{L^2}.$$

To prove the convexity of F and G it is enough to check that for a.e. $x \in \Omega$, the functions $u \mapsto f(x,u)$ and $p \mapsto g(x,p)$ are convex, indeed in this case from the convex inequality

$$f(x,\lambda u_1(x) + (1-\lambda)u_2(x)) \le \lambda f(x,u_1(x)) + (1-\lambda)f(x,u_2(x)),$$

holding at a.e. $x \in \Omega$ for $\lambda \in [0,1]$, we obtain the corresponding relation

$$F(\lambda u_1 + (1-\lambda)u_2) \le \lambda F(u_1) + (1-\lambda)F(u_2),$$

by integrating on Ω. In the same way, the convexity of $g(x,\cdot)$ for a.e. $x \in \Omega$ yields the convexity of G. We have that for every fixed $x \in \Omega$ the maps $f(x,\cdot)$ and $g(x,\cdot)$ are C^∞. Moreover

$$\partial_u f(x,u) = 2r(x)(r(x)u - q(x)),$$
$$\partial^2_{uu} f(x,u) = 2r^2(x) > 0,$$
$$\mathrm{grad}_p\, g(x,p) = 2Ap = \begin{pmatrix} 6p_1 + 4p_2 \\ 4p_1 + 8p_2 \end{pmatrix},$$

$$\mathrm{Hess}_p\, g(x,p) = 2A = \begin{pmatrix} 6 & 4 \\ 4 & 8 \end{pmatrix}$$

Since $\partial^2_{uu} f > 0$ and the eigenvalues of $\mathrm{Hess}_p g$ are $\left\{7+\sqrt{17}, 7-\sqrt{17}\right\}$, both stricly positive and so $\mathrm{Hess}_p g$ is positive definite, we have that $u \mapsto f(x,u)$ e $p \mapsto g(x,p)$ are proper and convex for every fixed $x \in \Omega$, hence F and G are convex. Since A is positive definite, we have $G(p) \geq 0$ and $G(0) = 0$, thus g is proper. Similarly, we have $F(u) \geq 0$ and $F(0) = \|q\|_{L^2} < +\infty$ according to the boundedness of Ω. Thus also F is proper. Since F and G are convex, the same holds for the composition $G \circ \Lambda$ of a convex proper function with a linear one and for the sum of convex functions $\mathscr{F}$. Moreover, $G \circ \Lambda(0) = G(0) < +\infty$ and so $\mathscr{F}(0) < +\infty$ thus also $G \circ \Lambda$ and $\mathscr{F}$ are proper. Finally, the strict convexity of G implies that $\mathscr{F}$ is strictly convex.

We prove now some regularity properties of F and G. Concerning G, we have $\lambda^- \|p\|^2_{L^2} \leq G(p) \leq \lambda^+ \|p\|^2_{L^2}$, where $\lambda^{\pm} = 7 \pm \sqrt{17}$. In particular, for any fixed $\bar{p} \in Y$, $\delta > 0$, we have for all $p \in B_{L^2}(\bar{p}, \delta)$

$$G(p) \leq \lambda^+ \|p\|^2_{L^2} = \lambda^+ (\|p - \bar{p}\|_{L^2} + \|\bar{p}\|_{L^2})^2 \leq \lambda^+ (\delta + \|\bar{p}\|_{L^2})^2,$$

in particular, there exists a neighborhood of $\bar{p}$ and the map G is uniformly upper bounded on that neighborhood. Thus G is continuous at $\bar{p}$. By the arbitrariness of $\bar{p}$, we conclude that G is continuous.

Concerning F, to prove its continuity in X it is enough to show that it is continuous in L^2 since the convergence in X implies the convergence in L^2. We have

$$\begin{aligned} F(u) =& \|ru - q\|^2_{L^2} \leq (\|ru\|_{L^2} + \|q\|_{L^2})^2 \leq \left(\sqrt{\|(ru)^2\|_{L^1}} + \|q\|_{L^2} \right)^2 \\ &\leq \left(\sqrt{\|r^2\|_{L^\infty} \cdot \|u^2\|_{L^1}} + \|q\|_{L^2} \right)^2 \\ &\leq \left(\|r\|_{L^\infty} \cdot \|u\|_{L^2} + \|q\|_{L^2} \right)^2, \end{aligned}$$

where we used the Hölder inequality to estimate

$$\|(ru)^2\|_{L^1} \leq \|r^2\|_{L^\infty} \|u^2\|_{L^1} = \|r\|^2_{L^\infty} \|u\|^2_{L^2}.$$

For any fixed $\bar{u} \in L^2$ and $\delta > 0$ we then have for all $u \in B_{L^2}(\bar{u}, \delta)$

$$\begin{aligned} F(u) &\leq (\|r\|_{L^\infty} \cdot \|u\|_{L^2} + \|q\|_{L^2})^2 \\ &\leq (\|r\|_{L^\infty} \cdot (\|u - \bar{u}\|_{L^2} + \|\bar{u}\|_{L^2}) + \|q\|_{L^2})^2 \\ &\leq \left(\|r\|_{L^\infty} \cdot (\delta + \|\bar{u}\|_{L^2}) + \|q\|_{L^2} \right)^2, \end{aligned}$$

and so, reasoning exactly as for G, we have that F is continuous in L^2 and so also in X.

In order to prove the existence of the solution, by the reflexivity of X, it remains to prove only that if $\|u\|_X \to +\infty$ we have $\mathscr{F}(u) \to +\infty$. Since $F \geq 0$, we have

$$\mathscr{F}(u) = F(u) + G(\nabla u) \geq G(\nabla u) \geq \lambda^- \|\nabla u\|_{L^2}.$$

By the boundedness of Ω, Poincaré's inequality yields the existence of a suitable constant $C > 0$ depending only on Ω such that $\|u\|_{L^2} \leq C \|\nabla u\|_{L^2}$, and so $\|u\|_X \leq (C+1)\|\nabla u\|_{L^2}$. We conclude that $\mathscr{F}(u) \geq \dfrac{\lambda^-}{C+1} \|u\|_X$, and so if $\|u\|_X \to +\infty$ we have $\mathscr{F}(u) \to +\infty$.

We have that $\mathscr{F}$ is convex, proper, l.s.c. and coercive on a reflexive space, thus admits a minimizer $\hat{u}$. Moreover, by the strict convexity, this minimizer is unique.

We want to compute now the conjugate of F and G. Since f, g are Carathéodory functions, to conjugate F, G as functions defined on L^2 it is enough to take the integral of the conjugate of $f(x, \cdot)$ and $g(x, \cdot)$, respectively. Thus we want to compute

$$f^*(x,u) = \sup_{w \in \mathbb{R}} \left\{ \langle w, u \rangle - f(x,u) \right\}, \qquad g^*(x,p^*) = \sup_{p \in \mathbb{R}^2} \left\{ \langle p^*, p \rangle - g(x,p) \right\}$$

(the conjugate is taken only w.r.t. u and p, respectively). The map $u \mapsto f(x,u)$ is superlinear for every fixed x, thus the supremum it is attained in $\mathbb{R}$. Moreover, $u \mapsto f(x,u)$ is smooth, thus to detect it we take the derivative in u of the argument of the supremum in f^*, and find the points where it vanishes.

$$\frac{d}{du}[\langle w,u\rangle - f(x,u)] = w - 2r(x)\,(r(x)u - q(x)) = 0,$$

and so

$$u = \frac{w - 2q(x)r(x)}{2r^2(x)}.$$

We keep track of the fact that, for every fixed $x \in \Omega$, the supremum in the formula defining $f^*(x,w)$ is attained uniquely at $u \in \mathbb{R}$ satisfying

$$w = 2r(x)\,(r(x)u - q(x))\,.$$

By substitution, we have

$$f^*(x,w) = \frac{w(4q(x)r(x)+w)}{4r^2(x)},$$

Since the map f was measurable in x and continuous w.r.t. u, we have that f is a normal integrand. Moreover, taken $u_0 \equiv 0$, we have $F(u_0) < +\infty$. Hence we can compute the conjugate of F by conjugating the integrand function. Thus if $w \in L^2(\Omega)$, we have

$$\begin{aligned} F^*(w) &= \int_\Omega f^*(x,w(x))\,dx = \int_\Omega \frac{w(x)(4q(x)r(x)+w(x))}{4r^2(x)}\,dx = \left\|\frac{w}{2r}\right\|^2_{L^2} + 2\langle \frac{w}{2r}, q\rangle_{L^2}, \\ &= \left\|\frac{w}{2r}+q\right\|^2_{L^2} - \|q\|_{L^2}. \end{aligned}$$

otherwise, since F is continuous w.r.t. the L^2-norm, we have that $F^*(w) = +\infty$ if $w \in X' \setminus L^2$.

Thanks to the previous computations on f, we have that the supremum in the definition of $F^*(w)$ with $w \in L^2$ is attained for $u \in L^2$ satisfying $w = 2r(ru - q)$ (equality in L^2).

In the same way, to compute g^* we notice that $p \mapsto g(x,p)$ is superlinear for every x (indeed, g does not depend on x), thus the supremum in the expression of g^* is attained, and by the smoothness of g to detect it it is enough to study the critical points of the argument of the supremum in g^*. Thus

$$\nabla_p[\langle p^*, p\rangle - g(x,p)] = p^* - 2Ap = 0,$$

which implies $p = \frac{1}{2}A^{-1}p^*$ and so

$$g^*(x,p^*) = \frac{1}{4}\langle A^{-1}p^*, p^*\rangle.$$

Thus we have for all $p^* \in L^2(\Omega;\mathbb{R}^d)$

$$\begin{aligned} G^*(p^*) &= \int_\Omega g^*(x,p^*(x))\,dx = \frac{1}{4}\int_\Omega \langle A^{-1}p^*(x), p^*(x)\rangle\,dx \\ &= \frac{1}{32}\int_\Omega \left(4p_1^{*2}(x) - 4p_1^*(x)p_2^*(x) + 3p_2^{*2}(x)\right)\,dx, \end{aligned}$$

since $A^{-1} = \begin{pmatrix} \frac{1}{2} & -\frac{1}{4} \\ -\frac{1}{4} & \frac{3}{8}\end{pmatrix} = \frac{1}{8}\begin{pmatrix} 4 & -2 \\ -2 & 3\end{pmatrix}$.

We notice that for any $w \in L^2$, if we define

$$u_w(x) = \frac{w(x) - 2q(x)r(x)}{2r^2(x)} \text{ for a.e. } x \in ,$$

we have that $u_w \in L^2(\Omega)$ (since r is bounded away from 0 in Ω) and for a.e. $x \in \Omega$ it holds $f(x,u_w(x)) + f^*(x,w(x)) - w(x)u_w(x) = 0$. Conversely, by the uniqueness of the point of supremum in the computation of f^*, for any given $w \in L^2(\Omega)$ we have that $f(x,u(x)) + f^*(x,w(x)) - w(x)u_w(x) = 0$ for a.e. $x \in \Omega$ only if $u = u_w$ in $L^2(\Omega)$.

Similarly, we have

$$G(p) + G^*(p^*) = \langle p^*, p\rangle_{L^2}$$

if and only if $p = \dfrac{1}{2}A^{-1}p^*$ in L^2.

The dual problem is

$$\sup_{\varphi\in Y'}[-F^*(\Lambda^*\varphi) - G^*(-\varphi)].$$

Since F, G, Λ are all continuous, the stability criterion is trivially satisfied: we have to find an element $u_0 \in X$ such that $F(u_0) < +\infty$ and G must be continuous at Λu_0. It is enough to take $u_0 = 0$. Thus the dual problem has a solution $\hat{\varphi}$. If $\Lambda^*\varphi \in H^{-1}(\Omega) \setminus L^2(\Omega)$ we have $-F^*(\Lambda^*\varphi) = -\infty$ while $G^*(-\hat{\varphi}) < +\infty$, hence since the supremum is attained at $\hat{\varphi}$, we must have $\Lambda^*\hat{\varphi} = -\operatorname{div}\hat{\varphi} \in L^2(\Omega)$. We have that $\varphi \mapsto F^*(\Lambda^*\varphi) + G^*(-\varphi)$ is clearly convex and l.s.c., and moreover it is strictly convex, since the eigenvalues of A^{-1} are the inverse of the eigenvalues of A, so both of them are strictly positive. In particular, the solution $\hat{\varphi}$ of the dual problem is unique.

The unique solution $\hat{u}$ of the primal problem and $\hat{\varphi}$ of the dual problem are linked by the extremality conditions

$$\begin{cases} F(\hat{u}) + F^*(\Lambda^*\hat{\varphi}) = \langle \Lambda^*\hat{\varphi}, \hat{u}\rangle_{X',X}, \\ \\ -G(\Lambda\hat{u}) - G^*(-\hat{\varphi}) = \langle \hat{\varphi}, \Lambda\hat{u}\rangle_{Y',Y}. \end{cases}$$

Since in our case we have $\Lambda^*\hat{\varphi} \in L^2$, these relations become

$$\begin{cases} \displaystyle\int_\Omega (f(\hat{u}(x)) + f^*(\Lambda^*\hat{\varphi}(x)) - \Lambda^*\hat{\varphi}(x)\cdot\hat{u}(x))\,dx = 0, \\ \\ \displaystyle\int_\Omega (g(\Lambda\hat{u}) + g^*(-\hat{\varphi}) - \langle -\hat{\varphi}(x), \Lambda\hat{u}(x)\rangle)\,dx = 0. \end{cases}$$

The integrand functions are always nonnegative, and they vanishes if and only if they vanish a.e. But this implies

$$\begin{cases} -\operatorname{div}\hat{\varphi} = 2r(ru - q), \\ -\hat{\varphi} = 2A\nabla\hat{u}, \end{cases}$$

hence we obtain for the minimizer the following PDEs, whose solution must be understood in the weak sense

$$\begin{cases} \operatorname{div}(2A\nabla\hat{u}) = 2r(ru - q), \\ u_{|\partial\Omega} = 0. \end{cases}$$

EXERCISE 9.2. Let Ω be an open bounded subset of $\mathbb{R}^2$, $q \in L^2(\Omega;\mathbb{R}^2)$. Find the projection of q on the set $G := \{\nabla v :\ v \in H^1_0(\Omega;\mathbb{R})\}$.

SOLUTION. We notice that G is a vector space of $X = L^2(\Omega;\mathbb{R}^2)$. Recalling that $u \mapsto \|\nabla u\|_{L^2}$ defines a norm on $H^1_0(\Omega)$ which, thanks to Poincaré's inequality, is equivalent to the norm of $H^1_0(\Omega)$, we have that G is closed in L^2. In particular, by the theorem on projection on closed convex sets in Hilbert spaces, there exists a projection $\nabla\bar{v} = \pi_G(q) \in G$, i.e., a minimizer of $\dfrac{1}{2}\|q - \nabla v\|^2_{L^2}$, which is characterized by

$$\langle q - \nabla\bar{v}, \nabla v - \nabla\bar{v}\rangle_{L^2} \le 0,$$

for all $\nabla v \in G$. Since G is a vector space, we have that equality holds. Thus we obtain

$$\langle q - \nabla\bar{v}, \nabla w\rangle = 0,$$

for all $w \in H^1_0$, and so

$$\langle q, \nabla w\rangle_{L^2} = \langle \nabla\bar{v}, \nabla w\rangle,$$

This means

$$\langle -\operatorname{div} q, w\rangle_{H^{-1},H^1_0} = \langle -\triangle\bar{v}, w\rangle_{H^{-1},H^1_0},$$

which implies that $\bar{v}$ is the unique solution in the weak sense to $\begin{cases} \triangle v = \operatorname{div} q, \\ v_{|\partial\Omega} = 0. \end{cases}$

EXERCISE 9.3. Let Ω be an open bounded subset of $\mathbb{R}^2$, $q \in L^2(\Omega;\mathbb{R}^2)$ be fixed. Set:

$$\mathscr{C} := \left\{ v \in H^1_0(\Omega;\mathbb{R}) : \|\nabla v - q\|_{L^2(\Omega;\mathbb{R}^2)} \le 1 \right\}.$$

Assume $\mathrm{int}\mathscr{C} \neq \emptyset$ and consider the problem

$$\inf_{u\in\mathscr{C}} \int_\Omega \frac{1}{2}|u(x) - \cos|x||^2\,dx.$$

(1) Prove that the problem admits a unique solution.
(2) Formulate the problem in the form $\mathscr{F}(u) = F(u) + G\circ\Lambda(u)$ where $F : X \to]-\infty,+\infty]$, $G : Y \to]-\infty,+\infty]$, and $\Lambda : X \to Y$, carefully precising the functional spaces X, Y and discussing the regularity of F, G, Λ.
(3) Write the dual problem and the extremality relations. Establish if the dual problem admits an unique solution.
(4) *(Not mandatory)* Using the previous results, write down a partial differential equation satisfied by the minimum.

SOLUTION. Set $X = H^1_0(\Omega;\mathbb{R})$, $X' = H^{-1}(\Omega;\mathbb{R})$, $Y = Y' = L^2(\Omega;\mathbb{R}^2)$, $\Lambda = \nabla : X \to Y$ is linear and continuous since it is clearly linear, and

$$\|\nabla u\|_{L^2} \le \|u\|_{L^2} + \|\nabla u\|_{L^2} = \|u\|_{H^1_0}.$$

Its adjoint is $\Lambda^* = -\mathrm{div} : Y' \to X'$ is also linear and continuous. Define $r(x) = \cos|x|$. We set $G(p) = I_{\overline{B_{L^2}(q,1)}}(p)$, $F(u) = \frac{1}{2}\|u - r\|^2_{L^2}$. Clearly F is strictly convex and continuous with respect to the norm of L^2, since we have $F(u) = \frac{1}{2}\|u - r\|^2_{L^2}$, and so also for the norm of X. Since $\overline{B_{L^2}(q,1)}$ is closed in L^2 and convex we have that G is convex and l.s.c.

We now verify that $\mathscr{F}$ is coercive on the reflexive space X. Indeed, it is enough to show that if $\|u\|_X \to +\infty$ we have $\mathscr{F}(u) \to +\infty$. Since Ω is bounded, we can use Poincaré inequality: there exists $C > 0$ depending only on Ω such that

$$\|u\|_{H^1_0} \le C \cdot \|\nabla u\|_{L^2}.$$

In particular, since $\mathscr{F}(u) \ge G(\nabla u)$, if $\|\nabla u\|_{L^2} > 1 + \|q\|_{L^2}$ we have $G(\nabla u) = +\infty$ and so coercivity follows.

We have

$$G^*(p^*) = I^*_{\overline{B_{L^2}(q,1)}}(p) = \sigma_{\overline{B_{L^2}(q,1)}}(p^*) = \sup_{p\in\overline{B_{L^2}(q,1)}} \langle p^*, p\rangle = \sup_{\eta\in\overline{B_{L^2}(0,1)}} \langle p^*, q+\eta\rangle = \langle p^*, q\rangle_{L^2} + \|p^*\|_{L^2}.$$

The supremum is attained for $p = q + \eta$ where $\eta = \dfrac{p^*}{\|p^*\|_{L^2}}$ if $p^* \neq 0$, otherwise η may be any element of $\overline{B_{L^2}(0,1)}$. Given $u^* \in H^{-1}\setminus L^2$, we can find a sequence $\{u_n\}_{n\in\mathbb{N}} \subseteq H^1_0$ such that $u_n \to u$ in L^2 and $\langle u^*, u_n\rangle_{X',X} \to +\infty$, thus

$$F^*(u^*) = \sup_{u\in H^1_0} \langle u^*, u\rangle_{X',X} - F(u) \ge \limsup_{n\to\infty} \langle u^*, u\rangle_{X',X} - F(u_n) = +\infty,$$

since $F(u_n) \to F(u)$ by continuity of F in L^2. Instead, if $u^* \in L^2 \subseteq H^{-1}$ we can conjugate F^* as it was defined in L^2, since F is continuous in L^2 and X is dense in L^2

$$\begin{aligned}
F^*(u^*) &= \sup_{u\in H^1_0} \langle u^*, u\rangle_{X',X} - F(u) = \sup_{u\in H^1_0} \langle u^*, u\rangle_{L^2} - F(u)\\
&= \sup_{u\in L^2} \langle u^*, u\rangle_{L^2} - F(u)\\
&= \sup_{u\in L^2} \left\{ \langle u^*, u\rangle_{L^2} - \frac{1}{2}\|u - r\|^2_{L^2} \right\}\\
&= \sup_{u\in L^2} \left\{ \frac{1}{2}\|u^*\|^2_{L^2} - \frac{1}{2}\|u^*\|^2_{L^2} + \langle u^*, u\rangle_{L^2} - \langle u^*, r\rangle_{L^2} + \langle u^*, r\rangle_{L^2} - \frac{1}{2}\|u - r\|^2_{L^2} \right\}
\end{aligned}$$

$$
\begin{aligned}
=&\frac{1}{2}\|u^*\|_{L^2}^2+\langle u^*,r\rangle_{L^2}+\sup_{u\in L^2}\left\{-\frac{1}{2}\|u^*\|_{L^2}^2+\langle u^*,u-r\rangle_{L^2}-\frac{1}{2}\|u-r\|_{L^2}^2\right\}\\
=&\frac{1}{2}\|u^*\|_{L^2}^2+\langle u^*,r\rangle_{L^2}+\sup_{u\in L^2}\left\{-\frac{1}{2}\|u^*-(u-r)\|_{L^2}^2\right\}=\frac{1}{2}\|u^*\|_{L^2}^2+\langle u^*,r\rangle_{L^2}\\
=&\frac{1}{2}\|u^*+r\|_{L^2}^2-\frac{1}{2}\|r\|_{L^2}^2,
\end{aligned}
$$

and the supremum is achieved for $u=u^*+r$.

The dual problem is

$$\sup_{\varphi\in Y'}[-F^*(\Lambda^*\varphi)-G^*(-\varphi)],$$

which in our case becomes

$$-\inf_{\substack{\varphi\in L^2\\ \mathrm{div}\varphi\in L^2}}\frac{1}{2}\|-\mathrm{div}\varphi+r\|_{L^2}^2-\frac{1}{2}\|r\|_{L^2}^2+\langle\varphi,q\rangle_{L^2}+\|\varphi\|_{L^2}.$$

We prove now the stability criterion: since if $\psi\in\mathrm{int}\mathscr{C}$, then there exists $\delta>0$ such that $\psi+\delta\xi\in\mathscr{C}$, i.e., $\|\nabla\psi+\delta\nabla\xi-q\|_{L^2}^2\le 1$, for all $\xi\in\overline{B_{H_0^1}(0,1)}$. By possibly replacing ξ by $-\xi$, we can always assume that $\langle\nabla\psi-q,\nabla\xi\rangle_{L^2}\ge 0$ and, by Poincaré's inequality, we have that

$$1=\|\xi\|_{H_0^1}\le C\cdot\|\nabla\xi\|_{L^2}.$$

Thus we have

$$
\begin{aligned}
\|\nabla\psi-q\|_{L^2}^2+\frac{\delta^2}{C^2}\le&\|\nabla\psi-q\|_{L^2}^2+\delta^2\|\nabla\xi\|_{L^2}^2\le\|\nabla\psi-q\|_{L^2}^2+\delta^2\|\nabla\xi\|_{L^2}+2\langle\nabla\psi-q,\nabla\xi\rangle\\
\le&\|\nabla\psi+\delta\nabla\xi-q\|_{L^2}^2\le 1,
\end{aligned}
$$

which implies that $\|\nabla\psi-q\|_{L^2}<1$, thus $\nabla\psi$ belongs to the interior of $\overline{B_{L^2}(q,1)}$, and so G is constantly zero in a suitable small neighborhood of $\nabla\psi$, in particular it is continuous at $\nabla\psi$. We have that F is continuous at ψ, hence finite at ψ, so stability criterion holds and a solution to the dual problem exists.

Extremality relations for a solution $\hat{u}$ of the primal problem and a solution $\hat{\varphi}$ of the dual problem are

$$\begin{cases}F(\hat{u})+F^*(\Lambda^*\hat{\varphi})=\langle\Lambda^*\hat{\varphi},\hat{u}\rangle_{X',X},\\[2ex] -G(\Lambda\hat{u})-G^*(-\hat{\varphi})=\langle\hat{\varphi},\Lambda\hat{u}\rangle_{Y',Y},\end{cases}.$$

The second extremality condition amounts to say that $\nabla\hat{u}-q\in\overline{B_{L^2}(0,1)}$ (to have $G(\nabla\hat{u})<+\infty$) and $\|-\hat{\varphi}\|_{L^2}=\langle-\hat{\varphi},q-\nabla\hat{u}\rangle$. More precisely, this conditions states that by Hölder's inequality

$$\|-\hat{\varphi}\|_{L^2}=\langle-\hat{\varphi},q-\nabla\hat{u}\rangle\le\|-\hat{\varphi}\|_{L^2}\cdot\|q-\nabla\hat{u}\|_{L^2}\le\|-\hat{\varphi}\|_{L^2},$$

since $\nabla\hat{u}-q\in\overline{B_{L^2}(0,1)}$. Thus the second extremality condition can be rewritten as

$$\hat{\varphi}=\lambda(\nabla\hat{u}-q)\text{ where }\begin{cases}\lambda=0,\text{ if }\|q-\nabla\hat{u}\|_{L^2}<1,\\[2ex] \lambda\ge 0,\text{ if }\|q-\nabla\hat{u}\|_{L^2}=1.\end{cases}$$

Since the first extremality condition provides $\hat{u}=-\mathrm{div}\,\hat{\varphi}+r$, we obtain the following PDE for $\hat{u}$, to be undestood in the weak sense

$$\begin{cases}\hat{u}-r=\lambda(\mathrm{div}\,q-\triangle\hat{u}),\text{ in }\Omega,\\ u_{|\partial\Omega}=0,\end{cases}\qquad\text{and so }\hat{\varphi}=\lambda(\nabla\hat{u}-q).$$

If the dual problem admits the solution $\varphi=0$, from the first extremality relation we must have $\hat{u}=r\in\mathscr{C}$ as unique solution of the primal problem, in particular $\|q-\nabla r\|_{L^2}=\|q-\nabla\hat{u}\|_{L^2}\le 1$, thus

- If $\|q-\nabla r\|_{L^2}<1$, the only solution of the dual problem is $\hat{\varphi}=0$.

- If $\|q - \nabla r\|_{L^2} = 1$, we have that $\hat{\varphi} = \lambda(\nabla r - q)$ is a nonzero solution of the dual problem if and only if it is divergence-free, i.e., $\operatorname{div} q = \triangle r$.

Summarizing, if the dual problem admits the null solution then $\|q - \nabla r\|_{L^2} \leq 1$ and the dual problem has 0 as its unique solution, unless $\operatorname{div} q = \triangle r$ and $\|q - \nabla r\|_{L^2} = 1$. In this latter case, $\hat{\varphi} = \lambda(\nabla r - q)$, $\lambda \geq 0$ are solutions of the dual problem and $\hat{u} = r \in \mathscr{C}$ is the unique solution of the primal problem.

If the dual problem do not admit the null solution, we have necessarily $\|q - \nabla \hat{u}\|_{L^2} = 1$.

CHAPTER 2

Second part

1. Lecture of 5 november 2018: Differentiation in infinite dimensional spaces (1h)

It is well known that given a differentiable function $g : \Omega \to \mathbb{R}$, where Ω is an open subset of $\mathbb{R}^n$, the point of minimum of f must be searched among the *critical points*, i.e., the points $x \in \Omega$ satisfying $df(x, y) = 0$. If the function $g \in C^2$, it is possible to study the Hessian matrix of g and if it is positive definite at a critical point $\bar{x}$, then $\bar{x}$ is a minimizer of f. We will now extend the notion from $\mathbb{R}^n$ to possibly infinite-dimensional spaces.

DEFINITION 1.1 (Frechét differential). Let X, Y be normed spaces, $\Omega \subseteq X$ be open X. A function $f : \Omega \to Y$ is called *Frechét differentiable* at $x_0 \in \Omega$ (shortly, F-differentiable at x_0) if there exists $A : X \to Y$ linear and continuous such that

$$\lim_{\|h\|_X \to 0} \frac{f(x_0 + h) - f(x_0) - Ah}{\|h\|_X} = 0.$$

In this case, the operator A is unique and called *Frechét differential* of f at x_0, and is denoted by $A = f'(x_0) = Df(x_0)$.

REMARK 1.2. We recall the following facts:

(1) If the Frechét differential exists, then it is unique.
(2) Every function which is Frechét differentiable at a point, it is also continuous at the same point.
(3) The differential is invariant if we change the norms on X and Y with equivalent norms.
(4) *chain's rule*: if X, Y, Z are normed space, Ω is an open subset of X, $x_0 \in \Omega$, V is open subset of Y, $f : \Omega \to V$, $g : V \to Z$, f differentiable at x_0 and g differentiable at $f(x_0)$, then $g \circ f : \Omega \to Z$ is differentiable at x_0 and
$$(g \circ f)'(x_0) = g'(f(x_0)) \circ f'(x_0),$$
which is a linear and continuous operator from X to Z.
(5) if f is constant in Ω then $f'(x) = 0$ for all $x \in \Omega$ (converse holds if Ω is connected)
(6) if $f : X \to Y$ is linear and continuous, then $f'(x) = f$ for all $x \in X$.
(7) if $f : X \times Y \to Z$ is bilinear and continuous, then it is differentiable at every point and $f'(x, y)(h, k) = f(h, y) + f(x, k)$ for all $(h, k) \in X \times Y$, indeed by definition
$$\begin{aligned}\frac{\|f((x,y) + (h,k)) - f(x,y) - f(h,y) - f(x,k)\|_Z}{\|h\|_X + \|k\|_Y} &\leq \frac{\|f(h,k)\|_X}{\|h\|_X + \|k\|_Y} \\ &\leq \frac{\|h\|_X \|k\|_Y}{\|h\|_X + \|k\|_Y} \left\| f\left(\frac{h}{\|h\|_X}, \frac{k}{\|k\|_Y}\right) \right\|_Z \\ &\leq C\|k\|_X \to 0,\end{aligned}$$
recalling that f maps bounded sets to bounded sets since it is bilinear and continuous.

DEFINITION 1.3. Let X, Y be normed space, and Ω be an open subset of X. Let $f : \Omega \to Y$ be a function which is Frechét differentiable at all $x_0 \in \Omega$. We can define a map $f' : \Omega \to \mathcal{L}(X, Y)$, where $\mathcal{L}(X, Y)$ is the space of linear and continuous functions from X to Y, defined by $x \mapsto f'(x)$. We say that $f \in C^1(\Omega; Y)$ if this map is continuous.

In the case that the dimension of X is greater than 1, the following concept of directional derivative plays an important role. Given $v \in X$, $\|v\|_X \neq 0$, we consider

$$\partial_v f(x_0) = \lim_{t \to 0} \frac{f(x_0 + tv) - f(x_0)}{t} \in Y$$

and study the map $v \mapsto \partial_v f(x_0)$. If this function is linear and continuous from X to Y, it is possible to define another concept of differential.

DEFINITION 1.4 (Gâteaux differential). Let X, Y be normed space, and let Ω be an open subset of X. A function $f : \Omega \to Y$ is called *Gâteaux differentiable* at $x_0 \in \Omega$ (shortly, G-differentiable at x_0) if there exists a linear and continuous operator $A : X \to Y$ such that

$$\lim_{t\to 0} \frac{f(x_0 + tv) - f(x_0)}{t} = Av.$$

In this case, the operator A is called the Gâteaux differential of f at x_0, and we will denote it by $A = f'_G(x_0)$. The map $x \mapsto f'_G(x)$ will be called *the Gâteaux derivative* of f.

REMARK 1.5. If a function admits Gâteaux differential at a point, then it is unique. Even in finite dimension it is possible to give examples of discontinuous G-differentiable functions (hence in particular G-differentiable functions that are not F-differentiable). However, from the definition it easy to see that if a function is F-differentiable then it is also G-differentiable and the two concepts coincides.

EXAMPLE 1.6. The ground space is $\mathbb{R}^2$. Consider the following map $f : \mathbb{R}^2 \to \mathbb{R}$

$$f(x,y) = \begin{cases} \dfrac{x^3 y}{x^6 + y^2}, & \text{if } (x,y) \neq 0, \\ \\ 0, & \text{if } (x,y) = (0,0). \end{cases}$$

We check the limit of f along the curve $\gamma(t) = (t, t^3)$. This curve tends to $(0,0)$ when $t \to 0$.

$$\lim_{t\to 0} f(\gamma(t)) = \lim_{t\to 0} \frac{t^6}{2t^6} = \frac{1}{2} \neq 0 = f(0,0).$$

So the function is not continuous at the origin, and so it cannot be F-differentiable at $(0,0)$. Along the axis, the function is identically zero, so the two partial derivatives vanishes at $(0,0)$. We compute the other directional derivatives along vectors $v = (v_x, v_y)$ with $v_x \neq 0$ and $v_y \neq 0$.

$$\begin{aligned} \partial_v f(0,0) &= \lim_{t\to 0} \frac{f((0,0) + t(v_x, v_y)) - f(0,0)}{t} = \lim_{t\to 0} \frac{f(t(v_x, v_y))}{t} = \lim_{t\to 0} \frac{1}{t} \frac{t^4 v_x^3 v_y}{t^6 v_x^6 + t^2 v_y^2} \\ &= \lim_{t\to 0} \frac{t v_x^3 v_y}{t^4 v_x^6 + v_y^2} = 0. \end{aligned}$$

So directional derivatives along every vector exist at $(0,0)$ and are all 0, so $f'_G(0,0) = 0$, but f is not F-differentiable at $(0,0)$.

THEOREM 1.7 (Lagrange's mean value theorem). *Let X, Y be normed spaces, Ω open subset of X, $f : \Omega \to Y$ a G-differentiable function in Ω. Let $x_1, x_2 \in \Omega$ be such that $\lambda x_1 + (1-\lambda)x_2 \in \Omega$ for all $\lambda \in [0,1]$. Then*

$$\|f(x_2) - f(x_1)\|_Y \leq \sup_{\lambda\in[0,1]} \|f'_G(\lambda x_2 + (1-\lambda)x_1)\|_{\mathcal{L}(X,Y)} \cdot \|x_2 - x_1\|_X,$$

where $\|\cdot\|_{\mathcal{L}(X,Y)}$ is the norm of the space of linear and continuous functions from X to Y.

PROOF. Fix $\varphi \in Y'$ and consider the map $F : [0,1] \to \mathbb{R}$ defined by

$$F(t) = \langle \varphi, f((1-t)x_1 + tx_2) \rangle_{Y',Y}.$$

This map is derivable, and we have

$$F'(t) = \langle \varphi, \langle f'_G((1-t)x_1 + tx_2), (x_2 - x_1) \rangle \rangle.$$

According to classical Lagrange's Mean Value Theorem, there exists $\theta \in]0,1[$ such that $F(1) - F(0) = F'(\theta)$, and so

$$\langle \varphi, f(x_2) \rangle_{Y',Y} - \langle \varphi, f(x_1) \rangle_{Y',Y} = \langle \varphi, \langle f'_G((1-\theta)x_1 + \theta x_2), (x_2 - x_1) \rangle \rangle$$

so we have

$$|\langle \varphi, f(x_2) - f(x_1)\rangle_{Y',Y}| \leq \|\varphi\|_{Y'} \|f'_G((1-\theta)x_1 + \theta x_2)\|_{\mathcal{L}(X,Y)} \|x_2 - x_1\|_X.$$

Since we can always choose $\varphi \in Y'$ such that $\|\varphi\|_{Y'} = 1$ and $\langle \varphi, f(x_2) - f(x_1)\rangle_{Y',Y} = \|f(x_2) - f(x_1)\|_Y$, the proof is concluded. □

THEOREM 1.8 (Total differential). *Let X, Y be normed spaces, Ω be an open subset of X, $f : \Omega \to Y$ be a G-differentiable function in Ω. If $f'_G : \Omega \to \mathcal{L}(X,Y)$ is continuous at $x_0 \in \Omega$ then f is F-differentiable at $x_0 \in \Omega$ and $f'(x_0) = f'_G(x_0)$.*

PROOF. Consider the function $\omega : X \to Y$ defined by

$$\omega(h) := f(x_0 + h) - f(x_0) - \langle f'_G(x_0), h\rangle.$$

This map is G-differentiable in a neighbourhood of 0 and $\omega'_G(h) = f'_G(x_0 + h) - f'_G(x_0)$. Since $\omega(0) = 0$, applying the Mean Value's Theorem, we have

$$\|\omega(h)\|_Y \leq \sup_{\lambda \in [0,1]} \|f'_G(x_0 + \lambda h) - f'_G(x_0)\|_{\mathcal{L}(X,Y)} \cdot \|h\|_X,$$

so by the continuity of f'_G, by letting $h \to 0$ we have

$$\frac{\|\omega(h)\|_Y}{\|h\|_X} \to 0,$$

the thesis follows. □

DEFINITION 1.9. Let X, Y be normed spaces, Ω open subset of X, $f : \Omega \to Y$ a G-differentiable function at $x_0 \in \Omega$. If $f'_G(x_0) = 0$ we say that x_0 is a *critical point* for f.

DEFINITION 1.10 (Higher order derivatives). Let X, Y be normed spaces, Ω be an open subset of X, $f : \Omega \to Y$ be a F-differentiable function. If the Frechét derivative $f' : \Omega \to \mathcal{L}(X,Y)$ is F-differentiable at $x_0 \in \Omega$, then the map $[f']'(x_0) : X \to \mathcal{L}(X,Y)$ will be denoted by $f''(x_0)$ and will be called *second-order differential* of f at x_0. If f' is F-differentiable in Ω, then $f'' : \Omega \to \mathcal{L}(X, \mathcal{L}(X,Y))$ is the *second order derivative* of f. If it is also continuous, then we will say that $f \in C^2(\Omega; Y)$.

PROPOSITION 1.11. *Let X, Y be normed spaces. Then the space $\mathcal{L}(X, \mathcal{L}(X,Y))$ is isometrically isomorphic to the space $\mathcal{L}^2(X \times X, Y)$ of the bilinear and continuous function from $X \times X$ to Y endowed with the norm*

$$\|\varphi\|_{\mathcal{L}^2} = \sup_{\substack{\|h\|_X \leq 1 \\ \|k\|_X \leq 1}} \|\varphi(h,k)\|_Y.$$

PROOF. Given $\psi \in \mathcal{L}(X, \mathcal{L}(X,Y))$ we set

$$\varphi(h,k) = \langle\langle \psi, h\rangle, k\rangle,$$

obtaining $\varphi \in \mathcal{L}^2(X \times X, Y)$.

Conversely, given $\varphi \in \mathcal{L}^2(X \times X, Y)$, for fixed h we have that the map $k \mapsto \varphi(h,k)$ is linear and continuous, and so the map $h \mapsto \varphi(h, \cdot)$ is linear and continuous from X to $\mathcal{L}(X,Y)$.

The other statements are trivial. □

PROPOSITION 1.12 (Taylor's Formula). *Let X, Y be normed spaces, Ω be an open subset of X, $f : \Omega \to Y$ of class $C^2(\Omega; Y)$. Let $x_0 \in \Omega$, $R > 0$ such that $B_X(x_0, R) \subseteq \Omega$. Then for every $h \in B_X(0, R)$ we have*

$$f(x_0 + h) = f(x_0) + \langle f'(x_0), h\rangle + \frac{1}{2}\langle\langle f''(x_0), h\rangle, h\rangle + \eta(h),$$

where $\eta : B_X(0, R) \to Y$ is a function satisfying

$$\lim_{\|h\|_X \to 0} \frac{\|\eta(h)\|_Y}{\|h\|_X^2} = 0.$$

PROOF. Fix $h \in B(0,r)$, $\psi \in Y'$ and define $g_h(t) = \langle \psi, f(x_0+th)\rangle_{Y',Y}$ for $t \in [-1,1]$. We have that $g_h \in C^2(]-1,1[)$ and

$$g_h'(t) = \langle \psi, f'(x_0+th)h\rangle_{Y',Y}$$
$$g_h''(t) = \langle \psi, \langle f''(x_0+th)h,h\rangle\rangle_{Y',Y}.$$

According to classical Taylor's formula applied to g we have:

$$\langle \psi, f(x_0+th)-f(x_0)\rangle_{Y',Y} = \langle \psi, f(x_0+th)\rangle_{Y',Y} - \langle \psi, f(x_0)\rangle_{Y',Y} = g_h(1)-g_h(0) = g_h'(0) + \frac{1}{2}g_h''(\xi_h),$$

for a suitable $\xi_h \in]0,1[$. This implies

$$\begin{aligned}\langle \psi, f(x_0+th)-f(x_0)\rangle_{Y',Y} &= \langle \psi, f'(x_0)h\rangle_{Y',Y} + \frac{1}{2}\langle \psi, \langle f''(x_0+\xi_h h)h,h\rangle\rangle_{Y',Y}\\ &= \langle \psi, f'(x_0)h\rangle_{Y',Y} + \frac{1}{2}\langle \psi, \langle f''(x_0)h,h\rangle\rangle_{Y',Y} + \frac{1}{2}\langle \psi, \langle [f''(x_0+\xi_h h)-f''(x_0)]h,h\rangle\rangle_{Y',Y}\\ &= \langle \psi, f'(x_0)h\rangle_{Y',Y} + \frac{1}{2}\langle \psi, \langle f''(x_0)h,h\rangle\rangle_{Y',Y} + \eta(h)\end{aligned}$$

where $\eta(h) := \frac{1}{2}\langle \psi, \langle [f''(x_0+\xi_h h)-f''(x_0)]h,h\rangle\rangle_{Y',Y}$. By continuity of f'', we have $\eta(h)/\|h\|_X^2 \to 0$ for $h \to 0$, moreover

$$\left| \psi, f(x_0+th) - f(x_0) - f'(x_0)h - \frac{1}{2}\langle f''(x_0)h,h\rangle\rangle_{Y',Y}\right| \le |\langle \psi, \eta(h)\rangle| \le \|\psi\|_{Y'}\|\eta(h)\|_Y.$$

Recalling the arbitrariness of ψ, we have

$$\left\| f(x_0+th) - f(x_0) - f'(x_0)h - \frac{1}{2}\langle f''(x_0)h,h\rangle \right\|_Y \le \|\eta(h)\|_Y,$$

which concludes the proof. □

PROPOSITION 1.13. *Let X be a normed space, Ω open subset of X, $f : X \to \mathbb{R}$ be a G-differentiable function.*

(1) *If $x_0 \in \Omega$ is a local maximum or minimum for f in Ω then $f_G'(x_0) = 0$.*
(2) *If $f \in C^2$ and $x_0 \in \Omega$ is a relative maximum then $\langle\langle f''(x_0), v\rangle, v\rangle \le 0$ for all $v \in X$.*
(3) *If $f \in C^2$ and $x_0 \in \Omega$ is a relative minimum then $\langle\langle f''(x_0), v\rangle, v\rangle \ge 0$ for all $v \in X$.*
(4) *If $f \in C^2$ and $x_0 \in \Omega$ satisfies $f'(x_0) = 0$ and there exists $c > 0$ such that $\langle\langle f''(x_0), v\rangle, v\rangle \le -c\|v\|_X^2$ for all $v \in X$, then x_0 is a relative maximum.*
(5) *If $f \in C^2$ and $x_0 \in \Omega$ satisfies $f'(x_0) = 0$ and there exists $c > 0$ such that $\langle\langle f''(x_0), v\rangle, v\rangle \ge c\|v\|_X^2$ for all $v \in X$, then x_0 is a relative minimum.*

PROOF.

(1) Let x_0 be a maximum and $v \in X$. By assumption, there exists $\delta > 0$ such that if $\|v\|_X < \delta$ then $f(x_0+tv) - f(x_0) \le 0$. Thus
$$\limsup_{t\to 0^+} \frac{f(x_0+tv)-f(x_0)}{t} \ge 0,$$
By the G-differentiabily assumption, the limit exists and is $\langle f_G'(x_0), v\rangle$, so we have $\langle f_G'(x_0), v\rangle \ge 0$ for all $v \in B_X(0,\delta)$. Since $v \in B_X(0,\delta)$ if and only if $-v \in B_X(0,\delta)$, we obtain $\langle f_G'(x_0), -v\rangle \ge 0$, thus $\langle f_G'(x_0), v\rangle = 0$. By the arbitrariness of v, the proof is conclude. Use a very similar argument when x_0 is a minimum.
(2) Let x_0 be a relative maximum. Then there exists $\delta > 0$ such that for all $v \in X$ with $\|v\|_X < \delta$ and for all $t \in [0,1]$ we have
$$0 \ge f(x_0+tv) - f(x_0) = t^2\left(\frac{1}{2}\langle\langle f''(x_0), v\rangle, v\rangle + \|v\|_X^2 \frac{\eta(tv)}{\|tv\|_X^2}\right),$$
where equality is given by Taylor's formula. Dividing by t^2 and letting $t \to 0$, we obtain the thesis.
(3) Use the above argument on $-f$.

(4) Assume that $f'(x_0) = 0$ e $\langle\langle f''(x_0), v\rangle, v\rangle \le -c\|v\|_X^2$ for all $x \in X$ and a suitable $c > 0$ independent of v. By Taylor's formula

$$f(x_0+v) - f(x_0) \le \frac{1}{2}\langle\langle f''(x_0), v\rangle, v\rangle + \eta(v) \le -\frac{c}{2}\|v\|_X^2 \left(1 - \frac{\eta(v)}{\|v\|_X^2}\right).$$

Since $\eta(v)/\|v\|_X^2 \to 0$ as $v \to 0$, there exists $\delta > 0$ such that for all $\|v\|_X < \delta$ we have $f(x_0+v) - f(x_0) < 0$, thus x_0 is a relative maximum.

(5) Use the above argument on $-f$.

□

Summary of Lecture 1

- In this lecture we introduced some concepts of derivatives and differentials in an infinite-dimensional space, namely
  - Frechét differential,
  - Directional derivatives,
  - Gateaux differential.
- It is important to stress that the continuity of the linear operator *must be assumed* in this setting, while in the finite-dimensional cases is authomatically granted.
- Directional derivatives amounts to property of the *slicing* function along lines.
- Gateaux differential amounts to have a linearity and continuity among the directional derivatives, while Frechét differentials gives the strongest conditions.
- We prove some results, like mean value theorem and Taylor's formula by a *scalarization* process, i.e., we consider suitable 1-dimensional functions obtained by applying elements of the dual to the functions, and applying on the scalarized function the classical calculus theorem in finite dimension.

2. Lecture of 12 november 2018: Implicit function in infinite-dimensional spaces. (2h)

THEOREM 2.1 (Dini's implicit function theorem). *Let X, Y, Z be Banach spaces, D be open subset of $X \times Y$, $f : D \to Z$ be a continuous function, $(x_0, y_0) \in D$ be such that $f(x_0, y_0) = 0$. Assume that in a neighborhood of (x_0, y_0) there exists $\partial_Y f(x, y)$ (i.e., the Frechét derivative of $y \mapsto f(x, y)$) and is continuous, and that $\partial_Y f(x_0, y_0)$ is an isomorphism of Y to Z. Then there exist $U \subset X$ and $V \subset Y$, which are neighborhood of x_0 and y_0 respectively, and an unique continuous function $\varphi : U \to V$ such that:*

$$\{(x, y) \in D : f(x, y) = 0\} \cap (U \times V) = \{(x, \varphi(x)) : x \in U\}, \qquad \varphi(x_0) = y_0.$$

We will say that φ locally explicits *f with respect to the variable x in a neighborhood of x_0. Moreover, if f is F-differentiable at (x_0, y_0), we have*

$$\varphi'(x_0) = -\Big(\partial_Y f(x_0, y_0)\Big)^{-1} \circ \partial_X f(x_0, y_0),$$

where $\partial_X f(x_0, y_0)$ is the Frechét derivative of $x \mapsto f(x, y_0)$ at x_0.

PROOF. Set $Q = \partial_Y f(x_0, y_0)$, by assumption $Q : Y \to Z$ is linear, continuous and bijective. Define the map $g : X \times Y \to Y$ by setting

$$g(x, y) = y - Q^{-1}(f(x, y)),$$

and notice that, given x, we have that y is a fixed point of $g(x, \cdot) : Y \to Y$ if and only if $f(x, y) = 0$. Indeed, from $y = y - Q^{-1}(f(x, y))$ we have $Q^{-1}(f(x, y)) = 0$. Since Q and Q^{-1} are isomorphisms by assumption, this can be true if and only if $f(x, y) = 0$.

For every fixed x, we have that $g(x, \cdot) : Y \to Y$ is Fréchet differentiable (thus also G-differentiable), denoted by $\partial_Y g(x, y) : Y \to Y$ the differential of $g(x, \cdot)$, we have

$$\partial_Y g(x, y) = \mathrm{Id}_Y - Q^{-1} \circ \partial_Y f(x, y),$$

and so $\partial_Y g(x_0, y_0) = 0$ by definition of Q.

By continuity of $\partial_Y f(x, y)$, we have that $\partial_Y g(x, y)$ is continuous at (x_0, y_0). Thus for every fixed $0 < \alpha < 1$ there exists a neighborhood $\overline{B(x_0, r_1)} \times \overline{B(y_0, r_1)}$ of (x_0, y_0) contained in D such that

$\|\partial_Y g(x,y)\|_{\mathcal{L}(Y,Y)} \leq \alpha$ for every $(x,y) \in \overline{B(x_0,r_1)} \times \overline{B(y_0,r_1)}$. Choosen $x \in \overline{B(x_0,r_1)}$ and $y_1, y_2 \in \overline{B(y_0,r_1)}$, by mean value theorem we have

$$\|g(x,y_1) - g(x,y_2)\|_Y \leq \sup_{t\in[0,1]} \|\partial_Y g(x, ty_1 + (1-t)y_2)\|_{\mathcal{L}(Y,Y)} \|y_1 - y_2\|_Y \leq \alpha \|y_1 - y_2\|,$$

i.e., g is Lipschitz continuous in the second variable, uniformly w.r.t. the first variable. Moreover, $g(x_0,y_0) = y_0$.

Given $u : B(x_0,r_1] \to B(y_0,r_1]$ continuous, we set $Tu(x) = g(x,u(x))$. We have

$$\begin{aligned}
\|Tu(x) - y_0\|_Y &= \|g(x,u(x)) - g(x_0,y_0)\|_Y \\
&\leq \|g(x,u(x)) - g(x,y_0)\|_Y + \|g(x,y_0) - g(x_0,y_0)\|_Y \\
&\leq \alpha r_1 + \|g(x,y_0) - g(x_0,y_0)\|_Y
\end{aligned}$$

Notice that by the continuity of g, there exist $\delta > 0$ such that if $|x - x_0| < \delta$ we have $\|g(x,y_0) - g(x_0,y_0)\|_Y \leq (1-\alpha)r_1$. Hence, set $E = C^0(B(x_0,\delta], B(y_0,r_1])$ we have for all $u \in E$ that $\|Tu - y_0\|_Y \leq r_1$, and so T maps E into E itself. Since E endowed with the norm of uniform convergece is a Banach space, and

$$\|Tu - Tv\|_\infty = \|g(x,u(x)) - g(x,v(x))\|_\infty \leq \alpha \|u - v\|_\infty,$$

with $0 < \alpha < 1$, we have that T is a contraction, thus it admits a unique fixed point φ. In particular, the graph of φ is the zero level set of f (possibly intersected with $B(x_0,\delta] \times B(y_0,r_1]$.

Suppose now that f is differentiable at (x_0,y_0) and set $P := \partial_X f(x_0,y_0)$. Recalling that $f(x_0,y_0) = 0$, we have

$$f(x,y) = P(x-x_0) + Q(y-y_0) + \sigma(x,y) \cdot (\|x - x_0\|_X + \|y - y_0\|_Y),$$

where $\sigma(x,y)$ vanishes for $(x,y) \to (x_0,y_0)$. Set $y = \varphi(x)$ we have $f(x,\varphi(x)) = 0$ and $y_0 = \varphi(x_0)$, so we obtain

$$0 = P(x-x_0) + Q(\varphi(x) - \varphi(x_0)) + \sigma(x,\varphi(x)) \cdot (\|x - x_0\|_X + \|\varphi(x) - \varphi(x_0)\|_Y),$$

i.e.,

$$\varphi(x) - \varphi(x_0) = -Q^{-1}P(x-x_0) - Q^{-1}(\sigma(x,\varphi(x)) \cdot (\|x - x_0\|_X + \|\varphi(x) - \varphi(x_0)\|_Y)),$$

and dividing by $\|x - x_0\|_X$ we obtain:

$$\frac{\varphi(x) - \varphi(x_0) + Q^{-1}P(x-x_0)}{\|x - x_0\|_X} = -Q^{-1}(\sigma(x,\varphi(x)) \cdot \left(1 + \frac{\|\varphi(x) - \varphi(x_0)\|_Y}{\|x - x_0\|_X}\right).$$

Since $\sigma(x,\varphi(x))$ vanishes for $x \to x_0$, and Q is an isomorphism, it is sufficient to prove that the function

$$\frac{\|\varphi(x) - \varphi(x_0)\|_Y}{\|x - x_0\|_X}$$

is bounded in a neighborhood of x_0. We have:

$$\begin{aligned}
\frac{\|\varphi(x) - \varphi(x_0)\|_Y}{\|x - x_0\|_X} &\leq \left\| Q^{-1}P\left(\frac{x - x_0}{\|x - x_0\|_X}\right)\right\|_Y + \\
&\quad + \|Q^{-1}(\sigma(x,\varphi(x))\|_Y \cdot \left(1 + \frac{\|\varphi(x) - \varphi(x_0)\|_Y}{\|x - x_0\|_X}\right) \\
&\leq \|Q^{-1}P\|_{\mathcal{L}(X,Y)} + \|Q^{-1}\|_{\mathcal{L}(Z,Y)} \cdot \|\sigma(x,\varphi(x))\|_Z + \\
&\quad + \|Q^{-1}(\sigma(x,\varphi(x)))\|_Y \cdot \frac{\|\varphi(x) - \varphi(x_0)\|_Y}{\|x - x_0\|_X}.
\end{aligned}$$

For x sufficiently near to x_0, we can assume $\|Q^{-1}\|_{\mathcal{L}(Z,Y)} \cdot \|\sigma(x,\varphi(x))\|_Z \leq 1/2$. So by substituting the previous inequality and rearranging the terms we have

$$\frac{\|\varphi(x) - \varphi(x_0)\|_Y}{\|x - x_0\|_X} \leq 2\|Q^{-1}P\|_{\mathcal{L}(X,Y)} + 1,$$

the thesis follows. □

THEOREM 2.2 (Inverse Function's Theorem). *Let X, Y be Banach spaces. Let Ω be an open subset of X, $x_0 \in \Omega$, $f : \Omega \to Y$ be of class C^1. If the differential of f at x_0 is an homeomorphism from X to Y, then there exist a neighborhood U of x_0 and an unique map $g : f(U) \to U$ such that $g(f(x)) = x$ for all $x \in U$ and $f(g(y)) = y$ for all $y \in f(U)$. Moreover, g is differentiable at $f(x_0)$ and $Dg(f(x_0)) = (Df(x_0))^{-1}$.*

PROOF. Apply Dini's Theorem to the map $F(x,y) = f(x) - y$. □

Implicit function theorem has the following consequence:

COROLLARY 2.3. *Let X, Y be Banach spaces. Let Ω be an open subset of X, $x_0 \in \Omega$, $f : \Omega \to Y$ be a C^1 map. If the differential of $Df(x_0) : X \to Y$ is an homeomorphism, then there exists an neighborhood U of x_0 and a neighborhood V of $y_0 = f(x_0)$ such that for every $y \in V$ the equation $y = f(x)$ admits a unique solution $x \in U$.*

PROOF. We have $x = \varphi(y)$ where $\varphi : V \to U$ is the implicit function defined by $F(x,y) = 0$ with $F(x,y) = f(x) - y$ by explicing the x-variable as a function of y-variable. □

REMARK 2.4.

(1) The formula providing the differential of the implicit function can be formally derived assuming the existence of a Frechét differentiable function φ implicitely defined by f in a neighborhood of x_0 and writing the Taylor's formula of the map $x \mapsto f(x, \varphi(x))$ around x_0, and then proceed as in the second part of the proof.

(2) The strategy of the first part of the proof can be interpreted as follows. We fix x sufficiently near to x_0 and we try to solve the equation $f(x,y) = 0$ in y. We use Taylor's formula
$$f(x,y) = f(x,y_0) + \partial_Y f(x,y_0)(y - y_0) + \omega(|y - y_0|),$$
where $\omega : [0,+\infty[\to [0,+\infty[$ is a strictly increasing continuous function satisfying $\omega(0) = 0$. On the other hands, continuity assumptions implies (by possibly changing the modulus ω)
$$f(x,y) = f(x_0,y_0) + \partial_Y f(x_0,y_0)(y - y_0) + \omega(|x - x_0| + |y - y_0|),$$
i.e.
$$Q^{-1} f(x,y) = y - y_0 + Q^{-1}\omega(|x - x_0| + |y - y_0|),$$
and so (by possibly changing again the modulus ω)
$$y \simeq y - Q^{-1} f(x,y) = g(x,y).$$
Thus for fixed x, $g(x,y)$ can be viewed as an *approximate solution* to $f(x,y) = 0$. All the other passages shows that indeed the fixed points y of $g(x,\cdot)$ are exactly the points for which $f(x,y) = 0$, and that we can collect all these solutions for fixed x in a graph of a continuous function of x.

We recall the following result in Functional Analysis.

THEOREM 2.5 (Open mapping theorem). *Let X, Y be Banach spaces and $A : X \to Y$ linear and continuous. The following are equivalent:*

(1) *A is surjective (i.e., $A(X) = Y$);*
(2) *A is open at every point, i.e., the image of open set is open;*
(3) *there exists a constant $M > 0$ such that for every $y \in Y$ there esists $x \in X$ with $y = Ax$ and $\|x\|_X \leq M\|y\|_Y$. In this case, we will define*

$$\operatorname{reg} A = \inf\{M > 0 : \text{ for all } y \in Y \text{ there exists } x \in X \text{ with } y = Ax \text{ and } \|x\|_X \leq M\|y\|_Y\}.$$

PROOF. Omitted. See Theorem II.5 p. 28 in **[3]**. □

REMARK 2.6. We can interpret $\operatorname{reg} A$ as follows. Given $y \in B_Y(0,1) \subseteq Y$, we consider the set $A^{-1}y$ of all $x \in X$ such that $Ax = y$. This set is nonempty by surjectivity, moreover it is closed by continuity, and convex by linearity. Then $\operatorname{reg} A := \sup\limits_{y \in B(0,1)} \inf\limits_{x \in A^{-1}y} \|x\|$.

Next theorem will allow us to relax the assumption $f \in C^1$ (we will follow the approach in **[8]**) in the Inverse Function Theorem.

THEOREM 2.7 (Graves). *Let X, Y be Banach spaces, $x_0 \in X$, $y_0 \in Y$, $\varepsilon > 0$, $f \in C^0(B_X(x_0, \varepsilon); Y)$ with $f(x_0) = y_0$. Let $A : X \to Y$ be a linear, continuous and surjective operator, and let $M > \operatorname{reg} A$. Suppose that there exists $0 < \delta < 1/M$ such that*

$$\|f(x_1) - f(x_2) - A(x_1 - x_2)\|_Y \le \delta \|x_1 - x_2\|_X,$$

for all $x_1, x_2 \in B_X(x_0, \varepsilon)$. Then the equation $y = f(x)$ admits a solution $x \in B(x_0, \varepsilon)$ for all $y \in B_Y(y_0, c\varepsilon)$ where $c = \dfrac{1}{M} - \delta$.

PROOF. without loss of generality, up to translation, we can assume $x_0 = 0$ and $y_0 = f(x_0) = 0$. Let $y \in B_Y(y_0, c\varepsilon)$ where $c, M, \delta, \varepsilon$ are as in the statement.

We define by induction a sequence as follows. Set $x_0 = 0$ and, by surjectivity, from open mapping theorem there exists $x_1 \in X$ such that $A(x_1) = y$ and $\|x_1\|_X \le M\|y\|_Y \le \varepsilon$.

Assume to have defined x_i, $i = 1, \dots, n-1$, such that for every $i = 1, \dots, n-1$ it holds

$$y - f(x_{i-1}) = A(x_i - x_{i-i}), \qquad \|x_i - x_{i-1}\|_X \le M(\delta M)^{i-1}\|y\|_Y.$$

In particular, we have $x_i \in B(x_0, \varepsilon)$, since

$$\begin{aligned}\|x_i\|_X &\le \sum_{j=1}^{i} \|x_i - x_{i-1}\| \le M\sum_{j=1}^{i}(\delta M)^{i-1}\|y\|_Y \le M\sum_{j=1}^{\infty}(\delta M)^{i-1}\|y\|_Y \\ &= \frac{M}{1 - M\delta}\|y\|_Y = \frac{\|y\|_Y}{c} \le \varepsilon.\end{aligned}$$

Define $x_n \in X$ as follows. By surjectivity, there exists x_n such that

$$y - f(x_{n-1}) = A(x_n - x_{n-i}),$$

since there exists ξ_n such that $A\xi_n = y - f(x_{n-1})$, and then it is enough to set $x_n = \xi_n + x_{n-1}$. Moreover, we have also (recalling the inductive step)

$$\begin{aligned}\|\xi_n\|_X = \|x_n - x_{n-1}\|_X &\le M\|y - f(x_{n-1})\|_Y = M\|A(x_{n-1}) - A(x_{n-2}) + f(x_{n-2}) - f(x_{n-1})\|_Y \\ &= M\|f(x_{n-2}) - f(x_{n-1}) - A(x_{n-2} - x_{n-1})\|_Y \le M\delta\|x_{n-2} - x_{n-1}\|_X \\ &\le M\delta M(\delta M)^{n-1}\|y\|_Y = M(\delta M)^n\|y\|_Y,\end{aligned}$$

thus we have as before $x_i \in B(x_0, \varepsilon)$.

Since $\displaystyle\sum_{n=1}^{\infty} \|x_i - x_{i-1}\|_X = \frac{M\|y\|_Y}{1 - \delta M} < +\infty$, we have that $\{x_n\}_{n\in\mathbb{N}}$ is a Cauchy sequence in X, thus converges to $x \in X$ and by passing to the limit in $y - f(x_{n-1}) = A(x_n - x_{n-1})$, we have $y = f(x)$. □

DEFINITION 2.8 (Strict differentiability). Let X, Y be Banach spaces, $\varepsilon > 0$, $f \in C^0(B_X(x_0, \varepsilon); Y)$. We say that f is *strictly differentiable* at x_0 if there exists a linear continuous and surjective operator $A : X \to Y$ such that

$$\lim_{\substack{x_1 \to x_0 \\ x_2 \to x_0}} \frac{\|f(x_1) - f(x_2) - A(x_1 - x_2)\|_Y}{\|x_1 - x_2\|_X} = 0.$$

A function satisfying the above relation is trivially differentiable at x_0, and $A = Df(x_0)$ is its differential at x_0.

REMARK 2.9.

(1) Not every differentiable function is strictly differentiable: indeed, if in $\mathbb{R}$ we consider $f(x) = x^2 \sin(1/x)$ for $x \neq 0$ and $f(0) = 0$, we have that f is differentiable at 0 with derivative equal to 0, however if we take the sequences $x_n = [(n + 1/2)\pi]^{-1}$ and $y_n = x_{n+1}$, we have that $x_n, y_n \to 0$, but $\dfrac{f(x_n) - f(y_n)}{|x_n - y_n|}$ has no limit for $n \to +\infty$.

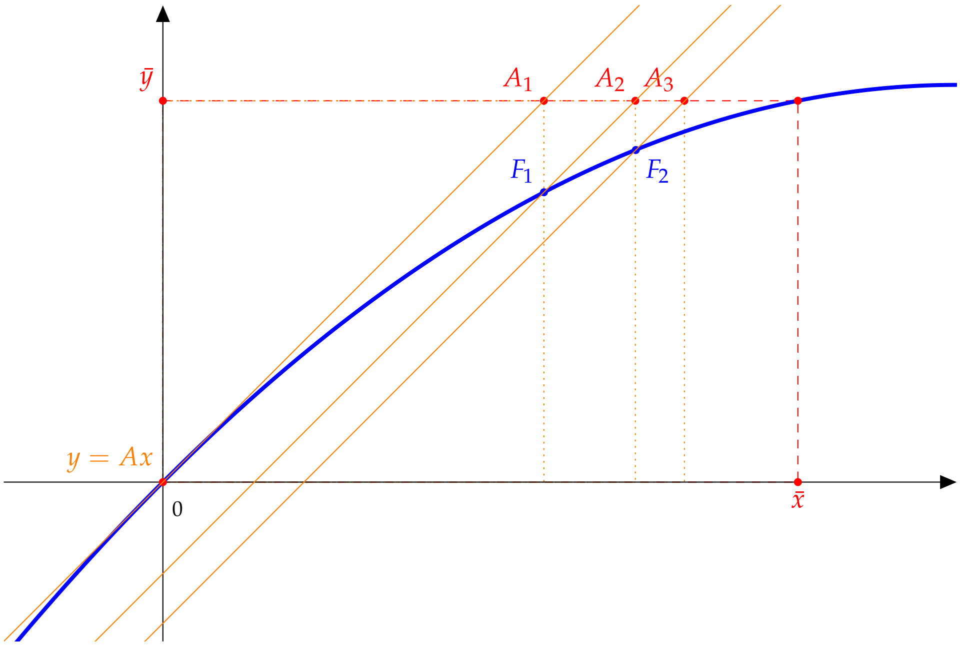


FIGURE 1. Geometrical idea of the proof of Graves' Theorem in $\mathbb{R}$: we consider a linear operator A satisfying the assumptions, then we construct a sequence as follows. The first point is the origin, then we look for the intersection of the line $y = Ax$ with $y = \bar{y}$ to determine A_1, and consider F_1 which is the point of graph f with the same first component of A_1. A_2 is defined as the intersection of the line parallel to $y = Ax$ and passing through F_1 with the line $y = \bar{y}$, and F_2 is the point of graph f with the same first component of A_2. We repeat the construction: given A_{n-1} and F_{n-1}, we define A_n as the intersection between the line parallel to $y = Ax$ and passing through F_{n-1} and $y = \bar{y}$, then F_n will be the point of graph f with the same first component of A_n. The assumptions grant that the sequence of the first components of A_n converges to limit point $\bar{x}$ which will satisfy $\bar{y} = f(\bar{x})$.

(2) C^1-regularity around x_0 is a sufficient condition for strict differentiability at x_0. However it is not necessary. For example, consider the function $f : [-1,1] \to \mathbb{R}$ whose epigraph is

$$\operatorname{epi} f := \overline{\operatorname{co}}\left(\{(\pm 1/k, \beta) : \beta \geq 1/k^2, k \in \mathbb{N}\} \cup \{0\} \times [0, +\infty[\right).$$

Clearly, this map is not differentiable at $1/k$ for every $k \in \mathbb{N}$, thus is not C^1 in any neighborhood of the origin. Given $x, y \in]-1, 1[$ and $\varepsilon > 0$, there is $\delta > 0$ such that if $x, y \in B(0, \delta)$ then

$$\frac{|f(x) - f(y)|}{|x - y|} \leq \varepsilon.$$

Indeed, it is easy to see that f is Lipschitz continuous with constant $\dfrac{1+2k}{k(1+k)}$ on $B(0, 1/k)$, thus for k sufficiently large $\dfrac{1+2k}{k(1+k)} < \varepsilon$ and $\delta = 1/k$. This implies that f is strictly differentiable at 0 and $Df(0) = 0$ even if f is not C^1 in any neighborhood of the origin.

(3) Graves' Theorem requires a sort of *uniform approximate differentiability* of the map f. Indeed, if we require that the condition in Graves' Theorem must hold *for every* $\varepsilon > 0$ and $\delta > 0$ then we have that f must be strictly differentiable.

COROLLARY 2.10. *Let X, Y be Banach spaces, $x_0 \in X$ and $f : X \to Y$ be strictly differentiable at x_0. Suppose that $Df(x_0) : X \to Y$ is surjective. Then there exist a neighborhood U of x_0 and a constant $c > 0$ such that for all $\tau > 0$ with $B(x, \tau) \subseteq U$ it holds*

$$B_Y(f(x), c\tau) \subseteq f(B_X(x, \tau)), \text{ for all } x \in U.$$

In other words, f is locally uniformly open *at every point in a neighborhood of x_0.*

PROOF. By strict differentiability, we can choose $\varepsilon > 0$ such that

$$\|f(x_1) - f(x_2) - A(x_1 - x_2)\|_Y \leq \delta \|x_1 - x_2\|_X,$$

is satisfied with $0 < \delta < 1/M$ and $A = Df(x_0)$ for all $x_1, x_2 \in B(x_0, \varepsilon)$. According to Graves' Theorem, $y = f(x)$ admits a solution $x \in B(x_0, \varepsilon)$ for all $y \in B(f(x_0), c\varepsilon)$, where $c = 1/M - \delta$. This means that

$$B_Y(f(x_0), c\varepsilon) \subseteq f(B_X(x_0, \varepsilon)).$$

Given $\bar{x} \in B_X(x_0, \varepsilon)$ and $\tau > 0$ auch that $B_X(\bar{x}, \tau) \subseteq B_X(x_0, \varepsilon)$, define $\tilde{f}(x) = f(x + \bar{x}) - f(\bar{x})$ for all $x \in B_X(0, \tau)$. We have that $\tilde{f}$ is strictly differentiable at 0 and $D\tilde{f}(0)$ is surjective. But then

$$B_Y(f(\bar{x}), c\tau) \subseteq f(B_X(\bar{x}, \tau)),$$

as desired. □

Now we can introduce a notion of *tangent space* to the zero-level set of a map $f : X \to Y$. To have a properly notion of tangent space to Z at x_0 with $f(x_0) = 0$, we must have an affine space containing x_0 and such that the distance between the points x of this affine space and Z if of higher order with respect to $\|x - x_0\|$.

THEOREM 2.11 (Lyusternik). *Let X, Y be Banach spaces, Ω be an open subset of X, $x_0 \in \Omega$ and $f : \Omega \to Y$ be an F-differentiable function with $f(x_0) = 0$. Assume that f' is continuous at x_0 and that $f'(x_0) : X \to Y$ is surjective with* $\ker f'(x_0) \neq \{0\}$. *Set*

$$Z := \{x \in \Omega : f(x) = 0\},$$

then for all $\varepsilon > 0$ there exists $\delta > 0$ such that $B(x_0, \delta) \subseteq \Omega$ and

$$\operatorname{dist}(x, Z) \leq \varepsilon \|x - x_0\|, \text{ for all } x \in \big(x_0 + \ker f'(x_0)\big) \cap B(x_0, \delta).$$

PROOF. Since f' is continuous at x_0, we have that f is strictly differentiable at x_0. Without loss of generality, we may assume $x_0 = 0$ and $f(x_0) = 0$. Let U be a neighborhood of $x_0 = 0$. For every $\varepsilon > 0$ there exists $\delta_1 > 0$ such that $B(x, \varepsilon \|x\|_X) \subseteq U$ for all $x \in B(0, \delta_1)$. Moreover, for all $x \in \ker f'(x_0) \cap B(0, \delta_1)$, recalling that $x_0 = 0$ and $f(x_0) = 0$, we have

$$0 = \lim_{\substack{x \to x_0 \\ x \in \ker f'(x_0)}} \frac{\|f(x) - f(x_0) - f'(x_0)(x - x_0)\|}{\|x - x_0\|} = \lim_{x \to x_0} \frac{\|f(x) - f'(x_0)(x)\|}{\|x\|} = \lim_{x \to x_0} \frac{\|f(x)\|}{\|x\|},$$

and so for all $c > 0$ there exists $0 < \delta_2 \leq \delta_1$ such that $\|f(x)\| \leq c\varepsilon \|x\|$ for all $x \in \ker f'(x_0) \cap B(0, \delta_2)$. In particular, we can take $0 < \delta \leq \delta_2$ so small that the Corollary of Graves' Theorem applies with $\tau = \varepsilon \|x\|$, so

$$B_Y(f(x), c\varepsilon\|x\|) \subset f(B(x, \varepsilon\|x\|)), \text{ for every } x \in \ker f'(x_0) \cap B(0, \delta),$$

thus, in particular, from

$$0 \in \overline{B_Y(f(x), c\varepsilon\|x\|)} \subset \overline{f(B(x, \varepsilon\|x\|))},$$

we obtain the existence of $\tilde{x} \in \overline{B(x, \varepsilon\|x\|)} \cap Z$ hence

$$\operatorname{dist}(x, Z) \leq \|\tilde{x} - x\| \leq \varepsilon \|x\| = \varepsilon \|x - x_0\|,$$

as desired. □

REMARK 2.12. We can give a smooth, finite-dimensional version of the Lyusternik's Theorem as follows. Let $X = \mathbb{R}^n$, $\Omega \subseteq X$ be open, $Y = \mathbb{R}^k$, $x_0 \in \Omega$, $f : \Omega \to \mathbb{R}^k$ be a C^1 function such that $f(x_0) = 0$. Assume that $k \leq \operatorname{rank}(\operatorname{Jac} f(x_0)) < n$ and set $Z := \{x \in \Omega : f(x) = 0\}$. Then

$$\lim_{\substack{x \to x_0 \\ x - x_0 \in \ker \operatorname{Jac} f(x_0)}} \frac{\operatorname{dist}(x, Z)}{\|x - x_0\|} = 0.$$

LEMMA 2.13 (Orthogonality relations). *Let X, Y be Banach spaces, $A : X \to Y$ be a linear, continuous and surjective operator. Then $A^*(Y') = (\ker A)^\perp$, where*

$$(\ker A)^\perp := \{\eta \in X' : \langle \eta, x\rangle_{X',X} = 0 \text{ for all } x \in \ker A\}.$$

PROOF. Let $\eta \in A^*(Y')$, in particular $\eta = A^*\psi$ for a suitable $\psi \in Y'$. Then

$$\langle \eta, x\rangle_{X',X} = \langle A^*\psi, x\rangle_{X',X} = \langle \psi, Ax\rangle_{X',X} = 0, \text{ for all } x \in \ker A,$$

thus $\eta \in (\ker A)^\perp$, and so $A^*(Y') \subseteq (\ker A)^\perp$.

Conversely, let $\eta \in (\ker A)^\perp$. Since $\langle \eta, x\rangle = 0$ for all $x \in \ker A$, we must have $\ker \eta \supseteq \ker A$. Given $y \in Y$, we prove that η is constant on $A^{-1}(y)$. Indeed, by surjectivity, there exists $\bar{x} \in X$ such that $A(\bar{x}) = y$ and so $A^{-1}(y) = \bar{x} + \ker A$. Since $\ker \eta \supseteq \ker A$, we have that $\eta(x) = \eta(\bar{x})$ for all $x \in A^{-1}(y)$.

So it is well defined the map $\psi : Y \to \mathbb{R}$, with $\psi(y) = \eta(x)$ for all $x \in A^{-1}(y)$.
The map ψ is linear: given $y_1, y_2 \in Y$, $\alpha \in \mathbb{R}$, by surjectivity there exist $x_1, x_2 \in X$ with $y_1 = Ax_1$ and $y_2 = Ax_2$. Then $y_1 + \alpha y_2 = A(x_1 + \alpha x_2)$, thus

$$\psi(y_1 + \alpha y_2) = \eta(x_1 + \alpha x_2) = \eta(x_1) + \alpha\eta(x_2) = \psi(y_1) + \alpha\psi(y_2).$$

We prove now that ψ is continuous. Let G be open in $\mathbb{R}$. We have

$$\begin{aligned}\psi^{-1}(G) &= \{y \in Y : \psi(y) \in G\} = \{Ax \in Y : x \in X, \eta(x) \in G\} = A(\{x \in X : \eta(x) \in G\})\\ &= A(\eta^{-1}(G)).\end{aligned}$$

By continuity, $\eta^{-1}(G)$ is open in X, and by open mapping theorem, A is open, hence $\psi^{-1}(G) = A(\eta^{-1}(G))$ is open, and so ψ is continuous. Finally, for every $x \in X$ we have $\psi(Ax) = \eta(x)$ hence $\eta = A^*\psi \in A^*(Y')$, so $A^*(Y') \supseteq (\ker A)^\perp$. □

The following theorem generalized the classical necessary condition for constrained minima given by Lagrange multipliers rule to the infinite-dimensional case.

THEOREM 2.14 (Lagrange multipliers). *Let X, Y be real Banach spaces, and let $F : X \to \mathbb{R}$, $\Phi : X \to Y$ be functions of class C^1. Define $Z = \{x \in X : \Phi(x) = 0\}$, fix $x_0 \in Z$, and assume that $\Phi'(x_0) : X \to Y$ has closed image in Y. If x_0 is a point of relative maximum or minimum for $F_{|Z}$, then there exist $(\lambda_0, \varphi) \in \mathbb{R} \times Y'$, $(\lambda_0, \varphi) \neq (0, 0)$, such that in X' if holds*

$$\lambda_0 F'(x_0) + (\Phi'(x_0))^*\varphi = 0,$$

moreover, if $\Phi'(x_0)$ is surjective, we can choose $\lambda_0 = -1$.

PROOF. Set $K = \Phi'(x_0)X$, i.e., the image of $\Phi'(x_0)$. By assumption, we have that K is closed.

(1) Assume $K \neq Y$, i.e., $\Phi'(x_0)$ is not surjective. In particular, there exists $\bar{y} \in Y \setminus K$, and by Hahn-Banach we can separate $\{\bar{y}\}$ from K by $\varphi \in Y'$, $\varphi \neq 0$, thus $\phi(\bar{y}) \leq \phi(y)$ for all $y \in K$. Since K is a vector space, we have $\phi(\bar{y}) \leq \phi(ky) = k\phi(y)$ for all $y \in K$, $k \in \mathbb{R}$. This implies $\phi_{|K} = 0$. Thus we choose $\lambda_0 = 0$ and we obtain for all $x \in X$

$$\langle \lambda_0 F'(x_0), x\rangle + \langle (\Phi'(x_0))^*\varphi, x\rangle = \langle \varphi, (\Phi'(x_0)x\rangle = 0,$$

as desired.

(2) Assume now $K = Y$, i.e., $\Phi'(x_0)$ is surjective, and $\ker \Phi'(x_0) = \{0\}$. Then $\Phi'(x_0) : X \to Y$ is an isomorphism, and so $(\Phi'(x_0))^* : Y' \to X'$ is isomorphism. Choosing $\lambda_0 = -1$ and $\varphi = [(\Phi'(x_0))^*]^{-1}F'(x_0)$ yields the thesis in this case.

(3) Assume now $K = Y$, i.e., $\Phi'(x_0)$ is surjective, but $\ker \Phi'(x_0) \neq \{0\}$. We apply Lyusternik's Theorem to Φ. Set $T_{x_0} = x_0 + \ker \Phi'(x_0)$, we have

$$\lim_{\substack{v \to x_0 \\ v \in T_{x_0}}} \frac{\operatorname{dist}(v, Z)}{\|v - x_0\|_X} = 0.$$

Choose $v = x_0 + th$, where $t \in \mathbb{R} \setminus \{0\}$ and $h \in \ker \Phi'(x_0)$, $\|h\|_X = 1$. We have

$$\lim_{t \to 0} \frac{\operatorname{dist}(x_0 + th, Z)}{|t|} = 0.$$

For every $t \neq 0$ there exists $x_t \in Z$ such that

$$\|x_0 + th - x_t\|_X \leq \operatorname{dist}(x_0 + th, Z) + t^2,$$

thus, by setting $r(t) = x_0 + th - x_t$, we have $r(0) = 0$ and

$$\|r'(0)\|_X = \lim_{t \to 0} \left\| \frac{r(t) - r(0)}{t} \right\|_X \leq \lim_{t \to 0} \frac{\operatorname{dist}(x_0 + th, Z)}{|t|} + |t| = 0.$$

Define now $f(t) = F(x_t) = F(x_0 + th - r(t))$. By assumption, f has a maximum or a minimum at $t = 0$, thus $f'(0) = 0$, i.e., $F'(x_0)h = 0$. Since this holds for every $h \in \ker \Phi'(x_0) \cap B_X(0,1)$, by linearity it holds for every $h \in \ker \Phi'(x_0)$, i.e., $F'(x_0) \in (\ker \Phi'(x_0))^\perp$. By the orthogonality relations, we have that $F'(x_0)$ belongs to $(\Phi'(x_0))^* Y'$, thus we can choose $\lambda_0 = -1$ and $\varphi \in [(\Phi'(x_0))^*]^{-1} F'(x_0) \subseteq Y'$.

□

We will introduce now a class of important infinite-dimensional operators. We recall that if V, W are sets, then V^W denotes the set of all maps $\psi : W \to V$.

DEFINITION 2.15 (Superposition operators). Let Ω be an open subset of $\mathbb{R}^d$, B a Borel subset of $\mathbb{R}^p$, $f : \Omega \times B \to \mathbb{R}$ a Carathéodory function. Given $u : \Omega \to B$, define a map $\Phi : B^\Omega \to \mathbb{R}^\Omega$ by setting $\Phi(u)(x) = f(x, u(x))$. The map Φ is called a *superposition operator*.

PROPOSITION 2.16 (First properties of superposition operators). *Let Ω be an open subset of $\mathbb{R}^d$ such that $\mathscr{L}^d(\Omega) < +\infty$, $f : \Omega \times \mathbb{R}^p \to \mathbb{R}$ be a Carathéodory function. Define $\Phi : (\mathbb{R}^p)^\Omega \to \mathbb{R}^\Omega$ by setting $\Phi(u)(x) = f(x, u(x))$.*

(1) *If $u : \Omega \to \mathbb{R}^p$ is measurable, then $\Phi(u)$ is measurable.*
(2) *If $u_n \to u$ in measure, then $\Phi(u_n) \to \Phi(u)$ in measure.*
(3) *If there exists $p, q \geq 1$ such that*

$$|f(x,t)| \leq a(x) + b|t|^{p/q},$$

where $a \in L^q(\Omega; [0, +\infty[)$ and $b \geq 0$, then $\Phi : L^p(\Omega; \mathbb{R}^p) \to L^q(\Omega; \mathbb{R})$ is continuous.

PROOF.

(1) Since u is measurable, there exists a sequence of simple functions $\{u_n\}_{n \in \mathbb{N}}$ pointwise converging to u. The map $x \mapsto f(x, u_n(x))$ is measurable, indeed, if $u_n = \sum_{j=1}^{N^{(n)}} c^{(n)} \chi_{A^{(n)}}$ we have

$$f(x, u_n(x)) = \sum_{j=1}^{N^{(n)}} f(x, c^{(n)}) \chi_{A^{(n)}}(x).$$

Since f is continuous w.r.t. the second variable, we can pass to the limit w.r.t. n thus $x \mapsto f(x, u(x))$ is measurable as pointwise limit of measurable function.

(2) By assumption, there is a set $\mathcal{N} \subseteq \Omega$ such that $\mathscr{L}^d(\mathcal{N}) = 0$ and for all $x \notin \mathcal{N}$ the map $f(x, \cdot)$ is continuous. We have to prove that for all fixed $\varepsilon > 0$, we have

$$\lim_{n \to +\infty} \mathscr{L}^d (\{x \in \Omega \setminus \mathcal{N} : |\Phi(u_n)(x) - \Phi(u)(x)| > \varepsilon\}) = 0.$$

For any $k \in \mathbb{N} \setminus \{0\}$ we define

$$\begin{aligned} \Omega_k &:= \left\{ x \in \Omega \setminus \mathcal{N} : f\left(x, \overline{B(u(x), 1/k)}\right) \subseteq \overline{B(f(x, u(x)), \varepsilon)} \right\} \\ &= \left\{ x \in \Omega \setminus \mathcal{N} : |f(x, \alpha) - f(x, u(x))| \leq \varepsilon \text{ for all } \alpha \in \overline{B(u(x), 1/k)} \right\}. \end{aligned}$$

Using the continuity of $f(x,\cdot)$ for all $x\in\Omega\setminus\mathcal{N}$, we have

$$\Omega_k=\left\{x\in\Omega\setminus\mathcal{N}:\ \sup_{h\in\mathbb{N}}|f(x,\alpha_h^{(k)})-f(x,u(x))|\le\varepsilon\right\},$$

where $\{\alpha_h^{(k)}\}_{h\in\mathbb{N}}$ is a countable set dense in $\overline{B(u(x),1/k)}$ (e.g. $\overline{B(u(x),1/k)}\cap\mathbb{Q}^p$). Since the map $x\mapsto|f(x,\alpha_h^{(k)})-f(x,u(x))|$ is measurable, its pointwise supremum on h is also measurable, and so Ω_k is measurable since it is sublevel of a measurable function.

We have $\Omega_k\subseteq\Omega_{k+1}$. Moreover, by continuity of $f(x,\cdot)$ on $\Omega\setminus\mathcal{N}$, we have $\bigcup_{k=1}^{\infty}\Omega_k=\Omega\setminus\mathcal{N}$. In particular, we have $\mathscr{L}^d(\Omega\setminus\Omega_k)\to0^+$ as $k\to+\infty$. Fix $\eta>0$, then there exists $\bar{k}>0$ such that for $k\ge\bar{k}$ we have $\mathscr{L}^d(\Omega\setminus\Omega_k)\le\eta/2$. Set $A_n:=\left\{x\in\Omega:\ u_n(x)\in\overline{B(u(x),1/\bar{k})}\right\}$ and, since u_n converges to u in measure, there exists $\bar{n}>0$ such that for all $n\ge\bar{n}$ we have $\mathscr{L}^d(\Omega\setminus A_n)<\eta/2$.

If $x\in A_n\cap\Omega_{\bar{k}}$ we have $u_n(x)\in\overline{B(u(x),1/\bar{k})}$ and so $|f(x,u_n(x))-f(x,u(x))|\le\varepsilon$, on the other hand $\mathscr{L}^d(\Omega\setminus(A_n\cap\Omega_{\bar{k}}))=\mathscr{L}^d(\Omega\setminus A_n)+\mathscr{L}^d(\Omega\setminus\Omega_{\bar{k}})\le\eta$.

Thus, for all $\eta>0$ there exists $\bar{n}>0$ such that for all $n\ge\bar{n}$

$$\mathscr{L}^d\left(\{x\in\Omega\setminus\mathcal{N}:\ |\Phi(u_n)(x)-\Phi(u)(x)|>\varepsilon\}\right)\le\mathscr{L}^d(\Omega\setminus(A_n\cap\Omega_{\bar{k}}))\le\eta,$$

which completes the proof.

(3) Assume now that there exists $p,q\ge1$, $a\in L^q(\Omega;[0,+\infty[)$, $b\ge0$, such that $|f(x,t)|\le a(x)+b|t|^{p/q}$. Given $u\in L^p(\Omega;\mathbb{R}^p)$ we have

$$|\Phi(u)(x)|=|f(x,u(x))|\le a(x)+b|u(x)|^{p/q},$$

hence

$$\|\Phi(u)\|_{L^q}=\|a(\cdot)\|_{L^q}+b\|u(\cdot)\|_{L^p}^{p/q}<+\infty.$$

Let now a sequence $\{u_n\}_{n\in\mathbb{N}}$ with $u_n\to u$ in L^p. We have[1]

$$\begin{aligned}
|\Phi(u_n)(x)-\Phi(u)(x)|^q&\le2^{q-1}\left[|\Phi(u_n)(x)|^q+|\Phi(u)(x)|^q\right]\\
&\le2^{q-1}\left[(a(x)+b|u_n(x)|^{p/q})^q+(a(x)+b|u(x)|^{p/q})^q\right]\\
&\le4^{q-1}\left[2a^q(x)+b^q|u_n(x)|^p+b^q|u(x)|^p\right]\\
&\le4^{q-1}\left[2a^q(x)+b^q|(u_n(x)-u(x))+u(x))|^p+b^q|u(x)|^p\right]\\
&\le4^{q-1}\left[2a^q(x)+b^q2^{p-1}|u_n(x)-u(x)|^p+(2^{p-1}+b^q)|u(x)|^p\right]\\
&\le C\left[a^q(x)+|u_n(x)-u(x)|^p+|u(x)|^p\right],
\end{aligned}$$

where $C>0$ is a suitable constant depending on p,q,b. Fix now $\varepsilon,\eta>0$. Since $\Phi(u_n)$ converges in measure to $\Phi(u)$, there exists $\bar{n}>0$ such that set

$$E_n:=\{x\in\Omega:\ |\Phi(u_n)(x)-\Phi(u)(x)|\ge\varepsilon^{1/q}\},$$

we have $\mathscr{L}^d(E_n)<\eta$ for all $n>\bar{n}$. Moreover, we can increase $\bar{n}$ in order to have also $\|u_n-u\|_{L^p}\le\varepsilon$ for all $n>\bar{n}$ We have

$$\begin{aligned}
&\int_\Omega|\Phi(u_n)(x)-\Phi(u)(x)|^q\,dx=\\
&\qquad=\int_{E_n}|\Phi(u_n)(x)-\Phi(u)(x)|^q\,dx+\int_{\Omega\setminus E_n}|\Phi(u_n)(x)-\Phi(u)(x)|^q\,dx
\end{aligned}$$

---

[1]During all this computation, we use the following fact. Given $v,w\in\mathbb{R}$, $w\neq0$, $s\ge1$, we have $\dfrac{|v-w|^s}{|v|^s+|w|^s}\le\dfrac{(|v|+|w|)^s}{|v|^s+|w|^s}=\dfrac{\left(\frac{|v|}{|w|}+1\right)^s}{\left(\frac{|v|}{|w|}\right)^s+1}$. By setting $t=\frac{|v|}{|w|}$, we have that $\dfrac{|v-w|^s}{|v|^s+|w|^s}\le\sup_{t\in[0,+\infty[}\dfrac{(t+1)^s}{t^s+1}\le2^{s-1}$, as it can easily verified by taking derivatives.

$$\begin{aligned}
&\leq C \int_{E_n} [a^q(x) + |u_n(x) - u(x)|^p + |u(x)|^p] \; dx + \varepsilon\eta \\
&\leq C \int_{E_n} [a^q(x) + |u(x)|^p] \; dx + C\|u_n - u\|^p_{L^p} + \varepsilon\eta \\
&\leq C \int_{E_n} [a^q(x) + |u(x)|^p] \; dx + (C + \eta)\varepsilon
\end{aligned}$$

Since $a^q(\cdot) + |u(\cdot)|^p \in L^1$, and $\mathscr{L}^d(E_n) < \eta$, for η sufficiently small we have $\int_{E_n} [a^q(x) + |u(x)|^p] \; dx < \varepsilon$, thus for n sufficiently large we have

$$\int_\Omega |\Phi(u_n)(x) - \Phi(u)(x)|^q \, dx \leq (2C + \eta)\varepsilon,$$

hence we have convergence in L^q.

□

We present now two results about the differentiability of the superposition operator. For further details and the proofs, we refer the reader to **[1]**.

PROPOSITION 2.17 (Differentiability of the superposition operator, case $p > 2$). *Let Ω be open in $\mathbb{R}^d$, $f : \Omega \times \mathbb{R} \to \mathbb{R}$ be a Carathéodory function. Suppose that there exists $p > 2$ such that:*

(1) *the following estimate holds*
$$|f(x,t)| \leq a(x) + b|t|^{p-1},$$
where $a \in L^{p/(p-1)}(\Omega; [0, +\infty[)$ and $b \geq 0$;

(2) *$\partial_t f$ exists and it is a Carathéodory function;*

(3) *the following estimate holds*
$$|\partial_t f(x,t)| \leq \alpha(x) + \beta|t|^{p-2},$$
where $\alpha \in L^{p/(p-2)}(\Omega; [0, +\infty[)$ and $\beta \geq 0$.

Then $\Phi : L^p(\Omega; \mathbb{R}) \to L^{p/(p-1)}(\Omega; \mathbb{R})$ is F-differentiable at every point u of its domain and
$$[\Phi'(u)v](x) = \partial_t f(x, u(x))v(x)$$

PROOF. Denote by $q = \frac{p}{p-1}$ the conjugate exponent to p.

Recalling the properties of the superposition operator, we can define a continuous map $\Psi : L^p \to L^{\frac{p}{p-2}}$ by setting
$$\Psi(u)(x) = \partial_t f(x, u(x)).$$
By Hölder's inequality, we have
$$\|\Psi(u)v\|_{L^q} \leq \|\Psi(u)\|_{L^{\frac{p}{p-2}}} \cdot \|v\|_{L^p},$$
in particular the map $v \mapsto \Psi(u)v$ is linear and continuous from L^p to L^q. Given $u, v \in L^p$, we define $\omega_u : L^p \to L^q$ by
$$\omega_u(v) := \Phi(u + v) - \Phi(v) - \Psi(u)v,$$
and recall that prove that Φ is F-differentiable at u with $\Phi'_F(u) = \Psi(u)$ amounts to prove that
$$\lim_{\|v\|_{L^p} \to 0} \frac{\|\omega_u(v)\|_{L^q}}{\|v\|_{L^p}} = 0.$$
We have

$$\begin{aligned}
\|\omega_u(v)\|^q_{L^q} &= \int_\Omega |f(x, u(x) + v(x)) - f(x, u(x)) - \partial_t f(x, u(x))v(x)|^q \; dx \\
&= \int_\Omega \left| \int_0^1 [\partial_t f(x, u(x) + \tau v(x)) - \partial_t f(x, u(x))] \, v(x) \, d\tau \right|^q \; dx \\
&= \int_\Omega |z_{uv}(x)v(x)|^q \; dx,
\end{aligned}$$

$$\leq \|v\|^q_{L^p} \cdot \|z_{uv}\|^q_{L^{\frac{p}{p-2}}},$$

where

$$z_{uv}(x) := \int_0^1 [\partial_t f(x, u(x) + \tau v(x)) - \partial_t f(x, u(x))] \, d\tau.$$

To conclude the proof we have to show that

$$\lim_{\|v\|_{L^p} \to 0} \|z_{uv}\|_{L^{\frac{p}{p-2}}} = 0.$$

Indeed,

$$\begin{aligned}\int_\Omega |z_{uv}(x)|^{\frac{p}{p-2}} \, dx &= \int_\Omega \left| \int_0^1 [\partial_t f(x, u(x) + \tau v(x)) - \partial_t f(x, u(x))] \, d\tau \right|^{\frac{p}{p-2}} dx \\ &\leq \int_0^1 \|\Psi(u + \tau v) - \Psi(u))\|^{\frac{p}{p-2}}_{L^{\frac{p}{p-2}}} \, d\tau\end{aligned}$$

The above integrand tends to 0 as $\|v\|_{L^p} \to 0$ by the continuity of Ψ, thus we have just to prove that Dominated Convergence Theorem applies to pass to the limit under the integral sign. To this aim, we use the growth condition on $\partial_t f$, hence

$$\begin{aligned}\|\Psi(u + \tau v) - \Psi(u))\|_{L^{\frac{p}{p-2}}} &\leq \left(2\|\alpha\|_{L^{\frac{p}{p-2}}} + \beta \| \, |u + \tau v|^{p-2} \|_{L^{\frac{p}{p-2}}} + \beta \| \, |u|^{p-2} \|_{L^{\frac{p}{p-2}}} \right) \\ &\leq \left(2\|\alpha\|_{L^{\frac{p}{p-2}}} + \beta \|u + \tau v\|^{\frac{1}{p-2}}_{L^p} + \beta \|u\|^{\frac{1}{p-2}}_{L^p} \right) \\ &\leq \left(2\|\alpha\|_{L^{\frac{p}{p-2}}} + \beta \left(\|u\|_{L^p} + \|\tau v\|_{L^p} \right)^{\frac{1}{p-2}} + \beta \|u\|^{\frac{1}{p-2}}_{L^p} \right).\end{aligned}$$

Thus for $\|v\|_{L^p} \leq 1$, since $\tau \in [0, 1]$, there exists a constant $C = C_{p,u,\beta,\alpha} > 0$ independent on v such that

$$\|\Psi(u + \tau v) - \Psi(u))\|_{L^{\frac{p}{p-2}}} \leq C_{p,u,\beta,\alpha},$$

for all $\|v\|_{L^p} \leq 1$ and $\tau \in [0, 1]$. Since we are integrating on $[0, 1]$ constants are integrable, thus Dominated Convergence Theorem applies and we obtain

$$\lim_{\|v\|_{L^p} \to 0} \|z_{uv}\|_{L^{\frac{p}{p-2}}} = 0,$$

as desired. □

When $p = 2$ the result is weaker.

PROPOSITION 2.18 (Differentiability of the superposition operator, case $p = 2$). *Let Ω be open in $\mathbb{R}^d$, $f : \Omega \times \mathbb{R} \to \mathbb{R}$ be a Carathéodory function. Assume that*

(1) *we have*
$$|f(x, t)| \leq a(x) + b|t|,$$
where $a \in L^2(\Omega; [0, +\infty[)$ and $b \geq 0$;
(2) *$\partial_t f$ exists and it is a Carathéodory function;*
(3) *we have*
$$|\partial_t f(x, t)| \leq M,$$
for a suitable $M \geq 0$.

Then $\Phi : L^2(\Omega; \mathbb{R}) \to L^2(\Omega; \mathbb{R})$ is G-differentiable at every point u of its domain and

$$[\Phi'_G(u)v](x) = \partial_t f(x, u(x))v(x).$$

In the same hypothesis, Φ is F-differentiable at u_0 if and only if $f(x, t) = h(x) + tk(x)$ where $h \in L^2(\Omega)$ and $k \in L^\infty(\Omega)$ are suitable functions. In that case, Φ is F-differentiable at every point of the domain.

PROOF. Omitted, see **[1]**. □

Summary of Lecture 2

- In this lecture we proved the Implicit Function Theorem in Banach spaces, deducing from it the inverse function theorem.
- We introduced also the concept of strict differentiability, which is an intermediate concept between Frechét differentiability and being C^1.
- We gave also a proper concept of tangent space in Lyusternik's theorem.
- We arrived to a Lagrange's Multiplier Rule for smooth constrained problems in Banach spaces.

3. Lecture of 19 november 2018: Necessary conditions in Calculus of Variations (3h)

We review now some results from functional Analysis, see [3] for further details.

Definition 3.1 (Convolution). Let $f \in L^1(\mathbb{R}^d)$, $\varphi \in C_c^\infty(\mathbb{R}^d)$. Then the map $f * \varphi : \mathbb{R}^d \to \mathbb{R}$ defined by

$$f * \varphi(x) := \int_{\mathbb{R}^d} f(y)\varphi(x-y)\,dy,$$

is well-defined and is call the *product of convolution* between f and φ. We have $f * \varphi \in C^\infty$. We have the following properties for $f \in L^1(\mathbb{R}^d)$, $\varphi, \psi \in C_c^\infty(\mathbb{R}^d)$.

(1) $f * \varphi = \varphi * f$;
(2) $\partial_j(f * \varphi) = f * (\partial_j \varphi)$;
(3) $\langle u, \psi * \varphi \rangle = \langle u * \psi, \varphi \rangle$.

Definition 3.2 (Mollifiers). A sequence of *mollifiers* on $\mathbb{R}^d$ is a sequence of functions $\{\rho_n\}_{n\in\mathbb{N}} \subseteq C_c^\infty(\mathbb{R}^d; [0, +\infty[)$ such that $\operatorname{supp} \rho_n \subseteq \overline{B(0, 1/n)}$ and $\displaystyle\int_{\mathbb{R}^d} \rho_n(x)\,dx = 1$ for all $n \in \mathbb{N}$.

Lemma 3.3 (Fundamental Lemma of Calculus of Variations). *Let $\Omega \subseteq \mathbb{R}^d$ be open, $\alpha \in L^1_{\mathrm{loc}}(\Omega)$. Suppose that for all $\varphi \in C_c^\infty(\Omega)$ it holds*

$$\int_\Omega \alpha(x)\varphi(x)\,dx = 0.$$

Then $\alpha = 0$ a.e. in Ω.

Proof. Let Ω' be an open bounded set such that $\overline{\Omega'} \subseteq \Omega$. In particular, there exists $\bar{n} > 0$ such that for $n < \bar{n}$ we have that $\overline{\Omega' + B(0, 1/n)} \subseteq \Omega$. For all $\varphi \in C_c^\infty(\Omega')$, we have that $\rho_n * \varphi \in C_c^\infty(\Omega)$. We have for all $n \geq \bar{n}$ and all $\varphi \in C_c^\infty(\Omega')$ that

$$0 = \langle \alpha, \rho_n * \varphi \rangle = \langle \alpha * \rho_n, \varphi \rangle.$$

Assume by contradiction that there exists $\bar{x} \in \Omega'$ such that $\alpha * \rho_n(\bar{x}) \neq 0$. Without loss of generality, we assume $\alpha * \rho_n(\bar{x}) > 0$ (the other case can be trated similarly). Thus there exists $\delta > 0$ such that $U = B(\bar{x}, \delta) \subseteq \Omega'$, $\alpha * \rho_n(y) > 0$ for all $y \in U$. If we take $\varphi(x) = \rho_k(x - \bar{x})$ with $k \in \mathbb{N}$, $k > 1/\delta$, then $\operatorname{supp} \varphi \subseteq U$, thus $\alpha * \rho_n(x) > 0$ on $\operatorname{supp} \varphi$. Moreover, there exists a set of positive measure $M \subseteq \operatorname{supp} \varphi$ such that $\varphi > 0$ on M. Hence

$$\langle \alpha * \rho_n, \varphi \rangle = \int_{\operatorname{supp} \varphi} \alpha * \rho_n(x)\varphi(x)\,dx \geq \int_M \alpha * \rho_n(x)\varphi(x)\,dx > 0,$$

contradicting $\langle \alpha * \rho_n, \varphi \rangle = 0$. So $\alpha_n * \rho_n = 0$ on Ω'. By letting $n \to \infty$, and recalling that $\alpha * \rho_n$ converges to α in $L^1(\overline{\Omega'})$, and hence pointwise a.e., we have that α is a.e. equal to a constant in Ω', thus by the arbitrariness of Ω', the same holds on Ω. $\square$

Corollary 3.4. *Let $\Omega \subseteq \mathbb{R}^d$ be open, $\alpha \in L^1_{\mathrm{loc}}(\Omega)$. Suppose that for all $\varphi \in C_c^\infty(\Omega)$ and $j \in \{1, \dots, d\}$ it holds*

$$\int_\Omega \alpha(x)\partial_j\varphi(x)\,dx = 0.$$

Then there exists $c \in \mathbb{R}$ such that $\alpha = c$ a.e. in Ω.

PROOF. Let Ω' be an open bounded set such that $\overline{\Omega'} \subseteq \Omega$. In particular, there exists $\bar{n} > 0$ such that for $n < \bar{n}$ we have that $\overline{\Omega' + B(0, 1/n)} \subseteq \Omega$. For all $\varphi \in C_c^\infty(\Omega')$, we have that $\rho_n * \varphi \in C_c^\infty(\Omega)$ and so also $\partial_j (\rho_n * \varphi) \in C_c^\infty(\Omega)$. We have for $n \geq \bar{n}$ and every $\varphi \in C_c^\infty(\Omega')$

$$0 = \langle \alpha, \partial_j(\rho_n * \varphi)\rangle = \langle \alpha * \rho_n, \partial_j \varphi\rangle = \langle \partial_j(\alpha * \rho_n), \varphi\rangle$$

Hence, by the Fundamental Lemma of Calculus of Variations, we have that the smooth map $\partial_j(\alpha * \rho_n)$ vanishes on Ω'. By the arbitrariness of j and the smoothness of $\partial_j(\alpha * \rho_n)$, we have that $\alpha * \rho_n$ is constant on Ω'. By letting $n \to \infty$, and recalling that $\alpha * \rho_n$ converges to α in $L^1(\overline{\Omega'})$, and hence pointwise a.e., we have that α is a.e. equal to a constant in Ω', thus by the arbitrariness of Ω', the same holds on Ω. □

COROLLARY 3.5 (Du Bois-Reymond Lemma). *Let $D = [a, b]$ be a compact interval in $\mathbb{R}$, $\alpha, \beta \in L^1(D; \mathbb{R}^d)$. Assume that for all $\varphi \in C_c^\infty(D; \mathbb{R}^d)$ we have*

$$\int_D [\alpha(x)\varphi(x) + \beta(x)\varphi'(x)]\, dx = 0.$$

Then $\beta \in W^{1,1}(D; \mathbb{R}^d)$ e $\beta' = \alpha$.

PROOF. After integration by parts, we have for all $\varphi \in C_c^\infty(D; \mathbb{R}^d)$

$$\begin{aligned}\int_D [\alpha(x)\varphi(x) + \beta(x)\varphi'(x)]\, dx &= -\int_D \left[\int_a^x \alpha(s)\, ds\right] \varphi'(x)\, dx + \int_D \beta(x)\varphi'(x)\, dx \\ &= \int_D \left[\beta(x) - \int_a^x \alpha(s)\, ds\right] \varphi'(x)\, dx.\end{aligned}$$

Thus there exists $c \in \mathbb{R}$ such that

$$\beta(x) - \int_a^s \alpha(s)\, ds = c,$$

hence

$$\beta(x) = c + \int_a^s \alpha(s)\, ds.$$

□

REMARK 3.6.

(1) Since the functions of $C_c^\infty(\mathbb{R}^d)$, are dense for the uniform convergence in $C_c^0(\Omega)$ and $C_c^1(\Omega)$ and the functions, the above results is still true if we replace in the statement $C_c^\infty(\Omega)$ with $C_c^1(\Omega)$ or , just in the case of the Fundamental Lemma of Calculus of Variations, even by $C_c^0(\Omega)$.

(2) The vector-valued case, i.e. $\alpha \in L^1_{\text{loc}}(\Omega; \mathbb{R}^m)$ can be treated in the same way. For example, let $\alpha \in L^1_{\text{loc}}(\Omega)$. Suppose that for all $\varphi \in C_c^\infty(\Omega)$ it holds

$$\int_\Omega \alpha(x)\varphi(x)\, dx = 0.$$

Since $\alpha(x) = (\alpha_1(x), \ldots, \alpha_m(x))$, and $\varphi(x) = (\varphi_1(x), \ldots, \varphi_m(x))$, if we choose $j \in \{1, \ldots, m\}$ and $\varphi_i \equiv 0$ if $i \neq j$ and $\varphi_j \in C_c^\infty(\Omega)$, we have that

$$0 = \langle \alpha, \varphi\rangle = \int_\Omega \alpha_j(x)\varphi_j(x)\, dx,$$

thus by the fundamental lemma in one dimension, we have $\alpha_j = 0$ a.e. By the arbitrariness of j, we obtain $\alpha = 0$ a.e.

For this part, we refer mainly to **[9]** and **[10]**.

DEFINITION 3.7 (Basic problem of Calculus of Variations). Let $I :=]a, b[\subseteq \mathbb{R}$ be an interval of $\mathbb{R}$, $L : I \times \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$ be a measurable function. We will denote the arguments of L by (t, x, v), where $t \in I$, $x \in \mathbb{R}^d$, $v \in \mathbb{R}^d$. Let X be a vector subspace of $C^0(\bar{I}; \mathbb{R}^d)$, and assume that all the maps of X are a.e. differentiable in I.

We consider the following problem (which will be called the *basic problem of Calculus of Variations*):

$$\inf_{x(\cdot) \in X} J(x), \qquad J(x) = \int_I L(t, x(t), \dot{x}(t))\, dt.$$

THEOREM 3.8 (Euler's Equations). *Consider the basic problem of C.o.V. with* $X = C^0(\bar{I};\mathbb{R}^d) \cap C^2(I;\mathbb{R}^d)$, $L \in C^2$. *If* $x(\cdot) \in X$ *is a solution, then satisfies* Euler's equations*:*

$$\begin{cases} \dfrac{d}{dt}\dfrac{\partial L}{\partial v_1}(t,x(t),\dot{x}(t)) = \dfrac{\partial L}{\partial x_1}(t,x(t),\dot{x}(t)), \\ \vdots \\ \dfrac{d}{dt}\dfrac{\partial L}{\partial v_j}(t,x(t),\dot{x}(t)) = \dfrac{\partial L}{\partial x_j}(t,x(t),\dot{x}(t)), \qquad j = 2,...,d-1, \\ \vdots \\ \dfrac{d}{dt}\dfrac{\partial L}{\partial v_d}(t,x(t),\dot{x}(t)) = \dfrac{\partial L}{\partial x_d}(t,x(t),\dot{x}(t)). \end{cases}$$

In general, we will call extremal *every map solving Euler's equation. Notice that an extremal is not necessarily a maximum or a minimum of* $J(\cdot)$.

PROOF. For every $\varphi \in C_c^\infty(I;\mathbb{R}^d)$ define $g_\varphi(\lambda) = J(x+\lambda\varphi)$ with $\lambda \in \mathbb{R}$. By assumption, g_φ has a minimum at $\lambda = 0$, moreover g_φ is differentiable at 0. Differentiating under the integral (which is possible since all the data are smooth) yields

$$\begin{aligned} g'_\varphi(\lambda) &= \frac{d}{d\lambda}\int_I L(t,x(t)+\lambda\varphi(t),\dot{x}(t)+\lambda\dot{\varphi}(t))\,dt \\ &= \int_I [\nabla_x L(t,x(t)+\lambda\varphi(t),\dot{x}(t)+\lambda\dot{\varphi}(t))\varphi(t) + \nabla_v L(t,x(t)+\lambda\varphi(t),\dot{x}(t)+\lambda\dot{\varphi}(t))\dot{\varphi}(t)]\,dt, \end{aligned}$$

where $\nabla_x = (\partial_{x_1},\dots,\partial_{x_d})$ and $\nabla_v = (\partial_{v_1},\dots,\partial_{v_d})$. Evaluating at $\lambda = 0$ and recalling that $g'_\varphi(0) = 0$ we have for all $\varphi \in C_c^\infty(I)$

$$0 = \int_I [\nabla_x L(t,x(t),\dot{x}(t))\varphi(t) + \nabla_v L(t,x(t),\dot{x}(t))\dot{\varphi}(t)]\,dt,$$

thus by Du Bois-Reymond Lemma

$$\nabla_x L(t,x(t),\dot{x}(t)) = \frac{d}{dt}\nabla_v L(t,x(t),\dot{x}(t)),$$

as desired. □

REMARK 3.9. Functions $\varphi(\cdot)$ used in the above proof were classically called *variations* of the solution $x(\cdot)$, this gave the names *Calculus of Variations* and, later, *Variational Analysis*.

The basic technique is to embed the infinite-dimensional problem in a one-dimensional problem by considering smooth perturbations of $x(\cdot)$ parameterized by the one-dimensional parameter λ and such that for $\lambda = 0$ we recover $x(\cdot)$. The perturbation $x(\cdot) \mapsto x(\cdot) + \lambda\varphi(\cdot)$ for fixed $\varphi \in C_c^\infty(I)$ is the simplest one. More generally, we can consider smooth perturbations $x(\cdot) \mapsto x_\lambda(\cdot)$, and the corresponding map $\lambda \mapsto L(t,x_\lambda(t),\dot{x}_\lambda(t))$.

This strategy has been implemented in the following theorem, that we present in a very simplified version.

THEOREM 3.10 (Noether). *Consider the basic problem of C.o.V. with* $X = C^0(\bar{I};\mathbb{R}^d) \cap C^2(I;\mathbb{R}^d)$, $L \in C^2$. *Assume that*

(1) *there exists a* C^1 *map* $S : \mathbb{R} \times \mathbb{R}^d \to \mathbb{R}^d$ *such that* $S(0,x) = x$;
(2) L *is invariant w.r.t. the action of* S, *i.e.,* $L\left(t, S(\lambda,x(t)), \dfrac{d}{dt}S(\lambda,x(t))\right) = L(t,x(t),\dot{x}(t))$.

Then given a solution $x(\cdot)$ *of the problem, there exists* $C \in \mathbb{R}$ *such that for all* $t \in I$

$$\nabla_v L(t,x(t),\dot{x}(t)) \cdot \frac{d}{d\lambda}\,[S(\lambda,x(t))]_{|\lambda=0} = C.$$

PROOF. We will write $x_\lambda(\cdot) = S(\lambda,x(\cdot))$. Straightforward computation yields:

$$\frac{d}{d\lambda}[L(t,x_\lambda(t),\dot{x}_\lambda(t))] = \left[\nabla_x L(t,x_\lambda(t),\dot{x}_\lambda(t))\frac{d}{d\lambda}x_\lambda(t) + \nabla_v L(t,x_\lambda(t),\dot{x}_\lambda(t))\frac{d}{d\lambda}\dot{x}_\lambda(t)\right]$$

$$= \left[\nabla_x L(t, x_\lambda(t), \dot{x}_\lambda(t)) \frac{d}{d\lambda} x_\lambda(t) + \nabla_v L(t, x_\lambda(t), \dot{x}_\lambda(t)) \frac{d}{d\lambda}\frac{d}{dt} x_\lambda(t)\right]$$
$$= \nabla_x L(t, x_\lambda(t), \dot{x}_\lambda(t)) \cdot \left[\frac{d}{d\lambda} x_\lambda(t)\right] + \nabla_v L(t, x_\lambda(t), \dot{x}_\lambda(t)) \cdot \frac{d}{dt}\left[\frac{d}{d\lambda} x_\lambda(t)\right],$$

recalling that, by smoothness, the derivatives in λ and in t can be switched.
Assume now that $x(\cdot)$ is a solution of Euler's equations, and that the Lagrangian is *invariant* with respect to the map $\lambda \mapsto x_\lambda(\cdot)$. This implies that $\frac{d}{d\lambda}[L(t, x_\lambda(t), \dot{x}_\lambda(t))] = 0$, moreover $\nabla_v L(t, x(t), \dot{x}(t)) = \nabla_v L(t, x_\lambda(t), \dot{x}_\lambda(t))$ and $\nabla_x L(t, x(t), \dot{x}(t)) = \nabla_x L(t, x_\lambda(t), \dot{x}_\lambda(t))$ by the smoothness of L.

$$0 = \frac{d}{d\lambda}[L(t, x_\lambda(t), \dot{x}_\lambda(t))]$$
$$= \frac{d}{dt}[\nabla_v L(t, x(t), \dot{x}(t))] \cdot \left[\frac{d}{d\lambda} x_\lambda(t)\right] + [\nabla_v L(t, x(t), \dot{x}(t))] \cdot \frac{d}{dt}\left[\frac{d}{d\lambda} \dot{x}_\lambda(t)\right]$$
$$= \frac{d}{dt}\left[\nabla_v L(t, x(t), \dot{x}(t)) \cdot \frac{d}{d\lambda} x_\lambda(t)\right].$$

Evaluating at $\lambda = 0$, we obtain that there exists $C \in \mathbb{R}$ such that for all $t \in I$

$$\nabla_v L(t, x(t), \dot{x}(t)) \cdot \frac{d}{d\lambda}[x_\lambda(t)]_{|\lambda=0} = C.$$

□

If the Lagrangian is invariant w.r.t. translation in time, we have the following result.

PROPOSITION 3.11 (Erdmann's condition). *Consider the basic problem of C.o.V. with* $X = C^0(\bar{I}; \mathbb{R}^d) \cap C^2(I; \mathbb{R}^d)$, $L \in C^2$. *Assume that L is* autonomous, *namely $\partial_t L = 0$, i.e., L does not depend explicitely on t. W Then if $x(\cdot)$ is an extremal,* Erdmann's condition holds*:*

$$L(x(t), \dot{x}(t)) - \dot{x}(t) \cdot \nabla_v L(x(t), \dot{x}(t)) = \text{cost}.$$

PROOF. Deriving, we have

$$\frac{d}{dt}[L(x(t), \dot{x}(t)) - \dot{x}(t) \cdot \nabla_v L(x(t), \dot{x}(t))] =$$
$$= \nabla_x L(x(t), \dot{x}(t)) \cdot \dot{x}(t) + \nabla_v L(x(t), \dot{x}(t)) \cdot \ddot{x}(t) - \ddot{x}(t) \cdot \nabla_v L(x(t), \dot{x}(t)) - \dot{x}(t) \cdot \frac{d}{dt} \nabla_v L(x(t), \dot{x}(t))$$
$$= \left[\nabla_x L(x(t), \dot{x}(t)) - \frac{d}{dt} \nabla_v L(x(t), \dot{x}(t))\right] \cdot \dot{x}(t) = 0,$$

where we used Euler's equation in the last step. □

Summary of Lecture 3

- In this lecture we introduced the basic problem of Calculus of Variation, i.e., the minimization of an integral functional $J(\cdot)$ with sufficiently smooth integrands (Lagrangian function) on a set of smooth functions. We will mainly treat the 1-dimensional case.
- We postponed the problem of the existence of minimizers, concentrating on the problem of deriving suitable necessary conditions.
- By the assumptions on the smoothness of the data, we can use the necessary condition $J'_G(\bar{x}(\cdot)) = 0$, where the set of *variations* used to compute the directional derivatives is given by compactly supported smooth functions.
- The above condition, together with Du Bois-Reymond Lemma, yields the celebrated Euler-Lagrange equations (in integral form).
- Smoothness of the data allows a further differentiation, obtaining classical Euler-Lagrange equations in differential form.
- In presence of *symmetries* in the lagrangian function, i.e., invariance of the lagrangian w.r.t. some transformation groups, we have the conservation of relevan quantities along the minimizers. For invariance w.r.t. action on the space variables, this is

expressed by Noether theorem, while for autonomous lagrangian (i.e., invariance in time) it is expressed by Erdmann's condition.

4. Lecture of 23 november 2018: Classical problems in C.o.V., conjugate points, sufficient conditions.(3h)

EXAMPLE 4.1 (Minimal surfaces). Consider the *soap bubble problem*. In $\mathbb{R}^3$, with variables denoted by (t, x, y), consider two rings in the planes $t = a$ e $t = b$ centeted on t-axis of radius A and B, respectively. Assume that a soap bubble surface joins the two rings. According to D'Alembert's principle, the equilibrium position minimizes the potential energy, which, in our case, amounts to minimize the area of the revolution surface joining the two rings. Assuming that the surface is generated by the rotation of the graph of a smooth map $t \mapsto x(t)$ around the t-axis, we have to minimize

$$J(x) = 2\pi \int_a^b x(t)\sqrt{1 + \dot{x}^2(t)}\, dt.$$

The Lagrangian is $L = L(x, v) = x\sqrt{1 + v^2}$, thus by Erdmann's condition,

$$x(t)\sqrt{1 + \dot{x}^2(t)} - \dot{x}(t)\frac{x(t)\dot{x}(t)}{\sqrt{1 + \dot{x}^2(t)}} = C.$$

Then

$$\frac{x(t)(1 + \dot{x}^2(t)) - x(t)\dot{x}^2(t)}{\sqrt{1 + \dot{x}^2(t)}} = C,$$

and so

$$\frac{x(t)}{\sqrt{1 + \dot{x}^2(t)}} = C.$$

Squaring and assuming $C \neq 0$ (otherwise we have the trivial solution $x(t) \equiv 0$, we have

$$\frac{x^2(t)}{C^2} = 1 + \dot{x}^2(t),$$

thus

$$\dot{x}(t) = \sqrt{\frac{x^2(t)}{C^2} - 1}.$$

This equation can be solved by separating the variables, and its solution is

$$x(t) = C \cosh\left(\frac{t + K}{C}\right),$$

where K, C are constant to be choosen in order to match $x(a) = A$ e $x(b) = B$ (for some choiches of A and B there may be no solutions). The curve $x(\cdot)$ is called *catenary*, and the surface of revolution obtained in this way are called *catenoids*.

EXAMPLE 4.2 (Brachistocrone). Consider the vertical plane xz, where the x-axis is horizonthal and the z-axis is vertical and oriented downwards. Consider the two points $O = (0, 0)$ and $P = (1, a)$ with $a > 0$. We look for a function $z : [0, 1] \to \mathbb{R}$ representing the trajectory along which a point particle starting from O with null initial velocity will arrive in P minimizing the travelling time (we assume there is no friction). In particular, it holds $z(0) = 0$ and $z(1) = a$. Without loss of generality, by physical reasons, we may assume that $z(\cdot)$ is not decreasing (recall that the axis z is pointing *downwards*). Assume moreover $z \in C^2(]0, 1[) \cap C^0([0, 1])$.

The length of an infinitesimal arc of the trajectory is $ds = \sqrt{1 + \dot{z}^2(x)}\, dx$. Denoted by $v(x)$ the velocity of the point particle at $(x, z(x))$, the conservation of mechanical energy yields that $E = \frac{1}{2}mv^2(x) - mgz(x)$ is costant along the motion. At $x = 0$ we have $v = 0$, since we have null initial velocity, and $z(0) = 0$, thus $E = 0$ for all $0 \leq x \leq 1$. Thus $v(x) = \sqrt{2gz(x)}$. The infinitesimal time needed to travel along the infinitesimal arc is then

$$dt = \frac{ds}{v} = \sqrt{\frac{1 + \dot{z}^2(x)}{2gz(x)}},$$

and so the total time of travel is

$$T(z) = \frac{1}{\sqrt{2g}} \int_0^1 \sqrt{\frac{1+\dot{z}^2(x)}{z(x)}}\, dx.$$

The integrand is the Lagrangian function

$$L(z, v) := \sqrt{\frac{1+v^2}{z}},$$

which does not depend on t, hence we are in the autonomous case, so Erdmann's condition holds, yielding

$$\sqrt{\frac{1+\dot{z}^2(x)}{z(x)}} - \dot{z}(x) \cdot \frac{\dot{z}(x)}{\sqrt{(1+\dot{z}^2(x))z(x)}} = C.$$

Recalling that $z(x) > 0$, we have

$$1 = C\sqrt{z(x)(1+\dot{z}^2(x))},$$

so $C > 0$ and we obtain the nonlinear ordinary differential equation

$$\frac{1}{C^2} = z(x)\left(1 + \left(\frac{dz}{dx}(x)\right)^2\right).$$

Recalling that $\dot{z} \geq 0$, we have

$$\frac{dz}{dx} = \frac{\sqrt{1 - C^2 z(x)}}{C\sqrt{z(x)}}.$$

This equation can be solved by separation of variables

$$\int \frac{C\sqrt{z}\, dz}{\sqrt{1-C^2 z}} = \int dx = x.$$

We make the following substitutions: first $w = \sqrt{1-C^2 z}$, $C^2 z = 1 - w^2$, $dz = -2w\, dw/C^2$, second $w = \cos\eta$, $dw = -\sin\eta\, d\eta$.

$$\begin{aligned} x &= \int \frac{C\sqrt{z}\, dz}{\sqrt{1-C^2 z}} = -\frac{2}{C^2}\int \sqrt{1-w^2}\, dw \\ &= -\frac{2}{C^2}\int \sqrt{1-w^2}\, dw = \frac{2}{C^2}\int \sin^2\eta\, d\eta = -\frac{2}{C^2}\int \frac{\cos 2\eta - 1}{2}\, d\eta = -\frac{\sin 2\eta}{2C^2} + \frac{\eta}{C^2} + \text{cost.} \end{aligned}$$

We have finally ($\theta = 2\eta$)

$$\begin{cases} x(\theta) &= \dfrac{1}{2C^2}(\theta - \sin\theta) + \text{cost} \\[2ex] z(\theta) &= \dfrac{1-w^2}{C^2} = \dfrac{1}{2C^2}(1-\cos\theta) \end{cases}$$

Since we want $x > 0$ for $\theta > 0$ and at $\theta = 0$ we want $x = 0$, we have

$$\begin{cases} x(\theta) &= \dfrac{1}{2C^2}(\theta - \sin\theta), \\[2ex] z(\theta) &= \dfrac{1-w^2}{C^2} = \dfrac{1}{2C^2}(1-\cos\theta), \end{cases}$$

i.e., a family of *cycloids*. The value of C is calculated by finding $\bar{\theta}$ such that $x(\bar{\theta}) = 1$, $z(\bar{\theta}) = a$, and substituting this value of θ in the equation for $z(\cdot)$. Clearly we have $\bar{\theta} \neq 0$, since $a \neq 0$, and $x(\theta) > 0$ for all $\theta > 0$. So we have $z(\bar{\theta})/x(\bar{\theta}) = a$, and finally

$$\begin{cases} a = \dfrac{1-\cos\bar{\theta}}{\bar{\theta} - \sin\bar{\theta}}, \\[2ex] C = \sqrt{\dfrac{\bar{\theta} - \sin\bar{\theta}}{2}}. \end{cases}$$

REMARK 4.3. The proof of Euler's equation is based on the single-variabile map $g_\varphi(\lambda) := J(x+\lambda\varphi)$ where φ is a fixed admissible variation, $\varphi \in C^1_0(I)$, which must satisfy the necessary condition $g'(0) = 0$ since $x(\cdot)$ is a minimum by assumption. However, since 0 is a minimum for g, when g is twice differentiable, we must have also $g''(0) \geq 0$ according to Taylor's formula, thus obtaining a *second-order necessary condition*.

PROPOSITION 4.4 (Legendre's necessary condition). *Consider the basic problem of C.o.V. with $X = C^0(\bar{I};\mathbb{R}^d) \cap C^2(I;\mathbb{R}^d)$ and $L \in C^2$. If $x(\cdot)$ is a solution, then $P(t) := \partial^2_{vv} L(t, x(t), \dot{x}(t)) \geq 0$ for all $t \in [a,b]$.*

PROOF. Let $x(\cdot)$ be a solution. Consider an admissible variation $\varphi \in C^1_c(\mathbb{R}^d)$, set $x_\lambda(t) = x(t) + \lambda\varphi(t)$, and let $g_\varphi(\lambda) = J(x_\lambda)$. By Euler's equations we have $g'_\varphi(0) = 0$. The second derivative of g_φ is

$$\begin{aligned} g''_\varphi(\lambda) &= \frac{d}{d\lambda}\int_I [\partial_x L(t, x_\lambda(t), \dot{x}_\lambda(t))\varphi(t) + \partial_v L(t, x_\lambda(t), \dot{x}_\lambda(t))\dot{\varphi}(t)]\, dt \\ &= \int_I \Big[\partial_{xx} L(t, x_\lambda(t), \dot{x}_\lambda(t))\varphi^2(t) + \partial_{xv} L(t, x_\lambda(t), \dot{x}_\lambda(t))\varphi(t)\dot{\varphi}(t) + \\ &\qquad + \partial_{vx} L(t, x_\lambda(t), \dot{x}_\lambda(t))\varphi(t)\dot{\varphi}(t) + \partial_{vv} L(t, x_\lambda(t), \dot{x}_\lambda(t))\dot{\varphi}^2(t)\Big]\, dt \\ &= \int_I \Big[\partial_{xx} L(t, x_\lambda(t), \dot{x}_\lambda(t))\varphi^2(t) + \partial_{xv} L(t, x_\lambda(t), \dot{x}_\lambda(t))2\varphi(t)\dot{\varphi}(t) + P(t)\dot{\varphi}^2(t)\Big]\, dt. \end{aligned}$$

Since $2\varphi(t)\dot{\varphi}(x) = \dfrac{d}{dt}[\varphi^2(t)]$, integration by parts yields

$$\begin{aligned} g''(0) &= \int_I \left[\partial_{xx} L(t, x(t), \dot{x}(t))\varphi^2(t) + \partial_{xv} L(t, x(t), \dot{x}(t))\frac{d}{dt}(\varphi^2(t)) + P(t)\dot{\varphi}^2(t)\right] dt \\ &= \int_I \left[\partial_{xx} L(t, x(t), \dot{x}(t))\varphi^2(t) - \frac{d}{dt}\partial_{xv} L(t, x(t), \dot{x}(t))\varphi^2(t) + P(t)\dot{\varphi}^2(t)\right] dt \\ &= \int_I \left[\left(\partial_{xx} L(t, x(t), \dot{x}(t)) - \frac{d}{dt}\partial_{xv} L(t, x(t), \dot{x}(t))\right)\varphi^2(t) + P(t)\dot{\varphi}^2(t)\right] dt \\ &= \int_I \Big[Q(t)\varphi^2(t) + P(t)\dot{\varphi}^2(t)\Big]\, dt \end{aligned}$$

Since 0 is a minimum for g, we have $g''(0) \geq 0$, so for every admissible variation

$$\int_I \Big[Q(t)\varphi^2(t) + P(t)\dot{\varphi}^2(t)\Big]\, dt \geq 0.$$

Assume by contradiction that there exists $\tau \in]a,b[$ such that $P(\tau) < 0$. By continuity, there exists $\delta > 0$ such that $B(\tau, 2\delta) \subseteq I$, $P < 0$ on $B(\tau, 2\delta)$, and we set $I_\delta = B(\tau,\delta)$. Given $\varepsilon > 0$, take a map $h_\varepsilon \in \mathrm{Lip}(\bar{I})$, $h_\varepsilon = 0$ in $\bar{I} \setminus I_\delta$, $|h'_\varepsilon(t)| = \chi_{I_\delta}(t)$ for a.e $t \in I$, and $\|h_\varepsilon\|_\infty \leq \varepsilon$. An example of such map can be constructed as follows: fix a set of $N > 2\delta/\varepsilon$ points $A_\varepsilon = \{t_1, \ldots, t_N\} \subset I_\delta$ such that $\max\limits_{i=1,\ldots,N} \min\limits_{\substack{j=1,\ldots,N \\ j\neq i}} |t_i - t_j| < \varepsilon$ and consider $h_\varepsilon(t) := \mathrm{dist}(t, A_\varepsilon \cup (\bar{I} \setminus I_\delta))$. By taking a sequence of mollifiers ρ_n, for n sufficiently large we have $\overline{I_\delta + B(0,1/n)} \subseteq B(\tau, 2\delta)$. Then, for n sufficiently large, we have

$$\int_I \Big[Q(t)[h_\varepsilon * \rho_n]^2(t) + P(t)[\dot{h}_\varepsilon * \rho_n]^2(t)\Big]\, dt \geq 0.$$

We have that $h_\varepsilon * \rho_n$ converges uniformly to h_ε on $\bar{I}$, thus for n sufficiently large $\|h_\varepsilon * \rho_n\|_\infty \leq 2\|h_\varepsilon\|_\infty = 2\varepsilon$, and $h_\varepsilon * \rho_n \to \dot{h}_\varepsilon = \chi_{I_\delta} 1$ in L^1. So for n sufficiently large

$$0 \leq \int_I \Big[Q(t)[h_\varepsilon * \rho_n]^2(t) + P(t)[\dot{h}_\varepsilon * \rho_n]^2(t)\Big]\, dt \leq 4\varepsilon^2 \|Q\|_\infty + \int_I P(t)[\dot{h}_\varepsilon * \rho_n]^2(t)]\, dt,$$

By letting $n \to \infty$,

$$0 \leq 4\varepsilon^2 \|Q\|_\infty + \int_{I_\delta} P(t)\, dt,$$

and by arbitrariness of $\varepsilon > 0$

$$\int_{I_\delta} P(t)\, dt \geq 0,$$

contradicting the fact that $P < 0$ on I_δ. □

REMARK 4.5. Under the assumptions of Legendre's necessary condition, we have seen that if $x(\cdot)$ is a solution, and set $g(\lambda) = J(x + \lambda\varphi)$ where $\varphi \in C^2_c(]a, b[)$ is an admissible variation,

$$\begin{aligned} g_\varphi(\lambda) &:= J(x + \lambda\varphi), \\ P(t) &:= \partial^2_{vv} L(t, x(t), \dot{x}(t)), \\ Q(t) &:= \partial^2_{xx} L(t, x(t), \dot{x}(t)) - \frac{d}{dt}\partial^2_{xv} L(t, x(t), \dot{x}(t)), \end{aligned}$$

where $\varphi \in C^1_c(]a, b[)$ is any admissible variation, we obtain $g'_\varphi(0) = 0$, which implies Euler's equations, and from the necessary condition $g''_\varphi(0) \geq 0$, i.e.,

$$g''_\varphi(0) = \int_I \left[Q(t)\varphi^2(t) + P(t)\dot{\varphi}^2(t)\right] dt \geq 0,$$

we obtain the necessary condition $P(t) \geq 0$, $t \in [a, b]$.

Since Taylor's formula for function of one variable yields the sufficient condition $g'(0) = 0$ and $g''(0) > 0$ for the function g to have a minimum at 0, in order to have a second-order necessary condition, a natural conjecture, firstly proposed by Legendre himself, would be: *assume that $x(\cdot)$ solves Euler's equations and that $P(t) > 0$, then $x(\cdot)$ is a local minimum in the C^1-norm for $J(\cdot)$.* However, Legendre's conjecture is false, as shown by the following counterexample.

EXAMPLE 4.6 (Conjugate points). In the plane (t, x), let $A = (0, 0)$ and $B = (T, 0)$, $T > 0$, two points. We want to travel from A to B minimizing the functional

$$J(x(\cdot)) := \int_0^T (\dot{x}^2(t) - x^2(t))\, dt.$$

Euler's equation yields $2\ddot{x}(t) = -2x(t)$, which has to be coupled with boundery conditions $x(0) = x(T) = 0$, so the trajectory $\bar{x}(t) \equiv 0$ is an extremal, moreover

$$P(t) = \partial_{vv} L(t, \bar{x}(t), \dot{\bar{x}}(t)) = 2 > 0,$$

thus may we conclude that $\bar{x}(t) \equiv 0$ solves the problem?

Assume $T \leq 1$. Hölder's inequality yields for every $0 \leq t \leq T$

$$|x(t)| \leq \left|\int_0^t \dot{x}(s)\, ds\right| \leq \int_0^t |\dot{x}(s)|^2\, ds \cdot t^{1/2} \leq \left(\int_0^T \dot{x}^2(s)\, ds\right)^{1/2}.$$

Thus, squaring and integrating in $[0, T]$

$$\int_0^T x^2(t)\, dt \leq \int_0^T \left(\int_0^T \dot{x}^2(s)\, ds\right) dt = T\int_0^T \dot{x}^2(s)\, ds \leq \int_0^T \dot{x}^2(t)\, dt.$$

Recalling that $T \leq 1$, we have

$$J(x) = \int_0^T \dot{x}^2(t)\, dt - \int_0^T x^2(t)\, dt \geq 0,$$

and since $J(\bar{x}) = 0$, we have proved that $\bar{x}$ is a minimum.

Assume now $T \geq 4$. We ask if also in this case it holds $J(x) \geq J(\bar{x}) = 0$. Consider $\tilde{x}(t) = t(T - t)$. this trajectory is smooth and satisfies boundary condition. Moreover

$$J(\tilde{x}) = \int_0^T (\dot{\tilde{x}}^2(t) - \tilde{x}^2(t))\, dt = \int_0^T (T - 2t)^2 - t^2(T - t)^2\, dt = \frac{T^3}{3} - \frac{T^5}{30} < 0,$$

thus $\bar{x}$ is *no longer* a minimum, hence the conjecture fails.

REMARK 4.7. The (wrong) proof that Legendre proposed for that conjecture was as follows. Given any differentiable map $w(\cdot)$, we must have for all $\varphi \in C^1_c(I)$

$$\int_a^b \frac{d}{dt}(w(t)\varphi^2(t))\, dt = 0.$$

Thus, assuming that $P(t) > 0$, we have

$$\begin{aligned}
g''(0) &= g''(0) + \int_a^b \frac{d}{dt}(w(t)\varphi^2(t))\,dt = \int_I \left[Q(t)\varphi^2(t) + P(t)\dot{\varphi}^2(t) + \frac{d}{dt}(w(t)\varphi^2(t))\right] dt \\
&= \int_I \left[Q(t)\varphi^2(t) + P(t)\dot{\varphi}^2(t) + w'(t)\varphi^2(t) + 2\varphi(t)\dot{\varphi}(t)w(t)\right] dt \\
&= \int_I P(t)\left[\frac{w'(t)+Q(t)}{P(t)}\varphi^2(t) + 2\frac{\varphi(t)w(t)}{P(t)}\dot{\varphi}(t) + \dot{\varphi}^2(t)\right] dt
\end{aligned}$$

If we choose $w(\cdot)$ to be a solution of

$$\frac{w'(t)+Q}{P(t)} = \left(\frac{w(t)}{P(t)}\right)^2,$$

we obtain

$$g''_\varphi(0) = \int_I P(t)\left[\frac{w(t)}{P(t)}\,\varphi(t) + \dot{\varphi}(t)\right]^2 dt.$$

so the sufficient condition $g''_\varphi(0) > 0$. This argument fails in supposing the *global* existence on $[a,b]$ of a solution of

$$\frac{w'(t)+Q}{P(t)} = \left(\frac{w(t)}{P(t)}\right)^2,$$

while in general this solution enjoys only local existence of solution.
Assuming $L \in C^3$, the change of variable $w(t) = -\dfrac{\dot{u}(t)P(t)}{u(t)}$ for $u \neq 0$ gives the *Jacobi's equation*

$$-\frac{d}{dt}(P(t)\dot{u}(t)) + Q(t)u(t) = 0,$$

the solution of this equation can be used to reconstruct $w(\cdot)$ if $u(\cdot)$ does not vanish in I.

EXAMPLE 4.8. In the previous example, we have $P(t) = -Q(t) = 2$ for every trajecetory, thus according to Legendre's argument, we have to solve

$$w'(t) = \frac{w^2(t)}{2} + 2,$$

which yields $w(t) = 2\tan(t+c)$, $c \in \mathbb{R}$, hence there is no possible choice of c to have such solution defined on the whole of $[0,T]$ for $T \geq \pi$, while for $0 < T < 1$ we can globally define $w(\cdot)$ and, indeed, making this argument working. In this case, the Jacobi's equation is simply $\ddot{u}(t) = u(t)$, whose solution is $u(t) = A\sin(t+\phi)$, $A, \phi \in \mathbb{R}$, and, again, for every choice of A, ϕ, we have that $u(t)$ vanishes in $I =]0,T[$ for $T \geq \pi$.

DEFINITION 4.9 (Conjugate points). Consider the basic problem of C.o.V. with $X = C^0(\bar{I};\mathbb{R}^d) \cap C^2(I;\mathbb{R}^d)$ and $L \in C^3$, $I =]a,b[$, and let $x(\cdot) \in X$ be an extremal. Define

$$\begin{aligned}
P(t) &:= \partial^2_{vv} L(t, x(t), \dot{x}(t)), \\
Q(t) &:= \partial^2_{xx} L(t, x(t), \dot{x}(t)) - \frac{d}{dt}\partial^2_{xv} L(t, x(t), \dot{x}(t)),
\end{aligned}$$

We say that $a < c \leq b$ is a *conjugate point* to a along $x(\cdot)$ if there exists a nonconstant solution of Jacobi's equation

$$-\frac{d}{dt}(P(t)\dot{u}(t)) + Q(t)u(t) = 0,$$

satisfying $u(a) = u(c) = 0$ (thus $\dot{u}(a) \neq 0$, otherwise we have the constant solution $u = 0$). Since $u = 0$ is solution, *all* nontrivial solution of the Jacobi's equations differs only for a multiplicative constant.

PROPOSITION 4.10 (Jacobi's necessay condition on conjugate points). *Consider the basic problem of C.o.V. with $X = C^0(\bar{I};\mathbb{R}^d) \cap C^2(I;\mathbb{R}^d)$ and $L \in C^3$, $I =]a,b[$, and let $x(\cdot) \in X$ be an extremal. Then there are no conjugate points to a in $]a,b[$.*

PROOF. Omitted. □

THEOREM 4.11 (Jacobi's sufficient condition). *Consider the basic problem of C.o.V. with $X = C^0(\bar{I};\mathbb{R}^d) \cap C^2(I;\mathbb{R}^d)$ and $L \in C^3$, $I =]a,b[$, and let $x(\cdot) \in X$ be an extremal. Define*

$$P(t) := \partial^2_{vv} L(t, x(t), \dot{x}(t)),$$
$$Q(t) := \partial^2_{xx} L(t, x(t), \dot{x}(t)) - \frac{d}{dt}\partial^2_{xv} L(t, x(t), \dot{x}(t)),$$

Suppose that there are no conjugate points to a in $]a,b]$ and $P(t) > 0$ for all $t \in [a,b]$. Then $\bar{x}(\cdot)$ is a local minimum.

PROOF. By assumption, given a solution of the Jacobi's equation u which never vanishes in $]a,b]$, we can use the change of variables $w(t) = -\dfrac{\dot{u}(t)P(t)}{u(t)}$, finding a global solution of

$$\frac{w'(t) + Q}{P(t)} = \left(\frac{w(t)}{P(t)}\right)^2.$$

Then Legendre's argument applies. Set $g_\varphi(\lambda) = J(x + \lambda\varphi)$ for all $\varphi \in C^1_c(I)$ we have

$$\begin{aligned}
g''_\varphi(0) &= g''(0) + \int_a^b \frac{d}{dt}(w(t)\varphi^2(t))\,dt = \int_I \left[Q(t)\varphi^2(t) + P(t)\dot{\varphi}^2(t) + \frac{d}{dt}(w(t)\varphi^2(t))\right] dt \\
&= \int_I \left[Q(t)\varphi^2(t) + P(t)\dot{\varphi}^2(t) + w'(t)\varphi^2(t) + 2\varphi(t)\dot{\varphi}(t)w(t)\right] dt \\
&= \int_I P(t)\left[\frac{w'(t) + Q(t)}{P(t)}\varphi^2(t) + 2\frac{\varphi(t)w(t)}{P(t)}\dot{\varphi}(t) + \dot{\varphi}^2(t)\right] dt \\
&= \int_I P(t)\left[\frac{w(t)}{P(t)}\,\varphi(t) + \dot{\varphi}(t)\right]^2 dt > 0.
\end{aligned}$$

$\square$

Up to now, we have searched solutions in $X = C^0([a,b]) \cap C^2(]a,b[)$. However it is easy to give very simple examples showing that this choice is too restrictive.

EXAMPLE 4.12. Consider the functional

$$J(x) := \int_{-1}^1 x^2(t)(\dot{x}^2(t) - 1)^2\,dt,$$

that we want to minimize with boundary condition $x(-1) = 0$ e $x(1) = 1$. Trivially, $J(x) \geq 0$. Let $x(\cdot) \in C^0([-1,1]) \cap C^1(]-1,1[)$ satisfying the boundary conditions. Then

$$1 = x(1) - x(-1) = \int_{-1}^1 \dot{x}(t)\,dt.$$

If $\dot{x}(t) > 1/2$ or $\dot{x}(t) < 1/2$ for every $t \in]-1,1[$ this clearly cannot happen. So there exist $t_1, t_2 \in]-1,1[$ where $\dot{x}(t_1) > 1/2$ and $\dot{x}(t_2) < 1/2$. Since $\dot{x}(\cdot) \in C^0(]-1,1[)$, according to the theorem of intermediate values, there exists $\tau \in]-1,1[$ with $\dot{x}(\tau) = 1/2$. In particular, there exists an open interval $V \subseteq]-1,1[$ such that $\tau \in V$ and $\dot{x}(t) \in [1/4, 3/4]$ for all $t \in V$. This implies that $x(\cdot)$ is strictly increasing in V, so it can vanish in V in at most one point. So in V we have $x(t) \neq 0$ a.e., and this implies $J(x) > 0$.

Define now the Lipschitz map $x(t) = 0$ if $t \in [-1,0]$ and $x(t) = t$ if $t \in [0,1]$. This map fulfills the boundary conditions and its derivative exists a.e. in $]-1,1[$ and satisfies $\dot{x}(t) = \chi_{]0,1[}(t)$ for a.e. $t \in [-1,1]$, so it make sense to compute $J(\cdot)$ on it (indeed, $x(\cdot) \in X := \mathrm{Lip}([-1,1])$), obtaining $J(x(\cdot)) = 0$. We conclude that this curves is a (nonsmooth) minimizer of J in $X := \mathrm{Lip}([-1,1])$. It is easy to see that in X there are infinite minimizers.

DEFINITION 4.13. Given a compact interval $[a,b] \subseteq \mathbb{R}$ we define

$$PWS([a,b]) := \left\{ x \in C^0([a,b]) : \begin{array}{l} \text{there exists a finite set } T \subseteq]a,b[\text{ such that } x \in C^1(]a,b[\setminus T) \\ \text{and there exist finite } \lim\limits_{t \to \bar{t}^\pm} \dot{x}(t) \text{ for all } \bar{t} \in T \end{array} \right\},$$

$$\mathrm{Lip}([a,b]) := \left\{ x \in C^0([a,b]) : \sup_{\substack{t,s\in[a,b]\\ t\neq s}} \frac{|x(s)-x(t)|}{|s-t|} < +\infty \right\},$$

$$AC([a,b]) := \left\{ x \in C^0([a,b]) : \begin{array}{c} \text{there exists } v \in L^1(]a,b[) \text{ such that} \\ x(t) = x(a) + \int_a^t v(s)\,ds \text{ for all } s \in [a,b] \end{array} \right\}.$$

The element of $PWS([a,b])$ are called *piecewise-smooth* functions, the elements of $\mathrm{Lip}([a,b])$ are called *Lipschitz continuous* function, the elements of $AC([a,b])$ are called *absolutely continuous* functions. Rademacher's Theorem, states that a Lipschitz continuous function defined on a finite-dimensional space is differentiable a.e., moreover we have $\mathrm{Lip}([a,b]) = W^{1,\infty}([a,b])$. We have also $W^{1,1}([a,b]) = AC([a,b]) \supset \mathrm{Lip}([a,b])$.

REMARK 4.14. On $PWS([a,b])$, $\mathrm{Lip}([a,b])$, and $AC([a,b])$, it make sense to write the functional of C.o.V., since all of them are space of continuous and a.e. differentiable functions on $[a,b]$. Performing the same computation of the derivation of Euler's equations, we obtain

$$\begin{aligned} 0 &= \int_I [\nabla_x L(t,x(t),\dot{x}(t))\varphi(t) + \nabla_v L(t,x(t),\dot{x}(t))\dot{\varphi}(t)]\,dt, \\ &= \int_I \left[-\left(\int_a^t \nabla_x L(s,x(s),\dot{x}(s))\,ds \right) \dot{\varphi}(t) + \nabla_v L(t,x(t),\dot{x}(t))\dot{\varphi}(t) \right] dt, \\ &= \int_I \left[\nabla_v L(t,x(t),\dot{x}(t)) - \int_a^t \nabla_x L(s,x(s),\dot{x}(s))\,ds \right] \dot{\varphi}(t)\,dt, \end{aligned}$$

so by Du Bois-Raymond Lemma we have that there exists $c \in \mathbb{R}$ such that for a.e. $t \in [a,b]$ the following *Euler's Equation in integral form* holds

$$\nabla_v L(t,x(t),\dot{x}(t)) = c + \int_a^t \nabla_x L(s,x(s),\dot{x}(s))\,ds.$$

When $x(\cdot)$ and L are sufficiently smooth, we can derive in t, obtaining the classical Euler's equation. The curves $x(\cdot)$ satisfying Euler's Equation in integral form are sometimes called weak extremals or extremals in the weak sense. Notice that being a solution of Euler's Equation in integral form is *stronger* than being an a.e. solution of the classical Euler's Equation.

EXAMPLE 4.15. Consider the minimization of the functional

$$J(x) = \int_0^1 \dot{x}^2(t)\,dt,$$

subject to $x(0) = 0$, $x(1) = 1$. According to Jensen's Inequality, by convexity of $r \mapsto r^2$, we have

$$\left(\int_0^1 \dot{x}(t)\,dt \right)^2 \leq \int_0^1 \dot{x}^2(t)\,dt,$$

and, since if $x(\cdot) \in AC([0,1])$ the right hand side is $(x(1)-x(0))^2$, we have $J(x) \geq 1$ for every $x \in AC([0,1])$.
Classical Euler's equation is $\ddot{x}(t) = 0$, and so the only extremal in $C^0([0,1]) \cap C^2(]0,1[)$ satisfying the boundary conditions is $x(t) = t$, and we have that this is a minimum according to Jensen's inequality.
In $PSW([0,1])$, Euler's equation in the weak sense is $2\dot{x}(t) = c$, for a suitable constant c, hence we obtain again as unique solution $x(t) = t$. Instead, we notice that if we consider the a.e. solutions of the classical Euler's requation, i.e., the piecewise smooth functions satisfying $\ddot{x}(t) = 0$ a.e. and respecting boundary condition, we obtain that *any* piecewise linear map satisfying the boundary condition is an a.e. solution of the classical Euler's equation. In particular, for $0 < \varepsilon < 1$ if we take $x_\varepsilon(t) = \dfrac{1}{\varepsilon}\chi_{[1-1/\varepsilon,1]}(t)(x-1+\varepsilon)$, we have that $\ddot{x}_\varepsilon = 0$ a.e., the boundary condition are fulfilled, and $J(x_\varepsilon) = 1/\varepsilon > 1$.

THEOREM 4.16. *Consider the basic problem of C.o.V. with $L \in C^1$. If $(x,v) \mapsto L(t,x,v)$ is convex, then every weak extremal is a global minimum. If moreover it is strictly convex, then the minimum is unique.*

PROOF. By smoothness assumptions on L, $\operatorname{epi} L$ admits a supporting hyperplane at each point, moreover

$$L(t,x_2,v_2)-L(t,x_1,v_1)\geq\langle\nabla_{x,v}L(t,x_1,v_1),(x_2-x_1,v_2-v_1)\rangle.$$

Let $\bar{x}$ be a weak extremal and y any other admissible curve. We define the *generalized moment*

$$p(t)=\nabla_v L(t,\bar{x}(t),\dot{\bar{x}}(t)),$$

thus Euler's equation in the weak form gives (in the weak sense)

$$\nabla_x L(s,x(s),\dot{x}(s))=\dot{p}(t).$$

Thus

$$\begin{aligned}J(y)-J(\bar{x})&=\int_a^b[L(t,y(t),\dot{y}(t))-L(t,\bar{x}(t),\dot{\bar{x}}(t))]\,dt\\&\geq\int_a^b\langle\nabla_{x,v}L(t,\bar{x}(t),\dot{\bar{x}}(t)),(y(t)-\bar{x}(t),\dot{y}(t)-\dot{\bar{x}}(t))\rangle\,dt\\&=\int_a^b\langle(\dot{p}(t),p(t)),(y(t)-\bar{x}(t),\dot{y}(t)-\dot{\bar{x}}(t))\rangle\,dt\\&=\int_a^b(\langle\dot{p}(t),y(t)-\bar{x}(t)\rangle+\langle p(t),\dot{y}(t)-\dot{\bar{x}}(t)\rangle)\,dt\\&=[\langle p(t),y(t)-\bar{x}(t)\rangle]_{t=a}^{t=b}=0.\end{aligned}$$

□

EXAMPLE 4.17. Consider the minimization of the functional

$$J(x):=\int_0^1(x^2(t)+(\dot{x}^2(t)-1)^2)\,dt,$$

subject to $x(0)=0$, $x(1)=0$. Trivially, for every $x(\cdot)\in AC([0,1])$ we have $J(x)>0$. We prove that infimum on $\operatorname{Lip}([a,b])$ is actually 0. A minimizing sequence is given by the triangular wave

$$x_n(t)=\int_0^t\operatorname{sign}(\sin(2n\pi s))\,ds,$$

so $|\dot{x}_n|=1$ a.e. in $[0,1]$, and $\|x_n\|_\infty=\frac{1}{2n}$, so $J(x_n)\leq\frac{1}{2n}$. The infimum is not achieved, and notice that the Lagrangian is not convex in v.

EXAMPLE 4.18. Consider the minimization of the functional

$$J(x):=\int_0^1(x^2(t)+g(\dot{x}))\,dt,$$

where $g(v)=v+\chi_{]-\infty,0]}(v)v^2$, and subject to $x(0)=0$, $x(1)=1$. We have for all $x(\cdot)\in AC$

$$J(x)\geq\int_0^1\left[\dot{x}(t)+\chi_{]-\infty,0]}(\dot{x}(t))\dot{x}^2(t)\right]\,dt\geq x(1)-x(0)=1,$$

and to have equality we must have $x(t)\equiv 0$, which is not an admissible curve. Thus $J(x)>1$ for all $x(\cdot)$ admissible. A minimizing sequence in $\operatorname{Lip}([0,1])$ such that $J(x_n)\to 1$ can be obtained by taking

$$x_n(t)=\chi_{[0,1-1/n]}(t)\frac{1}{n}\left(x-1+\frac{1}{n}\right).$$

REMARK 4.19. When we take $X=AC([a,b])$, unless to what happens in $\operatorname{Lip}([a,b])$, $PSW([a,b])$, $C^2(]a,b[)\cap C^0([a,b])$, the functional can be no longer G-differentiable. Thus we may have no longer validity of Euler's equation not even in the weak form. This implies that if may occurs that minima in AC cannot be detected by using Euler's equation.

DEFINITION 4.20. We say that the functional

$$J(x)=\int_a^b L(t,x(t),\dot{x}(t))\,dt,$$

to be minimized with boundary conditions $x(a) = x_a$ and $x(b) = x_b$, exhibits *Lavrent'ev phenomenon* if

$$\inf_{\substack{x\in AC([a,b])\\ x(a)=x_a\\ x(b)=x_b}} J(x) < \inf_{\substack{x\in \mathrm{Lip}([a,b])\\ x(a)=x_a\\ x(b)=x_b}} J(x).$$

REMARK 4.21. In particular, if Lavrent'ev phenomenon occurs, the usual numerical methods of minimization with finite elements (that are functions in $W^{1,\infty}$) will not be able to approximate the minimum. Lavrent'ev phenomenon may occurs even for *polynomial* Lagrangians, thus is not related to Lagrangian's smoothness.

EXAMPLE 4.22. The following smooth functional exhibit Lavrent'ev phenomenon (proved in 1926):

(1) Mania, (1934): $L(t,x,v) = (t-x^3)^2v^6$, $x(0) = 0$, $x(1) = 1$;
(2) Ball-Nizel, (1984): $L(t,x,v) = rv^3 + (x^2-t^3)^6v^{14}$, for $r > 0$, coupled with $x(0) = 0$, $x(1) = k$ admits as minimizer in $AC([0,1])$ $x(t) = kt^{2/3} \notin \mathrm{Lip}([a,b])$. But then $\partial_x L(t,x(t),\dot{x}(t)) \notin L^1$, thus not even the weak form of Euler's equation holds.

PROPOSITION 4.23 (C^1 smoothness of minimizers). *Consider the basic problem of C.o.V. with* $X = \mathrm{Lip}$. *Assume that che* $L(t,x,\cdot)$ *is strictly convex and* C^1, *then every minimizer* $\bar{x}$ *is* C^1.

PROOF. By assumption we have

$$\bar{x}(t) = \bar{x}(a) + \int_a^t \dot{\bar{x}}(s)\,ds,$$

where $\dot{\bar{x}} \in L^\infty$. We want to construct a map $\bar{v} \in C^0$ such that

$$\bar{x}(t) = \bar{x}(a) + \int_a^t \bar{v}(s)\,ds.$$

Define

$$S := \{t \in [a,b] : \dot{\bar{x}}(t) \text{ exists and weak Euler's equation holds}\}.$$

We have $\mathrm{meas}(S) = b-a$, and so S is dense in $[a,b]$. Consider now the limit

$$\lim_{\substack{\tau\to t\\ \tau\in S}} \dot{\bar{x}}(t).$$

If this limit exists at every t, we can define

$$\bar{v}(t) = \lim_{\substack{\tau\to t\\ \tau\in S}} \dot{\bar{x}}(t),$$

noticing that $\dot{\bar{x}}(t) = \dot{\bar{x}}(t)$ for every $t \in S$ and $\bar{v} \in C^0(]a,b[)$, as desired.

By contradiction, let $t \in]a,b[$ be such that the above limit does not exists. Since $\dot{\bar{x}}$ is bounded in S (by Lipschitz continuity of $\bar{x}(\cdot)$), there are sequences $\{t_i\}_{i\in\mathbb{N}}$ and $\{s_i\}_{i\in\mathbb{N}}$ in S such that $t_i \to t$, $s_i \to t$ but

$$v_1(t) := \lim_{i\to\infty} \dot{\bar{x}}(t_i) \neq \lim_{i\to\infty} \dot{\bar{x}}(s_i) =: v_2.$$

At t_i, s_i weak Euler's equation holds, thus

$$\partial_v L(t_i, x(t_i), \dot{\bar{x}}(t_i)) = C + \int_a^{t_i} \partial_x L(s, x(s), \dot{\bar{x}}(s))\,ds,$$
$$\partial_v L(s_i, x(s_i), \dot{\bar{x}}(s_i)) = C + \int_a^{s_i} \partial_x L(s, x(s), \dot{\bar{x}}(s))\,ds,$$

by passing to the limit for $i \to \infty$ and recalling the continuity of $\partial_v L$, we have $\partial_v L(t, x(t), v_1) = \partial_v L(t, x(t), v_2)$. Since $L(t,x,\cdot)$ is strictly convex, we have

$$\begin{aligned} L(t,\bar{x}(t),v_1) - L(t,\bar{x}(t),v_2) &> \langle \partial_v L(t,\bar{x}(t),v_2), v_1 - v_2\rangle = \langle \partial_v L(t,\bar{x}(t),v_1), v_1 - v_2\rangle \\ &= -[\langle \partial_v L(t,\bar{x}(t),v_1), v_2 - v_1\rangle] > -[L(t,\bar{x}(t),v_2) - L(t,\bar{x}(t),v_1)] \\ &= L(t,\bar{x}(t),v_1) - L(t,\bar{x}(t),v_2), \end{aligned}$$

which leads to a contradiction. □

THEOREM 4.24 (Hilbert-Weierstrass, 1890). *Consider the basic problem of C.o.V. with $X =$ Lip. Assume that $L \in C^2$, and $\partial^2_{vv} L > 0$ globally. Then every minimizer is C^2. if $L \in C^r$, $r > 2$ then the minimizer is C^r.*

PROOF. Applying the previous result, we have that $\bar{x} \in C^1$. So

$$p(t) := \partial_v L(t, \bar{x}(t), \dot{\bar{x}}(t)) = C + \int_a^t \partial_x L(s, \bar{x}(s), \dot{\bar{x}}(s))\, ds$$

is of class C^1. Since $\partial_{vv} L > 0$ Dini's Implicit Function Theorem applies, thus we obtain $\dot{\bar{x}}(t)$ as a C^1 function of the other variables, so $\bar{x} \in C^2$. The remaining part of the statement can be proved by induction. $\square$

THEOREM 4.25 (Tonelli's Existence Theorem). *Consider the basic problem of C.o.V. with $X = AC([a, b])$, L continuous, $v \mapsto L(t, x, v)$ convex and such that there exist $\alpha > 0$, $\beta \in \mathbb{R}$, $p > 1$ with $L(t, x, v) \geq \alpha|v|^p + \beta$. Then the minimization problem admits a solution $\bar{x}(\cdot) \in AC([a, b])$.*

PROOF. Let p' be such that $1/p + 1/p' = 1$. To avoid triviality, we assume $\inf\limits_{x\in X} J(x) < +\infty$. Let $\{x_i\}_{i\in\mathbb{N}}$ be a minimizing sequence. Since for i sufficiently large

$$0 \leq \int_a^b \alpha|\dot{x}(t)|^p\, dt + \beta(b - a) \leq \int_a^b L(t, x_i(t), \dot{x}_i(t)) = J(x_i) \leq 2 \inf_{x\in X} J(x) < +\infty,$$

we have that $\{\dot{x}_i\}_{i\in\mathbb{N}}$ is bounded in the reflexive space L^p (here we use $p > 1$), thus we may assume up to passing a subsequence that there exists $\bar{v} \in L^p$ such that $x_i \rightharpoonup \bar{v}$ weakly in L^p. Since $x_i(a) = x_a$, $x_i(b) = x_b$ for all $i \in \mathbb{N}$ and

$$x_i(t) = x_a + \int_0^b \chi_{[0,t]}(s)\, \dot{x}_i(s)\, ds = x_a + \langle \chi_{[0,t]}, \dot{x}_i \rangle_{L^{p'}, L^p},$$

by passing to the limit and using weak convergence, we have

$$\lim_{i\to\infty} x_i(t) = x_a + \langle \chi_{[0,t]}, v \rangle_{L^{p'}, L^p} = x_a + \int_0^t v(s)\, ds,$$

thus $x_i(\cdot)$ converges pointwise to

$$\bar{x}(t) = x_a + \int_0^t v(s)\, ds,$$

and we have that $\bar{x}(\cdot) \in AC([a, b])$ and satisfies the boundary conditions.

Given $a \leq s \leq t \leq b$, we have by Hölder inequality

$$\begin{aligned}|x_i(t) - x_i(s)| &\leq \int_t^s |\dot{x}_i(w)|\, dw \leq \int_a^b \chi_{[s,t]}(w)|\dot{x}_i(w)|\, dw \leq |t - s|^{1/p'} \|\dot{x}_i\|_{L^p} \\ &\leq \left(2 \inf_{x\in X} J(x)\right) |t - s|^{1/p'}.\end{aligned}$$

We obtain that the sequence $\{x_i\}_{i\in\mathbb{N}}$ is equibounded and equicontinuous (since it is equi-Hölder continuous), thus it converges uniformly to its pointwise limit $\bar{x}(\cdot)$.

We prove now that $\bar{x}(\cdot)$ is a solution. Given $M > 0$, define

$$S_M := \{t \in [a, b] : |\dot{\bar{x}}(t)| \leq M\},$$

and introduce the convex conjugate of L w.r.t. v, i.e. the Hamiltonian function defined by

$$H(t, x, p) = \sup_{v\in\mathbb{R}^n} \langle p, v\rangle - L(t, x, v).$$

Since L was continuous, recalling that the integrand is normal, and that we have weak convergence, by Fenchel-Moreau we have also

$$\begin{aligned}\int_{S_M} L(t, \bar{x}(t), \dot{\bar{x}}(t))\, dt &= \int_{S_M} \sup_{p\in\mathbb{R}^n} [\langle p, \dot{\bar{x}}(t)\rangle - H(t, x(t), p)]\, dt \\ &= \sup_{p\in L^\infty} \int_{S_M} [\langle p(t), \dot{\bar{x}}(t)\rangle - H(t, x(t), p(t))]\, dt\end{aligned}$$

$$
\begin{aligned}
&= \sup_{p\in L^\infty} \lim_{i\to\infty} \int_{S_M} [\langle p(t), \dot{x}_i(t)\rangle\, dt - \int_{S_M} \lim_{i\to\infty} H(t, x_i(t), p(t))\, dt \\
&= \sup_{p\in L^\infty} \lim_{i\to\infty} \int_{S_M} [\langle p(t), \dot{x}_i(t)\rangle\, dt - \limsup_{i\to\infty} \int_{S_M} H(t, x_i(t), p(t))\, dt \text{ (Lemma di Fatou)} \\
&= \sup_{p\in L^\infty} \liminf_{i\to\infty} \int_{S_M} [\langle p(t), \dot{x}_i(t)\rangle - H(t, x_i(t), p(t))]\, dt \\
&\leq \liminf_{i\to\infty} \sup_{p\in L^\infty} \int_{S_M} [\langle p(t), \dot{x}_i(t)\rangle - H(t, x_i(t), p(t))]\, dt \\
&= \liminf_{i\to\infty} \int_{S_M} L(t, x_i(t), \dot{x}_i(t))]\, dt \\
&= \liminf_{i\to\infty} J(x_i) - \int_{[a,b]\setminus S_M} L(t, x_i(t), \dot{x}_i(t))\, dt \\
&\leq \liminf J(x_i) - C\,\mathrm{meas}([a,b]\setminus S_M)
\end{aligned}
$$

recalling that $L(t,x,v) \geq C$. By letting $M \to +\infty$, we have

$$\int_a^b L(t, \bar{x}(t), \dot{\bar{x}}(t))\, dt \leq \liminf J(x_i),$$

which concludes the proof. □

THEOREM 4.26 (Clarke - Vinter, 1985). *Consider the basic problem of C.o.V. with $X = AC([a,b])$, L continuous, $v \mapsto L(t,x,v)$ convex and such that there exist $\alpha > 0$, $\beta \in \mathbb{R}$, $p > 1$ with $L(t,x,v) \geq \alpha|v|^p + \beta$. If L does not depend on t, then every solution is Lipschitz continuous (and so the Lavrent'ev phenomenon does not occurr).*

PROOF. Omitted. □

REMARK 4.27. We point out the following particular case of Lagrange's multiplier theorem applied to constrained optimization problem. Suppose that we want to minimize $\int_a^b L(t, x(t), \dot{x}(t))\, dt$ subject to the integral constraint $\int_a^b f(t, x(t), \dot{x}(t))\, dt = 0$ Then if $\bar{x}(\cdot)$ is a solution, it is an extremal w.r.t. the new Lagrangian $\lambda_0 L + \lambda f$ where $(\lambda_0, \lambda) \neq (0,0)$ and $\lambda_0 \in \{0,1\}$.

EXAMPLE 4.28 (Dido's problem). The legend say that Dido was the first-born daugther of Belus, king of Tyre, and married Acerbas (called also Sychaeus), who was a priest of Heracles, the wealthier of all the Phoenicians. Dido's brother, Pygmalion, blinded by greed, caught by surprise Sychaeus in the temple, during a sacrifice, and murdered him in front of the altar. For much time he kept hidden his murder, letting his sister to hope for the return of her husband. But Sychaeus' ghost, dead without an honorable burial, came in a Dido's dream showing her the altar where was murdered, and advicing her to flee away keeping with her the treasure that he had hidden in a secret place. Dido left Tyre with many followers and start a long journey, passing also by Cyprus and Malta.

The goddess Juno had promised them a new land where to establish a new city. Arrived on Lybic coasts, Dido obtained from Berber king Iarbas the permit to settle down, occupying as much land "as could be encompassed by an oxhide"; indeed the ancient name of Chartage was "Byrsa", that in greek means "oxhide" and in phoenician "fortified place".

Dido chose a peninsula, cut the oxhide in many fine strips, and in this way was able to encompass a large portion of land that will be the first settlement of the powerful town of Carthage. She encompassed an area of about 22 stadion (a stadion is about $185, 27\ m^2$).

Dido's problem (or *isoperimetric* problem) ask the shape that Dido had to choose in order to encompass the largest area possible, i.e., among all curves of fixed length with the extremals on a line, which is the one encompassing the largest area?

Dido's problem can be reformulated as problem of C.o.V. as follows: minimize

$$J(x) = -A(x) := -\int_a^b x(t),$$

where $x(a) = x(b) = 0$, $x(t) > 0$ with the constraint

$$\int_a^b \sqrt{1 + \dot{x}(t)}\, dt = \ell > b - a.$$

We consider the extremals of the Lagrangian

$$\mathcal{L}(t, x, v) := -\lambda_0 x + \lambda\sqrt{1 + v^2}.$$

If $\ell > b - a$ then $\lambda_0 \neq 0$, thus we can choose $\lambda_0 = 1$. Euler's equation gives

$$\frac{d}{dt}\left[\lambda \frac{\dot{x}}{\sqrt{1 + \dot{x}^2}}\right] = -1,$$

thus

$$\frac{\lambda \dot{x}}{\sqrt{1 + \dot{x}^2}} = -t + C.$$

Squaring, we have

$$\dot{x} = \frac{c - t}{\sqrt{\lambda^2 - (c - t)^2}},$$

so $x(t) = \sqrt{\lambda^2 - (c - t)^2} + K$ and finally

$$(x - K)^2 + (t - c)^2 = \lambda^2.$$

This implies that the solution is the half circle joining $x(a)$ a $x(b)$.

So among all the curves of the same length, the one encompassing the largest area is the circle.

THEOREM 4.29 (Karush-Kuhn-Tucker). *Let $f : \mathbb{R}^n \to \mathbb{R}$, $g_i : \mathbb{R}^n \to \mathbb{R}$, $h_j : \mathbb{R}^n \to \mathbb{R}$ be C^1 functions, $i = 1, \dots, m$, $j = 1, \dots, l$. Let $\bar{x}$ be a local minimum for f constrained to*

$$C := \{x \in \mathbb{R}^n : g_i(x) \leq 0,\ h_j(x) = 0, \text{ per ogni } i = 1, \dots, m,\ j = 1, \dots, l\}.$$

*Suppose that one of these condition holds (*constraint qualification*):*

(1) *f convex, g_i convex, h_i affine (Slater's condition);*
(2) *for every subset of constraints, the rank of the matrix built by the active inequality constraints and equality constraints has constant rank.*

Define $L : \mathbb{R}^n \times \mathbb{R}^{m+l} \to \mathbb{R}$ as

$$L(x, \lambda) = f(x) + \sum_{i=1}^m \lambda_i g_i(x) + \sum_{j=1}^l \lambda_j h_j(x),$$

then there exists $\bar{\lambda} \in \mathbb{R}^{m+l}$ such that:

(1) $\nabla L(\bar{x}, \bar{\lambda}) = 0$ *stationarity;*
(2) $\bar{x} \in C$ *admissibility;*
(3) $\bar{\lambda}_i \geq 0$, $i = 1, \dots, m$;
(4) $\bar{\lambda}_i g_i(\bar{x}) = 0$, $i = 1, \dots, m$.

PROOF. Omitted. □

THEOREM 4.30 (Fritz John). *Let X be a Banach space, $f : X \to \mathbb{R}$, $g_i : X \to \mathbb{R}$, $h_j : X \to \mathbb{R}$ be C^1 functions, $i = 1, \dots, m$, $j = 1, \dots, l$. Let $\bar{x}$ be a local minimum for f constrained to*

$$C := \{x \in X : g_i(x) \leq 0,\ h_j(x) = 0, \text{ per ogni } i = 1, \dots, m,\ j = 1, \dots, l\}.$$

Then there exist $\lambda_0 \in \{0, 1\}$, $\bar{\lambda} \in \mathbb{R}^{m+l}$ con $(\lambda_0, \bar{\lambda}) \neq 0$ such that if we define

$$L(x, \lambda) = \lambda_0 f(x) + \sum_{i=1}^m \lambda_i g_i(x) + \sum_{j=1}^l \lambda_j h_j(x),$$

it holds

(1) $\nabla L(\bar{x}, \bar{\lambda}) = 0$ *stationarity;*

(2) $\bar{x} \in C$ *admissibility;*
(3) $\bar{\lambda}_i \geq 0$, $i = 1, \ldots, m$;
(4) $\bar{\lambda}_i g_i(\bar{x}) = 0$, $i = 1, \ldots, m$.

PROOF. Omitted. □

Summary of Lecture 4

- If we ask for sufficiently smoothness, we can move from first-order necessary condition to higher order necessary condition (Legendre condition), however try to strenghten this condition to a sufficient condition in general fails due to the possible presence of *conjugate points*.
- Roughly speaking, conjugate points are points where the minimizers solution to EL equation loose their optimality as minimizers (e.g. antipodal points on a sphere for the geodesic distance). Conjugate points in physical systems have a strong physical meaning.
- If it is possible to exclude the presence of conjugate points, then Jacobi sufficient condition applies.
- In this lecture we meet also three classical problem in Calculus of Variation, nemely: Soap bubble problem (minimal revolution surface), Brachistocrone problem, Dido's problem (isoperimetric problem).
- They can be solved explicitely by using Euler-Lagrange equations or Erdmann's conditions.
- However in some cases, the restriction to work in spaces of smooth curves *prevents* to detect the *true physical minimizer*.
- Indeed, existence of minimizers heavly relies on the space of curves we are interested in. The usual way to face the problem is to prove existence in the broad space of AC curves by mean of the Tonelli-Weierstrass theorem, and then prove some *regularity* property of the minimizers.
- Unfortunately, it may happen that the infimum on AC curves is *stricly less* than infimum on Lipschitz or smoother curves. This is called Lavrentiev' phenomenon.
- Some convexity or growth condition could help to improve the regularity of the minimizers, also with the help of the necessary conditions.

CHAPTER 3

Third part

1. Lecture of 3 december 2018: Generalized gradients (3h)

REMARK 1.1. Up to now we considered cases where L is assumed to be smooth. But the situation dramatically change if the differentiability of L is dropped. A basic example of this situation can be given by adding to the problem a constraint such as $\dot{x}(t) \in V$ for a.e. $t \in [a,b]$, where V is given. In this case we can define a new Lagrangian $\tilde{L}(t,x,v) = L(t,x,v) + I_V(v)$, which does not fulfills the smoothness assumptions required up to now. This motivates the need for other differential tools.

DEFINITION 1.2 (Bouligand tangent cone). Let X be a normed space, $E \subseteq X$ be a set, $x \in E$. We define the *Bouligand tangent cone* to E at x by setting

$$T_E^F(x) := \left\{ \lambda w \in X : \lambda \geq 0 \text{ and } \exists \{y_n\}_{n\in\mathbb{N}} \subset E,\, y_n \to x,\, w = \lim_{n\to+\infty} \frac{y_n - x}{\|y_n - x\|} \right\}.$$

The set $T_E^F(x)$ is a cone, i.e., given $\lambda > 0$, $w \in T_E^F(x)$ we have $\lambda w \in T_E^F(x)$. However it is not necessarily convex. Another characterization can be given as follows: $w \in T_E^F(x)$ if and only if

$$\liminf_{t\to 0} \frac{d_E(x+tw)}{t} = 0.$$

DEFINITION 1.3 (Polar cone). Given a (possibly nonconvex) cone $C \subset X$, the *polar cone* of C is

$$C^\circ := \{x^* \in X' : \langle x^*, x\rangle_{X',X} \leq 0 \text{ for all } x \in C\}.$$

We have always that C° is w^*-closed and convex.

REMARK 1.4. If C is convex then $[T_C^F(x)]^\circ = N_C(x)$.

We recall here some differentiability properties of convex functions.

PROPOSITION 1.5. *Let X be a real normed space, $f : X \to \mathbb{R}$ be a convex function which is G-differentiable at x_0. Then $\partial f(x_0) = \{f'_G(x_0)\}$. Conversely, if f is convex and continuous at x_0 and $\partial f(x_0) = \{\varphi_0\}$, then f is G-differentiable at x_0 and $f'_G(x_0) = \varphi_0$.*

PROOF. For all $\xi \in \partial f(x_0)$, $v \in X$ and $h \in \mathbb{R}$ we have

$$f(x_0 + hv) - f(x_0) \geq h\langle \xi, v\rangle.$$

Hence, by G-differentiability at x_0 we can divide by $h > 0$ and take the limit

$$\langle f'_G(x_0), v\rangle_{X',X} = \lim_{h\to 0^+} \frac{f(x_0+hv) - f(x_0)}{h} \geq \langle \xi, v\rangle_{X',X}.$$

On the other hand, dividing by $h < 0$ and taking the limit

$$\langle f'_G(x_0), v\rangle_{X',X} = \lim_{h\to 0^-} \frac{f(x_0+hv) - f(x_0)}{h} \leq \langle \xi, v\rangle_{X',X},$$

and so $\langle f'_G(x_0), v\rangle_{X',X} = \langle \xi, v\rangle_{X',X}$. By the arbitrariness of v, we have $f'_G(x_0) = \xi$.

We omit the proof of the second part. □

THEOREM 1.6 (Mazur's differentiability theorem). *Let X be a separable Banach space, $D \subseteq X$ be a nonempty open convex subset of X, and $f : D \to \mathbb{R}$ be a convex continuous function. Then f is G-differentiable on $D_0 \subseteq D$, where D_0 is dense in D and it is a countable union of open dense sets.*

PROOF. Omitted. □

PROPOSITION 1.7. *Let X be a normed real space, $K \subseteq X$ be a nonempty conves set, and $f : K \to \mathbb{R}$ be a G-differentiable function. The following are equivalent:*

(1) *f is convex;*
(2) *$f(\xi) \geq f(x) + f'_G(x)(\xi - x)$ for all $\xi, x \in K$;*
(3) *the G-differential $f'_G : K \to X'$ is a monotone operator, i.e., for all $u, v \in K$ we have*

$$[f'_G(u) - f'_G(v)](u - v) \geq 0.$$

PROOF. Omitted. □

One of the most important aims of any extension of differentiability theory beyond the smooth, or the convex case, is to adapt to the nonsmooth case the usual necessary condition for the minimizers $F'(\bar{x}) = 0$ in something like $0 \in \partial F(\bar{x})$. Moving from convex analysis, in the past 30 years a new field of Analysis, called *Nonsmooth Analysis* or *Variational Analysis*, has been developed, in order to face optimization problems by mean of generalized differentiation tools. In the case of *nonsmooth and nonconvex problems*, others generalized gradients have been introduced, in many cases deeply related to the structure of the functional spaces in which the problem is stated.

Curiosity: Alexander Davidovich Ioffe, one of the worldwide most important experts of Nonsmooth Analysis, during a conference in Rome in 2009 said that up to his knowledge there were about 60 non-equivalent definitions of "generalized gradients". We will not dare to aim at it, and modestly will give an insight only on the most application-oriented used definitions. The recent treatise by Boris Mordukhovich, "Variational Analysis", 2006, Springer-Verlag, in two handy volumes, of 595 and 627 pages, respectively, is an attempt of systematic exposition of the topic (references amounts to 1379 items between articles, books and various publications).

Following Ioffe's approach, we start listing some "desiderable" properies that a "reasonable" subdifferential ∂ should enjoy. For simplicity we restrict ourselves to Banach spaces.

DEFINITION 1.8 (Subdifferential axioms). Let X, Y be Banach spaces, $f, g : X \to]-\infty, +\infty]$, $h : Y \to]-\infty, +\infty]$ be functions. Then ∂ must satisfy:

(S0) *substantiality*: $\partial f(x) = \emptyset$ if $x \notin \operatorname{dom} f$.
(S1) *localizability*: $\partial f(x) = \partial g(x)$ if $f = g$ in a neighborhood of x.
(S3a) *contiguity* I: if f is convex, then

$$\partial f(x) = \{x^* \in X' : f(x+h) - f(x) \geq \langle x^*, h\rangle \text{ for every } h \in X\}.$$

(S3b) *contiguity* II: if f is $C^{1,1}$ in a neighborhood of x then $\partial f(x) = \{f'(x)\}$.
(S4) *optimality*: if x is a minimum for f, then $0 \in \partial f(x)$.
(S5a) *calculability I*: if $g(x) = \lambda h(Ax + b) + \langle \ell, x\rangle + \alpha$, where $\lambda > 0$, $A : X \to Y$ is linear and surjective, $b \in Y$, $\ell \in X'$, $\alpha \in \mathbb{R}$, then $\partial g(x) = \lambda A^* \partial h(Ax + b) + \ell$.
(S5b) *calculability II*: if $p : X \times Y \to]-\infty, +\infty]$ is a function such that $p(x, y) = f(x) + h(y)$, then $\partial p(x, y) \subseteq \partial f(x) \times \partial h(y)$.
(S6) *boundedness*: if f is Lipschitz continuous in a neighborhood of x with Lipschitz constant K, then $\|x^*\|_{X'} \leq K$ for all $x^* \in \partial f(x)$.

These axioms are shared by a broad class of subdifferential of common use, and by all the subdifferential that we subject of our study. Finer classifications can be done, but they will be not matter of this course. We will introduce some subdifferential commonly used, by startinf from some possible more or less natural generalization of the subdifferential (or equivalentely of the normal cone) in the senso of convex analysis, and we will restrict ourselves to the Hilbert space case.

REMARK 1.9. Let X be an Hilbert space, $C \neq \emptyset$ a closed convex set. Given $x \in X$ there exists a unique $y \in C$ such that $\|x - y\| = d_C(x) := \min\{\|x - z\| : z \in C\}$. Such a y is the projection of x on C and it will be denoted by $\pi_C(x)$. This geometrical concept can be expressed also as follows: the ball centered at

$$x = \pi_C(x) + d_C(x)\,\frac{x - \pi_C(x)}{\|x - \pi_C(x)\|}$$

e di raggio $r_x = d_C(x)$ intersects C in a unique point $y = \pi_C(x)$, i.e. $\overline{B(x, d_C(x))} \cap C = \{\pi_C(x)\}$. Given $v \in X \setminus \{0\}$, and $r \geq 0$ such that $rv/\|v\| = x - \pi_C(x)$, we know that $\dfrac{v}{\|v\|} \in N_C(\pi_C(x))$ and so, by cone property, $v \in N_C(\pi_C(x))$. We can also say that the ball centered at $y + d_C\left(y + r\dfrac{v}{\|v\|}\right)$ of radius $r = d_C\left(y + r\dfrac{v}{\|v\|}\right)$ intersects C in the unique point y, i.e.

$$\overline{B\left(y + r\frac{v}{\|v\|}, r\right)} \cap C = \{y\}.$$

This implies that the distance of the points $z \in C$ from the center of the sphere is larger than r, and equality holds only at y: $\left\|z - \left(y + r\dfrac{v}{\|v\|}\right)\right\|^2 \geq r^2 = \left\|r\dfrac{v}{\|v\|}\right\|^2$ and so $\left\|z - y - r\dfrac{v}{\|v\|}\right\|^2 - \left\|r\dfrac{v}{\|v\|}\right\|^2 \geq 0$, hence

$$0 \leq \langle z - y, z - y - 2rv/\|v\|\rangle = -2r\left\langle \frac{v}{\|v\|}, z - y\right\rangle + \|z - y\|^2$$

that for $r > 0$ can be rewritten as

$$\langle v, z - y\rangle \leq \frac{1}{2r}\|v\| \cdot \|z - y\|^2$$

for all $z \in C$. We set $\sigma_{y,v} = 1/(2r)$. In the convex case, this construction holds for all $x = r\dfrac{v}{\|v\|} + y$, so it holds for every $r > 0$, and so passing to the limit for $r \to +\infty$ we recover the formula

$$\langle v, z - y\rangle \leq 0$$

for every $z \in C$, $v \in N_C(y)$. In the nonconvex case, the above relation may no longer hold true for all $r > 0$.

DEFINITION 1.10 (Proximal normal cone). Let X be an Hilbert space, K be a nonempty closed set. Given $y \in K$ we say that v is a *proximal normal* to K at y if there exists $\sigma_{y,v} \geq 0$ such that

$$\langle v, z - y\rangle \leq \sigma_{y,v}\|v\| \cdot \|z - y\|^2$$

for all $z \in C$. The set of all proximal normal to K at y will be called the proximal normal cone to K at y and will be denoted by $N^P_K(y)$. It is a convex cone in $X' = X$. We notice that the points $y + rv$ such that $r < 1/(2\sigma_{y,v})$ have unique projection on K and such projection is exactly y.

Exactly as in the convex case, we define:

DEFINITION 1.11 (Proximal subdifferential). Let X be an Hilbert space, $f : X \to]-\infty, +\infty]$ be a l.s.c. function. We say that $\xi_x \in X' = X$ is a *proximal subdifferential* of f at x and we will write $\xi_x \in \partial_P f(x)$ if $(\xi_x, -1) \in N^P_{\mathrm{epi}\, f}(x, f(x))$, i.e. if there exist $\delta \geq 0$ and $\sigma \geq 0$ such that

$$\beta - f(x) \geq \langle \xi_x, y - x\rangle - \sigma(\|y - x\|^2 + \|\beta - f(x)\|^2),$$

for all $y \in \mathrm{dom}\, f \cap B(x, \delta)$, $\beta \geq f(y)$. In particular

$$f(y) - f(x) \geq \langle \xi_x, y - x\rangle - \sigma(\|y - x\|^2 + \|f(y) - f(x)\|^2),$$

for all $y \in \mathrm{dom}\, f \cap B(x, \delta)$.

PROPOSITION 1.12 (Localization of proximal subdifferential). *Let X be an Hilbert space, $f : X \to]-\infty, +\infty]$ is a l.s.c. function. We have that $\xi_x \in \partial_P f(x)$ if there exist $\delta \geq 0$ and $\sigma \geq 0$ such that*

$$f(y) - f(x) \geq \langle \xi_x, y - x\rangle - \sigma\|y - x\|^2,$$

for all $y \in \mathrm{dom}\, f \cap B(x, \delta)$.

PROOF. Omitted, see **[7]**. □

THEOREM 1.13 (Clarke's Density Thereom). *Let X be an Hilbert space, $f : X \to]-\infty, +\infty]$ be a l.s.c. function. There exists a set D which is dense in* $\mathrm{dom}\, f$ *such that $\partial f(x) \neq \emptyset$ for all $x \in D$.*

PROOF. Omitted, see [7]. □

THEOREM 1.14 (*Fuzzy* sum rule). *Let X be an Hilbert space, $f_1, f_2 : X \to]-\infty, +\infty]$ be two l.s.c. function, $x \in \operatorname{dom} f_1 \cap \operatorname{dom} f_2$, $\zeta \in \partial(f+g)(x)$. Always we have $\partial f(x) + \partial g(x) \subseteq \partial(f+g)(x)$. If we suppose that at least one of these conditions holds*

(1) *f_1, f_2 are weakly l.s.c.;*
(2) *f_1 is Lipschitz continuous in a neighborhood of x,*

then for every $\varepsilon > 0$ there exist $\delta > 0$, $x_1, x_2 \in B(x, \delta)$ with $|f(x_i) - f(x)| < \varepsilon$, $i = 1, 2$ satisfying

$$\zeta \in \partial_P f_1(x_1) + \partial_P f_2(x_2) + \varepsilon B(0, 1).$$

PROOF. Omitted, see [7]. □

REMARK 1.15. Under suitable assumptions it is also possible to prove a *fuzzy* chain rule, see [7].

REMARK 1.16. The proximal normal cone is an intresting and significative geometrical object, however:

(1) it is strictly related to the Hilbertian scalar product (inducing a norm which is smooth in $X \setminus \{0\}$);
(2) without extra assuptions, the normal cone at a point is convex, but it may be not closed. Moreover limits of normals may fail to be normals (i.e. the proximal normal cone has not closed graph);
(3) $f(x) = -|x|^{3/2}$ is a function in $f \in C^1(\mathbb{R}) \setminus C^{1,1}(\mathbb{R})$, however $\partial_P f(0) = \emptyset$;
(4) *fuzzy* sum and chain rules are not intuitive and difficult to use: in general we are not allowed to pass to the limit as $\varepsilon \to 0^+$ since $\partial f_1(x)$ and $\partial f_2(x)$ may be even empty!

To circumvent (some of) these difficulties, the following definition is quite natural.

DEFINITION 1.17 (Limiting normal cone). Let X be an Hilbert space, C be closed and nonempty, $x \in C$. The *limiting (or Mordukhovich) normal cone* to C at x is defined by

$$N_C^L(x) = \{\zeta : \text{ exist } x_i \to x,\, \zeta_i \rightharpoonup \zeta, \text{ with } \zeta_i \in N_C^P(x_i)\}.$$

In the same way as before, it is possible to define the limiting subgradient: we say that $\xi_x \in X' = X$ is a *limiting subdifferential* to $f : X \to \mathbb{R} \cup \{+\infty\}$ at $x \in \operatorname{dom} f$ and we will write $\xi_x \in \partial_L f(x)$ if $(\xi_x, -1) \in N^L_{\operatorname{epi} f}(x, f(x))$. Calculus rules have been developed for this subdifferential, among which we recall the following: given a l.s.c. function f and a Lipschitz function g it holds $\partial(f+g)(x) \subseteq \partial f(x) + \partial g(x)$ (exact sum rule). Of course we have $N_C^P(x) \subset N_C^L(x)$. When C is convex all these cones reduces to the normal cone in the sense of convex analysis.

Nowadays this cone is the most used in particular in a special class of Banach spaces called (*Asplund* space), whose tractation is outside the matters of this course. To have a subdifferential working in a general Banach space it is necessary to introduce another object. The definition of this object is quite complicated, thus we will give a simplified definition in the Hilbert space case.

DEFINITION 1.18 (Clarke's normal cone). Let X be an Hilbert space, C be closed and nonempty, $x \in C$. We define the *Clarke's normal cone* $N_C^C(x) = \overline{\operatorname{co}} N_C^L(x)$. Given $f : X \to \mathbb{R} \cup \{+\infty\}$ and $x \in \operatorname{dom} f$, we say that $\xi_x \in X' = X$ is a *Clarke's generalized gradient* and write $\xi_x \in \partial f(x)$ if $(\xi_x, -1) \in N^C_{\operatorname{epi} f}(x, f(x))$.

THEOREM 1.19 (Clarke's generalized gradient in finite dimension). *Let $X = \mathbb{R}^n$, $f : \mathbb{R}^n \to]-\infty, +\infty]$ be a locally Lipschitz continuous function in a neighborhood of $x \in \mathbb{R}^n$. Then*

$$\partial f(x) = \overline{\operatorname{co}}\{v \in \mathbb{R}^n : y_i \in \operatorname{dom}(\nabla f(x)),\, \lim_{i \to +\infty} \nabla f(y_i) = v\}.$$

PROOF. Omitted, see [7] □

The advantages of Clarke's cone are its closedness and convexity, moreover it is possible to give simple and intuitive calculus rules. Moreover, for Lipschitz continuous functions it captures all the relevant information. The main drawback is that it can be very large, becoming useless for practical purpouses.

THEOREM 1.20 (*Nonsmooth* Euler's equations). *Let $\ell : \mathbb{R}^n \to \mathbb{R}$ be a locally Lipschitz continuous function, $L : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ be a globally Lipschitz continuous function. Consider the problem of minimizing the functional*

$$\ell(x(b)) + \int_a^b L(x(t), \dot{x}(t))\, dt,$$

on functions $x \in AC([a,b])$ with $x(a) = x_a$. Then if $x(\cdot)$ solves the problem, there exists $p \in AC([a,b])$ satisfying for a.e $t \in]a,b[$ $(\dot{p}(t), p(t)) \in \partial_C L(x(t), \dot{x}(t))$ and $-p(b) \in \partial_L \ell(x(b))$.

PROOF. Omitted, see [6] or [7]. □

Summary of Lecture 1

- Various kinds of generalized differentiation have been proposed in order to deal with *nonconvex* and *nonsmooth* minimization problems.
- Although many possible generalization of subdifferentability are possible, there are some *basic* properties which every *reasonable* subdifferential must enjoy in order to be considered *useful* for the optimization purpouses.
- A (possible) way to construct *new* subdifferentials or normal cones from old ones is to proceed by successive operations of closure and convexification, i.e., taking all cluster points of sequences in the graph of the subdifferential computed in the neighboring points of the point of interests, and then considering its convex hull.
- The starting point, leading to the construction of the *proximal* normal cone, is to consider a concept of *local* projection in a suitable neighborhood of a closed set (compare with the convex case, where the projection is *globally* well defined).

2. Lecture of 7 december 2018: Introduction to Control Theory and Differential Inclusions (3h)

EXAMPLE 2.1. Assume to have a trolley cart of mass $m = 1$ which is free to slide along a straight line truck without friction. Suppose that at the time $t_0 = 0$ the trolley is at the position x_0 with speed v_0.

If no external forces are acting on the system, the equation of the motion is $\ddot{x} = 0$, and so $x(t) = x_0 + v_0 t$. Suppose now that there is an external agent (controller), who is able to pull or push the cart with a time-depending force $u(t) \in [-1, 1]$. The motion equation becomes $\ddot{x} = u$, which can be reduced to a first-order system of differential equations by setting $\dot{x}_1 = x_2$, $\dot{x}_2 = u$. If the force u changes, the time law $x_u(t)$ (and possibly also the trajectory) of the trolley will change. We may consider the problem to minimize a certain functional of the trajectory $J = J(x_u(\cdot))$ *by acting* on the *control* u.

EXAMPLE 2.2. Consider a fish population in a lake. A model describing the variation in time of the number of fishes in the lake is given by logistic equation $\dot{x} = x(\alpha - x)$, where α is the maximum number of fishes that can be supported by lake resources

Suppose to manage a fishery, and be intentioned to send a certain number of fishers (in number u) to catch some fishes. The evolution of the systemcan be described to $\dot{x} = x(\alpha - x) - kxu$, where $k \geq 0$ is a parameter measuring the efficiency of fishers. We can have the goal to maximize the amount of fish catched in a fixed period (in this case, maybe, to catch everything immediately is not an optimal strategy) or similar goals.

EXAMPLE 2.3. Assume to have to park a car in a free parking space on the side of a road. We can act on steer and on the spped, however the steer angle is bounded. Common experience tells us that we can solve the problem by *multiple manouvers*. But can we mathematically prove that it is always possible to park the car if the size of the park space is enough?

Control theory related problems arise in many fields of human activities: engineering, economics, logistics and transportation, biology and even social sciences. A control system is a system subjected to an external influence possibly varying in time. The input variables are elaborated to produce the output variables, and the external agent acts in the elaboration influencing output in order to achieve some goals (for example, maximization or minimization of functionals depending on output variables).

We will treat the case in which the system is ruled by an ordinary differential equation $\dot{x}(t) = f(t, x(t), u(t))$ where $u(t) \in U$ is the variable controlled by the external agent. The aim will be to find an optimal strategy (or *optimal control*) u^* minimizing a certain cost related to the trajectory of the ODE starting from a suitable initial condition.

In the real-world systems, frequently arise situation in which the measurements are affected by errors (in time or on the position), thus a natural question is how much the optimal strategy is *robust*, i.e. sensitive to errors.

We will recall some basic results from ODE theory.

THEOREM 2.4 (Parametric contraction lemma). *Let X, T be complete metric spaces, $0 < \alpha < 1$, $\phi : T \times X \to X$ be continuous and such that*

$$d_X(\phi(t, x_1), \phi(t, x_2)) \leq \alpha d_X(x_1, x_2),$$

for all $x_1, x_2 \in X$, $t \in T$. Then for every t there exists a unique $x = x(t)$ such that $x(t) = \phi(t, x(t))$. The map $t \mapsto x(t)$ is continuous, and

$$d_X(y, x(t)) \leq \frac{1}{1-\alpha} d_X(y, \phi(t, y)).$$

PROOF. Fix $t \in T$, and let $x_0 \in X$. Set $x_1 = \phi(t, x_0)$ and $x_n = \phi(t, x_{n-1})$ for $n \geq 1$. We prove that the sequence $\{x_n\}_{n\in\mathbb{N}}$ is a Cauchy sequence in the complete space X: given $m, n \in \mathbb{N}$, $m \geq n$

$$d_X(x_{n+1}, x_n) = d_X(\phi(t, x_n), \phi(t, x_{n-1})) \leq \alpha d_X(x_n, x_{n-1}) \leq \alpha^n d_X(x_1, x_0)$$

$$d_X(x_m, x_n) \leq \sum_{j=n}^{m-1} d_X(x_{j+1}, x_j) \leq \sum_{j=n}^{m-1} \alpha^j d_X(x_1, x_0) \leq d_X(x_1, x_0) \sum_{j=n}^{\infty} \alpha^j = d_X(x_1, x_0) \frac{\alpha^n}{1-\alpha}$$

thus for $n, m \to \infty$ we have $d_X(x_m, x_n) \to 0$. So $\{x_n\}_{n\in\mathbb{N}}$ converges in X to an element denoted by $x = x(t)$. Recalling the continuity of ϕ, we have

$$\phi(t, x(t)) = \phi(t, \lim_{n\to\infty} x_n) = \lim_{n\to\infty} \phi(t, x_n) = \lim_{n\to\infty} x_{n+1} = x(t).$$

Uniqueness: given $x_1(t), x_2(t)$ satisfying $x_i(t) = \phi(t, x_i(t))$, $i = 1, 2$ we have

$$d_X(x_1(t), x_2(t)) = d_X(\phi(t, x_1(t)), \phi(t, x_2(t))) \leq \alpha d_X(x_1(t), x_2(t)),$$

and so $d_X(x_1(t), x_2(t)) = 0$ because $\alpha > 0$.

Choose now $y = x_0$ and construct the sequence $x_1 = \phi(t, x_0)$ and $x_n = \phi(t, x_{n-1})$ for $n \geq 1$. Then

$$d_X(x_n, y) \leq \sum_{j=0}^{n-1} \alpha^j d_X(x_1, x) \leq \sum_{j=0}^{\infty} \alpha^j d_X(y, \phi(t, y))$$

Since $x_n \to x(t)$ as $n \to +\infty$, by passing to the limit in the above equation we obtain the thesis.

We prove now that $t \mapsto x(t)$ is continuous. Let $t_n \to t$. The above relation becomes

$$d_X(y, x(\lambda)) \leq \frac{1}{1-\alpha} d_X(y, \phi(\lambda, y)).$$

Choose $y = x(t)$ and $\lambda = t_n$. We have

$$d_X(x(t), x(t_n)) \leq \frac{1}{1-\alpha} d_X(x(t), \phi(t_n, x(t))).$$

Since $t_n \to t$ and ϕ is continuous, we have $\phi(t_n, x(t)) \to \phi(t, x(t)) = x(t)$, and so for $n \to \infty$ the right hand side tends to 0, thus $x(t_n) \to x(t)$. □

THEOREM 2.5 (Brouwer's fixed point theorem). *Let $B \subseteq \mathbb{R}^n$ be homeomorphic to the unit ball in $\mathbb{R}^n$, and $f : B \to B$ continuous. Then there exists $\bar{x} \in B$ such that $f(\bar{x}) = \bar{x}$.*

LEMMA 2.6 (Gronwall). *Let $I = [a, b[$ be an interval of $\mathbb{R}$ with $a < b \leq +\infty$. Suppose that there are given $u \in C^0(I;\mathbb{R})$, $\alpha, \beta \in L^1(I;\mathbb{R})$, $\beta(t) \geq 0$ for a.e. $t \in I$, satisfying*

$$u(t) \leq \alpha(t) + \int_a^t \beta(s)u(s)\,ds, \qquad \forall t \in I.$$

Then

$$u(t) \leq \alpha(t) + \int_a^t \alpha(s)\beta(s)\exp\left(\int_s^t \beta(r)\,dr\right)ds, \qquad t \in I.$$

PROOF. Define the following absolutely continuous function

$$v(s) = \exp\left(-\int_a^s \beta(r)\,\mathrm{d}r\right)\int_a^s \beta(r)u(r)\,dr, \qquad s \in I.$$

Deriving, and recalling the sign of β, we have for a.e. $s \in I$

$$v'(s) = \left(u(s) - \int_a^s \beta(r)u(r)\,dr\right)\beta(s)\exp\left(-\int_a^s \beta(r)\,dr\right) \leq \alpha(s)\beta(s)\exp\left(-\int_a^s \beta(r)\,dr\right).$$

Since $v(a) = 0$, integrating we obtain

$$v(t) \leq \int_a^t \alpha(s)\beta(s)\exp\left(-\int_a^s \beta(r)\,dr\right)ds.$$

We notice that

$$\begin{aligned}\int_a^t \beta(s)u(s)\,ds &= \exp\left(\int_a^t \beta(r)\,dr\right)v(t)\\ &\leq \int_a^t \alpha(s)\beta(s)\exp\left(\underbrace{\int_a^t \beta(r)\,dr - \int_a^s \beta(r)\,dr}_{=\int_s^t \beta(r)\,dr}\right)\mathrm{d}s,\end{aligned}$$

and so the thesis follows recalling the inequality given in the assumptions. □

THEOREM 2.7 (Existence of solutions). *Let Ω be an open subset of $\mathbb{R} \times \mathbb{R}^n$, $(t_0, x_0) \in \Omega$, $g : \Omega \to \mathbb{R}^n$ a function such that*

(E1) *for all x the map $t \to g(t, x)$ defined on $\Omega_x := \{t \in \mathbb{R} : (t, x) \in \Omega\}$ is measurable;*
(E2) *for a.e. t the map $x \to g(t, x)$ defined on $\Omega_t := \{x \in \mathbb{R}^n : (t, x) \in \Omega\}$ is continuous;*
(E3) *for every compact set $K \subset \Omega$ there exist C_K, L_K such that $\|g(t, x)\| \leq C_K$, $\|g(t, x_1) - g(t, x_2)\| \leq L_K\|x_1 - x_2\|$ for all (t, x), (t, x_1), $(t, x_2) \in K$.*

Then there exists $\varepsilon > 0$ such that the Cauchy problem $\dot{x} = g(t, x)$, $x(t_0) = x_0$ admits a solution $x(\cdot)$ defined for $t \in [t_0, t_0 + \varepsilon[$.

If moreover $\Omega = \mathbb{R} \times \mathbb{R}^n$ and there exist constants C, L such that $\|g(t, x)\| \leq C$, $\|g(t, x_1) - g(t, x_2)\| \leq L\|x - y\|$ for all (t, x), (t, x_1), $(t, x_2) \in \mathbb{R} \times \mathbb{R}^n$, then the above Cauchy problems admits a unique solution $x(\cdot)$ defined on $[t_0, +\infty[$. Moreover, the solution depends countinously on the initial data x_0.

PROOF. Omitted. □

COROLLARY 2.8 (Uniqueness). *Same assumptions of the previous result. If $x_1(\cdot)$ and $x_2(\cdot)$ are solutions of the Cauchy problem $\dot{x} = g(t, x)$, $x(t_0) = x_0$ defined on $[t_0, t_1[$ and $[t_0, t_2[$ respectively, then $x_1(t) = x_2(t)$ for all $t \in [t_0, \min\{t_1, t_2\}]$.*

PROOF. Omitted. □

3. Lecture of 10 december 2018: Differential inclusions (3h)

DEFINITION 3.1 (Hausdorff distance). Let X be a Banach space, $A, A' \subseteq X$ be compact, nonempty sets. The *Hausdorff distance* between A and A' is given by

$$\begin{aligned} d_H(A, A') &= \max\{\operatorname{dist}(x, A'), \operatorname{dist}(x', A) : x \in A, x' \in A'\} \\ &= \inf_{\rho > 0}\{\rho : A \subset B(A', \rho) \text{ e } A' \subset B(A, \rho)\}, \end{aligned}$$

where $B(K, r) := \{y \in X : \operatorname{dist}(y, K) \le r\}$. The Hausdorff distance is a metric on the sets of nonempty compact subsets of X.

DEFINITION 3.2 (Set-valued maps). Let X, Y be Banach spaces. A *multifunction* or *set-valued function* F from X to Y is a map associating to every $x \in X$ a set $F(x) \subset Y$, i.e. a map $F : X \to 2^Y$. We will write also $F : X \rightrightarrows Y$. The *domain* of F is the set $\operatorname{dom} F := \{x \in X : F(x) \neq \emptyset\}$, while the *graph* of F is the set

$$\operatorname{graph} F := \{(x, y) : x \in \operatorname{dom} F, y \in F(x)\}.$$

We say that F

(1) is *closed valued* if $F(x)$ is closed for every x;
(2) is *compact valued* if $F(x)$ is compact and nonempty for every x;
(3) is *bounded* if there exists a bounded set $B \subseteq Y$ such that $F(x) \subseteq B$ for all $x \in X$;
(4) has *closed graph* if $\operatorname{graph} F$ is closed in $X \times Y$;
(5) is *upper semicontinuous* (u.s.c) at $x_0 \in X$ if for every open set A containing $F(x_0)$ there exists a neighborhood Ω of x_0 such that $F(x) \subseteq A$ for all $x \in \Omega$;
(6) is *lower semicontinuous* (l.s.c) at $x_0 \in X$ if for all $y_0 \in F(x_0)$ and for every neighborhood M of y_0 there exists a neighborhood Ω of x_0 such that $F(x) \cap M \neq \emptyset$ for all $x \in \Omega$;
(7) is *continuous* if it is both u.s.c. and l.s.c. If F is compact valued, this is equivalent to say that
$$\lim_{y \to x} d_H(F(y), F(x)) = 0,$$
i.e., F is continuous w.r.t. the topology induced by Hausdorff metric;
(8) is *Lipschitz continuous* if there exists $K > 0$ such that
$$F(x_1) \subseteq F(x_2) + K\|x_1 - x_2\| B_X(0, 1),$$
for all $x_1, x_2 \in X$.

DEFINITION 3.3 (Measurability of set-valued maps). Let $F : \mathbb{R}^m \rightrightarrows \mathbb{R}^n$ be a set-valued map. We say that F is *measurable* if for every $C \subseteq \mathbb{R}^n$ closed, the set

$$\{x \in \mathbb{R}^m : F(x) \cap C \neq \emptyset\}$$

is measurable.

PROPOSITION 3.4 (U.s.c. and closedness of the graph).

- *The graph of an u.s.c. set-valued map $F : X \rightrightarrows Y$ with closed domain and closed images is closed. The converse is true if there exists a compact $S \subseteq Y$ such that $F(x) \subseteq S$ for all $x \in X$.*
- *If X is locally compact, and the restriction $F_{|K}$ of F to K has compact graph for every $K \subseteq X$, then F is u.s.c.*

PROOF. Omitted. $\square$

DEFINITION 3.5 (Selections). Let $F : \mathbb{R}^m \rightrightarrows \mathbb{R}^n$ be a set-valued map. A *selection* of F is a map $f : \mathbb{R}^m \to \mathbb{R}^n$ such that $f(x) \in F(x)$ for all $x \in X$.

LEMMA 3.6. *Let U be a compact subset of $\mathbb{R}^m$ and $f : \mathbb{R}^n \times U \to \mathbb{R}^n$ be a continuous function. Then $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ defined by $F(x) := \{f(x, u) : u \in U\}$ is continuous.*

PROOF. Let $x \in \mathbb{R}^n$ and $\varepsilon > 0$ be fixed. For all $u \in U$ we notice that the map

$$(y, v) \mapsto f(y, v) - f(x, u)$$

is continuous and vanishes when $(y, v) = (x, u)$. Thus there exists $\delta_u > 0$ such that

$$|f(y, v) - f(x, u)| < \frac{\varepsilon}{2}$$

for every $|y - x| < \delta_u$. By continuity, there exists $\rho_u > 0$ such that $|f(y, v) - f(x, u)| < \varepsilon$ for $|y - x| < \delta_u$, $|v - u| < \rho_u$. By compactness, U can be covered by finitely many balls $B(u_i, \rho_{u_i})$, $i = 1, \dots, N$. Set $\delta = \min\{\delta_{u_1}, \dots, \delta_{u_N}\} > 0$, we have $d_H(F(y), F(x)) \leq \varepsilon$ for all $|y - x| < \delta$. Hence F is continuous at x and, by arbitrariness of x, the proof is concluded. □

THEOREM 3.7 (Carathéodory). *Let $A \subseteq \mathbb{R}^n$ be closed. Then for every $x \in \operatorname{co} A$ there exist $\lambda_i = \lambda_i(x) \in [0, 1]$, and $x_i = x_i(x) \in A$, $i = 1, ..., n+1$ such that $\sum_{i=1}^{n+1} \lambda_i x_i = x$.*

PROOF. Omitted. □

THEOREM 3.8 (Ekeland's Variational Principle). *Let X be a complete metric space, $\psi : X \to]-\infty, +\infty]$ be a l.s.c. function not identially equal to $+\infty$ and bounded from below (i.e. there exists $m \in \mathbb{R}$ such that $\psi(x) \geq m$ for every $x \in X$). Then for all $\varepsilon > 0$, $x_0 \in X$ there exists $x_\infty \in X$ such that*

(1) $\psi(x_\infty) + \varepsilon d(x_0, x_\infty) \leq \psi(x_0)$;

(2) $\psi(x_\infty) < \psi(x) + \varepsilon d(x, x_\infty)$ *for all* $x \in X \setminus \{x_\infty\}$.

PROOF. For every $x, y \in X$ we define the partial order

$$y \preceq_X x \text{ if and only if } \psi(y) + \varepsilon d(x, y) \leq \psi(x),$$

noticing that for all $x, y, \in X$ we have $x \preceq_X x$, if $x \preceq_X y$ and $y \preceq_X x$ then $y = x$ and by triangle inequality if $x \preceq_X y$ and $y \preceq_X z$ then $x \preceq_X z$. Notice that $\psi : X \to]-\infty, +\infty]$ is *order preserving*, i.e., if $x_1 \preceq_X x_2$ we have $\psi(x_1) \leq \psi(x_2)$.

To prove the theorem, given $x_0 \in X$ we must find $x_\infty \in X$ such that $x_\infty \preceq_X x_0$ and x_∞ is minimal w.r.t. $\preceq_X$, i.e. if $y \neq x_\infty$ then $y \not\preceq_X x$, i.e.,

$$\psi(y) > \psi(x_\infty) - \varepsilon d(x_\infty, y).$$

Set

$$S(x) := \{y \in X : \psi(y) \leq \psi(x) - \varepsilon d(x, y)\} = \{y \in X : y \preceq_X x\} =:]-\infty, x],$$

where in the right hand side we use the order notation.
Since $x \in S(x)$ for every $x \in X$, such sets are nonempty, moreover they are also closed since $S(x)$ is the $\psi(x)$-sublevel[1] of the l.s.c. function $y \mapsto \psi(y) + \varepsilon d(x, y)$.

The theorem is proved if given $x_0 \in X$, we can find $x_\infty \in S(x_0)$ such that $S(x_\infty) = \{x_\infty\}$, because this means that if $y \neq x_\infty$ we cannot[2] have $y \preceq_X x_\infty$, and so x_∞ is minimal.

Let x_0 be arbitrary, we construct a sequence by recurrence. Given x_n, we define x_{n+1} as follows: We choose $x_{n+1} \in S(x_n)$ such that

$$\psi(x_{n+1}) \leq \inf_{x \in S(x_n)} \psi(x) + \frac{1}{2^n}.$$

The existence of such an element comes from the definition of inf.

For any $y \in S(x_{n+1})$ (i.e., $y \preceq_X x_{n+1}$), since $x_{n+1} \in S(x_n)$ (i.e., $x_{n+1} \preceq_X x_n$), we have $y \preceq_X x_n$ and so $y \in S(x_n)$, thus $S(x_n) \supseteq S(x_{n+1})$ for every $n \in \mathbb{N}$.

Given $y \in S(x_n)$, and recalling that $y, x_n \in S(x_{n-1})$, we have

$$\psi(y) + \varepsilon d(y, x_n) \underbrace{\leq}_{y \preceq_X x_n} \psi(x_n) \underbrace{\leq}_{\text{by def. of } x_n} \inf_{x \in S(x_{n-1})} \psi(x) + \frac{1}{2^{n-1}} \underbrace{\leq}_{y \in S(x_{n-1})} \psi(y) + \frac{1}{2^{n-1}},$$

[1] Indeed given a sequence $\{y_n\}_{n \in \mathbb{N}} \subseteq S(x)$ converging to $y \in X$ then we can pass to the lim inf in $\psi(y_n) \leq \psi(x) - \varepsilon d(x, y_n)$, and, recalling that ψ is l.s.c., we have $\psi(y) \leq \liminf_{n \to \infty} \psi(y_n) \leq \psi(x) - \varepsilon d(x, y)$, and so $y \in S(x)$.

[2] Be careful: $y \notin S(x_\infty)$ means $y \not\preceq_X x$, but this *does not mean* $y \succ_X x$, since the order relation is not total.

hence for every $y \in S(x_n)$ we have $d(y, x_n) \leq \dfrac{1}{\varepsilon 2^{n-1}}$, thus the diameter of $S(x_n)$ tends to zero. This implies that $\{x_n\}_{n\in\mathbb{N}}$ is a Cauchy sequence, since $x_n, x_m \in S(x_n)$ for all $m \geq n$, thus it converges to an element $x_\infty \in \bigcap_{n=1}^{\infty} S(x_n)$, moreover if $y \in \bigcap_{n=1}^{\infty} S(x_n)$ then by passing to the limit in $d(y, x_n) \leq 1/(\varepsilon 2^{n+1})$ we have $y = x_\infty$, so $S(x_\infty) \subseteq \{x_\infty\} = \bigcap_{n=1}^{\infty} S(x_n)$, and since $x_\infty \in S(x_\infty)$ equality holds. □

COROLLARY 3.9 (Localized form of Ekeland's Variational Principle). *Let X be a complete metric space, $\psi : X \to]-\infty, +\infty]$ be a l.s.c. function not identially equal to $+\infty$ and bounded from below. Then for all $\varepsilon, \delta > 0$ and $x \in X$ satisfying $\psi(x) \leq \inf_{x\in X} \psi(x) + \varepsilon$ there exist $x \in X$ and $x_\infty \in X$ such that*

(1) *$\psi(x_\infty) \leq \psi(x_0)$ and $d(x_\infty, x_0) < \delta$;*

(2) *$\psi(x_\infty) < \psi(x) + \dfrac{\varepsilon}{\delta} d(x, x_\infty)$ for all $x \in X \setminus \{x_\infty\}$.*

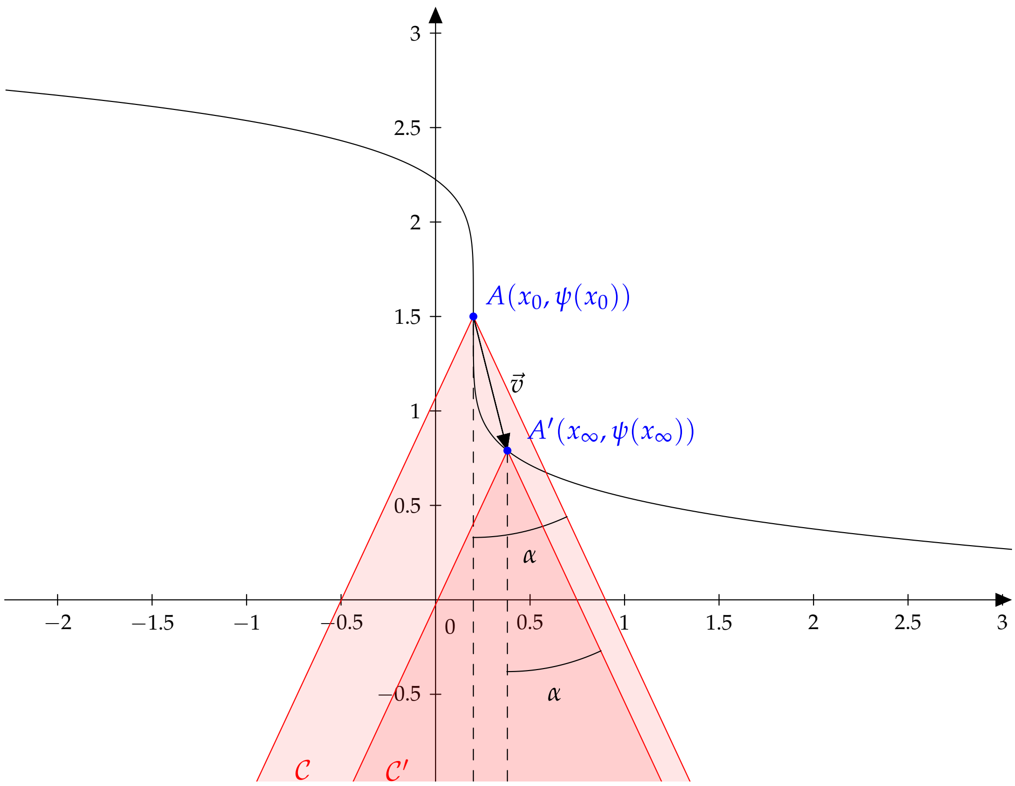
PROOF. Apply Ekeland's Variational Principle with ε/δ instead of δ: from

$$\psi(x_\infty) + \frac{\varepsilon}{\delta} d(x_0, x_\infty) \leq \psi(x_0)$$

by the choice of x_0 we deduce

$$\psi(x_\infty) + \frac{\varepsilon}{\delta} d(x_0, x_\infty) \leq \inf_{x\in X} \psi(x) + \varepsilon \leq \psi(x_\infty) + \varepsilon,$$

i.e. $d(x_0, x_\infty) \leq \delta$. All the other conclusions are the same. □



REMARK 3.10. Ekeland's Variational Principle's statement in $\mathbb{R}^2$ may be also expressed as follows: given a point A of the graph of ψ and an angle $0 < \alpha < \pi/2$, let $\mathcal{C}$ be the cone of vertex A, half-wideness α, axis parallel to y axis, and having A as point of maximum for the y-coordinate. Then there exists $A' \in \operatorname{graph} \psi \cap \mathcal{C}$ such that the graph of ψ, without the point A', is entirely contained in the complementar of the cone $\mathcal{C}' = \mathcal{C} + A' - A$, i.e., the translated of $\mathcal{C}$

along the vector $\vec{v} = A' - A$. For a comparison with the statement of the theorem, in this case we have $\tan\alpha = 1/\varepsilon$.

DEFINITION 3.11. Let Ω be an open subset of $\mathbb{R} \times \mathbb{R}^n$, U be a compact subset of $\mathbb{R}^m$, $f : \Omega \times \mathbb{R}^m \to \mathbb{R}^n$ be continuous, The *set of admissible controls* is

$$\mathcal{U} := \{u : \mathbb{R} \to U : \ u \text{ measurable}\}.$$

We say that an absolutely continuous function $x : [a, b] \to \mathbb{R}^n$ is an *admissible trajectory* for the *control system*

$$\dot{x} = f(t, x, u), \qquad\qquad u(\cdot) \in \mathcal{U},$$

if $\{(t, x(t)) : \ t \in [a, b]\} \subseteq \Omega$ and there exists $u(\cdot) \in \mathcal{U}$ such that $\dot{x}(t) = f(t, x(t), u(t))$ for a.e. $t \in [a, b]$.

To the control system we associate the *differential inclusion* $\dot{x} \in F(t, x)$ where $F : \Omega \rightrightarrows \mathbb{R}^n$ is defined by

$$F(t, x) := \{f(t, x, v) : \ v \in U\}.$$

We recall the following result.

THEOREM 3.12 (Lusin). *Let $h \in L^1([a, b])$. Then*

(1) *for every $\varepsilon > 0$ there exists a compact set $K \subseteq [a, b]$ and $g \in C^0(K)$ such that $h = g$ on K and* $\text{meas}([a, b] \setminus K) < \varepsilon$*;*
(2) *there exists a sequence of pairwise disjoint compacts $\{K_h\}_{h\in\mathbb{N}}$ such that $g \in C^0(K_h)$ for all $h \in \mathbb{N}$, $g = h$ on every K_h and*

$$\text{meas}\left([a, b] \setminus \bigcup_{h\in\mathbb{N}} K_h\right) = 0.$$

The link between admissible trajectories for the control system and trajectories of the differential inclusions is clarified in the following result.

THEOREM 3.13 (Filippov's Lemma on Implicit Functions). *Let Ω be an open subset of $\mathbb{R} \times \mathbb{R}^n$, U be a compact subset of $\mathbb{R}^m$, $f : \Omega \times \mathbb{R}^m \to \mathbb{R}^n$ be continuous. A curve $x \in AC([a, b]; \mathbb{R}^n)$ satisfies $\dot{x}(t) \in F(t, x(t))$ for a.e. $t \in [a, b]$, where $F(t, x) := \{f(t, x, v) : \ v \in U\}$, if and only if there exists $u(\cdot) \in \mathcal{U}$ such that $\dot{x}(t) = f(t, x(t), u(t))$ for a.e. $t \in [a, b]$.*

PROOF. Clearly, if there exists $u(\cdot) \in \mathcal{U}$ such that $\dot{x}(t) = f(t, x(t), u(t))$ for a.e. $t \in [a, b]$, then $\dot{x}(t) = F(t, x(t))$ for a.e. $t \in [a, b]$. We prove the converse implication. Let $\bar{w} \in U$ be fixed and define the set-valued map

$$W(t) := \begin{cases} \{w \in U : \ \dot{x}(t) = f(t, x(t), w)\}, & \text{if } \dot{x}(t) \in F(t, x), \\ \{\bar{w}\}, \text{ otherwise.} \end{cases}$$

Notice that the sets $W(t)$ are compact. Indeed, in the case $\dot{x} \notin F(t, x)$ this is trivial, otherwise since $W(t) = [\dot{x}(t) - f(t, x(t), \cdot)]^{-1}(0)$ with f continuous, they are closed. Moreover, $W(t) \subseteq U$ so they are also bounded, thus compact. Define the following map $\psi : L^1([a, b]; \mathbb{R}^m) \to \mathbb{R}$:

$$\psi(v) := \int_a^b d(v(t), W(t))\, dt.$$

We prove that this is a well-posed definition, i.e., $t \mapsto d(v(t), W(t))$ is an element of $L^1([a, b]; \mathbb{R}^m)$. After having proved this, it will be enough to prove that there exists $v_\infty \in L^1([a, b]; \mathbb{R}^m)$ with $\psi(v_\infty) = 0$. Indeed, in this case $v_\infty(t) \in W(t)$ for a.e. $t \in [a, b]$, thus proving the theorem.

Step 1: There exists a sequence of compact sets $\{K_m\}_{m\in\mathbb{N}}$ with $K_m \cap K_j = \emptyset$ and $\text{meas}\left([a, b] \setminus \bigcup_{m\in\mathbb{N}} K_m\right) = 0$ *such that $\dot{x}_{|K_m}, v_{|K_m} \in C^0(K_m)$ for every $m \in \mathbb{N}$.*

Proof of Step 1: Since $x(\cdot) \in AC([a,b])$, we have that $\dot{x}$ and v belong to $L^1([a,b])$. According to Lusin's Theorem, there exist $\{I_h\}_{h\in\mathbb{N}}$ and $\{J_k\}_{k\in\mathbb{N}}$ sequence of compact sets contained in $[a,b]$ such that $\dot{x} \in C^0(I_h)$, $v \in C^0(J_h)$ and

$$\operatorname{meas}\left([a,b] \setminus \bigcup_{h\in\mathbb{N}} I_h\right) = \operatorname{meas}\left([a,b] \setminus \bigcup_{k\in\mathbb{N}} J_k\right) = 0.$$

In particular, there exist $N_1, N_2 \subseteq [a,b]$ with $\operatorname{meas}(N_1) = \operatorname{meas}(N_2) = 0$ such that

$$[a,b] = N_1 \cup \bigcup_{h\in\mathbb{N}} I_h = N_2 \cup \bigcup_{k\in\mathbb{N}} J_k.$$

By taking the intersection, we have

$$\begin{aligned}
[a,b] &= \left(N_1 \cup \bigcup_{h\in\mathbb{N}} I_h\right) \cap \left(N_2 \cup \bigcup_{k\in\mathbb{N}} J_k\right) \\
&= \left(N_1 \cap \bigcup_{k\in\mathbb{N}} J_k\right) \cup \left(N_2 \cap \bigcup_{h\in\mathbb{N}} I_h\right) \cup (N_1 \cap N_2) \cup \left(\bigcup_{h\in\mathbb{N}} I_h \cap \bigcup_{k\in\mathbb{N}} J_k\right) \\
&\subseteq N_1 \cup N_2 \cup \bigcup_{h,k\in\mathbb{N}} I_h \cap J_k \\
&\subseteq N_1 \cup N_2 \cup \bigcup_{\substack{h,k\in\mathbb{N} \\ I_h\cap J_k\neq\emptyset}} I_h \cap J_k.
\end{aligned}$$

For every $m = (h,k) \in \mathbb{N}^2$ set $K_m = I_h \cap J_k$. If K_m is nonempty, then it is compact and $\dot{x}_{|K_m}, v_{|K_m} \in C^0(K_m)$ for every $m \in \mathbb{N}$. After having expunged by the list all the indexes m for which $K_m = \emptyset$, Step 1 is proved. $\diamond$

Step 2: There exists $L > 0$ such that $d(v(t), W(t)) \leq \|v(t)\| + L$ for every $t \in K_m$, $m \in \mathbb{N}$.

Proof of Step 2: If $t \in K_m$ then $v(t)$ is continuous. Since $W(t) \subseteq U$ and U is compact, there exists $L > 0$ such that

$$\sup\{\|w\| : w \in W(t)\} \leq \sup\{\|w\| : w \in U\} \leq L,$$

and so

$$d(v(t), W(t)) = \inf_{w\in W(t)} \|v(t) - w\| \leq \inf_{w\in W(t)} \|v(t) - 0\| + \|0 - w\| = \|v(t)\| + \inf_{w\in W(t)} \|w\| = \|v(t)\| + L.$$

This ends the proof of Step 2. $\diamond$

Step 3: The map $t \mapsto d(v(t), W(t))$ is measurable.

Proof of Step 3: To prove the measurability of $t \mapsto d(v(t), W(t))$, it is enough to prove that its restriction to every set K_m is measurable. To this aim, we will prove that $d(v(\cdot), W(\cdot))$ is l.s.c. on K_m. Let $t_\infty \in K_m$ and $\{t_n\}_{n\in\mathbb{N}}$ be a sequence in K_m such that

$$\liminf_{\substack{t\to t_\infty \\ t\in K_m}} d(v(\cdot), W(\cdot)) = \lim_{n\to\infty} d(v(t_n), W(t_n)).$$

For every n, let $w_n \in W(t_n)$ be such that $d(v(t_n), W(t_n)) = \|v(t_n) - w_n\|$. Such an element exists since the map $h(p) := \|v(t_n) - p\|$ is continuous (notice that here $v(t_n)$ is fixed) and $W(t_n)$ is compact. The sequence $\{w_n\}_{n\in\mathbb{N}}$ is contained in the compact set U, thus, up to possibly taking a subsequence, we have $w_n \to w_\infty$. Moreover, for every $n \in \mathbb{N}$ we have $\dot{x}(t_n) = f(t_n, x(t_n), w_n)$. Recalling that the map $\dot{x}(\cdot)$ is continuous in K_m, by passing to the limit $n \to \infty$ we have $\dot{x}(t_\infty) = f(t_\infty, x(t_\infty), w_\infty)$, and so $w_\infty \in W(t_\infty)$. But then

$$\liminf_{\substack{t\to t_\infty \\ t\in K_m}} d(v(t), W(t)) = \lim_{n\to\infty} d(v(t_n), W(t_n)) = \|v(t_\infty) - w_\infty\| \geq d(v(t_\infty), W(t_\infty)),$$

which proves l.s.c. at t_∞. $\diamond$

Step 4: The map $\psi : L^1([a,b]) \to \mathbb{R}$ satisfies Ekeland's Variational Principle's assumptions.

Proof of Step 4: Clearly ψ is bounded from below, moreover

$$\psi(v) = \int_a^b d(v(t), W(t))\, dt \le \int_a^b (\|v(t)\| + L)\, dt \le \|v\|_{L^1} + L(b-a) < +\infty.$$

We prove that ψ is l.s.c. Let $\{v_n\}_{n\in\mathbb{N}}$ an arbitrary sequence in L^1 with $v_n \to v$. By Fatou's Lemma, since $d(v_n(\cdot), W(\cdot))$ is measurable and nonnegative, we have

$$\int_a^b \liminf_{n\to\infty} d(v_n(t), W(t))\, dt \le \liminf_{n\to\infty} \int_a^b d(v_n(t), W(t))\, dt = \liminf_{n\to\infty} \psi(v_n).$$

Fix $\varepsilon > 0$. For a.e. $t \in [a,b]$ we have $v_n(t) \to v(t)$ and there exists $w_n \in W(t)$ such that $d(v_n(t), W(t)) > \|v_n(t) - w_n\| - \varepsilon$. By compactness, up to possibly taking a subsequence, $w_n \to w_\infty \in W(t)$, and so

$$\liminf_{n\to\infty} d(v_n(t), W(t)) \ge \|v(t) - w_\infty\| - \varepsilon \ge d(v(t), W(t)) - \varepsilon.$$

By the arbitrariness of $\varepsilon > 0$ we have $\liminf_{n\to\infty} d(v_n(t), W(t)) \ge d(v(t), W(t))$, and so

$$\psi(v) = \int_a^b d(v(t), W(t))\, dt \le \int_a^b \liminf_{n\to\infty} d(v_n(t), W(t))\, dt \le \liminf_{n\to\infty} \psi(v_n).$$

Since this holds true for every subsequence, we have

$$\psi(v) \le \liminf_{v_n\to v} \psi(v_n),$$

and so ψ is l.s.c. $\diamond$

Step 5: There exists $v_\infty \in L^1([a,b];\mathbb{R}^m)$ with $\psi(v_\infty) = 0$, and this concludes the proof of the theorem.

According to Ekeland's Variational Principle, there exists $v_\infty \in L^1$ such that $\psi(v_\infty) < \psi(v) + \frac{1}{2}\|v - v_\infty\|_{L^1}$ for all $v \in L^1$. Assume by contradiction that $\psi(v_\infty) > 0$. Then there exists m such that

$$\int_{K_m} d(v_\infty(t), W(t))\, dt > 0.$$

Let $\{q_j\}_{j\in\mathbb{N}}$ be a sequence dense in $\mathbb{R}^m$, and set

$$A_j := \{t \in K_n : d(v_\infty(t), q_j) < 2/3 \cdot d(v_\infty(t), W(t)) \text{ e } d(q_j, W(t)) \le 2/3 \cdot d(v_\infty(t), W(t))\}.$$

Clearly, $\bigcup_{j\in\mathbb{N}} A_j = \{t \in K_n : v_\infty(t) \notin W(t)\}$. The sets A_j are measurable, since $d(v_\infty(t), q_j) = |v_\infty(t) - q_j|$ and $d(v_\infty(t), W(t))$ are measurable functions (see Step 3). We will construct now a map $v(t)$ contradicting Ekeland's Variational Principle. Since $\psi(v_\infty) > 0$, there is at least a $\bar{\jmath} \in \mathbb{N}$ with $A_{\bar{\jmath}}$ of strictly positive measure. Set $v(t) = v_\infty(t)$ if $t \notin A_{\bar{\jmath}}$ and $v(t) = q_{\bar{\jmath}}$ if $t \in A_{\bar{\jmath}}$. We have that $v(t)$ is measurable, we will prove that this contradicts Ekeland's Variational Principle. We have

$$\|v - v_\infty\|_{L^1} = \int_{A_{\bar{\jmath}}} |q_j - v_\infty(t)|\, dt \le \frac{2}{3} \int_{A_{\bar{\jmath}}} d(v_\infty(t), W(t))\, dt.$$

Thus we obtain

$$\begin{aligned}
\psi(v) - \psi(v_\infty) &= \int_a^b \left(d(v(t), W(t)) - d(v_\infty(t), W(t))\right) dt = \int_{A_{\bar{\jmath}}} \left(d(q_{\bar{\jmath}}, W(t)) - d(v_\infty(t), W(t))\right) dt \\
&< \int_{A_{\bar{\jmath}}} \frac{2}{3} d(v_\infty, W(t))\, dt - \int_{A_{\bar{\jmath}}} d(v_\infty(t), W(t))\, dt \\
&= -\frac{1}{3} \int_{A_{\bar{\jmath}}} d(v_\infty(t), W(t))\, dt < -\frac{1}{2}\|v - v_\infty\|_{L^1}.
\end{aligned}$$

However, according to Ekeland's Variational Principle, we have $\psi(v) - \psi(v_\infty) > -\frac{1}{2}\|v - v_\infty\|_{L^1}$. $\square$

4. Lecture of 14 december 2018: Closure of the set of admissible trajectories (3h)

REMARK 4.1. We are going to treat now the problem of the *closure* of the set of admissible trajectories, i.e., providing sufficient conditions ensuring that the uniform limit of a sequence of admissible trajectories will be an admissible trajectory. In general this property fails.

EXAMPLE 4.2. Consider $\dot{x}(t) = u(t)$ for a.e. $t \in [0,1]$ where $u(t) \in \{-1,1\}$. It is easy to construct a sequence of admissible trajectories $\{x_n\}_{n\in\mathbb{N}}$ such that $\|x_n\|_\infty \to 0$, however $x_\infty(t) \equiv 0$ *is not* an admissible trajectory.

THEOREM 4.3 (Closedness of the set of admissible trajectories). *Assume that the set-valued map $F : [a,b] \times \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ associated to the system is continuous with compact and convex values. Then the set of trajectories $x(\cdot) \in AC([a,b];\mathbb{R}^n)$ such that $\dot{x}(t) \in F(t,x(t))$ for a.e. $t \in [a,b]$ is closed in $C^0([a,b];\mathbb{R}^n)$.*

PROOF. Let $\{x_n(\cdot)\}_{n\in\mathbb{N}} \subset AC([a,b];\mathbb{R}^n)$ be a sequence of AC curves satisfying $\dot{x}_n(t) \in F(t,x_n(t))$ for a.e. $t \in [a,b]$ and uniformly convergent to $x(\cdot) \in C^0([a,b];\mathbb{R}^n)$. In particular, for n sufficiently large there exists a compact set $K \subseteq [a,b] \times \mathbb{R}^n$ such that $(t,x_n(t)) \in K$ per ogni $t \in [a,b]$. Thus we have equiboundedness of the trajectories. By continuity of F, this implies boundedness of F on $[a,b] \times K$, so equi-Lipschitz continuity of $x_n(\cdot)$. This implies that $x(\cdot)$ is Lipschitz continuous, and so a.e. differentiable in $[a,b]$.

To end the proof, is thus enough to show that for all $\tau \in]a,b[$, where $\dot{x}(\tau)$ exists, we have $\dot{x}(\tau) \in F(\tau,x(\tau))$. By contradiction, let $\tau \in]a,b[$ be such that this property is not true. We strictly separate the compact and convex sets $F(\tau,x(\tau))$ and $\{\dot{x}(\tau)\}$ by an affine hyperplane, thus there exists $\varepsilon > 0$, $p \in \mathbb{R}^n$, $\|p\| = 1$ such that

$$\langle p, y\rangle \leq \langle p, \dot{x}\rangle - 4\varepsilon,$$

for all $y \in F(\tau,x(\tau))$. By continuity of F, there exists $\delta > 0$ such that if $|t-\tau| < \delta$ and $|x' - x(\tau)| < \delta$ then

$$\langle p, y\rangle \leq \langle p, \dot{x}\rangle - 3\varepsilon,$$

for all $y \in F(t,x')$.

Recalling that the map $t \mapsto x(t)$ is differentiable at τ, we can choose $\tau' \in [\tau, \tau+\delta[$ such that

$$\left|\frac{x(\tau')-x(\tau)}{\tau'-\tau} - \dot{x}(\tau)\right| < \varepsilon.$$

Moreover, by uniform convergence, we can choose n sufficiently large such that

$$\left|\frac{x(\tau')-x(\tau)}{\tau'-\tau} - \frac{x_n(\tau')-x_n(\tau)}{\tau'-\tau}\right| < \varepsilon.$$

This implies

$$\begin{aligned}
\langle p, \frac{x_n(\tau')-x_n(\tau)}{\tau'-\tau}\rangle &= \langle p, \dot{x}(\tau)\rangle + \langle p, \frac{x_n(\tau')-x_n(\tau)}{\tau'-\tau} - \dot{x}(\tau)\rangle \\
&\geq \langle p, \dot{x}(\tau)\rangle - \left|\frac{x_n(\tau')-x_n(\tau)}{\tau'-\tau} - \dot{x}(\tau)\right| \\
&= \langle p, \dot{x}(\tau)\rangle - \left|\frac{x_n(\tau')-x_n(\tau)}{\tau'-\tau} + \frac{x(\tau')-x(\tau)}{\tau'-\tau} - \frac{x(\tau')-x(\tau)}{\tau'-\tau} - \dot{x}(\tau)\right| \\
&\geq \langle p, \dot{x}(\tau)\rangle - \left|\frac{x_n(\tau')-x_n(\tau)}{\tau'-\tau} - \frac{x(\tau')-x(\tau)}{\tau'-\tau}\right| - \left|\frac{x(\tau')-x(\tau)}{\tau'-\tau} - \dot{x}(\tau)\right| \\
&\geq \langle p, \dot{x}(\tau)\rangle - 2\varepsilon.
\end{aligned}$$

However, recalling that $\langle p, y\rangle \leq \langle p, \dot{x}\rangle - 3\varepsilon$ for all $y \in F(t,x')$, we have also

$$\langle p, \frac{x_n(\tau')-x_n(\tau)}{\tau'-\tau}\rangle = \frac{1}{\tau'-\tau}\int_\tau^{\tau'} \langle p, \dot{x}_n(s)\rangle\, ds \leq \langle p, \dot{x}(\tau)\rangle - 3\varepsilon.$$

and so $\langle p, \dot{x}(\tau)\rangle - 2\varepsilon \leq \langle p, \dot{x}(\tau)\rangle - 3\varepsilon$, which is a contradiction. $\square$

REMARK 4.4. Notice that the statement requires the convexity of $F(t,x)$, not the convexity of U. These two facts in general are not equivalent.

THEOREM 4.5 (Continuity of Input-Output Map). *Let U be a compact subset of $\mathbb{R}^m$, $f : \mathbb{R} \times \mathbb{R}^n \times U \to \mathbb{R}^n$ be a continuous function such that $x \mapsto f(t,x,u)$ is of class C^1. Assume also that there exist $C, L > 0$ such that $|f(t,x,u)| \le C$ and $\|\partial_x f(t,x,u)\| \le L$ for all $(t,x,u) \in \mathbb{R} \times \mathbb{R}^n \times U$. Fix $\bar{x} \in \mathbb{R}^n$. The* input-output map *is the function associating to every control $u(\cdot) \in L^1([0,T];U)$ the unique solution $x_u(\cdot) \in C^0([0,T];\mathbb{R}^n)$ of $\dot{x}(t) = f(t,x(t),u(t))$ with $x(0) = \bar{x}$. In the above assumptions, the input-output map is continuous.*

PROOF. In the above assumption we have existence and uniqueness for the solution of the Cauchy problem for every $T > 0$. Set $\Lambda = L^1([0,T];U)$ and $X = C^0([0,T];\mathbb{R}^n)$, define $\Phi : \Lambda \times X \to X$ by setting

$$\Phi(u,w)(t) = \bar{x} + \int_0^t f(s,w(s),u(s))\,ds.$$

The theorem is proved if we show that Φ satisfies the assumptions of the parametric contraction lemma, in this case $x_u(\cdot)$ is the fixed point of $w \mapsto \Phi(u,w)$. We endow X with the equivalent norm $\|w\|_X = \max\{e^{-2Lt}|w(t)| : t \in [0,T]\}$.

We prove that $u \mapsto \Phi(u,w)$ is continuous. Let $\{u_k\}_{k\in\mathbb{N}}$ be a sequence in Λ converging to $u \in \Lambda$. We have

$$|\Phi(u_k,w)(t) - \Phi(u,w)(t)| \le \int_0^T |f(s,w(s),u_k(s)) - f(t,w(s),u(s))|\,ds.$$

From every subsequence $\{u_{k'}\}$ of $\{u_k\}_{k\in\mathbb{N}}$ is possible to extract a subsequence $\{u_{k''}\}$ a.e. pointwise converging to u, and by the boundedness of f is possible to use the Dominated Convergence Theorem

$$\lim_{k''\to\infty} |\Phi(u_{k''},w)(t) - \Phi(u,w)(t)| \le \int_0^T \lim_{k''\to\infty} |f(s,w(s),u_{k''}(s)) - f(s,w(s),u(s))|\,dt = 0,$$

By arbitrariness of $\{u_{k'}\}$, we conclude that

$$\lim_{k\to\infty} \|\Phi(u_k,w) - \Phi(u,w)\|_\infty = 0.$$

Suppose now that $w, w' \in X$ and set $\|w - w'\|_X = \delta$. Then $|w(s) - w'(s)| \le \delta e^{2Ls}$ for every $0 \le s \le T$. We thus have

$$\begin{aligned} e^{-2Lt}|\Phi(u,w)(t) - \Phi(u,w')(t)| &\le e^{-2Lt}\int_0^t |f(s,w(s),u(s)) - f(s,w'(s),u(s))|\,ds \\ &\le e^{-2Lt}\int_0^t L|w(s) - w'(s)|\,ds \le e^{-2Lt}\int_0^t L\delta e^{2Ls}\,ds < \frac{\delta}{2}. \end{aligned}$$

So $\|\Phi(u,w) - \Phi(u,w')\|_X \le \dfrac{1}{2}\|w - w'\|_X$. The thesis follows by parametric contraction lemma. □

REMARK 4.6. In the above assumptions, the input-output map is continuous, but not Lipschitz continuous. It becomes Lipschitz contininuous if we endow Λ with the metric of the convergence in measure, i.e. for every $u_1, u_2 \in \Lambda$ we define $d(u_1,u_2) = \text{meas}\{t \in [0,T] : u_1(t) \ne u_2(t)\}$.

We recall now some results about ordinary differential equations.

DEFINITION 4.7 (Adjoint system). Consider the system

$$\begin{cases} \dot{x}(t) &= A(t)x(t), \\ x(0) &= x_0, \end{cases}$$

where $A(t) \in \text{Mat}_{n\times n}(\mathbb{R})$. The *adjoint system* is $\dot{p}(t) = -p(t)A(t)$. We will call *fundamental matrix associated to the system* the matrix $M(t,s) \in \text{Mat}_{n\times n}(\mathbb{R})$ defined as the solution of

$$\begin{cases} \partial_t M(t,s) &= A(t)M(t,s), \\ M(s,s) &= \text{Id}_{\mathbb{R}^n}, \end{cases}$$

PROPOSITION 4.8. *In the previous notation, we have:*

(1) *Suppose* $\|A(t)\| \leq L$. *then the solutions of the system satisfy* $|x(t)| \leq e^{L(t-t_0)}|x_0|$.
(2) *We have* $M(t,s)M(s,t) = \mathrm{Id}_{\mathbb{R}^n}$.
(3) *Given the system*
$$\begin{cases} \dot{x}(t) &= A(t)x(t) + h(t), \\ x(\tau) &= x_\tau, \end{cases}$$
its solution is
$$x(t) = M(t,\tau)x_\tau + \int_\tau^t M(t,s)h(s)\,ds.$$
(4) *If* $A(t) \equiv A$ *is constant, then* $M(t,s) = e^{A(t-s)}$.
(5) *The i-th column of* $M(t,\tau)$ *solves*
$$\begin{cases} \dot{x}(t) &= A(t)x(t), \\ x(\tau) &= \vec{e}_i, \end{cases}$$
(6) *The map* $A(\cdot) \mapsto M(\cdot,\tau)$ *is continuous from* $L^1([a,b];\mathrm{Mat}_{n\times n}(\mathbb{R}))$ *to* $C^0([a,b];\mathrm{Mat}_{n\times n}(\mathbb{R}))$.

PROOF. Omitted. □

We will study now further regularity properties of the input-output map. In particular, we will study the differentiability w.r.t. the initial point and w.r.t. the controls. In the first case, we will set $g(t,x) := f(t,x,u(t))$.

THEOREM 4.9 (Differentiability w.r.t. initial state). *Consider the equation* $\dot{x}(t) = g(t,x(t))$ *with* $x(t_0) = x_0$ *and denote with* $x(t,x_0)$ *its solution. Let* $g \in C^0(\mathbb{R}\times\mathbb{R}^n)$ *be of class* C^1 *in the x-variable and such that* $|g| \leq M$, $|\partial_x g| \leq L$. *Let* $v_0 \in \mathbb{R}^n$ *be fixed, with* $|v_0| = 1$. *Let* $v_{x_0}(t)$ *be the solution of* $\dot{v}(t) = \partial_x g(t,x(t,x_0))v(t)$ *with initial condition* $v(t_0) = v_0$. *Then* $x(t,x_0)$ *is differentiable at* x_0 *and the directional derivatives satisfy*
$$v_{x_0}(t) := \lim_{\varepsilon\to 0} \frac{x(t,x_0+\varepsilon\bar{v}) - x(t,x_0)}{\varepsilon}$$
with uniform convergence in $[t_0,T]$.

PROOF. For sufficiently small $\varepsilon \geq 0$, define $x_\varepsilon(t,v_0) = x(t,x_0+\varepsilon v_0)$ and $y_\varepsilon(t,v_0) = x(t,x_0) + \varepsilon v_{x_0}(t)$. To prove the theorem, it is enough to show that
$$\lim_\varepsilon \frac{x_\varepsilon(t,v_0) - y_\varepsilon(t,v_0)}{\varepsilon} = 0.$$
Notice that $x_\varepsilon(\cdot,v_0)$ is a fix point of $w \mapsto \Phi(x_0+\varepsilon v_0, w)$ defined by
$$\Phi(x_0+\varepsilon v_0, w) = x_0 + \varepsilon v_0 + \int_{t_0}^t g(s,w(s))\,ds,$$
which is a contraction ($\alpha = 1/2$) w.r.t. to the previously defined norm $\|\cdot\|_X$. According to parametric contraction lemma
$$\frac{1}{\varepsilon}\|y_\varepsilon - x_\varepsilon\|_X \leq \frac{2}{\varepsilon}\|y_\varepsilon - \Phi(y_\varepsilon, u+\varepsilon\Delta u)\|_X$$
Thus it is enough to prove that
$$\lim_{\varepsilon\to 0}\sup_{t\in[t_0,T]} \frac{1}{\varepsilon}\left| x_0 + \varepsilon v_0 + \int_{t_0}^t g(s,y_\varepsilon(t,v_0))\,ds - y_\varepsilon(t,v_0)\right| = 0$$
Notice that
$$\begin{aligned} g(s,y_\varepsilon(s,v_0)) &= g(s,x(s,x_0)+\varepsilon v_{x_0}(s)) = g(s,x(s,x_0)) + \int_0^1 \frac{d}{d\sigma}\left(g(s,x(s,x_0)+\varepsilon\sigma v_{x_0}(s))\right)d\sigma \\ &= g(s,x(s,x_0)) + \int_0^1 \partial_x g(s,x(s,x_0)+\varepsilon\sigma v_{x_0}(s))\varepsilon v_{x_0}(s)\,d\sigma, \end{aligned}$$
$$v_{x_0}(t) - v_{x_0}(0) = \int_{t_0}^t \partial_x g(s,x(s,x_0))v_{x_0}(s)\,ds.$$

Thus

$$
\begin{aligned}
&\left|x_0+\varepsilon v_0+\int_{t_0}^t g(s,y_\varepsilon(t,v_0))\,ds-y_\varepsilon(t,v_0)\right|=\\
=&\left|x_0+\varepsilon v_0+\int_{t_0}^t g(s,x(s,x_0))\,ds+\int_{t_0}^t\int_0^1 \partial_x g(s,x(s,x_0)+\varepsilon\sigma v_{x_0}(s))\varepsilon v_{x_0}(s)\,d\sigma\,ds+\right.\\
&\left.-x(t,x_0)-\varepsilon v_{x_0}(t)\right|\\
=&\left|x_0+\int_{t_0}^t g(s,x(s,x_0))\,ds-x(t,v_0)+\varepsilon(v_0-v_{x_0}(t))+\right.\\
&\left.+\int_{t_0}^t\int_0^1 \partial_x g(s,x(s,x_0)+\varepsilon\sigma v(s))\varepsilon v_{x_0}(s)\,d\sigma\,ds\right|\\
=&\varepsilon\left|-\int_{t_0}^t \partial_x g(s,x(s,x_0))v_{x_0}(s)\,ds+\int_{t_0}^t\int_0^1 \partial_x g(s,x(s,x_0)+\varepsilon\sigma v_{x_0}(s))v_{x_0}(s)\,d\sigma\,ds\right|\\
=&\varepsilon\left|-\int_{t_0}^t\int_0^1 \partial_x g(s,x(s,x_0))v_{x_0}(s)\,d\sigma\,ds+\int_{t_0}^t\int_0^1 \partial_x g(s,x(s,x_0)+\varepsilon\sigma v_{x_0}(s))v_{x_0}(s)\,d\sigma\,ds\right|\\
=&\varepsilon\left|\int_{t_0}^t\int_0^1 \left[\partial_x g(s,x(s,x_0)+\varepsilon\sigma v_{x_0}(s))-\partial_x g(s,x(s,x_0))\right]v_{x_0}(s)\,d\sigma\,ds\right|
\end{aligned}
$$

Consedered a compact neighborhood of the trajectory $x(t,x_0)$, we can pass to the limit (uniformly) under the integral sign by the regularity of g applying the Dominated Convergence Theorem, the statement on the directional derivatives thus follows.

To prove the differentiability, is is sufficient to apply the Total Differential Theorem proving that the directional derivatives are continuous st x_0. Consider a sequence $\{\xi_n\}_{n\in\mathbb{N}}$ with $\xi_n\to x_0$. We prove that $v_{\xi_n}(t)\to v_{x_0}(t)$. By parametric contraction lemma, we already know that $x(\cdot,\xi_t)$ uniformly converges to $x(\cdot,x_0)$ in $[t_0,T]$. Set $A_n(t)=\partial_x g(t,x(t,\xi_n))$ and $A(t)=\partial_x g(t,x(t,x_0))$. For n sufficiently large, we have $\|A_n(t)-A(t)\|\le K$ and so, by dominated convergence

$$\lim_{n\to\infty}\|A_n-A\|_{L^1}=\lim_{n\to\infty}\int_{t_0}^T\|A_n(t)-A(t)\|\,dt=\int_{t_0}^T\lim_{n\to\infty}\|A_n(t)-A(t)\|\,dt=0.$$

By the properties of fundamental matrix, $M_n(t,\tau)$ uniformly converges to $M(t,\tau)$ where M_n and M are the fundamental matrices of the systems solved by v_{ξ_n} and v_{x_0}, respectively (i.e., ruled by $g(t,x(t,\xi_n))$ and $g(t,x(t,x_0))$ respectively). Therefore, $v_{\xi_n}(\cdot)$ uniformly converges to $v_{x_0}(\cdot)$. □

REMARK 4.10. In the same way as above, it can be proved that the input-output map is differentiable also w.r.t. the initial time t_0

5. Lecture of 17 december 2018: Dependence w.r.t. controls. Density. (3h)

PROPOSITION 5.1. *Assume the hypothesis on f to grant local existence and uniqueness, consider $\dot{x}(t)=f(t,x(t),u(t))$, $x(0)=x_0$, and denote by $x(t,u)$ its solution evaluated at time t. Suppose that $(x,u)\to f(t,x,u)$ is of class C^1 and that there exists $L>0$ such that $\|\partial_x f\|_\infty+\|\partial_u f\|_\infty\le L$. Let $\Delta u\in L^\infty$. Then there exists*

$$\lim_{\varepsilon\to0}\frac{x(t,u+\varepsilon\Delta u)-x(t,u)}{\varepsilon}=\frac{\partial}{\partial\varepsilon}x(t,u+\varepsilon\Delta u)_{|\varepsilon=0}=\int_0^t M(t,s)\partial_u f(s,x(s,u),u(s))\Delta u(s)\,ds$$

where $M(t,s)$ is the fundamental matrix of $\dot{v}(t)=\partial_x f(t,x(t,u),u(t))v(t)$.

PROOF. Let $z(t)$ be the solution of

$$\begin{cases}\dot{z}(t)=A(t)z(t)+\partial_u f(t,x(t,u),u(t))\Delta u(t),\\ z(0)=0,\end{cases}$$

where $A(t):=\partial_x f(t,x(t,u),u(t))$. We have

$$z(t)=\int_0^t M(t,s)\partial_u f(s,x(s,u),u(s))\Delta u(s)\,ds.$$

Thus it is enough to prove that

$$\frac{\partial}{\partial \varepsilon} x(t, u + \varepsilon \Delta u)_{|\varepsilon=0} = z(t)$$

uniformly in $[0, T]$, i.e.,

$$\lim_{\varepsilon \to 0} \frac{x(t, u + \varepsilon \Delta u) - x(t, u) - \varepsilon z(t)}{\varepsilon} = 0.$$

Set $x_\varepsilon(t) = x(t, u + \varepsilon \Delta u)$ and $y_\varepsilon(t) = x(t, u) + \varepsilon z(t)$, we estimate $|x_\varepsilon(t) - y_\varepsilon(t)|$. The map $x_\varepsilon(\cdot)$ is a fixed point of

$$w \mapsto \Phi(w, u + \varepsilon \Delta u) = x_0 + \int_0^t f(s, w(s), u(s) + \varepsilon \Delta u(s))\, ds.$$

and Φ is a contraction ($\alpha = 1/2$) w.r.t. the norm $\|\cdot\|_X$ previously defined. By parametric contraction lemma

$$\frac{1}{\varepsilon} \|x_\varepsilon - y_\varepsilon\|_X \leq \frac{2}{\varepsilon} \|\Phi(x_0 + \varepsilon v_0, y_\varepsilon) - y_\varepsilon\|_X.$$

So it is sufficient to prove that

$$\lim_{\varepsilon \to 0} \sup_{t \in [0,T]} \frac{1}{\varepsilon} \left| x_0 + \int_{t_0}^t f(s, y_\varepsilon(t, v_0), u(s) + \Delta u(s))\, ds - y_\varepsilon(t) \right| = 0.$$

We have

$$\begin{aligned}
&\frac{1}{\varepsilon} \left| x_0 + \int_0^t f(s, y_\varepsilon(s), u(s) + \varepsilon \Delta u(s))\, ds - y_\varepsilon(t) \right| = \\
&= \frac{1}{\varepsilon} \left| x_0 + \int_0^t f(s, y_\varepsilon(s), u(s) + \varepsilon \Delta u(s))\, ds - x(t, u) - \varepsilon z(t) \right| \\
&= \frac{1}{\varepsilon} \left| x_0 + \int_0^t f(s, x(s, u) + \varepsilon z(s), u(s) + \Delta u(s))\, ds - x_0 - \int_0^t f(s, x(s, u), u(s))\, ds - \varepsilon z(t) \right| \\
&= \frac{1}{\varepsilon} \left| \int_0^t [f(s, x(s, u) + \varepsilon z(s), u(s) + \varepsilon \Delta u(s)) - f(s, x(s, u), u(s))]\, ds - \varepsilon z(t) \right| \\
&= \frac{1}{\varepsilon} \left| \int_0^t \int_0^1 \frac{d}{d\sigma} [f(s, x(s, u) + \sigma \varepsilon z(s), u(s) + \sigma \varepsilon \Delta u(s))]\, ds - \varepsilon z(t) \right| \\
&= \frac{1}{\varepsilon} \left| \int_0^t \int_0^1 [\partial_x f(s, x(s, u) + \sigma \varepsilon z(s), u(s) + \sigma \varepsilon \Delta u(s))] \varepsilon z(s)\, ds + \right. \\
&+ \int_0^t \int_0^1 [\partial_u f(s, x(s, u) + \sigma \varepsilon z(s), u(s) + \sigma \varepsilon \Delta u(s))] \varepsilon \Delta u(s)\, ds + \\
&\left. - \int_0^t A(s) \varepsilon z(s) + \partial_u f(s, x(s, u), u(s)) \varepsilon \Delta u(s)\, ds \right| \\
&\leq \int_0^t \int_0^1 |\partial_x f(s, x(s, u) + \sigma \varepsilon z(s), u(s) + \sigma \varepsilon \Delta u(s)) - A(s)| \cdot |z(s)|\, ds + \\
&+ \int_0^t \int_0^1 |\partial_u f(s, x(s, u) + \sigma \varepsilon z(s), u(s) + \sigma \varepsilon \Delta u(s)) - \partial_u f(s, x(s, u), u(s))| \cdot |\Delta u(s)|\, ds
\end{aligned}$$

By Dominated Convergence Theorem, the limit is 0. The assumption of boundedness of the derivatives of f can be relaxed, requiring only their continuity. □

We will consider now the following problem: given a control system with admissible control value set U, is it possible to construct another system *approximatively equivalent* to the give one with *smaller* control value set $U' \subseteq U$?

THEOREM 5.2 (Density of the trajectories). *Assume the hypothesis on f to grant local existence and uniqueness, let $\dot{x}(t) = f(t, x(t), u(t))$, where $u(t) \in U$ and $U \subseteq \mathbb{R}^m$ is compact.*

(1) *The set of solutions generated by piecewise constant controls is dense (w.r.t. uniform convergence) in the set of solutions.*

(2) *Let $U' \subseteq U$ be closed and such that for every t, x we have*

$$\mathrm{co}\{f(t,x,u) : u \in U'\} \supseteq \{f(t,x,u) : u \in U\}.$$

Then every trajectory of the original system generated by a measurable control $u(\cdot)$ satisfying $u(t) \in U$ a.e. can be approximated in the uniform convergence norm by a trajectory generated by a measurable control $u'(\cdot)$ satisfying $u'(t) \in U'$.

PROOF. Let $L > 0$ be such that $|f(t,x_1,u) - f(t,x_2,u)| \leq L|x_1 - x_2|$.

The first assertion follows from the result about continuous dependence of the input-output map on controls: indeed, the piecewise constant functions are dense in L^1 according to the L^1-norm, thus given a control $u(\cdot)$, generating the trajectory $x_u(\cdot)$, it is possible to construct a sequenceof piecewise constant controls $\{u_n\}_{n\in\mathbb{N}}$ converging in L^1 to u. The corresponding solutions $x_n(\cdot)$ are uniformly convergent to $x_u(\cdot)$.

Suppose now to have a trajectory x_u of the original system, generated by a control $u(\cdot) \in \mathscr{U}$, and let us prove that we can uniformly approximate it by trajectories generated by controls taking values in U'.

Define the function $\psi(t) := e^{Lt} - 1$ and for every $\varepsilon > 0$ we define the tubular neighborhood of the trajectory $x_u(\cdot)$ by setting

$$\Gamma_\varepsilon := \{(t,x) : t \in [0,T], |x - x(t)| \leq \varepsilon\psi(t)\}.$$

Consider the set

$$\mathcal{F} := \{u' : [0,\tau] \to U' \text{ measurable } : (t, x_{u'}(t)) \in \Gamma_\varepsilon \text{ for all } t \in [0,\tau]\}.$$

On $\mathcal{F}$ we define the following partial order: $u'_1 \prec u'_2$ if $\mathrm{dom}(u'_1) \subseteq \mathrm{dom}(u'_2)$ and $u'_{2|\mathrm{dom}(u'_1)} = u'_1$. Given a totally ordered chain $\{u'_i : i \in I\}$, set $\mathrm{dom}(u'_\infty) = \bigcup_{i\in I} u'_i$ and if $t \in \mathrm{dom}(u'_\infty)$, set $u'(t) = u'_i(t)$ where i satisfies $t \in \mathrm{dom}(u'_i)$. Such u'_∞ is an upper bound of the chain, so by Zorn's lemma there are maximal elements.

Let $\bar{u}$ be a maximal in $\mathcal{F}$. If $\mathrm{dom}(\bar{u}) = [0,T]$ the proof is finished. Otherwise, suppose by contradiction that $\mathrm{intdom}(\bar{u}) =]0,\tau[$ with $\tau < T$, for all $t \in]0,\tau[$ holds $|x_u(t) - x_{\bar{u}}(t)| \leq \varepsilon\psi(t)$ with equality at $t = \tau$. Set $w = f(\tau, x_u(\tau), u(\tau))$, $v = \dfrac{x_u(\tau) - x_{\bar{u}}(\tau)}{|x_u(\tau) - x_{\bar{u}}(\tau)|}$ We prove that we can choose $u_\tau \in U'$ such that

$$\langle f(\tau, x_u(\tau), u_\tau), v\rangle > \langle w, v\rangle - \frac{\varepsilon L}{2}.$$

Indeed, by convexity assumption, $w = \sum_{i=0}^{n} \alpha_i f(\tau, x_u(\tau), v_i)$ with $\alpha_i \in [0,1]$, $v_i \in U'$, $i = 0,\dots,n$ and $\sum_{i=0}^{n} \alpha_i = 1$. If by contradiction for every $i = 0,\dots,n$ we had

$$\langle f(\tau, x_u(\tau), v_i), v\rangle < \langle w, v\rangle,$$

by multiplying for α_i and summing up

$$\langle w, v\rangle = \sum_{i=0}^{n} \alpha_i \langle f(\tau, x_u(\tau), v_i), v\rangle < \sum_{i=0}^{n} \alpha_i \langle w, v\rangle = \langle w, v\rangle,$$

which leads to a contradiction. So the choiche of u_τ is always possible.

We prove now that there exists $\delta > 0$ such that if we set $\bar{u}(t) = u_i$ in $[\tau, \tau+\delta[$, the solution $x_{\bar{u}}$ is contained in Γ_ε. This will conclude the proof contradicting the maximality of $\bar{u}$, indeed it will be a proper extension of $\bar{u}$ to $[0, \tau+\delta[$. To this aim, we estimate the right derivative at τ of $|x_u(t) - x_{\bar{u}}(t)| - \delta\psi(t)$: it will be enough to show that such a derivative is negative. Indeed,

$$\begin{aligned}\frac{d}{dt}\left[|x_{\bar{u}}(t) - x_u(t)| - \varepsilon\psi(t)\right]_{t=\tau^+} &= \langle w - f(\tau, x_{\bar{u}}(\tau), u_\tau), v\rangle - \varepsilon L e^{L\tau} \\ &= \langle w - f(\tau, x_u(\tau), u_\tau), v\rangle + \langle f(\tau, x_u(\tau), u_\tau) - f(\tau, x_{\bar{u}}(\tau), v\rangle - \varepsilon L e^{L\tau}\end{aligned}$$

$$
\begin{aligned}
&\leq \frac{\varepsilon L}{2} + |f(\tau, x_u(\tau), u_\tau) - f(\tau, x_{\bar{u}}(\tau)| - \varepsilon L e^{L\tau} \\
&\leq \frac{\varepsilon L}{2} + L|x_u(\tau) - x_{\bar{u}}(\tau)| - \varepsilon L e^{L\tau} \\
&\leq \frac{\varepsilon L}{2} + \varepsilon L \psi(\tau) - \varepsilon L e^{L\tau} \\
&= \frac{\varepsilon L}{2} + \varepsilon L(e^{L\tau} - 1) - \varepsilon L e^{L\tau} \\
&\leq -\frac{\varepsilon L}{2} \leq 0.
\end{aligned}
$$

□

REMARK 5.3. The assumptions of the previous theorem require

$$\mathrm{co}\,\{f(t,x,u) :\, u \in U'\} \supseteq \{f(t,x,u) :\, u \in U\}.$$

In general *it is not true that* this is granted by taking $U' \subseteq U$ such that $\mathrm{co}(U') = \mathrm{co}(U)$. In other words, to convexify the set of admissible velocities is not enough to convexify the set of admissible control values.

However if the system is *affine* w.r.t. u, i.e., $f(t,x,u) = A(t,x) + B(t,x)u$ then this is enough: given $U' \subseteq U$ such that $\mathrm{co}(U') = \mathrm{co}(U)$ we have

$$\mathrm{co}\{f(t,x,u) :\, u \in U'\} = \{f(t,x,u) :\, u \in \mathrm{co}(U')\} = \{f(t,x,u) :\, u \in \mathrm{co}(U)\} \supseteq \{f(t,x,u) :\, u \in U\}.$$

DEFINITION 5.4. Consider the control system $\dot{x}(t) = f(t, x(t), u(t))$, $x(0) = \bar{x}$, where $u(\cdot) \in \mathscr{U} := \{u : [0, +\infty[\to U \text{ measurable}\}$ Define the *reachable set* from $\bar{x}$ at time t:

$$R_{\bar{x}}(t) = \{x(t) :\, x(\cdot) \text{ is solution of the system }\}.$$

In general, $R_{\bar{x}}(t)$ is not compact, not even if U is compact.

PROPOSITION 5.5 (Compactness of the reachable set). *Assume the hypothesis on f to grant local existence and uniqueness. Consider the control system $\dot{x}(t) = f(t, x(t), u(t))$, $x(0) = \bar{x}$, with $u(\cdot) \in \mathscr{U} := \{u : [0, +\infty[\to U$ measurable$\}$ and set $F(t,x) := \{f(t, x(t), u(t)) :\, u \in \mathscr{U}\}$. Suppose that $F(t,x)$ is compact and convex, and that the graphs of the solutions are all contained up to time t in a common compact K. Then $R_{\bar{x}}(t)$ is compact.*

PROOF. Since $R_{\bar{x}}(t) \subseteq K$, such a set is bounded. We prove its closedness. Suppose to have a sequence $\{\xi_n\}_{n\in\mathbb{N}}$ in $R_{\bar{x}}(t)$. Then $\xi_i = x_{u_i}(t)$ where $x_{u_i}(\cdot)$ is the solution generated by the control $u_i \in \mathscr{U}$. Since all the trajectories are contained in a common compact set, the set of their velocities must be bounded, by smoothness of f. So these trajectories are equibounded and equiLipschitz continuous by continuity of F and compactness of $F(t,x)$. Thus up to a subsequence they uniformly converges to $x_\infty(\cdot)$, and in particular $\xi_i \to \xi_\infty = x_\infty(t)$. By convexity, we have that $x_\infty(\cdot)$ is a solution, and so $\xi_\infty \in R_{\bar{x}}(t)$. □

DEFINITION 5.6 (Chattering controls). Assume the hypothesis on f to grant local existence and uniqueness. Consider the control system $\dot{x}(t) = f(t, x(t), u(t))$, $x(0) = \bar{x}$, with $u(\cdot) \in \mathscr{U} := \{u : [0, +\infty[\to U \text{ measurable}\}$ and set $F(t,x) := \{f(t, x(t), u(t)) :\, u \in \mathscr{U}\}$. Let $F^\sharp(t,x) := \mathrm{co}(F(t,x))$. Consider a new set of controls

$$U^\sharp := \{u^\sharp = (\theta_0, \dots, \theta_n, u_0, \dots, u_n) \in [0,1]^{n+1} \times U^{n+1}\},$$

and set for every $u^\sharp \in U^\sharp$

$$f^\sharp(t,x,u^\sharp) = \sum_{i=0}^{n} \theta_i f(t,x,u_i).$$

By Carathéodory's Theorem, we have

$$F^\sharp(t,x) := \{f^\sharp(t, x(t), u^\sharp(t)) :\, u^\sharp \in \mathscr{U}^\sharp\},$$

where $\mathscr{U}^\sharp := \{u^\sharp : [0, +\infty[\to U^\sharp \text{ measurable}\}$ is called the set of *chattering controls*.

COROLLARY 5.7. *Assume the hypothesis on f to grant local existence and uniqueness. Consider the control system $\dot{x}(t) = f(t, x(t), u(t))$, $x(0) = \bar{x}$, where $u(\cdot) \in \mathscr{U} := \{u : [0, +\infty[\to U \text{ measurable}\}$. Suppose that U is compact and that the graphs of solutions up to time t are all contained in a common compact K. Then $R^\sharp_{\bar{x}}(t) = \overline{R}_{\bar{x}}(t)$, where $R^\sharp_{\bar{x}}(t)$ is the reachable set for $\dot{x}(t) = f^\sharp(t, x(t), u^\sharp(t))$, $x(0) = \bar{x}$, where $u^\sharp(\cdot) \in \mathscr{U}^\sharp$.*

PROOF. We can identify U with a subset of $U^\sharp$, since the map $u \mapsto (1, 0, \dots, 0, u, u, \dots, u)$ is bijective. This implies that we can identify $\mathscr{U}$ with a subset of $\mathscr{U}^\sharp$ and $R_{\bar{x}}(t)$ with a subset of $R^\sharp_{\bar{x}}(t)$. We will always made this identification in the following of the proof. Trivially $U \subseteq U^\sharp$, and so $R_{\bar{x}}(t) \subseteq R^\sharp_{\bar{x}}(t)$. According to the previous result, we have that $R^\sharp_{\bar{x}}(t)$ is closed, so $\overline{R_{\bar{x}}(t)} \subseteq R^\sharp_{\bar{x}}(t)$. Moreover, we have $\mathrm{co}(F(t, x)) \supseteq F^\sharp(t, x)$, so the trajectories generated by $\mathscr{U}$ are dense in the set of the trajectories of the system generated by $\mathscr{U}^\sharp$, thus $\overline{R_{\bar{x}}(t)} \supseteq R^\sharp_{\bar{x}}(t)$, the thesis follows. $\square$

COROLLARY 5.8 (Bang-bang Theorem). *In the above assumptions, suppose $\dot{x}(t) = A(t)x(t) + h(t, u(t))$. If $A(\cdot)$ and h are continuous and U is compact then $R^\sharp_x(t) = R_x(t)$.*

PROOF. Omitted. $\square$

REMARK 5.9. In the three previous results, the assumtpions requiring that the graphs of the considered trajectories must be contained in a common compact can be replaced by the following condition of *uniform growth*: there exists $C > 0$ such that $|f(t, x, u)| \leq C(1 + |x|)$ for all (t, x, u). Indeed, in this case we have $|\dot{x}| \leq C(1 + |x|)$, so if $x(t) \neq 0$ we have

$$\frac{d}{dt}|x(t)| \leq \left|\frac{d}{dt}|x(t)|\right| \leq C(1 + |x|)$$

Thus if $x(t) \neq 0$ we have $|x(t)| \leq r(t)$ where $\dot{r}(t) = C(1 + r(t))$, $r(0) = |x(0)|$. Solving this equation, we have that if $|x(t)| \neq 0$ we have $|x(t)| \leq e^{Ct}(|\bar{x}| + 1) - 1 \leq e^{CT}(|\bar{x}| + 1) - 1 := R$. Thus $x(t) \in \overline{B(0, R)}$, which is compact.

6. Lecture of 21 december 2018: Pontryagin's Maximum Principle and Dynamic Programming Principle (3h)

DEFINITION 6.1 (Mayer and Bolza problems). Suppose to have a control system $\dot{x}(t) = f(t, x(t), u(t))$, $x(0) = \bar{x}$ with $\mathscr{U} := \{u : [0, T] \to U \text{ measurable }\}$, a cost function $\psi : \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}$, and a set $\tilde{S} \subseteq \mathbb{R}^{n+1}$ called *target set*.

The *Mayer's problem* is to determine a control (called *optimal control*) realizing

$$\inf_{u \in \mathscr{U}} \psi(T, x(T))$$

among all the admissible trajectories of the system satisfying $(T, x(T)) \in \tilde{S}$.
In many cases, we have $\tilde{S} = \mathbb{R} \times S$, where $S \subseteq \mathbb{R}^n$ is a given closed set. In this case the endpoint constraint $(T, x(T)) \in \mathcal{S}$ simply becomes $x(T) \in S$. With a slightly abuse of terminology, in this case we will called also S the *target set*.

The *Bolza's problem* is to determine a control realizing

$$\inf_{u \in \mathscr{U}} \int_0^T L(t, x(t), u(t))\, dt + \psi(T, x(T)),$$

where $L : \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$ is a function called *current cost* among all the admissible trajectories of the system satisfying $(T, x(T)) \in \tilde{S}$.

Bolza's problem can be reformulated as Mayer's problem by considering the new auxiliary variable x_0, and adding to the system the equation $\dot{x}_0(t) = L(t, x(t), u(t))$ with $x_0(0) = 0$. Defined a new cost $\Psi(x_0, x) = x_0 + \psi(x)$, and a new target set $\mathcal{S} = \mathbb{R} \times \tilde{S}$, the problem set up in $\mathbb{R}^{n+2}$ is a Mayer's problem.

The existence of optimal controls in many cases can be proved easily through standard arguments of l.s.c. and compactness.

THEOREM 6.2 (Existence of optimal controls). *Consider a Mayer's problem with closed target $\tilde{S} \subseteq \in \mathbb{R}^{n+1}$ and l.s.c. cost function $\psi(\cdot)$. Assume one of the following conditions*

(1) *$R_{\bar{x}}(T)$ is compact and $(\{T\} \times R_{\bar{x}}(T)) \cap \tilde{S} \neq \emptyset$;*
(2) *the dynamics f is continuous, satisfying $|f(t,x,u)| \leq C(1+|x|)$ for a certain $C > 0$, and the associated set-valued map has convex values.*
(3) *$R_{\bar{x}}(T)$ is closed, $(\{T\} \times R_{\bar{x}}(T)) \cap \tilde{S} \neq \emptyset$, and $\psi(\cdot)$ is coercive;*

Then there exists an optimal control.

PROOF.

(1) By compactness of $R_{\bar{x}}(T)$ and closedness of $\tilde{S}$, we have that $\{T\} \times R_{\bar{x}}(T) \cap \tilde{S}$ is compact. Since $\psi(\cdot)$ is l.s.c., it admits a minimum $(T, x_T) \in \{T\} \times R_{\bar{x}}(T) \cap \tilde{S}$. Recalling that $x_T \in R_{\bar{x}}(T)$, this implies that there exists a control $u^*(\cdot)$ generating an admissible trajectory whose endpoint is x_T.

(2) Consider a minimizing sequence $u_n : [0,T] \to U$ of admissible controls, generating the corresponding sequence of trajectories $x_n(\cdot)$, i.e., such that $\lim_{n\to+\infty} \psi(T, x_n(T)) = \inf_{u\in\mathscr{U}} \psi(T, x_u(T))$ and $(T, x_u(T)) \in \tilde{S}$. The growth condition on f ensures that there exists a compact set K such that $x_n(t) \in K$ for all $t \in [0,T]$, $n \in \mathbb{N}$, thus by the smoothness of f, we have that $x_n(\cdot)$ are equiLipschitz continuous and equibounded, thus, up to a not relabeled subsequence, we may assume that $\{x_n(\cdot)\}_{n\in\mathbb{N}}$ uniformly converges to $x_\infty(\cdot)$. The growth condition implies the compactness of $F(t,x)$, and by assumption we have that $F(t,x)$ is convex, thus $x_\infty(\cdot)$ is an admissible trajectory, and so it is generated by an admissible control $u_\infty(\cdot)$. Thus, recalling l.s.c. of ψ,
$$\inf_{u\in\mathscr{U}} \psi(T, x_u(T)) = \lim_{n\to+\infty} \psi(T, x_n(T)) \geq \psi(T, x_\infty(T)) \geq \inf_{u\in\mathscr{U}} \psi(T, x_u(T)),$$
we conclude that $x_\infty(\cdot)$ is an optimal trajectory and $u_\infty(\cdot)$ is an optimal control.

(3) Same as in item 1., recalling that a lower semicontinuous coercive function admits a minimum on every closed sets.

□

LEMMA 6.3 (Lebesgue points). *Let $g \in L^1([0,T];\mathbb{R})$. Then for a.e. $\tau \in [0,T]$ we have*
$$\lim_{\varepsilon\to0^+} \frac{1}{2\varepsilon} \int_{\tau-\varepsilon}^{\tau+\varepsilon} |g(t) - g(\tau)|\, dt = 0.$$
The points τ where the above limit exists and vanishes are called Lebesgue's point *of g.*

PROOF. Omitted. □

We will now state in a simplified form the following result, yielding *necessary conditions* enjoyed by an optimal control.

THEOREM 6.4 (Pontryagin' Maximum Principle). *Consider a Mayer's problem with $\phi(x(T)) = -\psi(x(T))$. Suppose $f \in C^0$, ϕ differentiable, and $x \mapsto f(t,x,u)$ of class C^1. Let $u^* \in L^\infty([0,T];\mathbb{R}^m)$ an optimal control generating the optimal trajectory $x^*(\cdot)$. Let $p^*(\cdot)$ be the solution of the adjoint system*
$$\begin{cases} \dot{p}(t) &= -p(t)\partial_x f(t, x^*(t), u^*(t)) \\ p(T) &= \nabla\phi(x^*(T)) \end{cases}$$
Then for a.e. $t \in [0,T]$ we have
$$\langle p^*(t), f(t, x^*(t), u^*(t))\rangle = \max_{u\in U} \langle p^*(t), f(t, x^*(t), u)\rangle.$$

PROOF. Since $t \mapsto g(t) := f(t, x^*(t), u^*(t))$ is in L^1, it is enough to prove the statement for all Lebesgue points of g. Let $\tau \in [0,T]$ be a Lebesgue's point of g. We will prove the result for $t = \tau$.

Given $\varepsilon > 0$, set $u_\varepsilon(t) = u^*(t)(1 - \chi_{[\tau-\varepsilon,\tau[}) + \omega\chi_{]\tau-\varepsilon,\tau]}$ where $\omega \in U$ is arbitrary. Let $x_\varepsilon(\cdot)$ be the trajectory generated by u_ε. Recalling that $x_\varepsilon(\tau-\varepsilon) = x^*(\tau-\varepsilon)$, we have

$$\begin{aligned}
x_\varepsilon(\tau) &= x_\varepsilon(\tau-\varepsilon) + \int_{\tau-\varepsilon}^{\tau} f(t, x_\varepsilon(t), \omega)\,dt,\\
x^*(\tau) &= x^*(\tau-\varepsilon) + \int_{\tau-\varepsilon}^{\tau} f(t, x^*(t), u^*(t))\,dt,\\
\frac{x_\varepsilon(\tau) - x^*(\tau)}{\varepsilon} &= \frac{1}{\varepsilon}\int_{\tau-\varepsilon}^{\tau} [f(t, x_\varepsilon(t), \omega) - f(t, x^*(t), u^*(t))]\,dt\\
&= \frac{1}{\varepsilon}\int_{\tau-\varepsilon}^{\tau} [f(t, x_\varepsilon(t), \omega) - f(\tau, x^*(\tau), u^*(\tau))]\,dt +\\
&+ \frac{1}{\varepsilon}\int_{\tau-\varepsilon}^{\tau} [f(\tau, x^*(\tau), u^*(\tau)) - f(t, x^*(t), u^*(t))]\,dt
\end{aligned}$$

By definition of Lebesgue's point, for $\varepsilon \to 0^+$ the second term in the right hand side vanishes. By Dominated Convergence Theorem,

$$\lim_{\varepsilon\to 0^+} \frac{x_\varepsilon(\tau) - x^*(\tau)}{\varepsilon} = f(\tau, x^*(\tau), \omega) - f(\tau, x^*(\tau), u^*(\tau)) =: \xi.$$

In the time interval $[\tau, T]$, we have that $x_\varepsilon(\cdot)$ satisfy the same differential equation of $x(\cdot)$, but with a different inital data $x_\varepsilon(\tau) \neq x^*(\tau)$. By the theorem about the differentiability of the input-output map w.r.t. initial data, denoted by $y_\varepsilon(\cdot)$ the trajectory generated by u^* starting from $x^*(\tau) + \varepsilon\xi$ at time τ, we have that for $t \geq \tau$ it holds

$$\lim_{\varepsilon\to 0} \frac{y_\varepsilon(t) - x^*(t)}{\varepsilon} = v(t)$$

where $\dot{v}(t) = \partial_x f(t, x^*(t), u^*(t))v(t)$ and $v(\tau) = \xi$. If $x^*(\cdot)$ is optimal, then $\phi(y_\varepsilon(T)) \leq \phi(x^*(\tau))$ for all $\varepsilon > 0$. Since ϕ is differentiable, we have

$$0 \geq \frac{d}{d\varepsilon}\phi(y_\varepsilon(T)) = \langle \nabla\phi(y_\varepsilon(T)), v(T)\rangle.$$

Moreover, we notice that

$$\frac{d}{dt}\langle p(t), v(t)\rangle = -\langle p(t)\partial_x f(t, x^*(t), u^*(t)), v(t)\rangle + \langle p(t), \partial_x f(t, x^*(t), u^*(t))v(t)\rangle = 0.$$

thus $\langle p(t), v(t)\rangle$ is costant at all $t \in [\tau, T]$ and $\langle p(T), v(T)\rangle = \langle p(\tau), v(\tau)\rangle$. By definition, $p(T) = \nabla\phi(x^*(T))$ and $v(\tau) = \xi = f(\tau, x^*(\tau), \omega) - f(\tau, x^*(\tau), u^*(\tau))$. Thus

$$\langle p(\tau), \xi\rangle = \langle p(\tau), f(\tau, x^*(\tau), \omega) - f(\tau, x^*(\tau), u^*(\tau))\rangle = \langle \nabla\phi(x^*(T)), v(T)\rangle \leq 0$$

so for all $\omega \in U$ we have

$$\langle p(\tau), f(\tau, x^*(\tau), \omega)\rangle \leq \langle p(\tau), f(\tau, x^*(\tau), u^*(\tau))\rangle,$$

the thesis follows. □

REMARK 6.5. In many applications and exercise, we will try to construct an optimal control candidate by solving the adjoint system and defining

$$u^*(t) = \arg\max_{u\in U}\langle p^*(t), f(t, x^*(t), u)\rangle.$$

The Pontryagin's Maximum Principle yields only necessary conditions, not sufficient ones. However in many cases it leads to isolate good candidates.

DEFINITION 6.6. Suppose that a Mayer's problem with target $\mathcal{S} \subseteq \mathbb{R}^{n+1}$ and terminal constraint $(T, x(T)) \in \mathcal{S}$ is given. Consider the admissible trajectories $x_{s,y}(\cdot)$ such that $x_{s,y}(s) = y$ and $(T, x_{s,y}(T)) \in \mathcal{S}$. We define the *value function* $V(s, y) = \inf_{u\in\mathscr{U}} \psi(T, x_{s,y}(T))$.

THEOREM 6.7 (Dynamic Programming Principle). *In the hypothesis granting existence and uniqueness of the trajectories, suppose that $\mathcal{S}$ is closed. Then the value function V is nondecreasing along the admissible trajectories, and it is constant along the optimal trajectories.*

PROOF. Let $x(\cdot)$ be an admissible trajectory generated by the control $u(\cdot)$, $t_0 < t_1$, $x(t_1) = x_1$ e $x(t_0) = x_0$. Suppose by contradiction that there exists $\varepsilon > 0$ such that $V(t_1, x_1) = V(t_0, x_0) - \varepsilon$. Thus by definition there exists an admissible control $v(\cdot)$ such that, for the trajectory $x_v(\cdot)$ generated by v and satisfying $x_v(t_1) = x_1$, we will have $\psi(T, x_v(T)) < V(t_1, x_1) + \varepsilon/2$. But in this case we can define a new control $\hat{u}(t) = u(t)$ for $t_0 \le t < t_1$ and $\hat{u}(t) = v(t)$ for $t_1 \le t \le T$, and let $\hat{x}(\cdot)$ be its associated trajectory satisfying $\hat{x}(t_0) = x_0$. Since $\hat{x}(t_1) = x_1$, we have

$$\psi(T, \hat{x}(T)) < V(t_1, x_1) + \frac{\varepsilon}{2} = V(t_0, x_0) - \frac{\varepsilon}{2},$$

contradicting the definition of $V(t_0, x_0)$.

If u^* is an optimal control and x^* is an optimal trajectory satisfying $x^*(t_0) = x_0$, then necessarily $V(t_0, x_0) = \psi(T, x^*(T))$. Since V is nondecreasing along the admissible trajectories, we have $V(T, x^*(T)) \ge V(t, x^*(t)) \ge V(t_0, x_0)$ for every $t_0 < t < T$, but since the first and the last terms of this inequality are equal, we have equality. □

THEOREM 6.8. *In the assumptions of the previous theorem, let $Q \subseteq R^{n+1}$ be an open set such that $Q \cap S \neq \emptyset$. If $V \in C^1(Q)$ then, defined the* Hamiltonian *function*

$$H(t, x, p) := \min_{u \in U} \langle p, f(t, x, u) \rangle,$$

the Hamilton-Jacobi-Bellman equation holds*:*

$$\begin{cases} \partial_t V(t, x) + H(t, x, \partial_x V) = 0, \\ \\ V(T, x) = \psi(x). \end{cases}$$

PROOF. By assumption, along the admissible trajectories we have

$$\frac{d}{dt} V(t, y(t)) \ge 0,$$

i.e., $\langle \partial_x V(t, y(t)), f(t, y(t), u(t)) \rangle \ge 0$ so the left hand side in the Hamilton-Jacobi-Bellman equation is nonnegative. By contradiction, assume that it is strictly positive, i.e., there exists $\theta > 0$ such that for all $\omega \in U$, t_0, x_0 it holds

$$\partial_t V(t_0, x_0) + \langle \partial_x V(t_0, x_0), f(t_0, x_0, \omega) > \theta.$$

By continuity, for every (t, x) in a neighborhood W of (t_0, x_0) we have

$$\partial_t V(t, x) + \langle \partial_x V(t, x), f(t, x, \omega) > \theta.$$

Let $u \in \mathscr{U}$ be an admissible control, and let $x_u(\cdot)$ be its corresponding trajectory such that $x_u(t_0) = x_0$. There exists $\delta > 0$ sufficiently small such that $(t, x_u(t)) \in W$ for all $u \in \mathscr{U}$. Thus for all $u \in \mathscr{U}$ we have

$$\begin{aligned} V(t + \delta, x_u(t + \delta)) - V(t_0, x_0) &= \int_{t_0}^{t_0 + \delta} \frac{d}{dt} V(t, x_u(t))\, dt \\ &= \int_{t_0}^{t_0 + \delta} [\partial_t V(t, x_u(t)) + \langle \partial_x V(t, x_u(t)), f(t, x_u(t), u(t))]\, dt > \delta\theta. \end{aligned}$$

By taking the infimum on u, we have

$$\inf_{u \in \mathscr{U}} V(t + \delta, x_u(t + \delta)) \ge V(t_0, x_0) + \delta\theta.$$

However, the infimum w.r.t. u of the left hand side is attained on optimal trajectories, and its value is $V(t_0, x_0)$ by the dynamic programming principle, and this leads to a contradiction. □

REMARK 6.9. If we have a Bolza problem with running cost $L(t, x, u)$, the Hamiltonian function becomes

$$H(t, x, p) = \min_{u \in U} [\langle p, f(t, x, u) \rangle + L(t, x, u)].$$

The *minimum time problem* corresponds to the case $\psi \equiv 0$, $L(t, x, u) \equiv 1$.

REMARK 6.10. The dynamic programming principle provides additional conditions to be used in addition to Pontryagin Maximum Principle in particular in the cases where disambiguation is needed (for instance, when $p(t) = 0$). It can be proved that the value function is *characterized* to be the solution of this equation, i.e., if the cost found by using a control $u(\cdot)$ coincides with the value of the solution of the HAmilton-Jacobi equation, then u is optimal.

REMARK 6.11. In almost all the case of interest, the value function is not C^1, thus a *classic* solution of Hamilton-Jacobi may not exists. Nonsmooth analysis allows to interpret such an equation by mean of generalized gradients, thus defining solutions (enjoying also uniqueness property) of such an equation, which are called *viscosity solutions*.

Linear Quadratic Regulator (LQR) is control system widely used in modeling, and it can be considered as a prototype of many problems.

DEFINITION 6.12 (LQR - finite time horizon). Let $Q, M \in \mathrm{Mat}_{n\times n}(\mathbb{R})$ be positive semidefinite symmetric matrices, $R \in \mathrm{Mat}_{m\times m}(\mathbb{R})$ be a positive definite symmetric matrix, $A \in \mathrm{Mat}_{n\times n}(\mathbb{R})$, $B \in \mathrm{Mat}_{n\times m}$ matrices, $t_0, T \in \mathbb{R}$, $t_0 < T$. Consider the control system in $\mathbb{R}^n$

$$\begin{cases} \dot{x}(t) = Ax(t) + Bu(t), \\ x(t_0) = x_0 \in \mathbb{R}^n, \end{cases}$$

with $u \in \mathbb{R}^m$.
Our aim is to minimize the following cost (also called *performance index*):

$$J_T(u(\cdot)) = \int_{t_0}^{T} \langle (Qx(t), x(t)\rangle + \langle Ru(t), u(t)\rangle)\, dt + \langle Mx(T), x(T)\rangle.$$

The matricex Q, M, R are called respectively the *current state cost*, the *final state cost*, and the *input cost* matrix.

In order to solve the problem, we use the dynamic programming principle. Indeed, for $t_0 \le t \le T$ we define

$$V_t(x) = \inf_{u(\cdot)} \left\{ \int_t^T (\langle Qx(s), x(s)\rangle + \langle Ru(s), u(s)\rangle)\, ds + \langle Mx(T), x(T)\rangle \right\},$$

subject to $\dot{x}(s) = Ax(s) + Bu(s)$ and $x(t) = x$. We have in particular that $V_T(x) = \langle Mx, x\rangle$ for all $x \in \mathbb{R}^n$.

We notice the following facts: given $\lambda \in \mathbb{R}$, we have $V_t(\lambda x) = \lambda^2 V_t(x)$, $x \mapsto V_t(x)$ is continuous, and

$$V_t(x_1) + V_t(x_2) = \frac{1}{2}[V_t(x_1 + x_2) + V_t(x_1 - x_2)].$$

Indeed, since the system is linear, if $x(\cdot)$ is a solution associated to some control $u(\cdot)$ and starting from x, then also $\lambda x(\cdot)$ is a solution associated to the control $\lambda u(\cdot)$ and starting from λx. Thus

$$V_t(\lambda x) = \inf_{u(\cdot)} \left\{ \int_t^T (\langle Q\lambda x(s), \lambda x(s)\rangle + \langle R\lambda u(s), \lambda u(s)\rangle)\, ds + \langle M\lambda x(T), \lambda x(T)\rangle \right\},$$

subject to $\dot{x}(s) = Ax(s) + Bu(s)$ and $x(t) = x$, hence $V_t(\lambda x) = \lambda^2 V_t(x)$. The other property can be verified similarly. Finally, the continuity w.r.t. x is given by the linearity of the system.

In general, a continuous map $W(\cdot) : \mathbb{R}^d \to \mathbb{R}$ satisfying $W(\lambda x) = \lambda^2 W(x)$ and

$$W(x_1) + W(x_2) = \frac{1}{2}[W(x_1 + x_2) + W(x_1 - x_2)],$$

satisfies also $W(x) = \langle Px, x\rangle$, where $P = (p_{ij})_{i,j=1,\dots,d} \in \mathrm{Mat}_{d\times d}(\mathbb{R})$ is the symmetric matrix defined by $p_{ii} = W(\vec{e}_i)$ and $p_{ij} = p_{ji} = \dfrac{1}{4}[W(x_i + x_j) - W(x_i - x_j)]$.

In this way we obtain that $V_t(x) = \langle P(t)x, x\rangle$ where $t \mapsto P(t)$ is a continuous map and $P(t)$ is a symmetric matrix for every $t \in [t_0, T]$. In particular, we have also $\partial_t V_t(x) = \langle \dot{P}(t)x, x\rangle$ and $\partial_x V_t(x) = 2P(t)x$.

Our problem is a special case of Bolza problem where $L(t,x,u) = \langle Qx, x\rangle + \langle Ru, u\rangle$, and so
$H(t,x,p) = \min_u [\langle p, Ax+Bu\rangle + \langle Qx,x\rangle + \langle Ru,u\rangle] = \langle p, Ax\rangle + \langle Qx,x\rangle + \min_u [\langle p, Bu\rangle + \langle Ru,u\rangle]$

The function to be minimized is strictly convex, smooth and coercive in u, thus the minimum is characterized bu putting the differential equal to zero, i.e. $B^T p + 2Ru = 0$, and so
$u = -\dfrac{1}{2}R^{-1}B^T p$. The Hamilton-Jacobi equation is given by $\partial_t V_t(x) + H(t,x,\partial_x V_t(x)) = 0$, i.e.

$$\langle \dot{P}(t)x, x\rangle + \langle 2P(t)x, Ax\rangle + \langle Qx,x\rangle - \langle 2P(t)x, BR^{-1}B^T P(t)x\rangle + \langle B^T P(t)x, R^{-1}B^T P(t)x\rangle = 0,$$

i.e. recalling that $P(t) = P^T(t)$

$$\begin{aligned}\langle \dot{P}(t)x, x\rangle + &\langle A^T P^T(t)x, x\rangle + \langle P(t)Ax, x\rangle + \\ &+ \langle Qx,x\rangle - \langle 2P(t)BR^{-1}B^T P(t)x, x\rangle + \langle P(t)BR^{-1}B^T P(t)x, x\rangle = 0,\end{aligned}$$

holding for every x, hence we obtain the *matrix Riccati equation*

$$\dot{P}(t) + A^T P(t) + P(t)A + Q - P(t)BR^{-1}B^T P(t) = 0,$$

coupled with terminal condition $P(T) = M$. Moreover, the optimal control is linear and is given by $u(t) = -R^{-1}B^T P(t)x(t)$.

All the previous consideraton easily extends to smooth time-depending matrices $A = A(t)$, $B = B(t)$, $R = R(t)$, $Q = Q(t)$.

Now we will discuss the case of infinite time horizon, i.e., $T \to +\infty$.

DEFINITION 6.13 (LQR - infinite time horizon). Let $Q, M \in \mathrm{Mat}_{n\times n}(\mathbb{R})$ be positive semidefinite symmetric matrices, $R \in \mathrm{Mat}_{m\times m}(\mathbb{R})$ be a positive definite symmetric matrix, $A \in \mathrm{Mat}_{n\times n}(\mathbb{R})$, $B \in \mathrm{Mat}_{n\times m}$ matrices, $t_0, T \in \mathbb{R}$, $t_0 < T$. Consider the control system in $\mathbb{R}^n$

$$\begin{cases}\dot{x}(t) = Ax(t) + Bu(t), \\ x(t_0) = x_0 \in \mathbb{R}^n,\end{cases}$$

with $u \in \mathbb{R}^m$.
Our aim is to minimize the following cost:

$$J_\infty(u(\cdot)) = \int_{t_0}^{+\infty} \langle (Qx(t), x(t)\rangle + \langle Ru(t), u(t)\rangle)\ dt.$$

It is clear in this case, since the integral is over an half line, that it may happen that $J(u(\cdot)) \equiv +\infty$, thus the minimization problems has no meaningful sense.

To tackle this difficulty, we make the following strong assumption (*complete controllability*): for every $\tau > 0$, $x \in \mathbb{R}^n$ there exists a control steering x to 0 in time τ. It can be proved (*Kalman rank controllability condition*) that this is equivalent to ask

$$\mathrm{rank}[B|AB|A^2B|\dots|A^{n-1}B] = n.$$

If $T_0 \le T_1$ we have

$$\int_{t_0}^{T_0} \langle (Qx(t), x(t)\rangle + \langle Ru(t), u(t)\rangle)\ dt \le \int_{t_0}^{T_1} \langle (Qx(t), x(t)\rangle + \langle Ru(t), u(t)\rangle)\ dt,$$

thus it make sense to approximate the infinite horizon problem with a sequence of finite-horizon problems: the map $T \mapsto J_T(u(\cdot))$ is monotone increasing, thus $J_\infty(u(\cdot))$ can be written as limit for $T \mapsto +\infty$ of $J_T(u(\cdot))$ (it can be $+\infty$).

We thus consider the matrix solution $P(t,T)$ of

$$\begin{cases}\dot{P}(t) + A^T P(t) + P(t)A + Q - P(t)BR^{-1}B^T P(t) = 0, \\ \\ P(T) = 0\end{cases},$$

and we notice that if we set

$$V_t(x,T) = \inf_{u(\cdot)} \left\{ \int_t^T (\langle Qx(s), x(s)\rangle + \langle Ru(s), u(s)\rangle)\ ds \right\},$$

we obtain $V_t(x,T) = \langle P(t,T)x, x\rangle$ and the map $T \mapsto V_t(x,T)$ is monotone increasing. Thus we have existence of the limit

$$\bar{P}(t) = \lim_{T\to+\infty} P(t,T).$$

By using standard flow properties for ODE, it can be proved that $\bar{P}(t)$ solves

$$\dot{\bar{P}}(t) + A^T\bar{P}(t) + \bar{P}(t)A + Q - \bar{P}(t)BR^{-1}B^T\bar{P}(t) = 0, \qquad t > t_0,$$

moreover $V_t(x,\infty) = \langle \bar{P}(t)x, x\rangle$.

Given a control $\tilde{u}$ defined on $[t, t+\tau]$ and steering $x(t)$ to 0 in time τ (the existence of such a control is due to complete controllability assumption), we can extend $\tilde{u}(\cdot)$ to all $[t, +\infty[$ by setting $\tilde{u}(s) = 0$ for $s > t+\tau$, thus we have that $J_\infty(\tilde{u}(t)) < +\infty$. This ensures that the infimum of $J_\infty(\cdot)$ is finite. Set now $\bar{u}(t) = -R^{-1}B^T P(t)x(t)$. We want to prove that it is optimal. Assume that there exists $\hat{u}(\cdot)$ such that $J_\infty(\hat{u}(\cdot)) < J_\infty(\bar{u})$. It turns out that there exists T such that we have for the restrictions

$$J_T(\hat{u}(\cdot)) < J_T(\bar{u})$$

but this is impossible, since $\bar{u}(\cdot)$ achieves the minimum of $J_T(\cdot)$.

APPENDIX A

Background remarks

1. Remarks on ordered set

DEFINITION 1.1 (Partial order relation). Let X be a set. A relation $\preceq_X$ on X is called a *partial order relation* on X if it is reflexive, antisymmetric and transitive, i.e., for all $x, y, z \in X$ we have

- $x \preceq_X x$,
- if $x \preceq_X y$ and $y \preceq_X x$ then $y = x$,
- if $x \preceq_X y$ and $y \preceq_X z$ then $x \preceq_X z$.

The pair $(X, \preceq_X)$ where X is a set an $\preceq_X$ is a partial relation on it, is called *partially ordered set* or poset. We will write also $y \succeq_X x$ to say $x \preceq_X y$.

DEFINITION 1.2 (Totally ordered chains and sets). Let $(X, \preceq_X)$ be a partially ordered set. A subset $S \subseteq X$ is *totally ordered* (or a *totally ordered chain*) if for every $x, y \in S$ with $x \neq y$ either $x \preceq_X y$ or $y \preceq_X x$, i.e., every pair of elements of S are *comparable* by $\preceq_X$. If the whole of X is a totally ordered chain, we will say that $\preceq_X$ is a total order relation on X, and X is called a totally ordered set.

DEFINITION 1.3 (Minimality, maximality, upper and lower bounds). Let $(X, \preceq_X)$ be a partially ordered set, $S \subseteq X$. We say that

- $m \in X$ is *minimal* for $\preceq_X$ if $x \preceq_X m$ implies $x = m$.
- $M \in X$ is *maximal* for $\preceq_X$ if $x \succeq_X M$ implies $x = M$.
- $x \in X$ is a *lower bound* for S if $x \preceq_X s$ for all $x \in S$.
- $x \in X$ is an *upper bound* for S if $x \succeq_X s$ for all $x \in S$.

LEMMA 1.4 (Zorn). *Let $(X, \preceq_X)$ be a partially ordered set. Assume that every totally ordered chain S of X admits an upper bound $x_S \in S$. Then in X there are maximal elements for $\preceq_X$.*

DEFINITION 1.5 (Infimum and supremum). Let $(X, \preceq_X)$ be a partially ordered set, $S \subseteq X$.

- We say that S has an *infimum* if there exists $z \in X$ such that
  (1) $z \preceq_X s$ for every $s \in S$,
  (2) for any $x \in X$ satisfying $x \preceq_X s$ for all $s \in S$ we have $x \preceq_X z$.
  If such a z exists then it is necessarily unique (assume to have $z_1, z_2 \in X$ satisfying the above properties, then we have $z_1 \preceq_X z_2$ and $z_2 \preceq_X z_1$, so $z_1 = z_2$), and it will be called the *infimum* of S w.r.t. $\preceq_X$ and denoted by $\inf S$.
- We say that S has an *supremum* if there exists $z \in X$ such that
  (1) $z \succeq_X s$ for every $s \in S$,
  (2) for any $x \in X$ satisfying $x \succeq_X s$ for all $s \in S$ we have $x \succeq_X z$.
  If such a z exists then it is necessarily unique (assume to have $z_1, z_2 \in X$ satisfying the above properties, then we have $z_1 \succeq_X z_2$ and $z_2 \preceq_X z_1$, so $z_1 = z_2$), and it will be called the *supremum* of S w.r.t. $\preceq_X$ and denoted by $\sup S$.

REMARK 1.6. Let X be a set, $(Y, \preceq_Y)$ be a partially ordered set, $f : X \to Y$ be a map. We will use the following notation

$$\inf_{x\in X} f(x) = \inf \{f(x) :\, x \in X\},$$
$$\sup_{x\in X} f(x) = \sup \{f(x) :\, x \in X\},$$

noticing that in both cases the set $\{f(x) :\, x \in X\} \subseteq Y$, and the infimum and supremum are considered w.r.t. $\preceq_Y$.

DEFINITION 1.7 (Lattices). Let $(X, \preceq_X)$ be a partially ordered set. We say that

- $(X, \preceq_X)$ is a *join-semilattice* if for every $x_1, x_2 \in X$ the set $\{x_1, x_2\}$ admits a supremum. In this case we define $x_1 \vee x_2 = \sup\{x_1, x_2\}$,
- $(X, \preceq_X)$ is a *meet-semilattice* if for every $x_1, x_2 \in X$ the set $\{x_1, x_2\}$ admits an infimum. In this case we define $x_1 \wedge x_2 = \inf\{x_1, x_2\}$,
- $(X, \preceq_X)$ is a *lattice* if is both a join-semilattice and a meet-semilattice.

By induction, $(X, \preceq_X)$ is a join-semilattice (resp. meet-semilattice, lattice) if and only if every *finite* subset of X admits a supremum (resp. an infimum, both a supremum and an infimum). We say furthermore that

- $(X, \preceq_X)$ is a *complete join-semilattice* if *every* $S \subseteq X$ admits a supremum.
- $(X, \preceq_X)$ is a *complete meet-semilattice* if *every* $S \subseteq X$ admits an infimum.
- $(X, \preceq_X)$ is a *complete lattice* if *every* $S \subseteq X$ admits both an infimum and a supremum.

EXAMPLE 1.8.

- The extended real line $\mathbb{R} \cup \{\pm\infty\} = [-\infty, +\infty]$ is a complete lattice with respect to the usual order relation, in view of the topological completeness of real line. We set $\inf \emptyset = +\infty$ and $\sup \emptyset = -\infty$.
- Given a set S, the power set $\mathcal{P}(S) := \{F : F \subseteq S\}$ is a complete lattice: given $\mathcal{A} \subseteq \mathcal{P}(S)$ we have
$$\sup \mathcal{A} = \bigcup_{A \in \mathcal{A}} A, \qquad \inf \mathcal{A} = \bigcap_{A \in \mathcal{A}} A.$$
If $\mathcal{A} = \emptyset$, we take $\sup \emptyset = \emptyset$ and $\inf \emptyset = S$.

DEFINITION 1.9 (Order topology). On a totally ordered set $(X, \preceq_X)$ we can define the *open intervals w.r.t. the order*
$$\begin{aligned}]x, y[&:= \{t \in X : x \preceq_X t \preceq_X y,\, t \neq x,\, t \neq y\}, \\ ]x, +\infty[&:= \{t \in X : x \preceq_X t,\, t \neq x\}, \\ ]-\infty, x[&:= \{t \in X : t \preceq_X x,\, t \neq x\}, \\ ]-\infty, +\infty[&:= X. \end{aligned}$$
The *order topology* is the topology generated by all the open intervals w.r.t. the order.

2. Remarks on weak topology

We refer to Chapter 3 of **[4]** for all the proofs and further details on this section.

DEFINITION 2.1 (Weak topology). Let X be a set, $(Y_i, \tau_i)_{i \in I}$ be a family of topological spaces and $\mathscr{F} = \{f_i\}_{i \in I}$ be a family of functions such that $f_i : X \to Y_i$. Since there exists at least one topology τ_X on X such that the functions $f_i : X \to Y_i$ become continuous for all $i \in I$ (it is sufficient to take the discrete topology $\tau_X = \mathcal{P}(X)$), and recalling that the intersection of an arbitrary family of topologies is still a topology, there exists a coarser topology τ_X on X such that the functions $f_i : X \to Y_i$ become continuous for all $i \in I$. This topology is the intersection of all the topologies whith such a property, and it will be called the weak topology $\sigma(X, \mathscr{F})$ w.r.t. $\mathscr{F}$.

PROPOSITION 2.2 (Basis for the weak topology). *Let X be a set, $(Y_i, \tau_i)_{i \in I}$ be a family of topological spaces and $\mathscr{F} = \{f_i\}_{i \in I}$ be a family of functions such that $f_i : X \to Y_i$. Then a basis for the weak topology w.r.t. $\mathscr{F}$ is given by*
$$\mathscr{B} := \left\{ \bigcap_{j \in J} f_j^{-1}(A_j) : J \subseteq I,\ J \text{ finite},\ A_j \text{ open subset of } Y_j \text{ for all } j \in J \right\},$$
in the sense that every open set of the weak topology w.r.t. $\mathscr{F}$ can be written as an arbitrary union of elements in $\mathscr{B}$.

PROPOSITION 2.3 (Continuity for the weak topology). *Let X be a set, $(Y_i, \tau_i)_{i \in I}$ be a family of topological spaces and $\mathscr{F} = \{f_i\}_{i \in I}$ be a family of functions such that $f_i : X \to Y_i$. Let (Z, τ_Z) be another topological space. Then a map $g : Z \to X$, where X is equipped with the weak topology w.r.t. $\mathscr{F}$, is continuous if and only if $f_i \circ g : Z \to Y_i$ is continuous for all $i \in I$.*

We consider now the following particular case:

- $Y_i = \mathbb{R}$ for all $i \in I$, thus $\mathscr{F} \subseteq Y^X := \{f : X \to Y : f \text{ function}\}$;
- X is a normed space, endowed to the topology given by its norm $\|\cdot\|_X$ (we will call it the *strong topology on* X);
- $\mathscr{F} = X' := \{\ell : X \to \mathbb{R} : \ell \text{ is linear and continuous w.r.t. the strong topology on } X.\}$

DEFINITION 2.4 (Weak topology on normed spaces). Let X be a normed space, the topology $\sigma(X, X')$ is called the *weak topology* on X (we omit to specify w.r.t. X'). Given a sequence $\{x_k\}_{k\in\mathbb{N}}$ in X and $x \in X$ we say that x_k *weakly converges* to x if if converges w.r.t. $\sigma(X, X')$. In this case, x is the *weak limit* of $\{x_k\}_{k\in\mathbb{N}}$ and we will write $x_k \rightharpoonup x$.

THEOREM 2.5 (Properties of the weak topology on a normed space). *Let X be a normed space. The weak topology $\sigma(X, X')$ enjoys the following properties:*

(1) *The weak topology is Hausdorff (equivalently, if the weak limit exists, then it is unique).*
(2) *$x_k \rightharpoonup x$ if and only if $\langle \ell, x_k\rangle_{X',X} \to \langle \ell, x\rangle_{X',X}$ in $\mathbb{R}$ for all $\ell \in X'$.*
(3) *Given $x_0 \in X$, a basis of neighborhoods of x_0 for $\sigma(X, X')$ is given by finite intersection of sets of the form*
$$V_{\ell,r} := \{x \in X : |\langle \ell, x - x_0\rangle_{X',X}| < r\}$$
where $r > 0$, $\ell \in X'$.
(4) *If $x_n \to x$ strongly (i.e., according to the norm of X, equivalently if $\|x_n - x\|_X \to 0$) then $x_n \rightharpoonup x$, the converse in general does not hold.*
(5) *If $x_n \rightharpoonup x$ then $\|x_n\|_X$ is bounded and $\|x\|_X \le \liminf_{n\to\infty} \|x_n\|_X$.*
(6) *If $x_n \rightharpoonup x$ and $\|f_n - f\|_{X'} \to 0$ then $f_n(x_n) \to f(x)$.*
(7) *The weak and the strong topology coincide on X if and only if X is finite-dimensional, otherwise the weak topology is strictly coarser than the strong topology.*
(8) *A convex set $C \subseteq X$ is closed w.r.t. the strong topology if and only if it is closed w.r.t. the weak topology $\sigma(X, X')$.*

PROPOSITION 2.6. *Let E and F be two Banach spaces and let $T : E \to F$ be a linear operator. Then T is continuous in the strong topologies on E and F if and only if it is continuous from E endowed with $\sigma(E, E')$ topology into F endowed with $\sigma(F, F')$.*

EXERCISE 2.7. Let X be an infinite-dimensional normed space.

- the set $S := \{x \in X : \|x\|_X = 1\}$ is not closed w.r.t. $\sigma(X, X')$ and its closure w.r.t. the weak topology is $\{x \in X : \|x\|_X \le 1\}$;
- the set $B(0, 1) := \{x \in X : \|x\| < 1\}$ is not open w.r.t. $\sigma(X, X')$ and its interior w.r.t. the weak topology is empty.

DEFINITION 2.8 (Bidual space). Let X be a normed space. We define its *bidual space* X'' by setting
$$X'' = (X')' = \{f : X' \to \mathbb{R} : f \text{ is linear and continuous }\}.$$
Given $x \in X$, we define the *evaluation on x* $\delta_x : X' \to \mathbb{R}$ by setting $\delta_x(\ell) = \ell(x)$ for all $\ell \in X'$

PROPOSITION 2.9 (Properties of the evaluation). *Let X be a normed space, for all $x \in X$ we have $\delta_x \in X''$. The map $J : X \to X''$ defined as $J(x) = \delta_x$ is linear and continuous w.r.t. the strong and weak topologies on X and X'', and we will write*
$$\delta_x(\ell) = \langle J(x), \ell\rangle_{X'',X} = \langle Jx, \ell\rangle_{X'',X}.$$

DEFINITION 2.10 (Reflexive spaces). We say that a normed space is *reflexive* if $J : X \to X''$ is an homeomorphism.

DEFINITION 2.11 (Weak* topology). Let X be a normed space. The *weak* topology $\sigma(X', X)$ on X'* is the weak topology on X' w.r.t. $\mathscr{F} = \{Jx : x \in X\} \subseteq X''$. Given a sequence $\{x_k^*\}_{k\in\mathbb{N}}$ in X' and $x^* \in X'$ we say that x_k^* *weakly* converges* to x^* if if converges w.r.t. $\sigma(X', X)$. In this case, x^* is the *weak* limit* of $\{x_k^*\}_{k\in\mathbb{N}}$ and we will write $x_k^* \rightharpoonup x^*$.

PROPOSITION 2.12. *Let X be a normed space. The following properties hold:*

(1) *the weak* topology is Hausdorff;*

(2) *given $f \in X'$, a basis for the set of neighborhoods of f in the weak* topology is given by*
$$V = \{f \in X' : |\langle f - f_0, x_i\rangle_{X',X}| < \varepsilon, \text{ for all } i \in I\},$$
where I is finite, $x_i \in X$ for all $i \in I$ and $\varepsilon > 0$;

Let $\{f_n\}_{n\in\mathbb{N}}$ be a sequence in X', $f \in X'$, $\{x_n\}_{n\in\mathbb{N}}$ be a sequence, $x \in X$. Then:

(1) *$f_n \rightharpoonup^* f$ (i.e., f_n weakly* converges to f) if and only if $\langle f_n, x\rangle_{X',X} \to \langle f, x\rangle_{X',X}$ for all $x \in X$;*
(2) *If $f_n \to f$ strongly, then $f_n \rightharpoonup f$ weakly in $\sigma(X', X'')$, and if $f_n \rightharpoonup f$ weakly in $\sigma(X', X'')$ then $f_n \rightharpoonup^* f$ (i.e., weakly*, or in $\sigma(X', X)$);*
(3) *If $f_n \rightharpoonup^* f$, then $\|f_n\|$ is bounded and $\|f\| \leq \liminf \|f_n\|$;*
(4) *If $f_n \rightharpoonup^* f$ and $x_n \to x$ strongly in X, then $\langle f_n, x_n\rangle \to \langle f, x\rangle$.*

We recall now the following fundamental theorems:

THEOREM 2.13 (Banach-Alaoglu-Bourbaki). *The closed unit ball $\mathbb{B}' := \{x^* \in X' : \|x^*\|_{X'} \leq 1\}$ of X' is weakly*-compact.*

THEOREM 2.14 (Kakutani). *The closed unit ball $\mathbb{B} := \{x \in X : \|x\|_X \leq 1\}$ of X is weakly compact if and only if X is reflexive.*

3. Remarks on Sobolev spaces

Let I be a nonempty open interval of $\mathbb{R}$. Assume that $\psi : I \to \mathbb{R}$ is of class $C^1(I)$. Then, according to the formula of integration by parts we have for every $\varphi \in C^1_c(I)$ (i.e. $\varphi \in C^1(I)$ is zero outside a compact subset of I):
$$\int_I \varphi(x)\psi'(x)\,dx = [\varphi(x)\psi(x)]_{x=\inf I}^{x=\sup I} - \int_I \varphi'(x)\,\psi(x)\,dx = -\int_I \varphi'(x)\psi(x)\,dx,$$
since $\lim\limits_{x\to\inf I} \varphi(x) = \lim\limits_{x\to\sup I} \varphi(x) = 0$. We notice that the last term require much less regularity on ψ to be defined, since it does not require ψ to be $C^1(I)$, but just $L^1_{\mathrm{loc}}(I)$ (i.e. for every K compact subset of I we have $\psi \in L^1(K)$.
This suggest the following:

DEFINITION 3.1 (Weak derivative in $\mathbb{R}$). Let I be a nonempty open interval of $\mathbb{R}$. Let $f : I \to \mathbb{R}$ be a function of $L^1_{\mathrm{loc}}(I)$. We say that $g \in L^1_{\mathrm{loc}}(I)$ is the *weak derivative* or *derivative in the weak sense* of f if for every $\varphi \in C^1_c(I)$ we have
$$\int_I g(x)\,\varphi(x)\,dx = -\int_I f(x)\,\varphi'(x)\,dx.$$
If a function g satisfies the above relation, then is unique and will be denoted by f'. For smooth functions, the weak derivative and the classical one coicide.

The above definition can be easily extended to the several variable's case:

DEFINITION 3.2 (Weak derivative). Let Ω be a nonempty open subset of $\mathbb{R}^n$. Let $f : \Omega \to \mathbb{R}$ be a function of $L^1_{\mathrm{loc}}(\Omega)$. We say that $g_i \in L^1_{\mathrm{loc}}(I)$ is the *i-th weak partial derivative* or *i-th partial derivative in the weak sense* of f if for every $\varphi \in C^1_c(I)$ we have
$$\int_I g_i(x)\,\varphi(x)\,dx = -\int_I f(x)\,\partial_{x_i}\varphi(x)\,dx.$$
If a function g_i satisfies the above relation, then is unique and will be denoted by $\partial_{x_i} f$. For smooth functions, the weak derivative and the classical one coincide.

DEFINITION 3.3 (Sobolev space). Let Ω be a nonempty open subset of $\mathbb{R}^n$, $1 \leq p \leq \infty$, we define the *Sobolev space* $W^{1,p}(\Omega)$:
$$W^{1,p}(\Omega; \mathbb{R}) := \{f \in L^p(\Omega; \mathbb{R}) : \nabla f := (\partial_{x_1} f, \dots, \partial_{x_n} f) \in L^p(\Omega; \mathbb{R}^n)\}.$$
If $p = 2$ we set $H^1(\Omega; \mathbb{R}) = W^{1,2}(\Omega; \mathbb{R})$. Sobolev spaces are Banach spaces when equipped with the norm:
$$\|f\|_{W^{1,p}(\Omega;\mathbb{R})} := \|f\|_{L^p(\Omega;\mathbb{R})} + \|\nabla f\|_{L^p(\Omega;\mathbb{R}^n)}.$$

In the case $1 < p < +\infty$ we can use also the equivalent norm:

$$\|f\|_{W^{1,p}(\Omega;\mathbb{R})} := \left(\|f\|^p_{L^p(\Omega;\mathbb{R})} + \sum_{i=1}^n \|\partial_{x_i} f\|^p_{L^p(\Omega;\mathbb{R})} \right)^{1/p}.$$

The space $H^1(\Omega;\mathbb{R})$ is an Hilbert space with the scalar product:

$$\langle f_1, f_2 \rangle_{H^1(\Omega;\mathbb{R})} := \langle f_1, f_2 \rangle_{L^2(\Omega;\mathbb{R})} + \langle \nabla f_1, \nabla f_2 \rangle_{L^2(\Omega;\mathbb{R}^n)}$$

The space $W^{1,p}(\Omega;\mathbb{R})$ is separable for $1 \le p < +\infty$ and reflexive for $1 < p < +\infty$.

THEOREM 3.4 (Approximation by smooth functions). *Let Ω be an open subset of $\mathbb{R}^n$, $u \in W^{1,p}(\Omega)$ with $1 \le p < +\infty$. Then for any open subset ω compactly contained in Ω there exists a sequence $\{u_n\}_{n\in\mathbb{N}} \subseteq C^\infty_c(\mathbb{R}^n)$ such that $u_{n|\Omega} \to u$ in $L^p(\Omega)$ and $\nabla u_{n|\omega} \to \nabla u_{|\omega}$ in $L^p(\omega;\mathbb{R}^n)$. If $\Omega = \mathbb{R}^n$ or Ω has Lipschitz continuous boundary, there exists a sequence $\{u_n\}_{n\in\mathbb{N}} \subseteq C^\infty_c(\mathbb{R}^n)$ such that $u_{n|\Omega} \to u$ in $W^{1,p}(\Omega)$.*

THEOREM 3.5 (Characterization of Sobolev space). *Let Ω be an open subset of $\mathbb{R}^n$, $u \in L^p(\Omega;\mathbb{R})$ with $1 < p \le +\infty$ and let p' be such that $1/p + 1/p' = 1$. The following are equivalent:*

(1) *$u \in W^{1,p}(\Omega;\mathbb{R})$;*
(2) *there exists $C > 0$ such that for every $\varphi \in C^\infty_c(\Omega;\mathbb{R})$ and $i = 1,\dots,n$ it holds:*

$$\left| \int_{|\Omega} u(x)\partial_{x_i}\varphi(x)\,dx \right| \le C\|\varphi'\|_{L^{p'}(\Omega;\mathbb{R})}.$$

(3) *There exists $C > 0$ such that for every open subset ω compactly contained in Ω and $h \in \mathbb{R}$ with $|h| \le \inf_{y\in\partial\Omega} \mathrm{dist}(y,\omega)$ it holds ($\tau_h u(x) = u(x+h)$)*

$$\|\tau_h u - u\|_{L^p(\omega)} \le C|h|$$

Moreover, we can take $C = \|\nabla u\|_{L^p(\Omega;\mathbb{R}^n)}$.

In the case $p = 1$, we have still that (2) and (3) are equivalent, and (1) implies both of them. Function satisfying just (2) or (3) are called *function of bounded variation* and form the set $BV(\Omega)$.

COROLLARY 3.6. *If $u \in W^{1,p}(\Omega)$ and $\nabla u = 0$ a.e., then u is constant on each connected component of Ω.*

PROPOSITION 3.7 (product rule). *Let Ω be an open subset of $\mathbb{R}^n$. If $u, v \in L^\infty(\Omega) \cap W^{1,p}(\Omega)$ with $1 \le p \le \infty$ then $uv \in W^{1,p}(\Omega) \cap L^\infty(\Omega)$ and $\partial_{x_i}(uv) = v\partial_{x_i}u + u\partial_{x_i}v$ for every $i = 1,\dots,n$.*

DEFINITION 3.8 (Higher order Sobolev spaces). Given $\Omega \subseteq \mathbb{R}^n$ open and nonempty, and $m > 1$ we define by induction the following spaces:

$$W^{m,p}(\Omega;\mathbb{R}) := \{u \in L^p(\Omega,\mathbb{R}) : \nabla u \in W^{m-1,p}(\Omega;\mathbb{R}^n)\}.$$

and set $H^m(\Omega;\mathbb{R}) = W^{m,2}(\mathbb{R})$. We norm $W^{m,p}(\Omega;\mathbb{R})$ by summing all the L^p norm af all the (mixed) derivatives of order from 0 up to m, obtaining also in this case a Banach space. In a similar way, H^m is an Hilbert space.

THEOREM 3.9 (Sobolev embedding theorem). *Let $1 \le p < n$, and set $\frac{1}{p^*} = \frac{1}{p} - \frac{1}{n}$. Then for every $q \in [p, p^*]$*

$$W^{1,p}(\mathbb{R}^n;\mathbb{R}) \subseteq L^q(\mathbb{R}^n;\mathbb{R})$$

with continuous injection, and there exists $C = C_{n,p} > 0$ such that

$$\|u\|_{L^{\frac{1}{p}-\frac{1}{n}}(\mathbb{R}^n;\mathbb{R})} \le C\|\nabla u\|_{L^p(\Omega;\mathbb{R}^n)}, \qquad \textit{for every } u \in W^{1,p}(\mathbb{R}^n;\mathbb{R}).$$

THEOREM 3.10 (Limit case $p = n$). *For every $q \in [n, \infty[$*

$$W^{1,n}(\mathbb{R}^n;\mathbb{R}) \subseteq L^q(\mathbb{R}^n;\mathbb{R})$$

with continuous injection.

THEOREM 3.11 (Morrey). *Let $p > n$, then*

$$W^{1,p}(\mathbb{R}^n;\mathbb{R}) \subseteq L^\infty(\mathbb{R}^n;\mathbb{R})$$

with continuous injection, and there exists $C = C_{n,p} > 0$ such that for every $W^{1,p}(\mathbb{R}^n;\mathbb{R})$ we have

$$|u(y) - u(x)| \leq C|x-y|^\alpha \|\nabla u\|_{L^p},$$

with $\alpha = 1 - (n/p)$. This means that in this case every function of $W^{1,p}(\mathbb{R}^n;\mathbb{R})$ admits an Hölder continuous representative.

THEOREM 3.12. *Assume that Ω is an open subset of $\mathbb{R}^n$ with $\partial\Omega$ bounded and of class C^1, or Ω is an half space. Let $1 \leq p \leq \infty$. Then*

$$W^{1,p}(\Omega) \subset L^{p^*}(\Omega), \textit{for } p < n;$$
$$W^{1,n}(\Omega) \subset L^{q}(\Omega), \textit{for } n \leq q < \infty;$$
$$W^{1,p}(\Omega) \subset L^{\infty}(\Omega), \textit{for } p > n;$$

where $\frac{1}{p^*} = \frac{1}{p} - \frac{1}{n}$, and all the injections are continuous. Moreover, if $p > n$ there exists $C = C_{n,p,\Omega} > 0$ such that for every $W^{1,p}(\mathbb{R}^n;\mathbb{R})$ we have

$$|u(y) - u(x)| \leq C|x-y|^\alpha \|u\|_{W^{1,p}},$$

with $\alpha = 1 - (n/p)$.

THEOREM 3.13 (Rellich-Kondrachov). *Assume that Ω is an open bounded subset of $\mathbb{R}^n$ of class C^1. Let $1 \leq p \leq \infty$. Then*

$$W^{1,p}(\Omega) \subset L^{p^*}(\Omega), \textit{for } p < n;$$
$$W^{1,n}(\Omega) \subset L^{q}(\Omega), \textit{for } n \leq q < \infty;$$
$$W^{1,p}(\Omega) \subset C^0(\overline{\Omega}), \textit{for } p > n;$$

where $\frac{1}{p^*} = \frac{1}{p} - \frac{1}{n}$, and all the injections are compact. In particular, for all p, N we have $W^{1,p}(\Omega) \subset L^p(\Omega)$ with compact injection.

In many cases there may arise the problem of giving value 0 at the boundary of a set. This notion must be handled with care, since the boundary of an open C^1 subset of $\mathbb{R}^n$ is a set of null measure in $\mathbb{R}^n$.
We proceed in a different way:

DEFINITION 3.14 (Null trace at the boundary). We denote by $W^{1,p}_0(\Omega)$ the closure of $C^1_c(\Omega)$ with respect to the norm of $W^{1,p}(\Omega)$. With this definition, we have that $u \in W^{1,p}(\Omega) \cap C^0(\bar{\Omega})$ with $1 \leq p < +\infty$ belongs to $W^{1,p}_0(\Omega)$ if and only if $u = 0$ on Ω, thus recovering the classical definition. We set $H^1_0(\Omega) = W^{1,2}_0(\Omega)$. The space $W^{1,p}_0(\Omega)$ inherits the norm of $W^{1,p}(\Omega)$ and is a Banach space. The space $H^1_0(\Omega)$ equipped with the scalar product of $H^1(\Omega)$ is an Hilbert space.

THEOREM 3.15 (Poincaré's inequality). *Assume $1 \leq p < \infty$ and that Ω is bounded. Then there exists $C = C(\Omega, p) > 0$ such that*

$$\|u\|_{L^p} \leq C\|\nabla u\|_{L^p(\Omega;\mathbb{R}^d)}, \qquad \textit{for all } u \in W^{1,p}_0(\Omega).$$

In particular, $\|\nabla u\|_{L^p}$ defines on $W^{1,p}_0(\Omega)$ an equivalent norm on $W^{1,p}_0$. In the case of $H^1_0(\Omega)$, the scalar product $\langle \nabla u_1, \nabla u_2 \rangle_{L^2}$ is a scalar product that induces on $H^1_0(\Omega)$ a norm equivalent to the norm of $H^1_0(\Omega)$.

REMARK 3.16. Poincaré inequality holds also if Ω has finite measure or if it has bounded projection to a straigth line.

DEFINITION 3.17 (Dual spaces). For $1 \leq p < +\infty$ we set $W^{-1,p'} = (W_0^{1,p})'$, where $\frac{1}{p} + \frac{1}{p'} = 1$, and by $H^{-1}(\Omega)$ the dual of $H_0^1(\Omega)$. We identify $L^2(\Omega)$ with its dual, but we do not identify H_0^1 with its dual. We have

$$H_0^1(\Omega) \subset L^2(\Omega) \subset H^{-1}(\Omega),$$

with continuous and dense injections.

PROPOSITION 3.18 (Characterization of the dual). *Let $F \in W^{-1,p'}(\Omega)$ with $1 \leq p < +\infty$. Then there exists $f_0, \dots f_n \in L^{p'}(\Omega)$ such that*

$$\langle F, v \rangle_{W^{-1,p'}, W_0^{1,p}} = \int_\Omega f_0(x)\, v(x)\, dx + \sum_{i=1}^n \int_\Omega f_i(x) \partial_{x_i} v(x)\, dx, \text{ for all } v \in W_0^{1,p},$$

and $\|F\|_{W^{-1,p'}} = \max_{i=1,\dots,n} \|f_i\|_{L^{p'}}$. Moreover, if Ω is bounded, we can take $f_0 = 0$.

Bibliography

[1] Jürgen Appell and Petr P. Zabrejko, *Nonlinear superposition operators*, Cambridge Tracts in Mathematics, vol. 95, Cambridge University Press, Cambridge, 1990. MR1066204 (91k:47168)

[2] Andrea Braides, *Gamma-Convergence for Beginners*, Oxford University Press, 2002.

[3] Haim Brezis, *Analisi funzionale. Teoria e applicazioni*, Liguori, 1986.

[4] ———, *Functional analysis, Sobolev spaces and partial differential equations*, Universitext, Springer, New York, 2011. MR2759829

[5] Ivar Ekeland and Roger Témam, *Convex analysis and variational problems*, Corrected reprint of the 1976 English edition, Classics in Applied Mathematics, vol. 28, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1999. Translated from the French. MR1727362 (2000j:49001)

[6] F. H. Clarke, *Optimization and nonsmooth analysis*, 2nd ed., Classics in Applied Mathematics, vol. 5, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1990. MR1058436 (91e:49001)

[7] F. H. Clarke, Yu. S. Ledyaev, R. J. Stern, and P. R. Wolenski, *Nonsmooth analysis and control theory*, Graduate Texts in Mathematics, vol. 178, Springer-Verlag, New York, 1998. MR1488695 (99a:49001)

[8] Asen L. Dontchev, *The Graves theorem revisited*, J. Convex Anal. **3** (1996), no. 1, 45–53. MR1422750 (97g:46055)

[9] Mariano Giaquinta and Stefan Hildebrandt, *Calculus of variations. I*, Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 310, Springer-Verlag, Berlin, 1996. The Lagrangian formalism. MR1368401 (98b:49002a)

[10] ———, *Calculus of variations. II*, Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 311, Springer-Verlag, Berlin, 1996. The Hamiltonian formalism. MR1385926 (98b:49002b)