CentraleSupélec

Polycopié du cours de la $3^{\text{ème}}$ année, MOA

“Analyse Fonctionnelle”

A. Rozanova-Pierrat

17 september 2015

Contents

Introduction

The Functional Analysis, apeared in the beginning of last century from more old mathematical fields as the variational calculus, the theory of partial differential equations, the numerical analysis and the theory of integral equations, can be considered as a generalization and a geometrical interpretation of the mathematical analysis (calculus). Functions with some chosen caracteristic properties are considered as points or vectors of "functional spaces", in the most cases, infinite dimensional. Therefore, the notion of covergence from the calculus was generalized for the abstract functional spaces and it was related with *topological properties* of the functional spaces.

Let X be a set. There are two equivalent ways to define topological properties of a set X:

1. to define a convergence on X (which will implies the corresponding definition of a topology on X)
2. to define, in an axiomatic way, a *topology* on X, *i.e.* to define all *open* sets on X.

In the framework of this course we will be essentially restricted on the theory of metric spaces and the topology induced by a distance. An introduction of abstract topological spaces is given in Appendix A and can be consulted as a suplemental material. However, the most usefull questions - comparison of topologies, convergence, continuity and compactness- are advised to be read for having a more complete vision of the Functional Analysis.

Once the open sets are defined in X, the closed sets are their completions:

$$U \subset X \text{ is open iff } X \setminus U \text{ is closed.}$$

In addition to the open and closed sets in X, we need to define a neighborhood of an element $x \in X$. For the clearity, we will consider only open neighborhoods of x (see Appendix A for Definition A.1.6 and for the discussion about), *i.e.* open sets in X containing the point (element) x. The set of all open neighborhoods of x is denoted by $\mathcal{O}(x)$.

We will keep in mind (see Problem A.1.2 Appendix A) that a subset M of X (a topological space) is open if and only if every point $x \in A$ has a open neighborhood contained in A.

Now, we can define the convergence in our topology:

Definition 0.0.1 *Let (x_n) be a sequence of elements of X. We say that (x_n) converges to $x \in X$ if*

$$\forall V \in \mathcal{O}(x), \exists N \in \mathbb{N} \textit{ such that } \forall n \geq N \Rightarrow x_n \in V.$$

We see that the convergence is uniquely defined by open sets, *i.e.* the topology on X. Therefore, if we define the convergence, we fixe a topology on X.

In Chapter 2 we will introduce on X a distance (a metric) and a norm. Each time, it will modify the definition of open sets, but not the definition of the convergence. Abstract topological spaces are the most general spaces, which include metric spaces. Again, the metric spaces include all normed spaces, which include all pre-Hilbert spaces. In other words, all inner products define a norm, all norms define a distance, all distances define a topology, but not inverse! We will see an example of a convergence (weak convergence) which defines a nonmetrizable topology (cannot be associated with a distance). An other example is the space of infinitely differentiable and continuous functions with compact support (see Section 7.1), which is a topological space which cannot be metrizable too.

After the convergence, the property almost the most studied and important in Functional Analysis is the compactness. We refere to Appendix B for the compactness in the topological spaces. From this abstract theory we need to know at least the main definitions and results, which will be updated and precised in the framework of metric spaces in Section 2.3. Actually, one of the main goals of this course is to clearly understand why and in what the compactness property is important and to be able to distinct it in the framework of different topologies as strong, weak and weak* (see Chapter 5).

We will be also particularly focused on the theory of linear operators on Banach and Hilbert spaces, considered in Chapters 3 and 4 in the aim to consider in Chapter 6 spectral properties of compact operators, which are important in various applications, for instance in solving boundary valued problems for partial differential equations (PDEs). To be able to solve them, it is also important to be familar with the general theory of distributions, Chapter 7, and Sobolev spaces, Chapter 8. Two typical examples for solving the Poisson equation is given in Appendix F and the spectral properties of the Laplacian is considered in the class (see "TD 7").

The L^p spaces are supposed to be assumed in the course "Analyse" of Lionel Gabet [4] and we also refere to Appendix D for the additional study of these spaces. The results on L^p spaces can be assumed without proof but need to be known and will be used in numerous examples.

In the framework of this course, the generalities on the Hilbert and Sobolev spaces for Appendixes C and E can be omitted.

Chapter 1

Notations

X is a set or a space.

$\varnothing$ is the empty set.

$\mathcal{T}$ is a topology.

$(X, \mathcal{T})$ is a topological space: a set X equiped with a fixed topology $\mathcal{T}$ (*i.e.* a defined family of open sets on X, see Definition A.1.1)) is called the topological space $(X, \mathcal{T})$.

$\mathcal{O}(x)$ is the set of all open neighborhoods of x.

$\overline{X}^G$ is the closure of a set X in the space G.

$\overline{M}$ if M a subset or subspace of X, $\overline{M}$ is the closure of M in the topology of X.

$\overline{a}$ if a is a complexe number or a complexe valued matrix, then $\overline{a}$ is its complexe conjugated, *i.e* for $a = 1 + 2i$ we have $\overline{a} = 1 - 2i$.

$A \subset B$ means that a set A is a subset of a set B and A can be equal to B.

$A \subsetneq B$ means a set A is a proper subset of B ($A \neq B$).

$\operatorname{Im} A$ is the image of an operator A.

$\operatorname{Ker} A$ is the kernel of an operator A.

$\rightrightarrows$ means the uniform convergence.

$\to$ means the strong convergence.

$\rightharpoonup$ means the weak convergence.

$\overset{*}{\rightharpoonup}$ means the weak* convergence.

$\operatorname{supp} f$ means the support of a function f (see Chapter 7).

(X, d_X) is a metric space: a set X equiped with a distance d_X.

$d_X(x, y)$ denotes a distance between two elements x and y in the metric space (X, d_X).

$B_r(a)$ is an open ball of radius r centered in the point a.

$B_r^c(a)$ is a closed ball of radius r centered in the point a. In a normed space $B_r^c(a) = \overline{B_r(a)}$, but it can be not true in a metric space.

$\|\cdot\|_X$ is a norm on a vector space X.

H is usualy used to denote a Hilbert space.

$\mathcal{L}(X,Y)$ is the space of all linear continuous operators from X to Y.

X^* is the dual space to X: $X^* = \mathcal{L}(X,\mathbb{R})$.

$\langle f,g\rangle$ can be an inner product if f, $g \in X$ and X is a Pre-Hilbert space. If X is not a Pre-Hilbert space, then it means that $f \in X^*$, $g \in X$ and $\langle f,g\rangle = f(g)$ (the value of the functional f on the element g).

$\mathrm{Span}(e)$ is the set of finite linear combinations of elements of e.

I is the identity operator $I(x) = x$.

$\mathcal{D}'$ is the dual space to the space $\mathcal{D} = C_0^\infty$, named the space of distributions.

T_f is a regular distribution defined by a function $f \in L^1_{loc}$.

D^α is the derivative in the sense of distribution ($\alpha \in \mathbb{N}^n$).

Chapter 2

Reminders on the topology in metric and normed spaces

2.1 Distance or metric

2.1.1 Definitions and examples

Definition 2.1.1 *Let X be a set and $d : X \times X \to \mathbb{R}$ be a function. d is a* ***distance (metric)*** *on X if:*

1. $\forall (x, y) \in X \times X, \quad d(x, y) \geq 0$; *(positivity)*
2. $\forall (x, y) \in X \times X, \quad d(x, y) = 0$ *if and only if* $x = y$; *(identity of indiscernibles)*
3. $\forall (x, y) \in X \times X, \quad d(x, y) = d(y, x)$; *(symmetry)*
4. $\forall (x, y, z) \in X \times X \times X, \quad d(x, y) \leq d(x, z) + d(z, y)$ *(triangular inequality).*

If point 2 does not hold, d is called a ***pseudodistance****. If point 3 does not hold, d is called a* ***quasidistance****.*

The first condition of the positivity follows from the last three conditions, since

$$\begin{aligned}
\forall x, y \in X \times X \quad & d(x, y) + d(y, x) \geq d(x, x) \text{ (by triangle inequality)} \\
& \Longrightarrow d(x, y) + d(x, y) \geq d(x, x) \text{ (by symmetry)} \\
& \Longrightarrow 2d(x, y) \geq 0 \text{ (by identity of indiscernibles)} \\
& \Longrightarrow d(x, y) \geq 0.
\end{aligned}$$

Definition 2.1.2 *Set X with a given distance defined on it, i.e. the pair (X, d), is a* ***metric space*** *(if there no ambiguity in the notations, we will also use simply X instead of (X, d)). If d is a pseudodistance on X, then (X, d) is a* ***pseudometric space****. If d is a quasidistance on X, then (X, d) is a* ***quasimetric space*** *(by a quasidistance d it is possible to define a distance d' by the formula $d'(x, y) = (d(x, y) + d(y, x))/2$).*

Remark 2.1.1 *If (X, d) is a metric space and if $M \subset X$ is a subset of X, we can consider the restriction of d to M, what allows to equipe M with the structure of a metric space. In this case the metric on M is called the* ***induced*** *metric on M.*

Example 2.1.1 *Setting*

$$d(x, y) = \begin{cases} 0 & \text{if } x = y, \\ 1 & \text{if } x \neq y, \end{cases}$$

where x and y are elements of an arbitrary set X, we obtain a metric space (X, d) (a discrete space or space of isolated points). Indeed, by definition of d, the first three points of Definition 2.1.1 are satisfied. For the last point we have

1. *If $x = y$ the triangle inequality becomes:*

$$0 \leq 2 \quad \text{if } z \neq x \quad \text{or} \quad 0 \leq 0 \quad \text{if } z = x.$$

2. *If $x \neq y$ the triangle inequality becomes:*

$$1 \leq 2 \quad \text{if } z \neq x \text{ and } z \neq y \quad \text{or} \quad 1 \leq 1 \quad \text{if } z = x \text{ or } z = y.$$

Therefore the triangle inequality is satisfied.

Example 2.1.2 *Consider the space $C([0, 1])$ of all continuous functions on $[0, 1]$. Let us verify that*

$$d(f, g) = \max_{x \in [0,1]} |f(x) - g(x)|$$

is a distance on $C([0, 1])$:

1. *As the modulus is a positive function on $\mathbb{R}$:*

$$\forall x \in \mathbb{R} \quad |x| \geq 0,$$

   we have that

$$\forall (f, g) \in C([0, 1]) \times C([0, 1]), \quad d(f, g) \geq 0.$$

2. *We have for all $(f, g) \in C([0, 1]) \times C([0, 1])$*

$$d(f, g) = 0 \iff \max_{x \in [0,1]} |f(x) - g(x)| = 0 \iff \forall x \in [0, 1] \quad 0 \leq |f(x) - g(x)| \leq 0$$
$$\iff \forall x \in [0, 1] \quad |f(x) - g(x)| = 0 \iff \forall x \in [0, 1] \quad f(x) = g(x).$$

3. *For all $(f, g) \in C([0, 1]) \times C([0, 1])$ we have*

$$d(f, g) = \max_{x \in [0,1]} |f(x) - g(x)| = \max_{x \in [0,1]} |-[f(x) - g(x)]| = \max_{x \in [0,1]} |g(x) - f(x)| = d(g, f).$$

4. *For all* $(f,g,h) \in C([0,1]) \times C([0,1]) \times C([0,1])$ *we have*

$$\begin{aligned} d(f,g) &= \max_{x\in[0,1]} |f(x)-g(x)| = \max_{x\in[0,1]} |f(x)-h(x)+h(x)-g(x)| \\ &\leq \max_{x\in[0,1]} (|f(x)-h(x)|+|h(x)-g(x)|) \leq \max_{x\in[0,1]} |f(x)-h(x)| + \max_{x\in[0,1]} |h(x)-g(x)| \\ &= d(f,h)+d(h,g). \end{aligned}$$

We conclude that $(C([0,1]),d)$ *is a metric space.*

Example 2.1.3 *Consider* $X = R^n$ *with* $n \in \mathbb{N}^*$ *and* $p \in [1,\infty[$. *Let's define for* $x = (x_1,\ldots,x_n) \in X$ *and* $y = (y_1,\ldots,y_n) \in X$

$$d_p(x,y) = \left(\sum_{i=1}^n |x_i - y_i|^p\right)^{\frac{1}{p}} \quad \text{and } d_\infty(x,y) = \max_{i\in[1,\ldots,n]} |x_i - y_i|.$$

1. *We can easily see that* d_∞ *satisfies the assertions of Definition 2.1.1 and hence,* d_∞ *is a metric in* $\mathbb{R}^n$.

2. *Let us prove that* d_p *is a metric. It is obvious that points 1-3 are true for* d_p *for all* $p \in [1,\infty[$. *We need to prove the triangle inequality 4.*

   Let x,y,z *be three points in* $\mathbb{R}^n$ *and let* $A = x - z$, $B = z - y$. *Then* $x - y = A + B$ *and the triangle inequality 4 takes the form of Minkowski's inequality*

$$\left(\sum_{i=1}^n |A_i + B_i|^p\right)^{\frac{1}{p}} \leq \left(\sum_{i=1}^n |A_i|^p\right)^{\frac{1}{p}} + \left(\sum_{i=1}^n |B_i|^p\right)^{\frac{1}{p}}. \tag{2.1}$$

   The inequality is obvious for $p = 1$. *Suppose that* $p > 1$. *To prove Minkowski's inequality (2.1), we use Hölder's inequality:*

$$\sum_{i=1}^n |A_i B_i| \leq \left(\sum_{i=1}^n |A_i|^p\right)^{\frac{1}{p}} \left(\sum_{i=1}^n |B_i|^{p'}\right)^{\frac{1}{p'}}, \tag{2.2}$$

   where $\frac{1}{p} + \frac{1}{p'} = 1$, *and the following indentity for any* a *and* b *in* $\mathbb{R}$ *or* $\mathbb{C}$:

$$(|a|+|b|)^p = |a|(|a|+|b|)^{p-1} + |b|(|a|+|b|)^{p-1}.$$

   Thus, we can write

$$\sum_{i=1}^n (|A_i|+|B_i|)^p = \sum_{i=1}^n |A_i|(|A_i|+|B_i|)^{p-1} + \sum_{i=1}^n |B_i|(|A_i|+|B_i|)^{p-1}.$$

   We apply Hölder's inequality to each sum in the right-hand part of the equality, using *the fact that* $(p-1)p' = p$*:*

$$\sum_{i=1}^{n} |A_i|(|A_i|+|B_i|)^{p-1} \leq \left(\sum_{i=1}^{n} |A_i|^p\right)^{\frac{1}{p}} \left(\sum_{i=1}^{n} (|A_i|+|B_i|)^{(p-1)p'}\right)^{\frac{1}{p'}}$$
$$= \left(\sum_{i=1}^{n} |A_i|^p\right)^{\frac{1}{p}} \left(\sum_{i=1}^{n} (|A_i|+|B_i|)^{p}\right)^{\frac{1}{p'}}, \quad (2.3)$$

from where with

$$\sum_{i=1}^{n} |B_i|(|A_i|+|B_i|)^{p-1} \leq \left(\sum_{i=1}^{n} |B_i|^p\right)^{\frac{1}{p}} \left(\sum_{i=1}^{n} (|A_i|+|B_i|)^{p}\right)^{\frac{1}{p'}},$$

we find that

$$\sum_{i=1}^{n} (|A_i|+|B_i|)^{p} \leq \left(\sum_{i=1}^{n} (|A_i|+|B_i|)^{p}\right)^{\frac{1}{p'}} \left(\left[\sum_{i=1}^{n} |A_i|^p\right]^{\frac{1}{p}} + \left[\sum_{i=1}^{n} |B_i|^p\right]^{\frac{1}{p}}\right).$$

Dividing both sides of this inequality by $(\sum_{i=1}^{n}(|A_i|+|B_i|)^p)^{\frac{1}{p'}}$*, and noticing that* $1 - \frac{1}{p'} = \frac{1}{p}$ *we finally obtain that*

$$\left(\sum_{i=1}^{n} (|A_i|+|B_i|)^{p}\right)^{\frac{1}{p}} \leq \left[\sum_{i=1}^{n} |A_i|^p\right]^{\frac{1}{p}} + \left[\sum_{i=1}^{n} |B_i|^p\right]^{\frac{1}{p}}.$$

To finish the proof, we use the fact that $|A_i+B_i| \leq |A_i|+|B_i|$ *for all* i *and consequently we have (2.1).*

3. *Consider* $X = \mathbb{R}^n$ *with the metric*

$$d_2(x,y) = \left(\sum_{i=1}^{n} (x_i - y_i)^2\right)^{\frac{1}{2}}.$$

This metric is the Euclidean distance function.

4. *Consider* $X = \{$ *functions from* $\mathbb{R}$ *to* $\mathbb{R}$ *defined in* $0\}$*. For* f *and* g *in* X*, we define*

$$d(f,g) = |g(0) - f(0)|.$$

Let us prove that d *is a pseudodistance. For* $f(x) = x^2$ *and* $g(x) = x^3$ $(f \neq g)$*, by definition of* d*,* $d(f,g) = 0$*, so point 2 of Definition 2.1.1 does not hold. Points 1, 3 and 4 are true thanks to the properties of the modulus. Hence we conclude that* d *is a pseudodistance.*

Problem 2.1.1 1. *Prove the inequality*

$$ab \geq \frac{a^p}{p} + \frac{b^{p'}}{p'},$$

where $a \geq 0$*,* $b > 0$*,* $p \in]0,1[$ *and* $p' = \frac{p}{p-1} < 0$*.*

2. *Consider d_p for $p \in]0,1[$. Is it a metric on $\mathbb{R}^n$?*

Definition 2.1.3 *Let (X,d) be a metric space. Let $A \subset X$.* ***The distance between a set*** *A* ***and a point*** *$x \in X$ is defined by $d(A,x) = \inf_{a\in A} d(a,x)$.*

Remark 2.1.2 *If A is closed ($\overline{A} = A$), then*

$$d(x,A) = 0 \quad \iff \quad x \in A.$$

Problem 2.1.2 *Let there be two subsets of a metric space (X,d).Then the number*

$$z(A,B) = \inf_{(a,b)\in A\times B} d(a,b)$$

is called the distance between A and B. Show that $z(A,B) = 0$ if $A \cap B \neq \varnothing$, but not conversely. Hence, $z(A,B)$ is not a distance on $\mathcal{P}(X)$, the set of all subsets of X. ($A \notin X$ but $A \subsetneq X$, thus $A \in \mathcal{P}(X)$).

Show that for A and B two non-empty closed subsets of a metric space (X,d), the following function

$$d_{\mathrm{H}}(A,B) = \max\{\sup_{a\in A}\inf_{b\in B} d(a,b), \sup_{b\in B}\inf_{a\in A} d(a,b)\},$$

is a distance on the set of all closed subsets in X (d_{H} is a pseudodistance in X). Note that $d_{\mathrm{H}}(A,B)$ is called ***Hausdorff distance****.*

2.2 Underlying topology to a metric space. Completeness

2.2.1 Topology in a metric space

Definition 2.2.1 *Let (X,d) be a metric space. Given x in X, define the* ***open ball*** *around x with radius $r > 0$ by*

$$B_r(x) = \{y \in X |\ d(x,y) < r\}.$$

The set $B_r^c(x) = \{y \in X |\ d(x,y) \leq r\}$ is called the ***closed ball*** *around x with radius $r > 0$.*

A subset U of X is called ***open*** *in (X,d), if for all $x \in U$ there exists a radius $r > 0$, such that $B_r(x) \subset U$. The set of all open sets in (X,d) is called* ***the topology associated to the metric*** *d, denoted by $\mathcal{T}_d$.*

Attention : The same notation $\mathcal{T}_d$ in Appendix A is used for the discrete topology.

Let us also recall (see Appendix A and [4]) that

- A finite intersection and an arbitrary union of open sets in (X,d) are open in (X,d),
- A finite union and an arbitrary intersection of closed sets in (X,d) (the completions in X of open sets) are closed in (X,d).

Problem 2.2.1 *Let $Y \subset X$ be a subset of a metric space (X,d). Prove that $V \subset Y$ is open in (Y,d) (i.e., in the reduced topology on Y) if and only if there exists a open set $U \subset X$ of (X,d) such that $V = U \cap Y$.*

Example 2.2.1 *Let us consider on $\mathbb{Z}$ the distance $d(x,y) = |x-y|$ induced by the usual distance on $\mathbb{R}$. For all $x \in \mathbb{Z}$ the set $\{x\}$ is in the same time a open and a closed set in (Z,d). More generally, all subset of $\mathbb{Z}$ is open and closed in the same time in (Z,d).*

We adopt the following definition of a closure of a set in a metric space:

Definition 2.2.2 *The **closure** of a subset $M \subset X$ in a metric space (X,d), denoted by $\overline{M}$, is called the smallest closed set containing M which can be also defined by the intersection of all closed sets containing M:*

$$\overline{M} = \cap_{V \ closed, M\subset V} V.$$

Its more general version is given in Definition A.1.6 and for its properties see Problem A.1.3 and Theorem A.1.1.

Attention : Since, by definition, the open ball $B_r(x) \subset B_r^c(x)$ we have all times the inclusion $\overline{B_r(x)} \subset B_r^c(x)$, but not necessary the equality, *i.e.* the closure of the open ball can be different to the closed ball taken in the same point with the same radius.
For example, if we take $(\mathbb{Z}, d)$, defined in Example 2.2.1, then $B_1(0) = \{0\}$ and $\overline{B_1(0)} = \{0\}$ but $B_1^c(0) = \{-1, 0, 1\}$.

Definition 2.2.3 *Let (X, d_x) and (Y, d_Y) be two metric spaces. A map $f : X \to Y$ is called an **isometry or distance preserving** if for any $a, b \in X$ it holds*

$$d_Y(f(a), f(b)) = d_X(a, b).$$

An isometry is automatically injective.

A global isometry, isometric isomorphism, is a bijective isometry.

Definition 2.2.4 *Two metric spaces (X, d_x) and (Y, d_Y) are called **isometric** if there is a bijective isometry from X to Y.*

Problem 2.2.2 *Give an example of two isometric metric spaces.*

Definition 2.2.5 *Given two metric spaces (X, d_X) and (Y, d_Y), the function $f : X \to Y$ is said to be **Lipschitz continuous** if there exists a constant $K \geq 0$, called **a Lipschitz constant** such that,*

$$\forall x_1, x_2 \in X, \quad d_Y(f(x_1), f(x_2)) \leq K d_X(x_1, x_2).$$

*If $0 \leq K < 1$ the function is called a **contraction**.*

Example 2.2.2 *If $f : X \to Y$ is a bijection of two metric spaces (X, d_X) and (Y, d_Y) and, in addition, f and f^{-1} are Lipschitz continuous (f is **bi-Lipschitz**):*

$$\exists K_1 > 0 \text{ and } K_2 > 0 : \quad K_1 d_X(x_1, x_2) \leq d_Y(f(x_1), f(x_2)) \leq K_2 d_X(x_1, x_2) \quad \forall x_1, x_2 \in X,$$

*then f is **isomorphism** of (X, d_X) on (Y, d_Y). If, in addition, $K_2 = K_1 = 1$, then f is an isometry.*

Definition 2.2.6 *Two distances are **equivalent** if they define the same topology (the same open and closed sets).*

Example 2.2.3 *Let us consider two different metrics on X: d_1 and d_2. Thus we have two metric spaces (X, d_1) and (X, d_2). If the mapping $f : (X, d_1) \to (X, d_2)$, for example $f(x) = 3x$, is bi-Lipschitz, then two metrics d_1 and d_2 are equivalent.*

Example 2.2.4 ***(Product of two metric spaces)*** *If (X, d_X) and (Y, d_Y) are two metric spaces, we can define on the product space $X \times Y$ the **sum distance** d_s given by*

$$d_s((x_1, y_1), (x_2, y_2)) = d_X(x_1, x_2) + d_Y(y_1, y_2),$$

*and the **product distance** d_p given by*

$$d_p((x_1, y_1), (x_2, y_2)) = \max(d_X(x_1, x_2), d_Y(y_1, y_2)).$$

We can see that d_s and d_p verify Definition 2.1.1 and allow to define the metric structure on $X \times Y$. In addition, they verify

$$\forall z, w \in X \times Y \quad d_p(z, w) \leq d_s(z, w) \leq 2d_p(z, w),$$

Then, by Examples 2.2.2 and 2.2.3, the metrics d_s and d_p are equivalent (define the same topology). In addition, the set $O_X \times O_Y$, the product of a open set O_X in (X, d_X) and of a open set O_Y in (Y, d_Y), is an open set in $(X \times Y, d_p)$.

2.2.2 Convergence and continuity

The definition of the convergence introduced for abstract topological spaces (see Definition 0.0.1) can be now updated for a topology induced by a metric:

Definition 2.2.7 *A sequence of points (x_n) in a metric space (X, d) converges to a point $x \in X$ if every open ball $B_\epsilon(x)$ of x contains all points x_n starting from a certain index:*

$$\forall \epsilon > 0 \quad \exists N \in \mathbb{N}, \quad n \geq N \quad \Rightarrow d(x_n, x) < \epsilon.$$

The continuity of an application between two metric spaces (see also Appendix A Section A.2) is given by:

Definition 2.2.8 *Let (X, d_X) and (Y, d_Y) be two metric spaces. A function $f : X \to Y$ is called **continuous at the point** $x_0 \in X$ if for all $\epsilon > 0$ there exists $\delta > 0$ such that*

$$d_X(x_0, x) < \delta \quad \Rightarrow \quad d_Y(f(x_0), f(x)) < \epsilon.$$

*The function f is called **continuous on** X if f is continuous at every point of X.*

Now, we refer to theorems of continuous mappings in topological spaces from Section A.2).

Clearly, in a metric space

Proposition 2.2.1 *1. (x_n) converges to x in a metric space (X, d) if and only if*

$$\lim_{n \to \infty} d(x_n, x) = 0.$$

2. *a function $f : (X, d_X) \to (Y, d_Y)$ is continuous, if from $x_n \to x$ for $n \to \infty$ in X follows that $f(x_n) \to f(x)$ for $n \to \infty$ in Y.*

It is an immediate consequence of the definition of a limit that

1. No sequence can have two distinct limits;
2. If a sequence (x_n) converges to a point x, then so does every subsequence of (x_n).

Problem 2.2.3 *Let (X, d_X) and (Y, d_y) be two metric spaces and $f : X \to Y$. Prove that*

1. *If f is an isometry, then f is continuous.*
2. *If f is a Lipschitz continous function, then f is continuous.*
3. *If f is a contraction, then f is continuous.*

Let us precise the closed sets and the closure of a set in a metric space using the notion of the convergent sequence:

Proposition 2.2.2 *A subset $M \subset X$ is closed in (X, d) if and only if the limit of any sequence of elements of M, which converges in X, belongs to M:*

$$M \text{ is closed in } (X, d) \iff \forall (x_n) \subset M \;: x_n \to x \text{ in } X, \quad x \in M.$$

Example 2.2.5 *To prove that $\mathbb{Q}$ is not closed in $\mathbb{R}$ equiped with the usual distance, we can construct the decimal approximations of $\sqrt{2}$, i.e. a sequence of rational numbers converging toward $\sqrt{2}$ in $\mathbb{R}$ and $\sqrt{2} \notin \mathbb{Q}$.*

Proposition 2.2.3 *The closure in (X, d) of a set $M \subset X$ is equal to:*

1. *the set of limits of all sequences of elements of M;*
2. *the set of $x \in X$ such that for all $\epsilon > 0$ $M \cap B_\epsilon(x) \neq \varnothing$.*

2.2.3 Dense subsets of metric spaces

The notion of the density is general and can be formulated in the case of abstract topological spaces (see A.1.3). Restricted to the metric spaces, we have

Definition 2.2.9 *Let $M \subset X$ be a subset of a metric space (X, d). The set M is **dense** in X if $\overline{M} = X$. Equivalently, M is dense in (X, d) if M meets all nonempty open sets of (X, d).*

Thanks to Proposition 2.2.3, we can caracterize the dense subsets of a metric space in the following way:

Proposition 2.2.4 *A subset $M \subset X$ is dense in (X, d) if and only if*

$$\forall x \in X, \quad \forall \epsilon > 0, \quad M \cap B_\epsilon(x) \neq \varnothing.$$

This means, that M is dense in (X, d) if and only if M meets all nonempty open balls of (X, d).

We also have

Proposition 2.2.5 *A subset $M \subset X$ is dense in (X, d) if and only if for all element $x \in X$ there exists a sequence (y_n) of elements of M such that $\lim_{n\to+\infty} d(x, y_n) = 0$. This means, that M is dense in (X, d) if and only if any element $x \in X$ is the limit of a sequence of elements of M.*

Example 2.2.6 *The set of all rational numbers $\mathbb{Q}$ is dense in $\mathbb{R}$ for the usual distance.*

During our course we will see a lot of examples of dense subsets in different metric spaces.

2.2.4 Complete metric spaces

Definition 2.2.10 *In a metric space (X, d), we call **Cauchy sequence**, a sequence (u_n) such that*

$$\forall \epsilon > 0,\ \exists N > 0,\ \forall m, n > N\ \Rightarrow d(u_m, u_n) < \epsilon.$$

Remark 2.2.1 *In the equivalent way, which sometimes more useful, a Cauchy sequence (u_n) satisfies*

$$\forall \epsilon > 0,\ \exists N > 0,\ \forall m, n > N\ \Rightarrow d(u_{n+m}, u_n) < \epsilon.$$

Definition 2.2.11 *A metric space (X, d) is called **complete** if all Cauchy sequences of elements of X converge in X.*

Example 2.2.7

1. *$\mathbb{R}$ is complete. $\mathbb{Q}$ isn't.*
2. *$(C([a, b]), d_\infty)$ is complete. $(C([a, b]), d_2)$ isn't. Here d_∞ is the distance from Example 2.1.2 and $d_2(x, y) = \sqrt{\int_a^b |x(t) - y(t)|^2 dt}$ (see also Example 2.1.3).*

Proposition 2.2.6

1. *Every convergent sequence (x_n) in (X, d) is a Cauchy sequence in (X, d).*
2. *If (x_n) is the Cauchy sequence in (X, d) and if there exists a subsequence (x_{n_k}) such that $x_{n_k} \to x$ for $k \to +\infty$ in (X, d), then $x_n \to x$ for $n \to +\infty$ in (X, d).*

Proof. Ideas for the proof are given in Figure 2.1.□

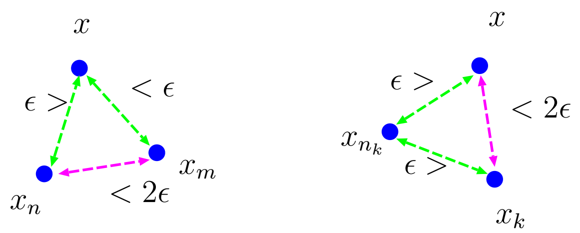

Figure 2.1 – As distances $d(x_n, x) < \epsilon$ and $d(x, x_m) < \epsilon$, then by the triangle inequality $d(x_n, x_m) < 2\epsilon$ (the left-hand schema). As $d(x_{n_k}, x_k) < \epsilon$ and $d(x_{n_k}, x) < \epsilon$, then by the triangle inequality $d(x_k, x) < 2\epsilon$ (the right-hand schema).

Lemma 2.2.1 *The product of two complete metric spaces (X, d_X) and (Y, d_Y) equiped with the sum distance d_s or the product distance d_p (see Example 2.2.4) is a complete metric space.*

Proof. Let d_p be the product distance on $X \times Y$. We notice that if $(x_n, y_n)_{n\in\mathbb{N}}$ is a Cauchy sequence in $(X \times Y, d_p)$, then the sequences $(x_n)_{n\in\mathbb{N}}$ and $(y_n)_{n\in\mathbb{N}}$ are Cauchy sequences in (X, d_X) and (Y, d_Y) respectively. Since (X, d_X) and (Y, d_Y) are complete, the sequences $(x_n)_{n\in\mathbb{N}}$ and $(y_n)_{n\in\mathbb{N}}$ converge to a limit:

$$x_n \to x \in X \quad \text{and} \quad y_n \to y \in Y.$$

Therefore, the sequence $(x_n, y_n)_{n\in\mathbb{N}}$ converges toward (x, y) in (X, Y). This finishes the proof. □

Let us prove the following very important two theorems:

Theorem 2.2.1 ***(Nested sphere theorem)*** *A metric space (X, d) is complete if and only if every nested sequence of closed balls in (X, d)*

$$B_1 = B^c_{r_1}(x_1) \supset B_2 = B^c_{r_2}(x_2) \supset \ldots \supset B_n = B^c_{r_n}(x_n) \supset \ldots,$$

such that $r_n \to 0$ as $n \to \infty$ has a nonempty intersection in X:

$$\cap_{n=1}^{\infty} B_n = \{x\} \quad \textit{for a } x \in X \quad (d(x_n, x) \to 0,\ n \to +\infty).$$

Proof $\Rightarrow$ Let (X, d) be complete and (B_n) be any nested sequence of closed balls in (X, d) such that for all $n \in \mathbb{N}$ r_n is the radius and x_n is the center of the ball B_n. Then the sequence (x_n) of centers of the balls is a Cauchy sequence, since $d(x_n, x_m,) < r_n$ for $m > n$ and $r_n \to 0$ as $n \to \infty$. Since (X, d) is complete, the Cauchy sequence (x_n) has a (unique!) limit in X, denoted by x. Then $x \in \cap_{n=1}^{\infty} B_n$. In fact, B_n contains every point of the sequence (x_n) except possibly the points $x_1, x_2, \ldots x_{n-1}$, and hence x is a limit point (see Definition A.1.6) of every ball B_n. But B_n is closed, and hence $x \in B_n$ for all n.

Let us proof the unicity of the point which belongs to the intersection of B_n. If x and y belong to $\cap_{n=1}^{\infty} B_n$, then for all n $r_n \geq d(x, y)$ and thus $d(x, y) = 0$ which implies that $x = y$.

$\Leftarrow$ Conversely, suppose every nested sequence of closed balls in (X, d) with radius converging to zero has a nonempty intersection. Let (x_n) be any Cauchy sequence in (X, d). Let us prove that then (x_n) converges in (X, d).

By definition of the Cauchy sequence, we can choose a term x_{n_1} of the sequence (x_n) such that

$$\forall n \geq n_1 \quad d(x_n, x_{n_1}) < \frac{1}{2}.$$

Let B_1 be the closed ball of the radius 1 with center x_{n_1}. Then we choose a term x_{n_2} of (x_n) such that

$$n_2 > n_1 \quad \text{and} \quad \forall n \geq n_2 \quad d(x_n, x_{n_2}) < \frac{1}{2^2}.$$

Let B_2 be the closed ball of the radius $\frac{1}{2}$ with center x_{n_2}. Continue this construction indefinitely, *i.e.*, once having chosen terms $x_{n_1}, x_{n_2}, \ldots, x_{n_k}$ $(n_1 < n_2 < \ldots < n_k)$, choose a term $x_{n_{k+1}}$ such that

$$n_{k+1} > n_k \quad \text{and} \quad \forall n \geq n_{k+1} \quad d(x_n, x_{n_{k+1}}) < \frac{1}{2^{k+1}},$$

which defines the center of the closed ball B_{k+1} of radius $\frac{1}{2^k}$, and so on. This construction gives a nested sequence (B_k) of closed balls with the raduis $r_k = \frac{1}{2^{k-1}}$ converging to zero. By hypothesis, these balls have a non-nonempty intersection, *i.e.*, there is a point $x \in \cap_{k=1}^{\infty} B_k$. This point is obviously the limit of the sequence (x_{n_k}). But if a Cauchy sequence contains a converging subsequence, then the sequence itself must converge to the same limit (see Proposition 2.2.6), *i.e.*, toward x. $\square$

Definition 2.2.12 *A subset M of a metric space (X, d) is said to be **nowhere dense** in (X, d) if it is dense in no (open) ball at all, or equivalently, if all open balls B in X contains an other non trivial ball S such that $S \cap M = \varnothing$ (check the equivalence).*

Remark 2.2.2 *If a subset M of a metric space (X, d) is not nowhere dense in (X, d), then there exists $r > 0$ and $x \in X$ such that the open ball $B_r(x) \subset M$. But if M is nowhere dense in (X, d), it means that $\overline{M}$ does not contain any non trivial open ball in (X, d).*

This concept plays an important role in Baire's Theorem:

Theorem 2.2.2 ***(Baire)*** *A complete metric space (X, d) cannot be represented as the union of a countable number of nowhere dense sets.*

Proof. Suppose to the contrary that

$$X = \cup_{n=1}^{\infty} M_n,$$

where every set M_n is nowhere dense in (X, d). Let B_0 be a closed ball of radius 1. Since M_1 is nowhere dense in B_0, being nowhere dense in X, there is a closed ball B_1 of radius less than $\frac{1}{2}$ such that

$$B_1 \subset B_0 \quad \text{and } B_1 \cap M_1 = \varnothing.$$

Since M_2 is nowhere dense in B_1, being nowhere dense in B_0, there is a closed ball B_2 of radius less than $\frac{1}{3}$ such that $B_2 \cap M_2 = \varnothing$, and so on. By this way, we get a nested sequence of closed balls (B_n) with radius converging to zero such that $B_n \cap M_n = \varnothing$ $(n = 1, 2, \ldots)$. By the nested sphere theorem, the intersection $\cap_{n=1}^{\infty} B_n$ contains a point $x \in X$. By construction, x cannot belong to any of the sets M_n, and thus $x \notin \cup_{n=1}^{\infty} M_n$. It follows that, contrary to the assumsion, $X \neq \cup_{n=1}^{\infty} M_n$. $\square$

We will see in Chapters 3 and 5 the importance of Baire's Theorem.

2.2.5 Completion of a metric space

Definition 2.2.13 *Let (X, d) be a metric space. A complete metric space (G, d) is called a **completion** of X if $X \subset G$ and its closure $\overline{X}^G = G$, **i.e.**, if X is a dense subset of G.*

Example 2.2.8 *The space of all real numbers $\mathbb{R}$ is the completion of the space of all rational numbers $\mathbb{Q}$.*

Theorem 2.2.3 *Every metric space (X, d) has a completion. This completion is unique in the following sense: if there are two completions E_1 and E_2, then they are isometric.*

For the proof see [7] (can be omitted).

2.2.6 Separable spaces

Definition 2.2.14 *A metric space is said to be separable if it has a countable dense subset.*

Example 2.2.9
- $\mathbb{R}^n$ *for* $n \in \mathbb{N}^*$ *contains the countable dense set of all points* $x = (x_1, \dots, x_n)$ *with rational coordinates.*
- *Let us consider the space of sequences* ℓ^2 *such that for all* $x = (x_1, \dots) \in \ell^2$ *and* $y = (y_1, \dots) \in \ell^2$ *the distance* $d_2(x, y)^2 = \sum_{i \geq 1} |x_i - y_i|^2 < \infty$*. The space* ℓ^2 *contains the countable dense set of all points* $x = (x_1, \dots)$ *with only finit number of nonzero coordiantes, which are rational.*
- *The space* $C([a, b])$ *of all continuous functions on* $[a, b]$ *with a metric*

$$d(g, f) = \max_{x \in [a,b]} |g(x) - f(x)|$$

  contains the countable dense set of polynomials with rational coefficients.
- *Let*

$$\ell^\infty = \{\text{bounded sequences } x = (x_1, \dots) |\ d_\infty(x, y) = \sup_k |x_k - y_k|\}.$$

  ℓ^∞ *is an example of a nonseparable space. Let us show that* ℓ^∞ *contains an uncountable dense set.*

  In fact, consider the set F *of all sequences consisting exclusively of zeros and ones. In this case,* F *has the power of the continuum, since there is a bijection between* F *and the set of all subsets of the set of natural numbers* $\mathbb{N}$*: for all* $A \subsetneq \mathbb{N}$ *we associate* $x_A = x_n$ *such that*

$$x_n = \begin{cases} 1, & \text{if } n \in A, \\ 0, & \text{if } n \notin A. \end{cases}$$

  We note that the distance between any two points of F *equals* 1*:*

$$d_\infty(x_A, x_B) = 1 \text{ if } A \neq B.$$

  Suppose we surround each point of F *by an open sphere of radius* $\frac{1}{2}$*, thereby obtaining an uncountably infinite family of pairwise disjoint spheres. Then if some set* M *is dense in* ℓ^∞*, there must be at least one point of* M *in each of the spheres. It follows that* M *cannot be countable and hence that* ℓ^∞ *cannot be separable.*

2.3 Compactness in metric spaces

Since metric spaces are topological spaces, all results and definitions of the compactness in the topological spaces (see Definition B.1.2) hold for metric spaces as well. Let us just detail the specific properties of compactness in metric spaces.

Since all open sets in a metric space are defined by the open balls (by their (arbitrary) union and by a finite intersection), we can update the notion of the open cover (see Definition B.1.1) in the framework of the metric spaces:

Definition 2.3.1 *Let (X, d) be a metric space containing a subset M and $\epsilon > 0$. A set $A \subset X$ is said to be an **ϵ-net** for the set M if,*

$$\forall x \in M \quad \text{there is at least one point } \ a \in A \ \text{ such that } \ d(x, a) \leq \epsilon.$$

In particular, M can be equal to X.

It is possible that $A \cap M = \varnothing$, but if A is an ϵ-net for M, it is possible to construct 2ϵ-set $B \subset M$.

Example 2.3.1 *The set of all points with integer coordinates is a $\frac{1}{\sqrt{2}}$-net of $\mathbb{R}^2$.*

Definition 2.3.2 *In a metric space (X, d) a subset M is called **totally bounded** if for all $\epsilon > 0$ there exists a finite ϵ-net of M.*

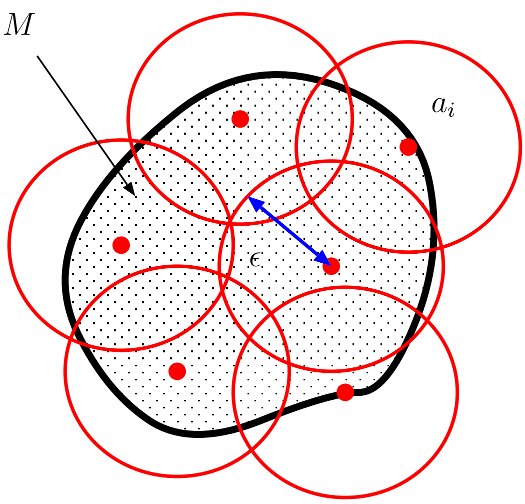

Figure 2.2 – An example of a finite ϵ-net for a set $M \subsetneq \mathbb{R}^2$.

Example 2.3.2 *Let us illustrate the existence of a finite ϵ-net using Fig. 2.2 with an example of a (compact) set M in $\mathbb{R}^2$. We see that $A_\epsilon = \{a_1, \ldots, a_6\}$ is the ϵ-net of M and in particular that $M \subset \cup_{i=1}^{6} \overline{B}_\epsilon(a_i)$.*

Remark 2.3.1 *If a metric space (X, d) is totally bounded, then (X, d) is separable.*

Indeed, for all $n \in \mathbb{N}$ we have a finite $\frac{1}{n}$-net, denoted by A_n. Thus $\cup_{n \in \mathbb{N}} A_n$ is a countable dense set in X.

We notice that:

1. If a set M is totally bounded, then its closure $\overline{M}$ is also totally bounded.
2. Every subset of a totally bounded set is itself totally bounded.

Every totally bounded set is bounded, being the union of a finite number of bounded sets. The converse is not true, as shown in the following example:

Example 2.3.3 *The unit sphere S in the space ℓ^2*

$$S = \{x = (x_1, \dots, x_n, \dots) \in \ell^2 | d_2(x, 0) = \sum_{n=1}^{\infty} x_n^2 = 1\}$$

is bounded but not totally bounded. In fact, let us consider in S the points

$$\begin{aligned} &e_1 = (1, 0, 0, \dots, 0, 0, \dots),\\ &e_2 = (0, 1, 0, \dots, 0, 0, \dots),\\ &\dots\dots\dots\dots\dots\dots\dots\dots\\ &e_n = (0, 0, 0, \dots, 1, 0, \dots),\\ &\dots\dots\dots\dots\dots\dots\dots\dots, \end{aligned}$$

where the nth coordinate of e_n is one and the others are all zero. The distance between any two points e_n and e_m $(n \neq m)$ is $\sqrt{2}$. Hence S cannot have a finite ϵ-net with $\epsilon < \frac{\sqrt{2}}{2}$.

Example 2.3.4 *In the Euclidean space $\mathbb{R}^n$, total boundedness is equivalent to boundedness. In fact, if M is bounded in $\mathbb{R}^n$, then M is contained in some sufficiently large cube Q. Partitioning Q into smaller cubes of side ϵ, we find that the vertices of the little cubes form a finite $(\frac{\sqrt{n}\epsilon}{2})$-net for Q and hence for any set contained in Q.*

Example 2.3.5 *Let P be the set of points $x = (x_1, x_2, \dots, x_n, \dots)$ in ℓ^2 satisfying the inequalities*

$$|x_1| \le 1, \quad |x_2| \le \frac{1}{2}, \quad \dots, \quad |x_n| \le \frac{1}{2^{n-1}}, \quad \dots$$

The set P, called the Hilbert cube, gives an example of an infinite dimensional totally bounded set. Let us prove it. Given any $\epsilon > 0$, we choose n such that

$$\frac{1}{2^{n-1}} < \frac{\epsilon}{2}.$$

We associate each point $x = (x_1, \dots, x_n, \dots)$ in P with the point

$$x^* = (x_1, x_2, \dots, x_n, 0, 0, \dots) \in P. \tag{2.4}$$

Then

$$d(x, x^*) = \sqrt{\sum_{i=n+1}^{\infty} x_i^2} \le \sqrt{\sum_{i=n}^{\infty} \frac{1}{4^i}} < \frac{1}{2^{n-1}} < \frac{\epsilon}{2}.$$

The set P^* of all points in P of the form (2.4) is a bounded set in the n-dimentional space, and, consequently, P^* is totally bounded. Let A be a finite $(\epsilon/2)$-net in P^*. Then A is a finite ϵ-net for the whole set P.

We give now the main theorems on the compactness in the metric spaces. The proofs can be found, for example, in [7].

In a metric space, *compact* (see Definition B.1.2) and *sequentially compact* (see Definition B.1.5) are equivalent:

Theorem 2.3.1 *Let (X, d) be a metric space. Then a subset M is compact if and only if it is sequentially compact.*

In particular for the finite dimensional spaces there is the Theorem of Heine-Borel

Theorem 2.3.2 ***(Heine-Borel)*** *In $\mathbb{R}^n$ (or $\mathbb{C}^n$) a set is compact if and only if it is bounded and closed.*

What can we say about the infinite dimensional case? To understand it, let us prove

Theorem 2.3.3 *Let (X, d) be a compact metric space. Then X is totally bounded.*

Proof. Let X be a compact. By Theorem 2.3.1, X is sequentially compact. Let us suppose that X is not totally bounded:

$$\exists \epsilon_0 > 0 : \nexists \text{ a finite } \epsilon_0 - \text{net } A_{\epsilon_0} \text{ of } X.$$

Let $a_1 \in X$. Thus, there exists (at least one point) $a_2 \in X$ such that $d(a_1, a_2) > \epsilon_0$ (otherwise $a_2 \in A_{\epsilon_0}$). Thus, there exists $a_3 \in X$ such that

$$d(a_1, a_3) > \epsilon_0 \quad \text{and} \quad d(a_2, a_2) > \epsilon_0 \quad (\text{otherwise, } a_1, a_2 \in A_{\epsilon_0}).$$

Given $a_1, \ldots, a_k$, we chose $a_{k+1} \in X$ such that

$$\forall i = 1, \ldots, k, \quad d(a_i, a_{k+1}) > \epsilon_0.$$

This construction gives an infinite sequence of distinct points $a_1, a_2, \ldots$ in X with no limit points, since

$$d(a_i, a_j) > \epsilon_0 \quad \text{if } i \neq j.$$

But then X cannot be sequentially compact. □

Example 2.3.6 *The total boundedness is a necessary condition for a metric space to be compact (see Theorem 2.3.3), but not sufficient. For example, let*

$$X = \{q \in \mathbb{Q} | q \in [0, 1]\}, \quad \forall q_1, q_2 \in X \quad d(q_1, q_2) = |q_1 - q_2|.$$

The metric space (X, d) is totally bounded, but not compact. In fact, the sequence of decimal approximations of the irrational number $\sqrt{2} - 1$

$$0, \quad 0.4, \quad 0.41, \quad 0.414, \quad 0.4142, \quad \ldots$$

is a sequence in (X, d), which does not converge in X (has no limit point in X).

Necessary and sufficient conditions for compactness of a metric space are given by

Theorem 2.3.4 *A metric space (X, d) is compact if and only if it is totally bounded and complete.*

Proof.

$\underline{\Rightarrow}$

Let (X, d) be compact. Then, by Theorem 2.3.3, (X, d) is totally bounded. Let us prove that (X, d) is complete. As (X, d) is compact, thus all sequences in (X, d) have a convergent subsequence. Let (x_n) be a Cauchy sequence in X. By the compactness of X, there exists a subsequence (x_{n_k}) which converges to a $x \in X$. Thanks to Proposition 2.2.6 point 2, it implies that $x_n \to x$ for $n \to \infty$ in X. Hence, all Cauchy sequences in X converge in X, and thus (X, d) is complete.

$\underline{\Leftarrow}$

Let (X, d) be totally bounded and complete. We want to prove that (X, d) is compact. Let (x_n) be any infinite sequence of distinct points in X.

Let us consider the 1-net of X (in X!) $A_1 = \{y_1, \ldots, y_{n_1}\}$. By definition of a total bounded set,

$$n_1 \text{ is finite} \quad \text{and} \quad X = \cup_{i=1}^{n_1} \overline{B}_1(y_i).$$

Thus, there exists i_0 $(1 \leq i_0 \leq n_1)$ such that the ball $\overline{B}_1(y_{i_0})$, denoted by S_1, contains an infinite subsequence $(x_k^{(1)})$ of the sequence (x_n).

As S_1 is a subset of a totally bounded set X, then S_1 is itself totally bounded.

Let $A_{\frac{1}{2}} = \{z_1, \ldots, z_{n_2}\}$ be the $\frac{1}{2}$-net of S_1. Then

$$n_2 \text{ is finite} \quad \text{and} \quad S_1 \subset \cup_{i=1}^{n_2} \overline{B}_{\frac{1}{2}}(z_i).$$

As previously, it follows, that there exists a closed ball $\overline{B}_{\frac{1}{2}}(z_{j_0})$ $(1 \leq j_0 \leq n_2)$, denoted by S_2, which contains an infinite subsequence $(x_k^{(2)})$ of the sequence $(x_k^{(1)})$.

Let $A_{\frac{1}{4}}$ be the $\frac{1}{4}$-net of S_2. Then there exists a closed ball S_3 of radius $\frac{1}{4}$ containing an infinite subsequence $(x_k^{(3)})$ of the sequence $(x_k^{(2)})$.

Continue this construction indefinitely, we find a sequence of closed balls S_n of radius $\frac{1}{2^{n-1}}$ containing an infinite number of terms of the sequence (x_n).

Let V_n be the closed ball with the same center as S_n but with a radius r_n twice as large (i.e., equal to $\frac{2}{2^{n-1}}$). Then clearly

$$V_1 \supsetneq V_2 \supsetneq V_3 \supsetneq \ldots \supsetneq V_n \supsetneq \ldots$$

and moreover $r_n = \frac{2}{2^{n-1}} \to 0$ as $n \to \infty$. Since X is complete, it follows (by the nested sphere theorem) that

$$\cap_{n=1}^{\infty} V_n \neq \varnothing$$

and thus there exists $x \in X$ such that

$$\cap_{n=1}^{\infty} V_n = \{x\}.$$

Consequently, x is a limit point of the original sequence (x_n), since every neighborhood of x contains some ball S_j and hence some infinite subsequence $(x_k^{(j)})$. Therefore every infinite sequence (x_n) of distinct points of X has a limit point in X. It follows that X is countably compact and hence sequentially compact and hence compact, by Theorem 2.3.1. □

For the relatively compactness (see Definition B.3.1) we have the following result:

Theorem 2.3.5 *A subset M of a complete metric space (X, d) is relatively compact if and only if it is totally bounded.*

Proof.

We notice that $\overline{M}$ is a closed subset of a complete metric space (X, d). Thus, $(\overline{M}, d)$ is a complete metric space. Consequently, according to Theorem 2.3.4, $\overline{M}$ is compact iff $\overline{M}$ is totally bounded. Moreover, $\overline{M}$ is totally bounded iff M is totally bounded. $\square$

Example 2.3.7 *Any bounded subset of $\mathbb{R}^n$ is totally bounded and hence relatively compact (this is a version of the Bolzano-Weierstrass theorem).*

Example 2.3.8 ***(Relatively compact sets in $C([a, b])$).*** *If $(X, d) = (C([a, b]), d_\infty)$ there is a criterion for relative compactness, called Arzela's theorem.*

Theorem 2.3.6 (Arzela's theorem) *A set Φ of continuous functions defined on a closed interval $[a, b]$ is relatively compact in $(C([a, b]), d_\infty)$ if and only if Φ is uniformly bounded:*

$$\exists C \geq 0 \text{ such that } \forall x \in [a, b] \text{ and } \forall \phi \in \Phi \quad |\phi(x)| < C,$$

and uniformly equicontinuous:

$$\forall \epsilon > 0\ \exists \delta > 0 \text{ such that } \begin{array}{ll} 1) & \forall \phi \in \Phi \\ 2) & \forall x_1, x_2 \in [a, b] \text{ such that } d(x_1, x_2) < \delta \end{array} \qquad |\phi(x_1) - \phi(x_2)| < \epsilon.$$

Example 2.3.9 *1. Arzela's theorem says that Φ can contain a subsequence, uniformly converging to a continuous function, but not necessary from Φ. For example, let*

$$\Phi = \{\phi_n(x) = \frac{1}{n} |\ x \in [0, 1],\ n \in \mathbb{N}\}.$$

Then $\phi_n(x) \rightrightarrows 0$, but $0 \notin \Phi$.

2. Let

$$\Phi = \{\cos(nx), \quad x \in [0, 1], \quad n \in \mathbb{N}\}.$$

Φ is uniformly bounded, since for all x and n $|\cos(nx)| \leq 1$, but not uniformly equicontinuous.

Let $\epsilon = 1$. Then

$$\forall \delta > 0 \quad \exists n \in \mathbb{N} \text{ and } \exists x_1, x_2 \in [0, 1] : \quad |\cos(nx_1) - \cos(nx_2)| = 2 > \epsilon.$$

3. Let $L > 0$ and $\alpha \in]0, 1]$ be fixed. The set

$$\Phi = \{\phi : [0, 1] \to \mathbb{R} |\ |\phi(x_1) - \phi(x_2)| \leq L|x_1 - x_2|^\alpha\}$$

contains all constant functions, thus Φ is not uniformly bounded. But Φ is uniformly equicontinuous: $\forall \epsilon > 0\ \exists \delta > 0$, which can be found from $L\delta^\alpha < \epsilon$, such that

$$\begin{array}{ll} 1) & \forall \phi \in \Phi \\ 2) & \forall x_1, x_2 \in [a, b] \text{ such that } d(x_1, x_2) < \delta \end{array} \qquad |\phi(x_1) - \phi(x_2)| < \epsilon.$$

Proof of Arzela's theorem.

$\Rightarrow$

Let Φ be a relatively compact set in $C([a,b])$. We recall that $(C([a,b]), d_\infty)$ is a complete metric space. By Theorem 2.3.5,

$$\forall \epsilon > 0 \quad \exists \ \text{ a finite } \frac{\epsilon}{3} - \text{net of } \Phi,$$

denoted by $A_\epsilon = \{\phi_1, \ldots, \phi_k\} \subsetneq C([a,b])$. Therefore, for all $i = 1, \ldots, k$ ϕ_i is bounded on $[a,b]$:

$$|\phi_i(x)| \leq K_i.$$

We denote

$$K = \max_{1 \leq i \leq k} K_i + \frac{\epsilon}{3}.$$

By definition of a $\frac{\epsilon}{3}$-net, we have

$$\forall \phi \in \Phi \quad \exists \phi_i : d_\infty(\phi, \phi_i) = \max_{x \in [a,b]} |\phi(x) - \phi_i(x)| \leq \frac{\epsilon}{3}. \tag{2.5}$$

Then

$$|\phi(x)| = |\phi(x) - \phi_i(x) + \phi_i(x)| \leq |\phi(x) - \phi_i(x)| + |\phi_i(x)| \leq |\phi_i(x)| + \frac{\epsilon}{3} \leq K_i + \frac{\epsilon}{3} \leq K.$$

The inequality $|\phi(x)| \leq K$ holds for all $\phi \in \Phi$ and all $x \in [a,b]$. Thus Φ is uniformly bounded.

Let us prove that Φ is uniformly equicontinuous.

As for all $i = 1, \ldots, k$ ϕ_i are continuous on a compact $[a,b] \subsetneq \mathbb{R}$ functions, it follows that ϕ_i are uniformly continuous on $[a,b]$:

$$\forall \epsilon > 0 \quad \exists \delta_i(\epsilon) > 0 : \quad |x_1 - x_2| < \delta_i \quad \Rightarrow \quad |\phi_i(x_1) - \phi_i(x_2)| < \frac{\epsilon}{3}.$$

We take $\delta = \min_{1 \leq i \leq k} \delta_i$ and for all i $(1 \leq i \leq k)$ we have

$$\forall \epsilon > 0 \quad \exists \delta(\epsilon) > 0 : \quad |x_1 - x_2| < \delta \quad \Rightarrow \quad |\phi_i(x_1) - \phi_i(x_2)| < \frac{\epsilon}{3}.$$

Using (2.5), we obtain for $|x_1 - x_2| < \delta$

$$\begin{aligned}
&|\phi(x_1) - \phi(x_2)| = |\phi(x_1) - \phi_i(x_1) + \phi_i(x_1) - \phi_i(x_2) + \phi_i(x_2) - \phi(x_2)| \\
&\leq |\phi(x_1) - \phi_i(x_1)| + |\phi_i(x_1) - \phi_i(x_2)| + |\phi_i(x_2) - \phi(x_2)| < \frac{\epsilon}{3} + \frac{\epsilon}{3} + \frac{\epsilon}{3} = \epsilon,
\end{aligned}$$

from where it follows that Φ is uniformly equicontinuous.

$\Leftarrow$

Let Φ be uniformly bounded and uniformly equicontinuous subset of $C([a,b])$. Let us prove that Φ is relatively compact in $C([a,b])$. By Theorem 2.3.5 we need to prove that Φ is totally bounded: for all $\epsilon > 0$ there exists a finite ϵ-net.

From the uniform boundedness of Φ we have

$$\forall \phi \in \Phi, \quad \forall x \in [a,b] \quad |\phi(x)| \leq K,$$

and from the uniform equicontinuity of Φ we have

$$\forall \epsilon > 0 \quad \exists \delta > 0 : \quad \forall \phi \in \Phi, |x_1 - x_2| < \delta \quad \Rightarrow \quad |\phi(x_1) - \phi(x_2)| \leq \frac{\epsilon}{5}.$$

We divide the interval $[a, b]$ along the x-axis into subintervals of length less than δ, by introducing points of subdivision $x_0, x_1, x_2, \dots, x_n$ such that

$$a = x_0 < x_1 < x_2 < \dots < x_n = b,$$

and then draw a vertical line through each of these points. Similarly, we divide the interval $[-K, K]$ along the y-axis into subintervals of length less than $\frac{\epsilon}{5}$, by introducing points of subdivision $y_0, y_1, y_2, \dots, y_m$ such that

$$-K = y_0 < y_1 < y_2 < \dots < y_m = K,$$

and then draw a horizontal line through each of these points. In this way, the rectangle $[a, b] \times [-K, K]$ is divided into nm cells of horizontal side length less than δ and vertical side length less than $\frac{\epsilon}{5}$.

We now associate with each function $\phi \in \Phi$ a polygonal line $y = \psi(x)$ which has vertices at points of the form (x_k, y_j) and differs from the function ϕ by less than $\frac{\epsilon}{5}$ at every point x_k (see Fig. 2.3):

$$\forall k \quad |\phi(x_k) - \psi(x_k)| < \frac{\epsilon}{5}.$$

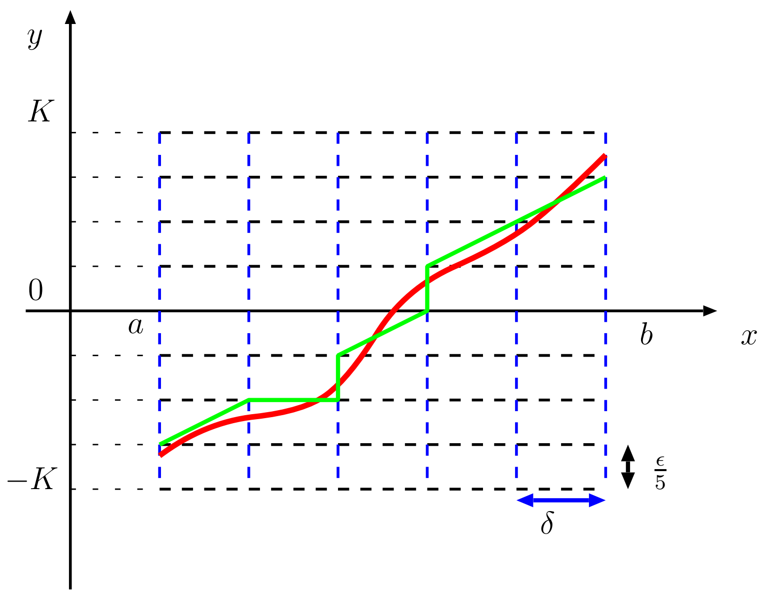

Figure 2.3 – An example of a $\phi \in \Phi$ in red and of its approximation ψ in green.

Since by our construction

$$|\phi(x_k) - \psi(x_k)| < \frac{\epsilon}{5}, \quad |\phi(x_{k+1}) - \psi(x_{k+1})| < \frac{\epsilon}{5}, \quad |\phi(x_k) - \phi(x_{k+1})| < \frac{\epsilon}{5},$$

we find that

$$|\psi(x_k) - \psi(x_{k+1})| < \frac{3\epsilon}{5}.$$

Since $\psi(x)$ is linear on $[x_k, x_{k+1}]$, then

$$\forall x \in [x_k, x_{k+1}] \quad |\psi(x_k) - \psi(x)| < \frac{3\epsilon}{5}.$$

Let $x \in [a, b]$ and k be such that $|x_k - x| = \min_{0 \le i \le n} |x_i - x|$, *i.e.*, let x_k be the nearest point on the left to x for $0 \le k \le n$. Then

$$|\phi(x) - \psi(x)| \le |\phi(x) - \phi(x_k)| + |\phi(x_k) - \psi(x_k)| + |\psi(x_k - \psi(x)| \le \frac{\epsilon}{5} + \frac{\epsilon}{5} + \frac{3\epsilon}{5} = \epsilon.$$

Consequently, the set of polygonal lines $(\psi(x))$ forms an ϵ-net for Φ. But there is exists only a finite number of such lines. Therefore Φ is totally bounded. □

Analysis of the proof of Arzela's Theorem shows that it is not really important that all functions are defined on an interval $[a, b]$, but it is crucial that the interval $[a, b]$ is compact in $\mathbb{R}$. Thus, it holds the following Theorem (the proof is almost the same as for Arzela's Theorem):

Theorem 2.3.7 ***(Ascoli-Arzela Theorem)*** *A set Φ of continuous complex-valued functions functions defined on a compact metric space K is relatively compact in $(C(K), d_\infty)$ if and only if Φ is uniformly bounded and uniformly equicontinuous.*

2.4 Normed vector spaces

2.4.1 Definition and examples

Definition 2.4.1 *Let X be a vector space and $N : X \to \mathbb{R}$ a function. N is called a* ***norm*** *on X if*

1. $\forall x \in X,\ N(x) = 0 \quad \iff \quad x = 0$
2. $\forall (x, \lambda) \in X \times \mathbb{C}, \quad N(\lambda x) = |\lambda| N(x)$
3. $\forall (x, y) \in X \times X, \quad N(x + y) \le N(x) + N(y)$ *(triangular inequality)*

If point 1 does not hold, (nevertheless, assertion 2 implies $N(0) = 0$) N is called a ***semi-norm****.*

Remark 2.4.1

1. *Assertions 1 and 3 imply $N(x) \ge 0$ for all x in X (take $y = -x$ in assertion 3).*
2. *By the mathematical induction from assertion 3, it follows that*

$$N(x_1 + \ldots + x_n) \le N(x_1) + \ldots + N(x_n).$$

From a geometrical point of view, the last inequality can be interpeted as the length of a segment between two points x_1 and $x_1 + x_2 + \ldots + x_n$ is smaller than the length of the broken line based on the points $y_i = x_1 + \ldots + x_i$ for $i = 1, \ldots, n$ (see Fig. 2.4).

Definition 2.4.2 *Let N be a norm on a vector space X. The couple (X, N) is called a* ***normed vector space****. For elements x in X $N(x)$ is usually noted $\|x\|_X$.*

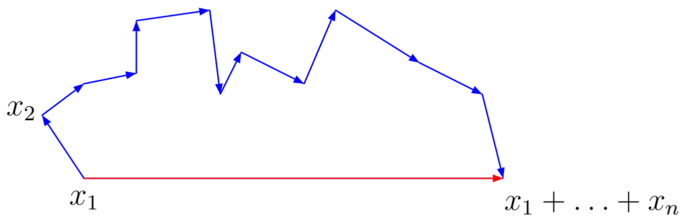

Figure 2.4 – Generalized triangular inequality

Example 2.4.1 *1. The modulus of a linear function $f : X \to \mathbb{R}$ is a seminorm in X:*

$$\text{for } x \in X \quad N(x) = |f(x)| \quad \text{satisfies 2) and 3), but not 1) in Definition 2.4.1.}$$

In particular, if $X = \mathbb{R}$ (or $\mathbb{C}$), the seminorm $N(x) = |x|$ is a norm in X (here it is important that the dimension of X equals to 1). Thus, the notion of a norm can be considered as a generalization of the notion of modulus.

2. In $\mathbb{R}^n$ $\|\cdot\|_{\ell^p}$ and $\|\cdot\|_{\ell^\infty}$ for $1 \le p < \infty$ also define the norms. For $p = 2$ we find the Euclidean norm

$$\|x\| = \sqrt{\sum_{k=1}^{n} |x_k|^2}.$$

3. Consider $X = \ell^\infty$ introduced in Example 2.2.9. For $u \in X$, define $N(u) = \sup_{i\in\mathbb{N}} |u_i|$. Then $N(u)$ is a norm on ℓ^∞.

4. Let us consider the space of sequences $\ell^p = \{x = (x_1, \ldots) |\ \sum_{i\ge1} |x_i|^p < \infty\}$ for $1 \le p < \infty$. Then $\|x\|_{\ell^p} = \left(\sum_{i\ge1} |x_i|^p\right)^{\frac{1}{p}}$ is a norm on ℓ^p.

5. For the space $C([a, b])$ of continuous functions on $[a, b]$ we can consider

$$N_\infty(f) = \|f\|_\infty = \max_{a\le x\le b} |f(x)|, \qquad N_p(f) = \|f\|_p = \left(\int_a^b |f(x)|^p \mathrm{d}x\right)^{\frac{1}{p}},$$

where $1 \le p < \infty$. Then the spaces $(C([a, b]), N_\infty)$ and $(C([a, b]), N_p)$ are normed spaces.

6. Let us consider the space Lip of all Lipschitz continuous functions f from a normed space X to $\mathbb{R}$. We can define the norm by

$$\|f\|_{Lip} = |f(x_0)| + \sup_{x\ne y} \frac{|f(x) - f(y)|}{\|x - y\|}.$$

We can directly verify that

Proposition 2.4.1 *Let $(X, \|\cdot\|)$ be a linear vector normed space. Then the ball*

$$B^c_r(0) = \{x \in X | \quad \|x\| \le r\}$$

is convex and symmetric set in X.

Problem 2.4.1 *Prove that a function $N(x) : X \to \mathbb{R}$ is a norm if it satisfies:*

1. *$N(x)$ is a positive homogeneous function :*
$$N(\lambda x) = |\lambda| N(x) \quad \forall \lambda \in \mathbb{R}.$$
2. *The set $\{x \in X | N(x) \leq 1\}$ is convex.*
3. *$N(x) > 0$ if $x \neq 0$.*

(If we don't have 3), then N is a semi-norm.)

Remark 2.4.2 *Using Problem 2.4.1, it is easy to verify that $\|\cdot\|_p$ is a norm on $C([a,b])$: $|x|^p$ for $p > 1$ is a positive homogeneous function, which is convex and strictly positive if $x \neq 0$, what implies that $\|\cdot\|_p$ is a norm.*

We can see that in the same linear vector space X we can define different norms. In the next Section we will answere the question: if there are two different norms $\|\cdot\|_1$ and $\|\cdot\|_2$ on X, what about the properties of $(X, \|\cdot\|_1)$ and $(X, \|\cdot\|_2)$?

2.4.2 Converging sequences and continuous applications

Let us precise the notions of the convergence and the continuity in the framework of normed spaces:

Definition 2.4.3 1. *$(X, \|\cdot\|)$ be a normed linear vector space and (x_n) be a sequence of elements of X. We say that (x_n) converges to $x \in X$ if*
$$d(x_n, x) = \|x_n - x\| \to 0 \quad \text{for } n \to +\infty.$$
2. *Let f be a mapping of a normed space $(X, \|\cdot\|)$ to a normed space $(Y, \|\cdot\|)$. The mapping f is continuous if from the convergence of x_n to x in X for $n \to +\infty$ follows that the sequence $(f(x_n))_{n\in\mathbb{N}}$ converge to $f(x)$ in Y:*
$$\|x_n - x\|_X \to 0 \text{ for } n \to +\infty \quad \Rightarrow \quad \|f(x_n) - f(x)\|_Y \to 0 \text{ for } n \to +\infty \tag{2.6}$$

Example 2.4.2 *Let $(X, \|\cdot\|_X)$ be a normed linear vector space. The norm $\|\cdot\|_X$ is a continuous function on X:*
$$\text{if} \quad \|x_n - x\|_X \to 0 \text{ for } n \to +\infty \quad \Rightarrow \quad \|x_n\|_X \to \|x\|_X \text{ for } n \to +\infty.$$
To prove it is sufficient to notice that for all x and y in X it holds the following inequality:
$$|\, \|x\|_X - \|y\|_X| \leq \|x - y\|_X.$$
Therefore, we have that
$$|\, \|x\|_X - \|x_n\|_X| \leq \|x - x_n\|_X$$
and, since $\|x - x_n\|_X \to 0$ for $n \to +\infty$, it implies that $\|x\|_X - \|x_n\|_X \to 0$ for $n \to +\infty$.

Remark 2.4.3 *Any linear continuous mapping $f : (X, \|\cdot\|_X) \to (Y, \|\cdot\|_Y)$ is Lipschitz continuous:*
$$\exists K \geq 0 \quad \|f(x) - f(y)\|_Y \leq K\|x - y\|_X \quad \forall x, y \in X.$$
If in addition, $K < 1$, f is a contraction.

2.5 Underlying metric and topology to a normed space

We can recognize in Example 2.4.1 the formulas of the metrics given in Example 2.1.3 written for a distance between x and 0.

2.5.1 Metric and a norm

Proposition 2.5.1 *Let $(X, \|\cdot\|)$ be a normed space. Then the function*

$$\rho(x,y) = \|x-y\| \quad \text{for } (x,y) \in X \times X \tag{2.7}$$

is a metric in X. Moreover, the metric ρ is absolutely homogenous, i.e.

$$\rho(\lambda x, \lambda y) = |\lambda| \rho(x,y) \quad \text{for } (x,y,\lambda) \in X \times X \times \mathbb{C}, \tag{2.8}$$

and invariant by translation (see Fig. 2.5)

$$\rho(x+z, y+z) = \rho(x,y). \tag{2.9}$$

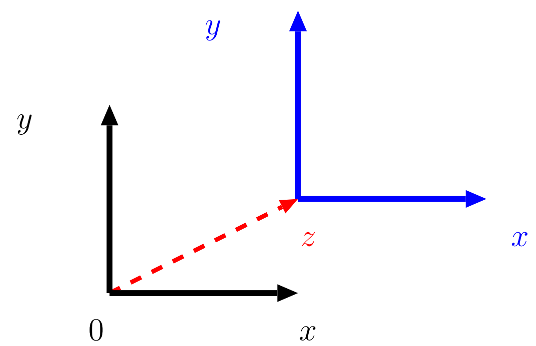

Figure 2.5 – Translation of two vectors x and y by a vector z.

Proposition 2.5.2 *A metric d in a space X is defined by a norm in X by formula (2.7) if and only if d satisfies (2.8) and (2.9).*

Problem 2.5.1 *If d satisfies (2.8) and (2.9), prove that $N(x) = d(x,0)$ is a norm.*

This time the closure of an open ball is equal to the coresponding closed ball:

Proposition 2.5.3 *Let $(X, \|\cdot\|)$ be a normed space. Then*

$$\overline{B}_r(x) = B_r^c(x).$$

Proof. See [5] p.19 (can be omitted).

2.5.2 Equivalent norms

Definition 2.5.1 *Let $\|\cdot\|_1$ and $\|\cdot\|_2$ be two norms in a vector space X. The norm $\|\cdot\|_2$ is called **stronger** than the norm $\|\cdot\|_1$ if there exists a positive constant $C > 0$ such that*

$$\forall x \in X \quad \|x\|_1 \leq C\|x\|_2.$$

Example 2.5.1 *Let us consider the space $C([a,b])$ of continuous functions on $[a,b]$ with two following norms:*

$$\|f\|_\infty = \max_{a\leq x\leq b} |f(x)|, \qquad \|f\|_2 = \left(\int_a^b |f(x)|^2 \mathrm{d}x\right)^{\frac{1}{2}}$$

We find that $\|\cdot\|_\infty$ is stronger than $\|\cdot\|_2$:

$$\|f\|_2 \leq \left(\int_a^b 1 \mathrm{d}x\right)^{\frac{1}{2}} \max_{a\leq x\leq b} |f(x)| = \sqrt{b-a}\|f\|_\infty.$$

A stronger norm will provide a stronger topology, *i.e.* more open/closed sets. If there are two norms $\|\cdot\|_1$ and $\|\cdot\|_2$ on X and $\|\cdot\|_2$ is stronger than $\|\cdot\|_1$, it means that

1. the convergence in $(X, \|\cdot\|_2)$ implies the convergence in $(X, \|\cdot\|_1)$ (but not converse!)
2. if a set F is dense in $(X, \|\cdot\|_2)$ then F is also dense in $(X, \|\cdot\|_1)$ (but not converse!)

Definition 2.5.2 *Let $\|\cdot\|_1$ and $\|\cdot\|_2$ be two norms in a vector space X. The norms are called **equivalent** if there exist two positive constants $c > 0$ and $C > 0$ such that*

$$\forall x \in X \quad c\|x\|_2 \leq \|x\|_1 \leq C\|x\|_2.$$

We notice that norms are equivalent iff associated balls can be included in one another (after a possible homothetic transformation).

Theorem 2.5.1 *If X is a finite-dimensional vector space $\dim(X) < \infty$, then all norms in X are equivalent.*

Proof. Since X is a finite-dimensional vector space, there exists a basis $\{u_i,\ 1 \leq i \leq n\}$ such that for all $x \in X$ there exist unique α_i $(i = 1, \ldots, n)$ such that

$$x = \sum_{i=1}^n \alpha_i u_i.$$

Let us denote by $\|\cdot\|_2$ the usual Euclidean norm:

$$\|x\|_2 = \|\sum_{i=1}^n \alpha_i u_i\|_2 = \left(\sum_{i=1}^n |\alpha_i|^2\right)^{\frac{1}{2}}.$$

We want to prove that any norm $\|\cdot\|$ in X is equivalent to the Euclidean norm $\|\cdot\|_2$. We start with the proof of the existence of $c > 0$ such that $c\|x\| \leq \|x\|_2$. Using the triangular inequality, we have that

$$\|x\| = \|\sum_{i=1}^n \alpha_i u_i\| \leq \sum_{i=1}^n |\alpha_i|\|u_i\|.$$

As for all i $\|u_i\|$ are positive numbers, we apply the Cauchy-Schwartz inequality in $\mathbb{R}^n$,

$$\|x\| \leq \sum_{i=1}^{n} |\alpha_i| \|u_i\| \leq \left(\sum_{i=1}^{n} |\alpha_i|^2\right)^{\frac{1}{2}} \left(\sum_{i=1}^{n} \|u_i\|^2\right)^{\frac{1}{2}} = \|x\|_2 \left(\sum_{i=1}^{n} \|u_i\|^2\right)^{\frac{1}{2}}.$$

We set $c = (\sum_{i=1}^{n} \|u_i\|^2)^{-\frac{1}{2}}$ and finally obtain that

$$c\|x\| \leq \|x\|_2.$$

Let us justify the existence of $C > 0$ such that $\|x\|_2 \leq C\|x\|$. The inequality $c\|x\| \leq \|x\|_2$ implies that if a sequence (x_n) converges with respect to $\|\cdot\|_2$, then (x_n) converges with respect to $\|\cdot\|$ too. Consequently, the norm $\|\cdot\|$ is continuous in $(X, \|\cdot\|_2)$ (as application from X to $\mathbb{R}^+$) and attains its minimum $m > 0$ on the unit sphere

$$S_1(0) = \{x \in X| \quad \|x\|_2 = 1\},$$

(which is compact by the Heine-Borel theorem):

$$\min_{\|x\|_2=1} \|x\| = m > 0.$$

We can thus write that for $\|x\|_2 = 1$

$$\|x\|_2 m = 1 \cdot m \leq \|x\|. \tag{2.10}$$

Suppose now that $x = \frac{y}{\|y\|_2}$, *i. e.* $\|x\|_2 = 1$, but y is not necessary in $S_1(0)$. Consequently, from (2.10) we find using the linear property of the norm that

$$\|y\|_2 \leq \frac{1}{m}\|x\|\|y\|_2 = \frac{1}{m}\left\|\frac{y}{\|y\|_2}\right\|\|y\|_2 = \frac{1}{m}\|y\|\frac{\|y\|_2}{\|y\|_2} = \frac{1}{m}\|y\|.$$

Now we choose $C = 1/m$. □

Corollary 2.5.1 *Let X be a finite-dimensional vector space. There is only one topology induced by the norms.*

2.5.3 Compactness

There is a very important statement about the compactness of unit balls in normed spaces:

Theorem 2.5.2 *Let $(X, \|\cdot\|)$ be a normed linear vector space of infinite dimension. Then the closed unit ball $B_1^c(0) = \{x \in X | \|x\| \leq 1\}$ is not compact.*

To prove it, we need the following Lemma of Riesz about the quasi-perpendicular:

Lemma 2.5.1 ***(Riesz, quasi-perpendicular)*** *Let E be a closed subspace of a normed space $(X, \|\cdot\|)$, such that $E \neq X$. Then*

$$\forall \epsilon \in]0,1[\quad \exists z_\epsilon \notin E, \quad \|z_\epsilon\| = 1 \quad \textit{such that} \quad d(z_\epsilon, E) > 1 - \epsilon,$$

where $d(z_\epsilon, E) = \inf_{u \in E} \|z_\epsilon - u\|$.

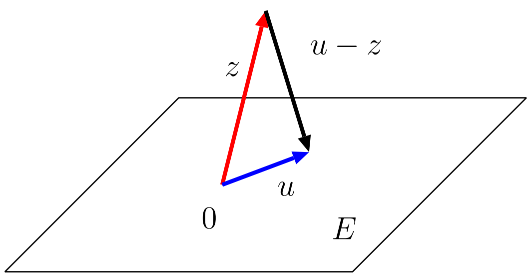

Figure 2.6 – For all vector u in E we have $\|u-z\| > \|z\| = 1$, since the length of the perpendicular z is smaller than the length of any vector $u-z$.

Remark 2.5.1 *Let $X = \mathbb{R}^3$, E be a plane containing zero and let $z \in X$ be a perpendicular vector to E with $\|z\| = 1$. Then for all $u \in E$ we have (see Fig. 2.6)*

$$\|u_z\| > \|z\| = 1.$$

Proof. Let $x \notin E$ (since $E \neq X$ then $\exists x \notin E$). We denote $\inf_{u\in E} \|x-u\| = d$. As $x \notin E$ and, in addition, E is a closed subspace of X, then $d > 0$. Moreover, by the definition of the infinum, we have

$$\forall \epsilon \in]0,1[\quad \exists u_\epsilon \in E \quad \text{such that} \quad d \leq \|u_\epsilon - x\| < \frac{d}{1-\epsilon}.$$

We consider

$$z_\epsilon = \frac{u_\epsilon - x}{\|u_\epsilon - x\|}, \quad \|z_\epsilon\| = 1.$$

Let us prove that $\underline{z_\epsilon \notin E}$:

if $z_\epsilon \in E$, then $u_\epsilon - x \in E$ and consequently, $x \in E$, what is the contradiction.

Finally, let us prove that $\underline{d(z_\epsilon, E) > 1-\epsilon}$:

for all $u \in E$

$$\|z_\epsilon - u\| = \left\| \frac{u_\epsilon - x}{\|u_\epsilon - x\|} - u \right\| = \frac{1}{\|u_\epsilon - x\|} \|x - (u_\epsilon - u\|u_\epsilon - x\|)\|,$$

and, since $u_\epsilon - u\|u_\epsilon - x\| \in E$ and, in addition, $\|x - w\| \geq d$ for all $w \in E$, we obtain

$$\|z_\epsilon - u\| > \frac{d(1-\epsilon)}{d} = 1 - \epsilon. \quad \square$$

Let now prove Theorem 2.5.2.

Proof (Theorem 2.5.2)

Let $y_1 \in B_1^c(0)$ such that $\|y_1\| = 1$. Suppose $y_1, \ldots, y_{n-1} \in B_1^c(0)$ such that

$$\forall i \quad \|y_i\| = 1 \quad \text{and} \quad E_{n-1} = \text{Span}(y_1, \ldots, y_{n-1}), \quad \dim E_{n-1} = n-1 < \infty.$$

Since E_{n-1} is finite dimensional, then it is a closed proper subspace of X ($E_{n-1} \neq X$). By Riesz Lemma, it follows that

$$\exists y_n \notin E_{n-1} \quad \|y_n\| = 1 : \quad \forall i = 1, \ldots n-1 \quad \|y_i - y_n\| > \frac{1}{2}.$$

Consequently, there exists a sequence $(y_i)_{i\in\mathbb{N}} \subset B_1^c(0)$ such that for all $i \neq j$ it holds $\|y_i - y_j\| > \frac{1}{2}$. Therefore, the sequence $(y_i)_{i\in\mathbb{N}}$ does not contain any convergent subsequence, what implies that $B_1^c(0)$ is not compact. □

The following proposition is a direct corollary of the non compactness of the balls in an infinite dimensional normed space:

Proposition 2.5.4 *A subspace E of a normed space $(X, \|\cdot\|)$ is locally compact (the intersection of E with any closed ball in X is compact) if and only if E is finite dimensional.*

We finish with the following corollary:

Corollary 2.5.2 *Let $(X, \|\cdot\|)$ be infinite dimensional normed space. Then all compact sets in $(X, \|\cdot\|)$ are nowhere dense in X:*

if M is compact, then for any open ball $B \subset X$ there exists a non trivial ball $S \subset B$ such that $S \cap M = \varnothing$.

Proof. Let M be a compact in an infinite dimensional normed space $(X, \|\cdot\|)$. Suppose the converse, that M is not nowhere dense in X: there exists a ball $B_r(a)$ in X such that any ball containing in $B_r(a)$ has a nonempty intersection with M. Let us show that in this case $\overline{M} \supset B_r(a)$.

We take $x_0 \in B_r(a)$ and consider a sequence of balls $B_{r_n}(x_0)$ with $r_n = \frac{r - \|x_0 - a\|}{n}$, $n \in \mathbb{N}^*$. If $\|x - x_0\| < r_n$, then

$$\|x - a\| \leq \|x - x_0\| + \|x_0 - a\| < r_n + \|x_0 - a\| \leq r.$$

Consequently, $B_{r_n}(x_0) \subset B_r(a)$ and then, by the assumption, for all $n \in \mathbb{N}^*$

$$B_{r_n}(x_0) \cap M \neq \varnothing.$$

Let $x_n \in B_{r_n}(x_0) \cap M$. Obviously, $x_n \to x_0$ for $n \to +\infty$ in $(X, \|\cdot\|)$. Since for all $n \geq 1$ $x_n \in M$, then $x_0 \in \overline{M}$. Thus, $B_r(a) \subset \overline{M}$ and hence $\overline{B}_r(a) \subset \overline{M}$. As M is compact, then $\overline{M}$ and $\overline{B}_r(a)$ are compact too. Therefore, using Proposition 2.5.4, the compacteness of the ball $\overline{B}_r(a)$ implies that X is finite dimensional, contrary to the assumption. □

2.6 Banach spaces

Definition 2.6.1 *A normed vector space that is complete is called a* ***Banach space****.*

Example 2.6.1

1. *$\mathbb{R}^n$ is a Banach space for any norm defined on it.*
2. *$(C([a,b], \|\cdot\|_{L^2})$ is not a Banach space.*
3. *$(C([a,b], \|\cdot\|_\infty)$ is a Banach space ($\|f\|_\infty = \max_{a\leq x\leq b} |f(x)|$).*
4. *for $p \in [1,\infty]$, $\Omega \subset \mathbb{R}^n$, $L^p(\Omega)$ with the L^p-norm is a Banach space (see Fischer-Riesz Theorem in Subsection D.1.2).*

See 'TD1' for the proofs and more examples.

Problem 2.6.1 *Prove that the space $C([0,2])$ of continuous function on $[0,2]$ equiped with the norm*

$$\|f\| = \left(\int_0^2 |f(t)|^2 \mathrm{d}t\right)^{\frac{1}{2}}$$

is not a Banach space.

Indication: *Consider the limit of the sequence (f_n) of elements of $C([0,2])$ defined by*

$$f_n(t) = \begin{cases} 0, & 0 \leq t < 1 - \frac{1}{n} \\ 1 + n(t-1), & 1 - \frac{1}{n} \leq t < 1 \\ 1, & 1 \leq t \leq 2 \end{cases}.$$

Definition 2.6.2 *We say that*

1. *$\sum_{k=1}^{\infty} x_k$ is **convergent in** X if there exists $x \in X$ such that $S_n \to x$ for $n \to \infty$ in X: $\|S_n - x\|_X \to 0 \quad n \to \infty$.*
2. *$\sum_{k=1}^{\infty} x_k$ is **absolutely convergent in** X if $\sum_{k=1}^{\infty} \|x_k\|_X < \infty$.*

Thus it holds

Theorem 2.6.1 *The normed space $(E, \|\cdot\|_E)$ is complete if and only if every absolutely convergent series in E,*

$$\text{i.e. } x_n \in E \ : \quad \sum_n \|x_n\|_E < \infty,$$

converges in E.

Proof. $\underline{\Rightarrow}$ If E is complete and (x_n) is absolutely convergent, let show that the sequence of partial sums $S_n = \sum_{k=1}^{n} x_k$ is a Cauchy sequence in E.

Indeed, thanks to the absolute convergence,

$$\begin{aligned} &\forall \epsilon > 0 \quad \exists n_0(\epsilon) \in \mathbb{N}: \quad \forall n, p > n_0(\epsilon) \\ &\|S_{n+p} - S_n\|_E = \|x_{n+1} + \ldots + x_{n+p}\|_E \leq \|x_{n+1}\|_E + \ldots + \|x_{n+p}\|_E \leq \epsilon. \end{aligned}$$

Thus, (S_n) is a Cauchy sequence in E. As E is complete, it implies that (S_n) converges in E.

$\underline{\Leftarrow}$ To show that if every absolutely convergent series $x_n \in E$ converges in E then E is complete, first, we see that (x_n) is a Cauchy sequence in E. Consequently, it is possible to extract for a fixed $\epsilon > 0$ a subsequence (x_{n_k}) such that

$$\|x_{n_k} - x_{n_{k+1}}\| \leq \frac{\epsilon}{2^k}.$$

Indeed, we define n_1 such that

$$\|x_{n_1} - x_n\| \leq \epsilon \quad \text{for all } n \geq n_1 = n_0(\epsilon),$$

we define $n_2 \geq n_1$ such that

$$\|x_{n_2} - x_n\| \leq \frac{\epsilon}{2} \quad \text{for all } n \geq n_2 = n_0(\frac{\epsilon}{2}),$$

and so on.

Let also introduce the following sequence:

$$\begin{aligned}
&y_1 = x_{n_1},\\
&y_2 = x_{n_2} - x_{n_1},\\
&\dots\ \dots\ \dots\\
&y_p = x_{n_p} - x_{n_{p-1}},\\
&\dots\ \dots\ \dots
\end{aligned}$$

Let show, thanks to the absolute convergence of (x_{n_p}), that the partial sums of the sequence (y_n) converge in E and consequently (x_{n_p}) converges in E.

Firstly we notice that by our construction

$$Y_p = y_1 + y_2 + \ldots + y_p = x_{n_p} \quad \text{and} \quad \|y_1\| \le \epsilon,\ \|y_2\| \le \frac{\epsilon}{2},\ \ldots,\ \|y_p\| \le \frac{\epsilon}{2^p}.$$

Thus, as (x_{n_p}) is absolutely convergent,

$$\sum_{k=1}^{\infty} y_k$$

is absolutely convergent too:

$$\sum_{k=1}^{\infty} \|y_k\| \le \sum_{k=1}^{\infty} \frac{\epsilon}{2^k},$$

where $(\frac{\epsilon}{2^k})$ is a geometrical sequence. By the assumption, it follows that $\sum_{k=1}^{\infty} y_k$ converges in E, which implies that

$$\sum_{k=1}^{m} y_k = x_{n_m} \to x \in E \quad m \to \infty.$$

To conclude, we use the fact that if (x_n) is a Cauchy sequence containing a convergent (in E) subsequence, then the sequence converges in E. $\square$

Chapter 3

Linear operators

3.1 Definition of a linear continuous and bounded operator in normed spaces

Definition 3.1.1 ***(Linear and densely defined operator)*** *Let $(X, \|\cdot\|)$ and $(Y, \|\cdot\|)$ be two normed vector spaces. Let D be a vector subspace of X.*

A mapping $A : D \subset X \to Y$ is called a ***linear operator*** *(or a linear function, or a linear mapping) if for all u and v in D and all λ in $\mathbb{R}$*

$$\begin{aligned} A(u+v) &= A(u) + A(v) \\ A(\lambda u) &= \lambda A(u) \end{aligned}$$

D is called the domain of A. We say that A is ***densely defined*** *if D is dense in X.*

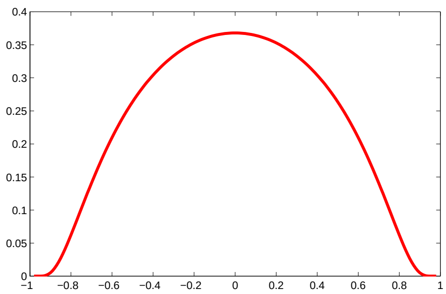

Figure 3.1 – Graphe of the function $g(t) = e^{-1/(1-t^2)}$ if $|t| < 1$ and 0 if $|t| \geq 1$.

Example 3.1.1 *Let $X = L^1(\mathbb{R})$ and $Y = L^\infty(\mathbb{R})$.*

Define $g(t) = e^{-1/(1-t^2)}$ if $|t| < 1$ and 0 if $|t| \geq 1$ (see Fig 3.1). The function g is in C^∞

We define the convolution operator

$$\begin{aligned} &F : X \to Y \\ &f \to \begin{array}{l} \mathbb{R} \to \mathbb{R} \\ x \mapsto \int_{\mathbb{R}} f(y) g(x-y) \mathrm{d}y \end{array} \quad , \end{aligned}$$

which can briefly be written as

$$F(f) = (f * g)(x) = \int_{\mathbb{R}} g(x-y)f(y)\mathrm{d}y. \tag{3.1}$$

F is a linear form or a linear functional.

Let us justify that $f * g \in L^\infty(\mathbb{R})$.

It is easy to see that $g \in C^\infty(\mathbb{R})$ and bounded. Thus, $g \in L^\infty(\mathbb{R})$. Therefore,

$$|(f * g)(x)| \leq \int_{\mathbb{R}} |g(x-y)||f(y)|\mathrm{d}y \leq \max_{x\in\mathbb{R}} |g(x)| \int_{\mathbb{R}} |f(y)|\mathrm{d}y \leq \|g\|_{L^\infty}\|f\|_{L^1}.$$

Definition 3.1.2 ***(Bounded operator)*** *A set X is bounded if there exists $r > 0$, such that $X \subset B_r(0)$.*

Let X and Y be two normed vector spaces. Let $A : D \subset X \to Y$ be a linear operator. We say that A is ***bounded*** *if the image of any bounded set in X is a bounded set in Y.*

Remark 3.1.1 *Linear operator $A : X \to Y$ is bounded if and only if A is bounded on a non-trivial ball $B_r(0)$ centered in 0 in X.*

Proof. If A is bounded, it is obvious that $AB_r(0)$ is bounded for all $r > 0$. (Note that $AB_r(0)$ means A applied to $B_r(0)$).

Suppose that A is bounded on $B_r(0)$ for a fixed $r \neq 0$. We proceed in two steps:

- Let $R > 0$ and $R \neq r$. For $B_R(0) = \lambda B_r(0)$ with $\lambda = \frac{R}{r}$, we see that $AB_R(0)$ is bounded.
- Let M be a bounded set in X. Thus

  $$\exists R > 0 : \quad M \subset B_R(0)$$

  and therefore, $AM \subset AB_R(0)$. Since $AB_R(0)$ is bounded, it implies that AM is bounded too. □

Proposition 3.1.1 *Let X and Y be two normed vector spaces. For a linear operator $A : X \to Y$ the following assertions are equivalent:*

1. *A is continuous,*
2. *A is bounded,*
3. *$\exists C \geq 0 : \quad \|Ax\|_Y \leq C\|x\|_X \quad \forall x \in X$.*

Proof. <u>(1) ⇒ (2)</u> Let A be continuous. Thus, in particular, A is continuous in 0:

$$\forall \epsilon > 0\ \exists \delta = \delta(\epsilon) > 0 : \quad \|x\|_X \leq \delta \ \Rightarrow\ \|Ax\|_Y < \epsilon,$$

where we have used $A0 = 0$, since A is linear. Thus, A is bounded on $B_\delta(0)$ and hence A is bounded.

<u>(2) ⇒ (3)</u> Let A be bounded. If $x = 0$ (3) is obvious. Let $x \neq 0$. We normalize it by introducing

$$x_0 = \frac{x}{\|x\|_X} \ \Rightarrow\ \|x_0\|_X = 1 \ \Rightarrow\ x_0 \in \overline{B_1(0)}.$$

Since $\overline{B_1(0)}$ is bounded, $A\overline{B_1(0)}$ is bounded too. Consequently,

$$\exists C \geq 0 : \quad \|Ax_0\|_Y \leq C,$$

or equivalently, by the linearity of A and the norm,

$$\|Ax_0\|_Y = \left\| A\left(\frac{x}{\|x\|_X}\right) \right\|_Y = \left\| \frac{1}{\|x\|_X} Ax \right\|_Y = \frac{\|Ax\|_Y}{\|x\|_X} \leq C,$$

which gives (3).

<u>(3) ⇒ (1)</u> Let $x_n \to x$ for $n \to \infty$ in X . We have, due to the linearity of A, that

$$\|Ax_n - Ax\|_Y = \|A(x_n - x)\|_Y \leq C\|x_n - x\|_X \to 0 \; n \to \infty,$$

what implies that $\|Ax_n - Ax\|_Y \to 0$ for $n \to \infty$, *i.e.* $Ax_n \to Ax$ for $n \to \infty$ in Y. $\square$

Remark 3.1.2 *Linear operator $A : (X, \|\cdot\|_X) \to (Y, \|\cdot\|_Y)$ is continuous if and only if A is continuous in* 0.

Problem 3.1.1 *Let X and Y be two normed vector spaces. Prove that a linear operator $A : X \to Y$ is continuous on X if and only if A is continuous at 0. (Note that $A0 = 0$!)*

Example 3.1.2 ***(Linear bounded operators, Fredholm operators)***

1. *Let $X = Y = C([a, b])$ with its usual norm: $\|f\|_{C([a,b]} = \max_{a \leq x \leq b} |f(x)|$. We define the operator A by the formula:*

$$f \in C([a, b]) \mapsto \left[t \mapsto y(t) = \int_a^b K(t, s)f(s)ds \in C([a, b]) \right]$$

*where $Af = y$ and K, called **kernel** of A, is a fixed continuous function on $[a, b] \times [a, b]$. The operator A is called the Fredholm operator.*

By its definition A is linear. Let us show that A is bounded, or equivalently, continuous:

$$|y(t)| \leq \int_a^b |K(t, s)f(s)|ds \leq M(b - a)\|f\|_{C([a,b]} \quad \forall t \in [a, b],$$

where we have used

- *As K is a continuous function on the compact $[a, b] \times [a, b]$, K is bounded on $[a, b] \times [a, b]$:*

$$\exists M > 0 \quad |K(t, s)| \leq M \quad \forall (t, s) \in [a, b] \times [a, b].$$

- $|f(s)| \leq \|f\|_{C([a,b]} \quad \forall s \in [a, b]$.

Thus, we obtain for $C = M(b - a)$ that

$$\|y\|_{C([a,b])} \leq C\|f\|_{C([a,b])}. \quad \square$$

2. *Let $X = Y = L^2(]a,b[)$. We define the Fredholm operator A such that*

$$y(t) = A(f(s)) = \int_a^b K(t,s)f(s)ds \in L^2(]a,b[),$$

where the kernel $K \in L^2(]a,b[\times]a,b[)$:

$$\int_{]a,b[\times]a,b[} |K(t,s)|^2 dt\, ds = M^2 < \infty.$$

Let us show that A is bounded, or equivalently, continuous. We use the Cauchy-Schwartz inequality (or the Hölder inequality for $p = 2$):

$$|y(t)| \leq \int_a^b |K(t,s)||f(s)|ds \leq \sqrt{\int_a^b |K(t,s)|^2 ds}\sqrt{\int_a^b |f(s)|^2 ds}.$$

Therefore

$$|y(t)|^2 \leq \sqrt{\int_a^b |K(t,s)|^2 ds}\|f\|^2_{L^2},$$

and integrating the last inequality over $]a,b[$ on t, we obtain

$$\|y\|^2_{L^2} \leq \int_a^b \int_a^b |K(t,s)|^2 dsdt \|f\|^2_{L^2}.$$

Since $\|K\|_{L^2(]a,b[\times]a,b[)} = M$, we conclude that

$$\|y\|_{L^2} \leq M\|f\|_{L^2}.$$

Example 3.1.3 ***(Unbounded operator)***

Let us consider

$$A = \frac{d}{dt} : D \subsetneq L^2(]0,2\pi[) \to L^2(]0,2\pi[)$$

with domain $D = C^1(]0,2\pi[)$. We note that A is not defined for all $f \in L^2(]0,2\pi[)$, but its domain D is dense in $L^2(]0,2\pi[)$.

Let us take $x_n \in D$ such that

$$x_n(t) = \frac{1}{\sqrt{2\pi}} e^{int}, \quad n \in \mathbb{Z}.$$

Then

$$\|x_n\|_{L^2(]0,2\pi[)} = \sqrt{\int_{]0,2\pi[} \left|\frac{1}{\sqrt{2\pi}} e^{int}\right|^2 d\mu} = 1, \Rightarrow x_n \in \overline{B_1(0)},$$

but

$$\|Ax_n\|_{L^2(]0,2\pi[)} = |n|\|x_n\|_{L^2(]0,2\pi[)} = |n| \to \infty \quad n \to \infty.$$

Therefore, A is not bounded.

3.2 $\mathcal{L}(X,Y)$: The space of linear continuous operators

Proposition 3.2.1 *Let X and Y be two normed vector spaces. For a linear operator $A : X \to Y$ we define:*

$$\alpha = \sup_{x\neq 0} \frac{\|Ax\|_Y}{\|x\|_X}, \qquad \beta = \sup_{\|x\|_X\leq 1} \|Ax\|_Y, \qquad \gamma = \sup_{\|x\|_X=1} \|Ax\|_Y,$$
$$\delta = \inf\{C \in \mathbb{R}^+ | \|Ax\|_Y \leq C\|x\|_X \ \forall x \in X\}.$$

Then $\alpha = \beta = \gamma = \delta$.

Proof. Let $M_r = \sup_{\|x\|_X=r} \|Ax\|_Y$. Then $M_1 = \gamma$ and $M_r = rM_1 = r\gamma$.

1. $\underline{\alpha = \gamma}$:
$$\alpha = \sup_{r>0} \sup_{\|x\|_X=r} \frac{\|Ax\|_Y}{\|x\|_X} = \sup_{r>0} \frac{M_r}{r} = M_1 = \gamma.$$

2. $\underline{\beta = \gamma}$:
$$\beta = \sup_{0<r\leq 1} M_r = M_1 \sup_{0<r\leq 1} r = M_1 = \gamma.$$

3. $\underline{\alpha = \delta}$: Let $\phi(x) = \frac{\|Ax\|_Y}{\|x\|_X}$. As $\phi \geq 0$, by definition of the supremum, we have
$$\alpha = \sup_{x\neq 0} \phi(x) = \inf\{C \in \mathbb{R}^+ | \phi(x) \leq C\} = \delta. \quad \square$$

Corollary 3.2.1 *An linear operator A is bounded if and only if one of α, β, γ or δ is finite.*

Definition 3.2.1 ***(Space** $\mathcal{L}(X,Y)$**)*** *Let X and Y be two normed vector spaces.*

The set of linear and continuous operators from X to Y is denoted $\mathcal{L}(X,Y)$.

It is easy to verify that $\mathcal{L}(X,Y)$ is a linear vector space.

Definition 3.2.2 *Let $A \in \mathcal{L}(X,Y)$. The norm of A in $\mathcal{L}(X,Y)$ is defined by*

$$\|A\|_{\mathcal{L}(X,Y)} = \sup_{x\neq 0} \frac{\|Ax\|_Y}{\|x\|_X} = \sup_{\|x\|_X\leq 1} \|Ax\|_Y = \sup_{\|x\|_X=1} \|Ax\|_Y$$
$$= \inf\{C \in \mathbb{R}^+ | \|Ax\|_Y \leq C\|x\|_X \ \forall x \in X\}.$$

Therefore, $(\mathcal{L}(X,Y), \|\cdot\|_{\mathcal{L}(X,Y)})$ is a normed vector space.

Problem 3.2.1 *Prove that $\|\cdot\|_{\mathcal{L}(X,Y)}$ is a norm in $\mathcal{L}(X,Y)$.*

Remark 3.2.1 *$\mathcal{L}(X,X)$ is a normed algebra: for a product AB of linear continuous operators, defined as $(AB)x = A(Bx)$, we have $\|AB\| \leq \|A\|\|B\|$ (if A and B are linear continuous, then AB is linear continuous as a composition of two linear continuous mappings).*

Theorem 3.2.1 *If Y is a Banach space, then $\mathcal{L}(X,Y)$ is a Banach space.*

Proof. Let (A_n) be a Cauchy sequence in $\mathcal{L}(X,Y)$, *i.e.*,

$$\forall \epsilon > 0 \quad \exists N(\epsilon) > 0 \quad \forall m, n \geq N(\epsilon) \quad \|A_n - A_m\|_{\mathcal{L}(X,Y)} < \epsilon. \tag{3.2}$$

We need to show that there exists $A \in \mathcal{L}(X,Y)$ such that $A_n \to A$ for $n \to \infty$ in $\mathcal{L}(X,Y)$. Let us use γ for the definition of the norm:

$$\|A_n - A_m\|_{\mathcal{L}(X,Y)} < \epsilon \quad \Leftrightarrow \quad \sup_{\|x\|_X=1} \|A_n x - A_m x\|_Y \leq \epsilon.$$

Moreover, we have for all $m, n \geq N(\epsilon)$

$$\sup_{\|x\|_X=1} \|A_n x - A_m x\|_Y \leq \epsilon \quad \Leftrightarrow \quad \|A_n x - A_m x\|_Y \leq \epsilon \quad \forall x : \|x\|_X = 1, \tag{3.3}$$

in other words, $(A_n x)$ is a Cauchy sequence in Y. Y is a Banach space, thus $(A_n x)$ is convergent in Y: there exists an element $y \in Y$ such that $y = \lim_{n\to\infty} A_n x$. Let us note it as Ax: $y = Ax$. Then we obtain:

1. A is linear: the limit $\lim_{n\to\infty}$ is linear and A_n are linear for all n.
2. $A_n \to A$ for $n \to \infty$ in $\mathcal{L}(X,Y)$: in (3.3) we fix x and n and pass to $m \to \infty$, thus we find with the limit that

   $$\forall n \geq N(\epsilon) \quad \|(A_n - A)x\|_Y = \|A_n x - Ax\| \leq \epsilon \quad \text{for } \|x\| = 1.$$

   Taking a supremum on x (note that N does not depend on x!) we obtain

   $$\|A_n - A\|_{\mathcal{L}(X,Y)} < \epsilon.$$

3. $A \in \mathcal{L}(X,Y)$: as $\|A_n - A\|_{\mathcal{L}(X,Y)} < \epsilon$ for $n \geq N(\epsilon)$ then $A_n - A$ is bounded for $n \geq N(\epsilon)$, therefore, $A = A_n + (A - A_n) \in \mathcal{L}(X,Y)$ as a sum of linear bounded operators. □

Corollary 3.2.2 *If X is a Banach space, then $\mathcal{L}(X,X)$ is complete.*

An important corollary from the equivalence of norms in a finite dimensional space is that, given X a finite dimensional normed vector space and Y a normed vector space, all linear mapping from X to Y is continuous.

Proposition 3.2.2 *Let $(X, \|\cdot\|_X)$ be finite dimensional normed vector space and $(Y, \|\cdot\|_Y)$ be normed vector space. Then the space $\mathcal{L}(X,Y)$ is equal to the space of all linear mappings from X to Y.*

Proof. As $(X, \|\cdot\|_X)$ is finite dimensional, then all norms on it are equivalent. Consequently, if $(e_1, \ldots, e_N)$ is a base of X, we can consider in X the norm:

$$\|x\|_X = \left\| \sum_{i=1}^{N} x_i e_i \right\|_X \stackrel{def}{=} \sup_{i=1,\ldots,N} |x_i|.$$

Let A be any linear mapping of X to Y. By the linearity of A and by the triangle inequality, we find

$$\|Ax\|_Y = \left\| A\left(\sum_{i=1}^{N} x_i e_i\right) \right\|_Y \leq \sum_{i=1}^{N} |x_i| \|Ae_i\|_Y \leq \left(\sum_{i=1}^{N} \|Ae_i\|_Y\right) \|x\|_X,$$

which proves the continuity of A. $\square$

We refer to [5] for more examples of linear but not continuous operators in infinite dimensional normed spaces (see p. 27 and 28).

We give now the BLT-theorem (see [2] and [9] p.31):

Theorem 3.2.2 ***(BLT: Bounded Linear Transformation)*** *Let X be a normed vector space, D be a dense subspace of X and Y be a Banach space. Let $B : D \subsetneq X \to Y$ be a densely-defined linear continuous operator. There exists a unique continuous linear operator $A : X \to Y$ that extends B and that has the same norm.*

For the proof of the BLT-theorem see TD2.

3.2.1 Definition of the dual space

Definition 3.2.3 *A linear mapping f of a normed space X to $\mathbb{R}$ is called* ***functional*** *or a* ***linear form*** *on X.*

Definition 3.2.4 *We note $X^* = \mathcal{L}(X, \mathbb{R})$ the space of linear continuous functionals on X. It is called the dual space of X.*

The norm on X^* is defined by

$$\forall f \in X^* \quad \|f\| = \sup_{x \neq 0} \frac{|f(x)|}{\|x\|}.$$

Remark 3.2.2 *X^* is always a Banach space (whether X is a Banach space or not).*

Definition 3.2.5 *We denote the value of $f \in X^*$ on $x \in X$ by $\langle f, x \rangle = f(x)$.*

3.3 Bounded operators and theorem of Banach-Steinhaus

Proposition 3.3.1 *Let X and Y be two normed spaces and $A : X \to Y$ be a linear operator. Then A is continuous if and only if $f_A(x) = \|Ax\|_Y \in \mathcal{L}(X, \mathbb{R})$, i.e. $f_A(x) \in X^*$ is a continuous functional on X.*

Proof. We see that

$$\begin{aligned} A \text{ is continuous} &\iff \exists C > 0 : \; \forall x \in X \; \|Ax\|_Y \leq C\|x\|_X \\ &\iff f_A(x) \text{ is bounded for } \|x\|_X \text{ bounded} \iff f_A(x) \in X^*. \quad \square \end{aligned}$$

Definition 3.3.1 ***(Graph of an operator)*** *Let X and Y be two normed vector spaces and $A : X \to Y$ be a linear operator. A set*

$$G_A = \{(x, Ax) | \; x \in X\} \subset X \times Y$$

is called ***graph*** *of the operator A.*

Definition 3.3.2 ***(Closed operator)*** *Let X and Y be two normed vector spaces and $A : X \to Y$ be linear. Then the operator A is called* ***closed*** *(operator), if its graph is closed (set):*

if $\Big((x_n, Ax_n)\Big)_{n\in\mathbb{N}} \subset G_A$ *and for* $n \to +\infty$ $(x_n, Ax_n) \to (x, y) \in G_A$, *then* $y = Ax$.

Remark 3.3.1 *Operator $A : (X, \|\cdot\|_X) \to (Y, \|\cdot\|_Y)$ is closed, then*

$$\begin{array}{l} x_n \to x \ \mathit{in}\ X \\ Ax_n \to y \ \mathit{in}\ Y \end{array} \quad \Rightarrow \quad y = Ax.$$

We give the following important theorem without proof:

Theorem 3.3.1 ***(Closed graph)*** *Let X, Y be Banach spaces. Linear operator $A : X \to Y$ is continuous iff A is closed.*

Using the definition of the graph of an operator we prove the theorem of the continuous inverse operator:

Theorem 3.3.2 ***(Banach, continuous inverse operator)*** *Let X and Y be Banach spaces, A be a linear bijection $X \to Y$ and, in addition, let A be continuous. Then its inverse $A^{-1} : Y \to X$ is also a (linear) continuous operator.*

Proof. Let us compare two graphs:

$$G_A = \{(x, Ax) |\ x \in X\} \subset X \times Y$$
$$G_{A^{-1}} = \{(y, A^{-1}y) |\ y \in Y\} \subset Y \times X.$$

We notice that, actually, $y = Ax$ and $A^{-1}y = x$, hence $(y, A^{-1}y) = (Ax, x)$.

In addition, we define the operator $J : X \times Y \to Y \times X$ by the formula

$$J(x, y) = (y, x)$$

in the way, that $G_{A^{-1}} = JG_A$. We recall (see Chapter 2) that the normed space $X \times Y$ has the norm $\|(x, y)\|_{X\times Y} = \|x\|_X + \|y\|_Y$. By its definition, J is a bijection of two normed spaces, J and J^{-1} are continuous, and consequently, J is a homeomorphism (see Definition A.2.3).

Since, A is continuous and X, Y are Banach spaces, then we apply the Closed graph theorem to conclude that A is closed. Therefore, G_A is closed, thus, as a homeomorphism maps closed sets to closed sets (see Theorem A.2.4), $G_{A^{-1}}$ is also closed. Then, A^{-1} is closed. Since X and Y are Banach spaces, by the Closed graph theorem, we obtain that A^{-1} is continuous. $\square$

The following notions of bounded operators which we will use in Chapter 5:

Definition 3.3.3 *Let X and Y be two normed spaces. A set of bounded linear operators $(A_i)_{i\in I} \subset \mathcal{L}(X, Y)$ is said to be* ***bounded on a vector*** *$x \in X$, if*

$$\|A_i x\| \le C \quad \forall i \in I, \quad \mathit{where}\ C = C(x) \ge 0.$$

Definition 3.3.4 *A set of linear bounded operators $(A_i)_{i\in I} \subset \mathcal{L}(X, Y)$ is* ***uniformly bounded*** *if*

$$\|A_i\|_{\mathcal{L}(X,Y)} \le C \quad \forall i \in I.$$

We notice that if $(A_i)_{i\in I} \subset \mathcal{L}(X,Y)$ is uniformly bounded then

$$\|A_i x\|_Y \leq C\|x\|_X \quad \forall x \in X \quad \forall i \in I.$$

Theorem 3.3.3 ***(Banach-Steinhaus)*** *Let X be a Banach space and Y be a normed space. Let $(A_i)_{i\in I} \subset \mathcal{L}(X,Y)$ be a set of linear bounded operators. $(A_i)_{i\in I}$ is uniformly bounded if and only if $(A_i)_{i\in I}$ is bounded on all vectors x from X.*

Proof. $\Rightarrow$ Obvious.

$\Leftarrow$ Let for all $i \in I$ and for all $x \in X$ $\|A_i x\|_Y \leq C(x)$, where by $C(x)$ is denoted the constant depending on x.

Let

$$X_n = \{x \in X| \ \forall i \in I \ \ \|A_i x\|_Y \leq n\} = \cap_{i\in I}\{x \in X| \ \ \|A_i x\|_Y \leq n\}$$

As A_i are continuous operators and by the continuity of the norm, X_n is an intersection of closed sets (by inverse image of continuous mappings), hence, X_n is closed in X. In addition, for all $x \in X$, $x \in X_n$ if $n \geq C(x)$. In particular, $\cup_{n\in\mathbb{N}} X_n = X$. Since X is a Banach space, by Baire's Theorem, there exists $n \in \mathbb{N}$ such that at least one X_n contains a ball of positive raduis: $B_\epsilon(x_0) \subset X_n$. We also have for all $y \in B_\epsilon(0) \subset X$ that, by the linearity of A_i,

$$\begin{aligned}\|A_i y\|_Y = \|A_i(y + x_0 - x_0)\|_Y = \|A_i(y + x_0) - A_i x_0\|_Y &\leq \|A_i(y + x_0)\|_Y + \|A_i x_0\|_Y \\ &\leq n + \|A_i x_0\|_Y \leq n + C(x_0).\end{aligned}$$

We set $y = \epsilon \frac{x}{\|x\|_X} \in X$ and find

$$\left\|A_i\left(\epsilon\frac{x}{\|x\|_X}\right)\right\|_Y = \frac{\epsilon}{\|x\|_X}\|A_i x\|_Y \leq n + C(x_0),$$

and thus

$$\forall x \in X \quad \forall i \in I \quad \|A_i x\|_Y \leq \|x\|_X \frac{n + C(x_0)}{\epsilon}. \quad \square$$

3.4 Hahn-Banach theorem and its corollaries

We give without proof the following theorem (see [2]):

Theorem 3.4.1 ***(Hahn-Banach)*** *Let X be a real vector space and $\phi : X \to \mathbb{R}^+$ be a semi-norm. In addition, let L be a subspace of X and $l : L \to \mathbb{R}$ be a linear functional on L such that*

$$l(x) \leq \phi(x) \quad \forall x \in L.$$

Then there exists a linear functional $\Lambda : X \to \mathbb{R}$ such that

1. $\Lambda(x) = l(x) \quad \forall x \in L \quad (\Lambda|_L = l)$
2. $\Lambda(x) \leq \phi(x) \quad \forall x \in X.$

There is a direct corollary of the Hahn-Banach theorem:

Corollary 3.4.1 *Given a real normed linear space X, let L be a subspace of X and l a bounded linear functional on L. Then l can be extended to a bounded linear functional Λ on the whole space X without increasing its norm, $\|l\|_{\mathcal{L}(L,\mathbb{R})} = \|\Lambda\|_{\mathcal{L}(X,\mathbb{R})}$.*

In what follows we use three corollaries from the Hahn-Banach theorem (see TD2 for the proof in the case of a separable normed space X):

Corollary 3.4.2 *Let X be a normed vector space, $x \in X$ and $x \neq 0$. Then there exists a linear continuous functional $f \in X^*$ such that*

$$\|f\|_{\mathcal{L}(X,\mathbb{R})} = 1 \quad \textit{and} \quad \langle f, x\rangle = \|x\|.$$

Proof. We apply Corollary 3.4.1 of the Hahn-Banach theorem for $L = \{tx | t \in R\} \subset X$ and $f_0 \in L^*$ defined as

$$f_0(tx) = t\|x\|.$$

We notice that

$$f_0(x) = \|x\|$$

and if $y = tx$ then

$$|f_0(y)| = |t|\|x\| = \|tx\| = \|y\| \quad \Rightarrow \quad \|f_0\|_{\mathcal{L}(L,\mathbb{R})} = 1.$$

Therefore, by Corollary 3.4.1 of the Hahn-Banach theorem, there exists $f \in X^*$ such that $f(x) = \|x\|$ and $\|f\| = 1$. $\square$

Corollary 3.4.3 *Let X be a normed vector space, L be a subspace of X and $x_0 \notin L$ such that $d(L, x_0) = d > 0$. Then there exists $f \in X^*$ such that*

1. $f(x) = 0 \quad \forall x \in L$,
2. $f(x_0) = 1$,
3. $\|f\| = \frac{1}{d}$.

Proof. Let us take $L^1 = L + < x_0 >$, where $< x_0 >$ is a linear space constructed on x_0 taking all its linear combinations. Thus,

$$\forall y \in L^1 \quad \exists! x \in L \text{ and } \exists! t \in \mathbb{R} : y = x + t x_0 \quad \forall y \in L^1.$$

We define $f_0 \in (L^1)^*$ by the formula:

$$f_0(y) = t.$$

Therefore, if $y \in L$ it follows that $f_0(y) = 0$ (thus point 1)), and we also have $f_0(x_0) = 1$ (thus point 2)).

Let us show that $\|f_0\| = \frac{1}{d}$. On the one hand,

$$|f_0(y)| = |t| = \frac{|t|\|y\|}{\|y\|} = \frac{\|y\|}{\left\|\frac{x}{t} + x_0\right\|} \leq \frac{\|y\|}{d},$$

since

$$\left\|\frac{x}{t} + x_0\right\| = \|x_0 - \left(-\frac{x}{t}\right)\| \geq d \quad (\text{as } -\frac{x}{t} \in L).$$

Hence, we find that $\|f_0\| \leq \frac{1}{d}$.

On the other hand, we show that $\|f_0\| \geq \frac{1}{d}$.

Since $d = \inf_{x \in L} \|x_0 - x\|$, it follows that

$$\exists (x_n) \in L \text{ such that } d = \lim_{n \to \infty} \|x_0 - x_n\|.$$

As

$$1 = f_0(x_0 - x_n) \leq \|f_0\| \|x_0 - x_n\|,$$

we obtain for $n \to \infty$ that $\|f_0\| \geq \frac{1}{d}$.

From $\|f_0\| \leq \frac{1}{d}$ and $\|f_0\| \geq \frac{1}{d}$ we conclude $\|f_0\| = \frac{1}{d}$. Thanks to Corollary 3.4.1 of the Hahn-Banach theorem, we expand f_0 to $f \in X^*$ which satisfies all three conditions. □

Corollary 3.4.4 *Let X be a Banach space, L be a subspace of X. L is not dense in X if and only if*

$$\exists f \in X^* \quad f \neq 0 \textit{ such that } f(x) = 0 \quad \forall x \in L.$$

Proof.$\Rightarrow$ Let L be not dense in X:

$$\overline{L} \neq X.$$

Thus, there exists $x_0 \in X$ such that $d(x_0, L) = d > 0$. Applying Corollary 3.4.3,

$$\exists f \in X^* : \quad f(x_0) = 1 \quad (f \neq 0) \quad \text{ and } \quad f(x) = 0 \quad \forall x \in L.$$

$\Leftarrow$ Let $\overline{L} = X$, *i.e.*,

$$\forall x \in X \quad \exists (x_n) \subset L : \quad x_n \to x \; n \to \infty.$$

Then, by the assumption,

$$\exists f \in X^* \quad (f \neq 0) \quad \text{ such that } f(y) = 0 \quad \forall y \in L,$$

from where we find, by the continuity of f, that

$$\forall x \in X \quad f(x) = \lim_{n \to \infty} f(x_n) = 0 \quad \Rightarrow \quad f = 0.$$

This is the contradiction with $f \neq 0$. Therefore, $\overline{L} \neq X$. □

Remark 3.4.1 *1.* $\langle f, x \rangle = 0 \quad \forall x \in X$ *implies* $f = 0$.

2. $\langle f, x \rangle = 0 \quad \forall f \in X^*$ *implies* $x = 0$.

Proof *Let's assume the converse. If $x \neq 0$, by Corollary 3.4.2, there exists $f \in X^*$ such that $f \neq 0$ and $\langle f, x \rangle = \|x\| \neq 0$. It is the contradiction.* □

Remark 3.4.2 *Can we consider the notation $\langle f, x \rangle$ as an "inner product" between elements of two spaces X^* and X? It is obvious bilinear and continuous with respect to $x \in X$. Is it continuous with respect to f too?*

3.5 Reflexivity

Definition 3.5.1 *The **bidual** space of X is noted by X^{**} and defined by*

$$X^{**} = (X^*)^*.$$

It is the normed space of a linear continuous functional from X^* to $\mathbb{R}$ with the norm:

$$\|F\|_{\mathcal{L}(X^*,\mathbb{R})} = \sup_{f \in X^*, \|f\|_{\mathcal{L}(X,\mathbb{R})} \leq 1} |\langle F, f\rangle|.$$

Let us fix $x \in X$. Then for $f \in X^*$ the mapping $F_x : X^* \to \mathbb{R}$ such that $f \mapsto \langle f, x\rangle$ is a linear continuous functional on X^*.

There is a natural injection Φ from X to X^{**}. Given $x \in X$, associate $\Phi(x) = F_x \in X^{**}$ defined by a linear continuous functional on X^*: $F_x : f \in X^* \mapsto \langle f, x\rangle \in \mathbb{R}$. Thus we have the equality:

$$\langle F_x, f\rangle_{X^{**},X^*} = \langle f, x\rangle_{X^*,X} \quad \forall x \in X, \quad \forall f \in X^*.$$

The notation $\langle x, y\rangle_{X,Y}$ means that $x \in X$, $y \in Y$ and $\langle x, y\rangle = x(y)$.

Proposition 3.5.1 *Let X be a normed space. Then the natural injection $\Phi : X \to X^{**}$ is a linear isometric (thus continuous) operator.*

Proof. The linearity is obvious. Let us prove that

$$\|F_x\|_{X^{**}} = \|x\|_X.$$

We find that

$$\|F_x\|_{X^{**}} = \sup_{\|f\| \leq 1} |\langle F_x, f\rangle| = \sup_{\|f\| \leq 1} |\langle f, x\rangle| = \|x\|_X.$$

We need to justify that $\sup_{\|f\|\leq 1} |\langle f, x\rangle| = \|x\|_X$.

Let $x \neq 0$. We see that

$$\sup_{\|f\|\leq 1} |\langle f, x\rangle| \leq \|x\|_X \iff \sup_{x \neq 0} \frac{|\langle f, x\rangle|}{\|x\|_X} \leq 1 \iff |\langle f, x\rangle| \leq \|x\|_X.$$

In addition, thanks to Corollary 3.4.2,

$$\forall x_0 \in X \; \exists f_0 \in X^* : \; \|f_0\| = 1 \; \langle f_0, x_0\rangle = \|x_0\|.$$

Thus, $\sup_{\|f\|\leq 1} |\langle f, x\rangle| = \|x\|_X$. $\square$

As Φ is isometric, it means that $X \approx \Phi(X) \subset X^{**}$, where $\Phi(X)$ is a subspace of X^{**}. If Φ is surjective (and then bijective), X and X^{**} can be identified.

Definition 3.5.2 *We say that X is **reflexive** if $\Phi(X) = X^{**}$. In this case we write $X = X^{**}$ understanding the isometric equivalence.*

Example 3.5.1 *1. $\mathbb{R}^n$ is a reflexive space.*

2. ℓ^p and L^p for $1 < p < \infty$ are reflexive spaces (see Appendix D for the proof).

3. Let c_0 be the space of all sequences $x = (x_1, \ldots, x_k, \ldots)$ converging to zero, with the norm (see Chapter 2)

$$\|x\| = \sup_k |x_k|.$$

Then the space c_0^* is isomorphic to the space ℓ^1 of all absolutely summable sequences. So $c_0^* = \ell^1$, and $(\ell^1)^* = \ell^\infty$. Therefore, $c_0 \neq c_0^{**}$.

4. L^∞ and L^1 are not reflexive spaces (see Appendix D for the proof).

Let us mention without proof the following results (see [2] for the proof):

Theorem 3.5.1 *Let X be a Banach space. X is reflexive if and only if X^* is reflexive.*

Definition 3.5.3 ***(Uniformly convex)*** *A Banach space X is* ***uniformly convex*** *if for all $\epsilon > 0$, there exists $\delta > 0$ such that*

$$\forall x \in X \text{ and } y \in X : \quad \|x\| \leq 1, \quad \|y\| \leq 1 \text{ and } \|x - y\| > \epsilon \quad \Rightarrow \quad \frac{\|x + y\|}{2} < 1 - \delta.$$

Theorem 3.5.2 *(**Milman-Pettis**) Let X be a uniformly convex Banach space. Then X is reflexive.*

Chapter 4

Hilbert spaces

4.1 Sesquilinear and bilinear forms

A vector space X can be defined on $\mathbb{C}$ or on $\mathbb{R}$. If it is not specified, that X is a real vector space, it is supposed that X is defined on $\mathbb{C}$. However, there is some difference in the terminology for the real and the complex spaces. For instance,

Definition 4.1.1 *A mapping $A : X \to Y$ of a complex vector space X to a complex vector space Y is called* ***antilinear*** *if A satisfies*

$$\forall \lambda, \mu \in C \quad x, y \in X \quad A(\lambda x + \mu y) = \overline{\lambda} A(x) + \overline{\mu} A(y).$$

Definition 4.1.2 ***(Real vector space)*** *Let X be a vector space on $\mathbb{R}$. We call* ***bilinear form*** *on X a function $a : X \times X \to \mathbb{R}$ such that, for all u, v, w in X and λ, μ in $\mathbb{R}$, we have:*

1. $a(\lambda u + \mu v, w) = \lambda a(u, w) + \mu a(v, w)$,
2. $a(u, \lambda v + \mu w) = \lambda a(u, v) + \mu a(u, w)$.

It is ***symmetric*** *if $a(u, v) = a(v, u)$ for all u, v in X.*

Definition 4.1.3 ***(Complex vector space)*** *Let X be a vector space on $\mathbb{C}$. We call* ***sesquilinear form*** *on X a function $a : X \times X \to \mathbb{C}$ such that, for all u, v, w in X and λ, μ in $\mathbb{C}$, we have:*

1. $a(\lambda u + \mu v, w) = \lambda a(u, w) + \mu a(v, w)$,
2. $a(u, \lambda v + \mu w) = \overline{\lambda} a(u, v) + \overline{\mu} a(u, w)$.

Definition 4.1.4 *Let X be a vector space on $\mathbb{C}$ or $\mathbb{R}$. A sesquilinear/bilinear form is* ***positive*** *if $a(u, u) \geq 0$ for all u in X.*

It is ***definite positive*** *if $a(u, u) > 0$ for all u in $X \setminus \{0\}$.*

It is ***hermitian*** *if a is sesquilinear and if $a(u, v) = \overline{a(v, u)}$ for all u, v in X. In particular, $a(u, u)$ is real for all $u \in X$.*

Example 4.1.1 *Let us consider the space $L^2(]a,b[)$. The function:*

$$a(u,v) = \int_{]a,b[} uv d\mu \quad \forall (u,v) \in L^2(]a,b[) \times L^2(]a,b[)$$

is a bilinear form, which is symmetric, positive and definite positive, since $\|u\|_{L^2}^2 = a(u,u)$.

Definition 4.1.5 *Let X be a vector space on $\mathbb{C}$ (respectively on $\mathbb{R}$).*

We call ***inner product*** *on X, a definite positive hermitian (respectively symmetric bilinear) form on X. We usually note it $\langle u, v\rangle$ or (u,v).*

We say that u and v are ***orthogonal*** *if $\langle u, v\rangle = 0$ and we note $u \perp v$.*

Problem 4.1.1 *Let $X = \mathbb{R}^n$ and $a(x,y) = \sum_{i,j=1}^n a_{ij}x_iy_j$ be a bilinear form associated to the real matrice $A = (a_{ij})_{i,j=1,\ldots,n}$:*

$$a(x,y) = \sum_{i,j=1}^n a_{ij}x_iy_j = \langle Ax, y\rangle,$$

where $\langle x, y\rangle = \sum_{i=1}^n x_iy_i$ is the inner product in $\mathbb{R}^n$ (prove it!).

Prove that

1. *a is symmetric iff A is symmetric, i.e. $A = A^t$.*
2. *a is an inner product in $\mathbb{R}^n$ iff A is strictly positive defined:*

$$\forall x \in \mathbb{R}^n \; \exists \alpha > 0 : \; \langle Ax, x\rangle \geq \alpha \langle x, x\rangle, \text{ and } \langle Ax, x\rangle = 0 \iff x = 0.$$

Remark 4.1.1 *If $\langle \cdot, \cdot\rangle$ is an inner product on a vector space X, then*

$$\|u\| = \sqrt{\langle u, u\rangle} \quad \forall u \in X$$

is a norm on X. All inner products $\langle \cdot, \cdot\rangle$ are associated with a norm $\|u\|$ defined as $\sqrt{\langle u, u\rangle}$.

The converse is not true at all times:

A norm $\|\cdot\|_X$ is associated with an inner product iff it satisfies the parallelogram law:

$$\|f+g\|_X^2 + \|f-g\|_X^2 = 2(\|f\|_X^2 + \|g\|_X^2) \quad \forall (f,g) \in X \times X. \tag{4.1}$$

In this case, the norm $\|\cdot\|_X$ defines the inner product which can be introduced by the formula:

$$\langle f, g\rangle = \frac{1}{4}(\|f+g\|_X^2 - \|f-g\|_X^2) + \frac{i}{4}(\|x+iy\|^2 - \|x-iy\|^2),$$

where $i = \sqrt{-1}$.

See, for example,[9] p.27.

We add the usual definitions of parallel vectors and normalized vectors:

Definition 4.1.6 *Let X be a vector space on $\mathbb{R}$ with an inner product $\langle\cdot,\cdot\rangle$.*

*We say that $u \in X$ and $v \in X$ are **collinear** if there exists $\lambda \in \mathbb{R}$ such that $u = \lambda v$.*

*We say that $u \in X$ is **normalized** or **unit** if $\|u\| = \sqrt{\langle u, u\rangle} = 1$.*

Theorem 4.1.1 *(**Pythagorean theorem**) Let X be a vector space with an inner product $\langle\cdot,\cdot\rangle$. If $\langle u, v\rangle = 0$, then for the norm defined by the inner product ($\|u\| = \sqrt{\langle u, u\rangle}$) we have*

$$\|u+v\|^2 = \|u\|^2 + \|v\|^2.$$

Proof. Using the properties of the inner product, we find

$$\begin{aligned}\|u+v\|^2 &= \langle u+v, u+v\rangle = \langle u, u+v\rangle + \langle v, u+v\rangle \\ &= \langle u,u\rangle + \langle u,v\rangle + \langle v,u\rangle + \langle v,v\rangle = \|u\|^2 + \|v\|^2,\end{aligned}$$

since $\langle u, v\rangle = \langle v, u\rangle = 0$. $\square$

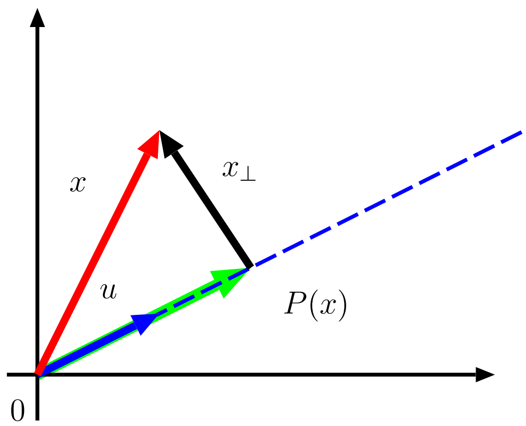

Figure 4.1 – Projection of x (in red) on the direction of u (in blue). $x_\perp$ is in black and $P(x)$ is in green.

Definition 4.1.7 *Let X be a vector space with an inner product $\langle\cdot,\cdot\rangle$.*

*Let u be a normalized vector in X. Then we say that $P(x)$ ($P : X \to X$, $x \mapsto P(x)$) is a **projection** of $x \in X$ on the direction defined by u (see Fig. 4.1) if:*

1. $x_\perp = x - P(x) \perp u$,
2. $P(x) = \lambda u$.

In addition, we find that $\lambda = \langle x, u\rangle$:

$$\langle u, x_\perp\rangle = \langle u, x - \lambda u\rangle = \langle u, x\rangle - \lambda\langle u, u\rangle = \langle u, x\rangle - \lambda = 0 \quad \Longleftrightarrow \quad \lambda = \langle u, x\rangle.$$

Let us also introduce the definition of an orthogonal subspace:

Definition 4.1.8 *Given U a subspace of X, the **orthogonal** of U is the set of vectors of X that are orthogonal to all of the vectors in U. It is noted $U^\perp$.*

*Given U and V two subspaces of X, we say theses spaces are **orthogonal** if for any u in U and any v in V, one has $u \perp v$.*

Remark 4.1.2 *We will see later that for all subsets U of X, its orthogonal $U^\perp$ is a closed vector subspace of X.*

4.2 Pre-Hilbert spaces

Definition 4.2.1 *A **Pre-Hilbert space** (or inner product space) is a vector space with an inner-product.*

Example 4.2.1 *1. Let us consider for $p \geq 1$ the normed vector space ℓ^p of infinite sequences $x = (a_1, a_2, \ldots)$ with the norm (see Chapter 2)*

$$\|x\|_{\ell^p} = \left(\sum_k |a_k|^p \right)^{\frac{1}{p}} < \infty.$$

Using the parallelogram law (4.1), we can show that the norm $\|\cdot\|_{\ell^p}$ is associated to an inner product only for $p = 2$:

Let us take two sequences in ℓ^p

$$f = 1, 1, 0, 0, \ldots, 0, \ldots \quad \text{and} \quad g = 1, -1, 0, 0, \ldots, 0, \ldots.$$

Thus,

$$\begin{array}{lll} f + g = 2, 0, 0, \ldots, 0, \ldots & \text{and} & f - g = 0, 2, 0, \ldots, 0, \ldots \\ \|f\|_{\ell^p} = \|g\|_{\ell^p} = 2^{\frac{1}{p}} & \text{and} & \|f + g\|_{\ell^p} = \|f - g\|_{\ell^p} = 2. \end{array}$$

Equation (4.1) becomes

$$4 + 4 = 4 \cdot 2^{\frac{2}{p}} \iff p = 2.$$

We conclude that only ℓ^2 can be a Pre-Hilbert space. In addition we verify that

$$\langle x, y \rangle = \sum a_i \overline{b_i}, \quad \text{where} \quad \begin{array}{l} x = (a_1, a_2, \ldots) \in \ell^2 \\ y = (b_1, b_2, \ldots) \in \ell^2 \end{array},$$

is an inner product. Let us also notice that ℓ^2 is complete.

2. In analogous way, $L^p([a, b])$ space for $p \geq 1$ is a Pre-Hilbert space iff $p = 2$. The inner product in $L_2([a, b])$ is given by

$$\langle f, g \rangle = \int_{[a,b]} f \overline{g} d\mu \quad \forall (f, g) \in L_2([a, b]) \times L_2([a, b]).$$

$L_2([a, b])$ is also an example of a complete Pre-Hilbert space.

3. *The space $C([0, \frac{\pi}{2}])$ of all continuous functions on $[0, \frac{\pi}{2}]$ with the norm*

$$\|f\| = \max_{0 \le t \le \frac{\pi}{2}} |f(t)|$$

is not a Pre-Hilbert space (take $f(t) = \cos t$ and $g(t) = \sin t$, then $\|f\| = \|g\| = 1$, $\|f + g\| = \sqrt{2}$, $\|f - g\| = 1$, and consequently, (4.1) fails), but it is a Banach space.

4. *The space $C([0, \frac{\pi}{2}])$ of all continuous functions on $[0, \frac{\pi}{2}]$ with the norm*

$$\|f\| = \left(\int_0^{\frac{\pi}{2}} |f(t)|^2 dt \right)^{\frac{1}{2}}$$

is not complete, but it is a Pre-Hilbert space with the inner product

$$\langle f, g \rangle = \int_0^{\frac{\pi}{2}} f(t)\overline{g(t)} dt.$$

Theorem 4.2.1 ***(Cauchy-Schwartz-Bunjakowski)*** *Let E be a Pre-Hilbert space. Then it holds*

$$|\langle x, y \rangle| \le \|x\|\|y\| \quad \forall (x, y) \in E \times E, \tag{4.2}$$

where $\|x\| = \sqrt{\langle x, x \rangle}$.

Proof. Let us prove it for a real Pre-Hilbert space E. The proof for a complex Pre-Hilbert space E can be found in [5] p. 187. Thus for all $\lambda \in \mathbb{R}$

$$\begin{aligned} &\langle x - \lambda y, x - \lambda y \rangle = \|x\|^2 - 2\lambda \langle x, y \rangle + \lambda^2 \|y\|^2 \ge 0 \\ \iff \quad &\Delta = |\langle x, y \rangle|^2 - \|x\|^2 \|y\|^2 \le 0, \end{aligned}$$

which gives directely that $|\langle x, y \rangle| \le \|x\|\|y\|$. Here we have considered

$$\|x\|^2 - 2\lambda \langle x, y \rangle + \lambda^2 \|y\|^2 \ge 0$$

as a quadratique function of λ. $\square$

The Cauchy-Schwartz-Bunjakowski inequality has two important corollaries:

Corollary 4.2.1 *Let E be a Pre-Hilbert space and for all $x \in E$ $\|x\| = \sqrt{\langle x, x \rangle}$. Then the $\| \cdot \|$ of $x \in E$ can be also found by the formula*

$$\|x\| = \max_{\|y\|=1} |\langle x, y \rangle|.$$

Proof. Set $\|y\| = 1$. By Theorem 4.2, we have

$$|\langle x, y \rangle| \le \|x\|\|y\| = \|x\| \quad \forall x \in X,$$

i.e., $|\langle x, y \rangle|$ is bounded by $\|x\|$.

Let us take now $y = \frac{x}{\|x\|}$ for $x \ne 0$. We find the equality:

$$\frac{\langle x, x \rangle}{\|x\|} = \|x\|.$$

Thus $\|x\|$ is the maximum of $|\langle x, y \rangle|$ over all y, such that $\|y\| = 1$. $\square$

Problem 4.2.1 *Let E be a Pre-Hilbert space. Prove (using Corollary 4.2.1) that the function*

$$\|x\| = \sqrt{\langle x, x\rangle}$$

is a norm in E. Therefore, each Pre-Hilbert space is a normed space.

Corollary 4.2.2 *The inner product is continuous as a function of variables of $\langle x, \cdot\rangle$, $\langle \cdot, y\rangle$ and as a function of two variables: $\langle \cdot, \cdot\rangle$, i.e. if $x_n \to x$ and $y_n \to y$ in E, then $\langle x_n, y_n\rangle \to \langle x, y\rangle$.*

Problem 4.2.2 *Prove Corollary 4.2.2.*

4.3 Hilbert spaces

Definition 4.3.1 *A **Hilbert** space is a complete Pre-Hilbert space.*

Subsequently, a Hilbert Space is a Banach space with an inner product. We recall the relations between different spaces in Fig. 4.2

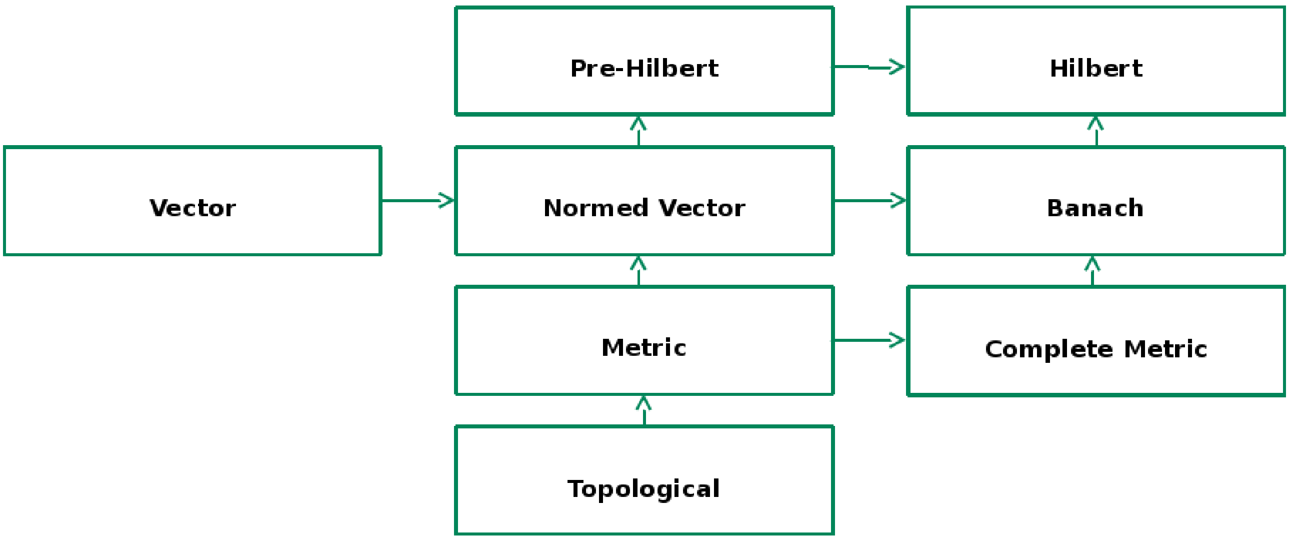

Figure 4.2 – Recap of the different types of spaces. The notation $A \to B$ means that A is more general that B and that B is a particular case of A.

Example 4.3.1 $\mathbb{R}^n$, L^2, ℓ^2 *are Hilbert spaces.*

Proposition 4.3.1 *Let X be a Hilbert space.*

A sesquilinear (respectively bilinear) form $a : X \times X \to \mathbb{C}$ (respectively $\mathbb{R}$) is continuous if there exists a constant $C > 0$ such that

$$\forall (x, y) \in X, \quad |a(x, y)| \leq C\|x\|\|y\|.$$

Problem 4.3.1 *Prove Proposition 4.3.1.*

Note that a sesquilinear (bilinear) form is always continuous if X has a finite dimension.

Definition 4.3.2 *Let X be a Hilbert space. We say that a sesquilinear (bilinear) form $a : X \times X \to \mathbb{C}$ $(\mathbb{R})$ is* ***coercive (or elliptic)*** *if there exists a constant $\alpha > 0$ such that*

$$\forall x \in X \quad a(x,x) \geq \alpha \|x\|^2.$$

Remark 4.3.1 *A coercive sesquilinear (bilinear) form is definite positive.*

4.4 Hilbertian basis

All results of this section are true for the complex Hilbert spaces. For the clearity, all proofs are given for the real case.

4.4.1 Fourier series and orthonormal systems in a Pre-Hilbert space

Definition 4.4.1 *Let X be a Pre-Hilbert space. Let $\{v_i, i \in I\}$ be a family of elements of X. We say this family is* ***orthogonal*** *if*

$$\forall i \in I, j \in I, i \neq j, \langle v_i, v_j \rangle = 0.$$

We say this family is ***orthonormal*** *if, additionally,*

$$\forall i \in I, \langle v_i, v_i \rangle = 1.$$

Example 4.4.1 *Let $X = \ell^2$. Let us define for $i \in \mathbb{N}$ the sequence $v_i = (0, \ldots, 0, 1, 0, \ldots)$, where only one coordinate of v_i, the coordinate number i is not zero and equal to 1. The sequence $(v_i)_{i \in \mathbb{N}}$ is an orthonormal family in ℓ^2.*

Definition 4.4.2 *Let $\{v_i, i \in I\}$ be an orthonormal system in a Pre-Hilbert space X. Let $f \in X$. The numbers*

$$c_i = \langle f, v_i \rangle, \quad i \in I, \tag{4.3}$$

are called ***Fourier coefficients*** *of f with respect to the system $\{v_i, i \in I\}$.*

The series $\sum_{i \in I} c_i v_i$ is called the ***Fourier series*** *of f with respect to the system $\{v_i, i \in I\}$.*

Remark 4.4.1 *In what follows we will answer to the question:*

When does the Fourier series of f converge to f in X?

Let us show that the Fourier coefficients of f minimise the distance between f and the finite dimentional subspace $S_n = \text{Span}(v_1, \ldots, v_n)$, *i.e.* $g = \sum_{i=1}^n c_i v_i$ is the orthogonal projection of f on S_n.

Any element of S_n can be written as $s_n = \sum_{i=1}^n \alpha_i v_i$ for some $\alpha_i \in \mathbb{R}$.

As $(v_i)_{i\in\mathbb{N}}$ is orthonormal system in X, we explicitly find

$$\begin{aligned}
\|f-s_n\|^2 &= \langle f-\sum_{k=1}^n \alpha_k v_k, f-\sum_{j=1}^n \alpha_j v_j\rangle \\
&= \langle f,f\rangle - 2\langle f, \sum_{k=1}^n \alpha_k v_k\rangle + \langle \sum_{k=1}^n \alpha_k v_k, \sum_{j=1}^n \alpha_j v_j\rangle \\
&= \|f\|^2 - 2\sum_{k=1}^n \alpha_k c_k + \sum_{k=1}^n |\alpha_k|^2 \\
&= \|f\|^2 - \sum_{k=1}^n |c_k|^2 + \sum_{k=1}^n |\alpha_k - c_k|^2,
\end{aligned}$$

from where follows the result

$$0 \le d(f,S_n)^2 = \min_{\alpha_k\in\mathbb{R}} \|f-\sum_{k=1}^n \alpha_k v_k\|^2 = \|f-\sum_{k=1}^n \langle f, v_k\rangle v_k\|^2 = \|f\|^2 - \sum_{k=1}^n |\langle f, v_k\rangle|^2.$$

Therefore,

$$\sum_{i=1}^n |c_i|^2 \le \|f\|^2$$

independently of n. Thus, for $n\to+\infty$ we obtain the Bessel inequality:

$$\sum_{i=1}^{+\infty} |c_i|^2 \le \|f\|^2. \tag{4.4}$$

Let us also notice

Corollary 4.4.1 *1. Let $c_1,\dots,c_n,\dots$ be a numerical sequence. The necessary condition for the numerical sequence to be a sequence of Fourier coefficients for an element $f\in X$, is that*

$$\sum_{i=1}^{+\infty} |c_i|^2 < \infty.$$

2. Let $(c_i)_{i\in\mathbb{N}}$ be a sequence of Fourier coefficients for an element $f\in X$ with respect to the orthonormal system $\{v_i, i\in\mathbb{N}\}$. Then $c_i\to 0$ for $i\to+\infty$.

Remark 4.4.2 *The second point of Corollary 4.4.1 follows from the convergence of $\sum_{i=1}^{+\infty}|c_i|^2 < \infty$. To prove the existence of the limit of the Fourier series, we need to have a complete Pre-Hilbert space, i.e., a Hilbert space.*

4.4.2 Hilbertian basis

Definition 4.4.3 ***(Hilbertian basis)*** *Let H be a Hilbert space. Let $e=\{e_i, i\in I\}$ be an orthonormal system in H. The system e is called a basis of H, if*

$$\forall u\in H \quad u=\sum_{i\in I}\langle u, e_i\rangle e_i,$$

where $\langle u, e_i\rangle = c_i$ are the Fourier coefficients of u with respect to e.

Remark 4.4.3 *In Definition 4.4.3 the sum is finite or countable.*

Contrary to the basis in an finite dimentional vector space, u does not need to be equal to a linear combination of elements of the basis. But u needs to be approached (as close as desired) by a linear combination of elements of the basis:

$$\left\| \sum_{i=1}^{n} \langle u, e_i \rangle e_i - u \right\|_H \to 0 \text{ for } n \to +\infty,$$

where $\|u\|_H = \sqrt{\langle u, u \rangle_H}$.

Corollary 4.4.2 *Let H be a Hilbert space. Let $e = \{e_i, i \in I\}$ be an orthonormal system in H. The system e is an orthonormal basis in H iff it holds the Parseval equality:*

$$\forall u \in H \quad \sum_{i \in I} |\langle u, e_i \rangle|^2 = \|u\|^2.$$

Proof. Let e be an othonormal basis in H, *i.e.*,

$$\forall u \in H \quad u = \sum_{i \in \mathbb{N}} \langle u, e_i \rangle e_i.$$

We take the inner product of the lust equality with u and thus obtain the Parseval equality. Let now e be an orthonormal system in H such that

$$\forall u \in H \quad \sum_{i \in \mathbb{N}} |\langle u, e_i \rangle|^2 = \|u\|^2.$$

It means (see Section) that

$$\|u - \sum_{i \in \mathbb{N}} \langle u, e_i \rangle e_i\|^2 = \|u\|^2 - \sum_{i \in \mathbb{N}} |\langle u, e_i \rangle|^2 = 0,$$

which implies that $u = \sum_{i \in \mathbb{N}} \langle u, e_i \rangle e_i$ in H. $\square$

Remark 4.4.4 *Let us notice that for all subset $e \subset H$ of a Hilbert space H we always have*

1. $e^\perp = \overline{Span(e)}^\perp$
2. $\overline{Span(e)} \oplus e^\perp = H$.

Definition 4.4.4 *Let H be a Hilbert space. Let $e = \{e_i, i \in I\}$ be an orthonormal system in H. We say that e is **total** if* Span(e) *is dense in H:*

$$\overline{\mathrm{Span}(e)} = H.$$

Theorem 4.4.1 *Let H be a Hilbert space.*

For an orthonormal system $e = \{e_i, i \in I\}$ in H the following assertions are equivalent:

1. *e is total,*
2. $e^\perp = \{0\}$:
$$\forall i \in I \quad \langle e_i, x \rangle = 0 \quad \Rightarrow \quad x = 0, \tag{4.5}$$
3. *e is an orthonormal basis.*

Proof. See the proof in the general case, presented by Theorem C.2.1. □

Definition 4.4.5 *Hilbert space H is called* ***isometric*** *(or unitary equivalent) to an Hilbert space Z if there exists a surjective isometry $B : H \to Z$.*

Remark 4.4.5 *For Hilbert spaces H and Z a surjective isometry B is a linear bijective operator which preserves the inner product:*

$$\forall (x, y) \in H \times H \quad \langle x, y\rangle_H = \langle Bx, By\rangle_Z.$$

Theorem 4.4.2 *Let H be a separable Hilbert space (see Section A.2.3). Then*

1. *There exists an orthonormal countable basis.*

2. *The real Hilbert space H is isometric to $\mathbb{R}^n$ if the dimension of H is finite and equal to n. If H is an infinite dimentional space, H is isometric to ℓ^2. (If the Hilbert space is complex, it is isometric to C^n for the finite dimentional case, or to the complex space ℓ^2 for the infinite dimentional case.)*

Proof.

1. <u>There exists an orthonormal countable basis</u>: Let $v = \{v_1, \ldots, v_n, \ldots\}$ be a countable set dense in H (it exists since H is separable.) Let us take a maximal subset $f = \{v_{n_1}, v_{n_2}, \ldots\}$ of linear independant elements of v ($\sum_i \alpha_i v_{n_i} = 0 \ \Rightarrow \forall i \ \alpha_i = 0$.)

   We can construct f in the following way:

   $$\begin{aligned} &n_1 = \min\{i : \ v_i \neq 0\}, \\ &n_2 = \min\{i : \ v_i \notin \text{Span}(v_{n_1})\}, \\ &n_3 = \min\{i : \ v_i \notin \text{Span}(v_{n_1}, v_{n_2})\}, \\ &\ldots \end{aligned}$$

   If f is a maximal subset of linear independant elements of v, then

   $$\text{Span}(v) = \text{Span}(f) \quad \text{and} \ \overline{\text{Span}(f)} = H,$$

   since v is dense in H. Let us orthogonalize f by the orthogonalization of Gramm-Schmitt. We denote $f_k = v_{n_k}$. Thus we define

   $$\begin{aligned} &e_1 = \frac{f_1}{\|f_1\|} \quad \Rightarrow \quad \|e_1\| = 1, \\ &e_2 = \frac{f_2 - \langle f_2, e_1\rangle e_1}{\|f_2 - \langle f_2, e_1\rangle e_1\|} \quad \Rightarrow \quad \|e_2\| = 1 \text{ and } e_2 \perp e_1, \\ &\ldots \\ &e_n = \frac{f_n - \sum_{j=1}^{n-1}\langle f_n, e_j\rangle e_j}{\|f_n - \sum_{j=1}^{n-1}\langle f_n, e_j\rangle e_j\|} \quad \Rightarrow \quad \|e_n\| = 1 \text{ and } e_n \perp \{e_1, \ldots, e_{n-1}\} \\ &\ldots \end{aligned}$$

   We obtain the <u>orthonormal</u> system $e = \{e_1, e_2, \ldots\}$, such that $\text{Span}(f) = \text{Span}(e)$, and therefore, $\overline{\text{Span}(e)} = H$. It means that e is a closed, thus total, orthonormal system in H. Consequently, e is an orthonormal basis of H.

2. <u>H is isometric to $\mathbb{R}^n$ or ℓ^2:</u> From Week 4 we know that ℓ^2 is a Hilbert space with the inner product
$$\langle x, y\rangle = \sum x_i y_i, \quad \text{where} \quad \begin{matrix} x = (x_1, x_2, \dots) \in \ell^2 \\ y = (y_1, y_2, \dots) \in \ell^2 \end{matrix} .$$
We suppose in our proof that ℓ^2 and H are real vector spaces.

Since H is separable, by the first point of the Theorem, there exists an orthonormal countable basis $e = \{e_1, e_2, \dots\}$ such that
$$\forall f \in H \quad f = \sum_{i=1}^{\infty} \langle f, e_i\rangle e_i.$$
The sequence of the Fourier coefficients of f, denoted by
$$c_f = (\langle f, e_i\rangle)_{i\in\mathbb{N}^*},$$
and which is uniquely defined for all $f \in H$, is an element of ℓ^2: by the Bessel inequality $\sum_{i=1}^{\infty} |\langle f, e_i\rangle|^2 < \infty$. Thus, for all $f \in H$ there exists unique $c_f \in \ell^2$.

The converse is also true: by Lemma C.1.3, for all elements c in ℓ^2 there exists a unique element $y \in H$.

Let us define an operator
$$F : H \to \ell^2 \quad \forall f \in H \quad F(f) = c_f.$$
We have proved that F is a bijection. Let us prove that
$$\forall (f, g) \in H \times H \quad \langle f, g\rangle_H = \langle F(f), F(g)\rangle_{\ell^2} = \langle c_f, c_g\rangle_{\ell^2} = \sum_i \langle f, e_i\rangle\langle g, e_i\rangle.$$
For $(f, g) \in H \times H$ we have
$$f = \sum_{i=1}^{\infty} \langle f, e_i\rangle_H e_i, \quad g = \sum_{j=1}^{\infty} \langle g, e_j\rangle_H e_j,$$
from where
$$\langle f, g\rangle_H = \langle \sum_{i=1}^{\infty} \langle f, e_i\rangle_H e_i, \sum_{j=1}^{\infty} \langle g, e_j\rangle_H e_j\rangle = \sum_{i=1}^{\infty} \langle f, e_i\rangle\langle g, e_i\rangle. \quad \square$$

Let us give some examples of Hilbertian basis:

Example 4.4.2 *1. In ℓ^2 the set of sequences*
$$\begin{aligned} &e_1 = (1, 0, 0, \dots, 0, 0, \dots), \\ &e_2 = (0, 1, 0, \dots, 0, 0, \dots), \\ &\dots\dots\dots\dots\dots\dots\dots\dots \\ &e_n = (0, 0, 0, \dots, 1, 0, \dots), \\ &\dots\dots\dots\dots\dots\dots\dots\dots, \end{aligned}$$
where the nth coordinate of e_n is one and the others are all zero, is an orthonormal basis.

2. *For $n \in \mathbb{N}^*$ functions*

$$1, \ \cos x, \ \sin x, \ \cos 2x, \ \sin 2x, \ \ldots, \ \cos nx, \ \sin nx, \ \ldots$$

is an orthogonal basis (which can be normalized) of $L^2([-\pi, \pi])$.

3. *Chebyshev's polynomials*

$$T_n(x) = \sqrt{\frac{2}{\pi}} \cos(n \arccos x), \quad n \in \mathbb{N}^*$$

form an orthonormal basis in the space $X = \{f | \ \frac{f^2}{\sqrt{1-x^2}} \in L^1([-1,1])\}$ endowed with the following inner product:

$$\langle f, g \rangle = \int_{[-1,1]} \frac{f(x)g(x)}{\sqrt{1-x^2}} dx.$$

4. *Hermite's polynomials*

$$H_n(x) = (-1)^n e^{x^2} \frac{d^n}{dx^n} \left(e^{-x^2} \right), \quad n \in \mathbb{N}^*$$

form an orthonormal basis in the space $X = \{f | \ e^{-x^2} f^2 \in L^1(\mathbb{R})\}$ endowed with the following inner product:

$$\langle f, g \rangle = \int_{\mathbb{R}} e^{-x^2} f g dx.$$

Remark 4.4.6 *When we write $\int_\Omega f(x)dx$, it means that it is the Lebesgue integral of f over Ω.*

Problem 4.4.1 *Prove that the set of all polynomials with rational coefficients on $]a, b[$ is dense in the set of all polynomials with real coefficients on $]a, b[$, which is dense in $L^2(]a, b[)$.*

Example 4.4.3 *Since L^2 is separable, L^2 is isometric to ℓ^2.*

4.5 Orthogonal projection

We refer to Definition 2.1.3 of the distance between a set and a point in a metric space. See also Remark 2.1.2.

We also recall:

Definition 4.5.1 *Let X be a vector space. $A \subset X$ is **convex** if*

$$\forall (x, y) \in A \times A \quad A_{x,y} = \{z = tx + (1-t)y : \ 0 \le t \le 1\} \subset A.$$

See Fig. 4.3 for an example of a convex and non convex sets.

Theorem 4.5.1 ***(Projection)*** *Let H be a Hilbert space and $A \subsetneq H$ be convex and a closed subset of H. Then for all $x \in H$ there exists a unique $x^* \in A$, called the projection of x on A, such that $\|x - x^*\| = d(x, A)$.*

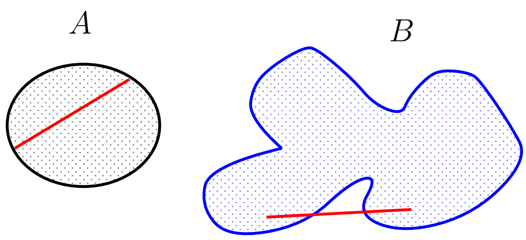

Figure 4.3 – Example of a convex set A and a non convex set B. Red lines represent $A_{x,y}$ and $B_{x,y}$ for fixed x and y.

Proof. Existence

Let $\delta = d(x, A) = \inf_{y\in A} d(x, y)$. By definition of infimum, we have

$$\exists (x_n)_{n\in\mathbb{N}} \in A : \quad d(x, x_n) \searrow \delta$$

or equivalently,

$$\exists (\epsilon_n)_{n\in\mathbb{N}} \in \mathbb{R}^+ : \quad d(x, x_n) = \delta + \epsilon_n, \quad \text{and } \epsilon_n \searrow 0.$$

Without loss of generality, let us suppose that $x = 0$. (We move all by the vector $-x$ and instead of A we consider $\tilde{A} = A - x$, as it is shown in Figure 4.4).

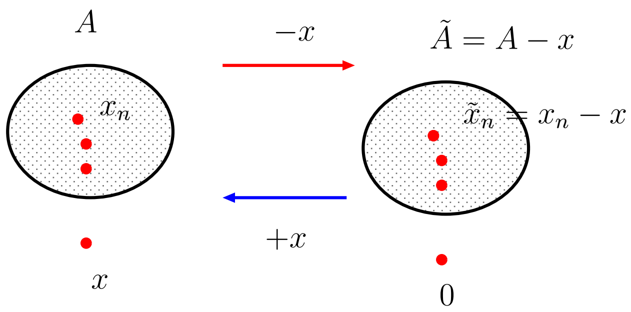

Figure 4.4 – Moving of A and x on the vector $-x$.

For $x = 0$ we have that $d(0, x_n) = \|x_n\| \searrow \delta$. Using the parallelogram law (4.1), we find that

$$\|x_n - x_m\|^2 = 2(\|x_n\|^2 + \|x_m\|^2) - \|x_n + x_m\|^2.$$

We can write

$$\|x_n\|^2 = \delta^2 + \epsilon'_n, \quad \|x_m\|^2 = \delta^2 + \epsilon'_m, \text{for } \epsilon'_n \searrow 0.$$

Let us now estimate the term $\|x_n + x_m\|^2$.

Since A is convex, it follows that

$$\forall (x, y) \in A \times A \quad \Rightarrow \quad z = \frac{x+y}{2} \in A.$$

Thus,

$$\left\|\frac{x_n + x_m}{2}\right\|^2 \geq \delta^2 \quad \Rightarrow -\|x_n + x_m\|^2 \leq -4\delta^2.$$

Therefore, we can estimate

$$\|x_n - x_m\|^2 = 2(\|x_n\|^2 + \|x_m\|^2) - \|x_n + x_m\|^2 \leq 4\delta^2 + 2\epsilon'_n + 2\epsilon'_m - 4\delta^2 = 2\epsilon'_n + 2\epsilon'_m \to 0 \; m, n \to +\infty.$$

Consequently, we obtain that $\|x_n - x_m\| \to 0$ for $m, n \to +\infty$, from where it follows that (x_n) is a Cauchy sequence in H.

H is complete, thus there exists $x^* \in H$ such that $\|x_n - x^*\| \to 0$ for $n \to +\infty$. But (x_n) is a convergent sequence of elements of A, which is closed, thus $x^* \in A$.

Moreover, by the continuity of the norm, we have

$$\|x^*\| = \lim_{n \to \infty} \|x_n\| = \delta.$$

Hence, we have proven that

$$\forall x \in H \quad \exists x^* \in A : \quad \|x - x^*\| = d(x, A).$$

Uniqueness

Let $x^* \in A$ and $y^* \in A$ be such that

$$\|x^*\| = \delta, \quad \|y^*\| = \delta.$$

Then we find by the parallelogram law (4.1) that

$$\|x^* - y^*\|^2 = 2(\|x^*\|^2 + \|y^*\|^2) - \|x^* + y^*\|^2 \leq 4\delta^2 - 4\delta^2 = 0.$$

Here we have used that A is convex, thus $\frac{x^*+y^*}{2} \in A$ which implies that

$$\left\| \frac{x^* + y^*}{2} \right\|^2 \geq \delta^2 \quad \Rightarrow -\|x^* + y^*\|^2 \leq -4\delta^2.$$

We have $0 \leq \|x^* - y^*\| \leq 0$ from where it follows that $\|x^* - y^*\| = 0$ and hence $y^* = x^*$. □

Therefore we can reformulate Theorem 4.5.1 in the following form:

Corollary 4.5.1 *Let X be a Hilbert space and $S \subset X$ be a closed subspace.*

For any $x \in X$, there exists a unique $x^* \in S$ such that $\|x - x^*\| = d(x, S)$.

In addition, x^* is the orthogonal projection of x on S:

$$\|x - x^*\| = d(x, S) \quad \text{iff} \quad \forall y \in S \quad (x - x^*) \perp y. \tag{4.6}$$

Proof. Direct: Let $x \in X$. Let $x^* \in S$ satisfy $\|x - x^*\| = d(x, S)$.

Let's take $y \in S$ and $\lambda > 0$, then, as S is a subspace, $x^* - \lambda y \in S$. We have

$$\begin{aligned} &\|x - x^*\|^2 \leq \|x - (x^* - \lambda y)\|^2 = \langle x - (x^* - \lambda y), x - (x^* - \lambda y)\rangle \\ &= \|x - x^*\|^2 + 2\lambda\langle x - x^*, y\rangle + \lambda^2\|y\|^2, \end{aligned}$$

from where, dividing by $2\lambda \neq 0$, we find

$$\frac{\lambda}{2}\|y\|^2 + \langle x - x^*, y\rangle \geq 0.$$

The term $\langle x - x^*, y\rangle$ does not depend on λ and the inequality holds for all $\lambda \neq 0$. Thus, we can consider the function $f(\lambda) = c_1\lambda + c_2$ with constant coefficients ($c_1 = \frac{1}{2}\|y\|^2$ and $c_2 = \langle x - x^*, y\rangle$) which is linear and continuous on λ. Therefore, passing to the limit for $\lambda \to 0$, we obtain that

$$\forall y \in S \quad \langle x - x^*, y\rangle \geq 0.$$

Since S is a subspace, if $y \in S$, then $-y \in S$:

$$\forall y \in S \quad \langle x - x^*, -y\rangle \geq 0.$$

This implies that

$$\forall y \in S \quad \langle x - x^*, y\rangle = 0 \quad \Longleftrightarrow \quad \forall y \in S \quad (x - x^*) \perp y.$$

<u>Converse:</u> Let $x \in X$. Let $x^* \in S$ be such that

$$\forall y \in S \quad (x - x^*) \perp y.$$

We have

$$\begin{aligned}\|x - y\|^2 &= \|x - x^* + x^* - y\|^2 = \langle x - x^* + x^* - y, x - x^* + x^* - y\rangle \\ &= \|x - x^*\|^2 + 2\langle x - x^*, x^* - y\rangle + \|x^* - y\|^2 \geq \|x - x^*\|^2,\end{aligned}$$

where we use the following facts:

1. Since x^* and y are in S, then $x^* - y \in S$ and consequently $2\langle x - x^*, x^* - y\rangle = 0$.
2. $\|x^* - y\|^2 \geq 0$.

Let us take $\inf_{y\in S}$ of the inequality $\|x-y\| \geq \|x-x^*\|$ (the right-hand part does not depend on y):

$$d(x, S) \geq \|x - x^*\|.$$

In addition, $x^* \in S$ implies that

$$\|x - x^*\| \geq d(x, S),$$

from where we conclude that $\|x - x^*\| = d(x, S)$. $\square$

Remark 4.5.1 *We can give another proof of Corollary 4.5.1: Let's take $y \in S$ and $\lambda \in \mathbb{R}$, then, as S is a subspace, $x^* - \lambda y \in S$. Let us define a function h of λ by the formula:*

$$h(\lambda) = \|x - (x^* - \lambda y)\|^2 = \|x - x^*\|^2 + 2\lambda\langle x - x^*, y\rangle + \lambda^2\|y\|^2.$$

Thus

$$\forall \lambda \in \mathbb{R} \quad h(0) = \|x - x^*\|^2 \leq h(\lambda),$$

which implies that h takes its minimum value at the point $\lambda = 0$. We also notice that $h(\lambda)$ is a quadratic function on λ.

Therefore, by the property of the extremal point (the minimum point here)

$$h'(0) = 2\langle x - x^*, y\rangle = 0,$$

which holds for all $y \in S$ and means that $x - x^* \perp S$.

Inversely, if for all $y \in S$ $\langle x - x^*, y\rangle = 0$, then for all $y \neq 0$ in S

$$h(0) < h(1).$$

It means that

$$\forall y \in S \setminus \{0\} \quad \|x - x^*\| < \|x - (x^* - y)\|,$$

or, since for all $y \in S \setminus \{0\}$ $\quad x^* - y$ defines an element $z \in S \setminus \{x^*\}$ (S is a linear space and $x^* \in S$ and $y \in S$), it means that x^* is the strict global (in S) minimum point and thus unique.

Proposition 4.5.1 *Let H be a Hilbert space and A be its subset. Then $A^\perp$ is a closed vector subspase in H.*

Proof $A^\perp$ is a vector subspace because all linear combinations of elements of $A^\perp$ keep the orthogonal property and thus gives an element of $A^\perp$ (by the linearity of the inner product).

$A^\perp$ is closed, since for any convergent sequence of elements in $A^\perp$ its limit is orthogonal to A by the continuity of the inner product. □

Definition 4.5.2 *Let $H_1, \ldots, H_n$ are Hilbert spaces. Set $H_1 \times \ldots \times H_n$ is denoted by $H_1 \oplus \ldots \oplus H_n$ endowed with the inner product given by the formula*

$$\forall x = (x_1, \ldots, x_n) \quad \forall y = (y_1, \ldots, y_n) \quad \langle x, y\rangle_{H_1 \oplus \ldots \oplus H_n} = \langle x_1, y_1\rangle_{H_1} + \ldots + \langle x_n, y_n\rangle_{H_n},$$

where if $x \in H_1 \oplus \ldots \oplus H_n$, it means that $x = (x_1, \ldots, x_n)$ and $x_i \in H_i$ for all i. Vector operations are defined for each coordinate of x.

Problem 4.5.1 *Prove that $H = H_1 \times \ldots \times H_n$ is a Hilbert space:*

- $\langle x, y\rangle_H = \langle x_1, y_1\rangle_{H_1} + \ldots + \langle x_n, y_n\rangle_{H_n}$ *is an inner product on H*
- *H is complete.*

Remark 4.5.2 *H is called* ***the orthogonal direct sum*** *of the spaces $H_1, \ldots, H_n$. Why is the sum is called orthogonal?*

Let $H = H_1 \oplus H_2$, $x \in H_1$ and $y \in H_2$. Then $x = (a, 0) \in H$ and $y = (0, b) \in H$. Thus

$$\langle x, y\rangle_H = \langle a, 0\rangle_{H_1} + \langle 0, b\rangle_{H_2} = 0,$$

and consequently $H_1 \perp H_2$. In the general case, $H_i \perp H_j$ for $i \neq j$.

Remark 4.5.3 *For all $x \in H = H_1 \times \ldots \times H_n$ there exists unique $x_i \in H_i$ ($i = 1, \ldots, n$) such that $x = (x_1, \ldots, x_n)$.*

Definition 4.5.3 *Let P be an operator from a normed space X to a normed space Y. The kernel of P, denoted by* $\operatorname{Ker} P$, *is called the set*

$$\operatorname{Ker} P = \{x \in X \mid Px = 0\}.$$

Let us prove the following theorem:

Theorem 4.5.2 *Let H be a real Hilbert space and $S \subset H$ be a closed subspace. The operator P from H to S defined by $P(x) = x^*$ (where x^* is the orthogonal projection of x on S) has these properties:*

1. *P is a linear operator.*
2. *$P^2 = P$: $\forall x \in H$ $P(P(x)) = P(x)$. In addition,* $\mathrm{Im}(P) = S$.
3. *If $P \neq 0$, then* $\|P\| = 1$.
4. *P is continuous.*
5. *Its kernel is* $\mathrm{Ker}\, P = S^\perp$.
6. *$H = S \oplus S^\perp$ and $S \neq H$ iff $S^\perp \neq \{0\}$.*

Proof.

1. P is a linear operator:

   Let us show that
   $$\forall \lambda \in \mathbb{R} \quad \forall x \in H \quad P(\lambda x) = \lambda P(x). \tag{4.7}$$
   Indeed, since, by definition of P, $P(x) = x^*$ is the orthogonal projection of x on S, then, thanks to Corollary 4.5.1,
   $$\forall y \in S \quad \langle x - P(x), y \rangle = 0.$$
   Thus, by linearity of the inner product:
   $$\forall \lambda \in \mathbb{R} \quad \forall y \in S \quad \langle \lambda x - \lambda P(x), y \rangle = 0.$$
   On the other hand, for $(\lambda x) \in H$ we have
   $$\forall y \in S \quad \langle \lambda x - P(\lambda x), y \rangle = 0.$$
   Since the orthogonal projection of λx is unique, we find (4.7). Let us show that
   $$\forall x_1 \in H \quad \forall x_2 \in H \quad P(x_1 + x_2) = P(x_1) + P(x_2). \tag{4.8}$$
   For all $x_1 \in H$ and for all $x_2 \in H$ we have
   $$\forall y \in S \quad \langle x_i - P(x_i), y \rangle = 0 \quad \text{for } i = 1, 2.$$
   By linearity of the inner product, we find
   $$\forall y \in S \quad \langle x_1 + x_2 - (P(x_1) + P(x_2)), y \rangle = 0.$$
   For the element $x_1 + x_2$ of H we also have
   $$\forall y \in S \quad \langle x_1 + x_2 - P(x_1 + x_2), y \rangle = 0.$$
   Thanks to the uniqueness of the orthogonal projection of $x_1 + x_2$ on S, we obtain (4.8).

2. $P^2 = P$: $\forall x \in H$ $P(P(x)) = P(x)$. $\mathrm{Im}(P) = S$:

   By definition of P, for all $x \in H$ $P(x) \in S$. Moreover, if $x \in S$, then $d(x, S) = 0$ and $P(x) = x$. Consequently, $P^2 = P$ and $\mathrm{Im}(P) = S$.

3. If $P \neq 0$, then $\|P\| = 1$:

   Let us notice that as for all $x \in H$ its projection $P(x) \in S$, we can take $y = P(x)$ in (4.6) and obtain that
   $$\forall x \in H \quad \langle x - P(x), P(x) \rangle = 0.$$
   Using Pythagorean Theorem (Theorem 4.1.1), we have
   $$\|x\|^2 = \|P(x)\|^2 + \|x - P(x)\|^2.$$
   Therefore,
   $$\forall x \in H \quad \|P(x)\| \leq \|x\|,$$
   which implies that $\|P\| \leq 1$.

   If $P \neq 0$, there exists $x \neq 0$ in S. Thus $P(x) = x$ and $\|P(x)\| = \|x\|$. Then we conclude that $\|P\| = 1$.

4. P is continuous:

   Since P is linear and its norm is finite (equal to 1), then P is continuous.

5. Its kernel is $S^{\perp}$:

   Thanks to Corollary 4.5.1, we directely find
   $$x \in \operatorname{Ker} P \quad \Longleftrightarrow \quad P(x) = 0 \quad \Longleftrightarrow \quad x^* = 0 \quad \Longleftrightarrow \quad x \perp S \quad \Longleftrightarrow \quad x \in S^{\perp}.$$

6. $H = S \oplus S^{\perp}$ and $S \neq H$ iff $S^{\perp} \neq \{0\}$:

   Thanks to the previous point, we can also write that $H = \operatorname{Im}(P) \oplus \operatorname{Ker} P$. Thus we directely see that
   $$\operatorname{Im}(P) = S \neq H \quad \Longleftrightarrow \quad S^{\perp} = \operatorname{Ker} P \neq \{0\}.$$
   Let us prove that $H = S \oplus S^{\perp}$. For all $x \in H$
   $$x = (x - P(x)) + P(x), \tag{4.9}$$
   where $P(x) \in S = \operatorname{Im}(P)$. Let us show that $(x - P(x)) \in S^{\perp} = \operatorname{Ker} P$:
   $$P(x - P(x)) = P(x) - P(P(x)) = P(x) - P(x) = 0.$$
   In addition $\operatorname{Im}(P) \perp \operatorname{Ker} P$ and decomposition (4.9) is unique by the uniqueness of the orthogonal projection. $\square$

4.6 Riesz representation theorem

Theorem 4.6.1 ***(Riesz representation)*** *Let H be a Hilbert space. Any linear continuous functional on H can be uniquely presented by the inner product in H:*
$$\forall f \in H^* \quad \exists!\, y \in H : \quad \forall x \in H \quad f(x) = \langle x, y \rangle, \tag{4.10}$$
and moreover, $\|f\|_{H^*} = \|x\|_H$. (In other words, any Hilbert space is isometric to its dual.)

Remark 4.6.1 *The Riesz representation theorem states that*

1. *For any (fixed) y in H, the linear form $f_y : H \to \mathbb{C}$ defined by*

$$\forall x \in H \quad f_y(x) = \langle x, y\rangle$$

is a linear continuous form on H ($f_y \in H^*$) and

$$\|f_y\|_{H^*} = \|y\|_H. \tag{4.11}$$

2. *The function $F : H \to H^*$ defined by $F(y) = f_y$ is an isometric isomorphism from H to H^*.*

Proof. Let us prove (4.10) for a real Hilbert space H.

Let f be a linear continuous function on H: $f \in H^*$. Let's consider

$$H_0 = \operatorname{Ker} f = \{x \in H : f(x) = 0\} = f^{-1}(\{0\}).$$

For H_0 we have:

- Since $\operatorname{Ker} f$ is a vector space (if $f(x) = f(y) = 0$ then for all α and β in $\mathbb{R}$ $f(\alpha x+\beta y) = \alpha f(x) + \beta f(y) = 0$), H_0 is a subset of H.
- Since f is continuous and the one point set $\{0\}$ is closed in $\mathbb{R}$, then the inverse image of $\{0\}$ is closed in H (see Apendix A.2). Therefore, H_0 is a closed subspace of H.

Let us prove that $\dim H_0^\perp = 1$.

We fix $x_0 \notin \operatorname{Ker} f$, *i.e.* $x_0 \in H_0^\perp$ (indeed, $\operatorname{Ker} f = H_0$ is a closed subspace of H, thus $H = H_0 \oplus H_0^\perp$ implies that $x_0 \in H_0^\perp$) and $f(x_0) \neq 0$. Such x_0 exists if $f \neq 0$ (if $f \equiv 0$ then f obviously has the representation (4.10) with $x = 0$ and in this case $\|x\| = \|f\| = 0$). We also notice that $x_0 \neq 0$ since $f(0) = 0$ (as f is linear continuous, it is continuous in 0).

Let $x \in H$. For $y = x - f(x)\frac{x_0}{f(x_0)}$ we find that

$$f(y) = f\left(x - f(x)\frac{x_0}{f(x_0)}\right) = f(x) - f(x)\frac{f(x_0)}{f(x_0)} = f(x) - f(x) = 0,$$

i.e. $y \in H_0$.

For a fixed x_0 in $H_0^\perp$ the representation of x by the formula

$$x = y + \alpha x_0, \qquad \text{where } y \in H_0, \text{ and } \alpha \in \mathbb{R},$$

is unique.

Indeed, let $x = y + \alpha x_0$ for $y \in H_0$ and $x = y' + \alpha' x_0$ for $y' \in H_0$. Then

$$(\alpha - \alpha')x_0 = y' - y.$$

If $\alpha = \alpha'$ it implies that $y' = y$. If $\alpha \neq \alpha'$, it implies that

$$x_0 = \frac{y' - y}{\alpha - \alpha'} \in H_0,$$

which is a contradiction with the assumption that $x_0 \in H_0^\perp$.

Consequently, for $\alpha = \frac{f(x)}{f(x_0)} \in \mathbb{R}$ we have that any vector $x \in H$ can be uniquely presented by

$$x = y + \alpha x_0, \quad y \in H_0, \quad x_0 \in H_0^\perp. \tag{4.12}$$

Thus $H = H_0 \oplus H_0^\perp$ and $H_0^\perp = \langle x_0 \rangle$ with $\dim H_0^\perp = 1$.

Let us show that there exists a unique $y \in H$ such that for all $x \in H$ $f(x) = \langle x, y\rangle$.

Let us consider (4.12) with a normalized vector x_0: $\|x_0\| = 1$.

We apply to (4.12) f:

$$f(x) = 0 + \alpha f(x_0),$$

and we take the inner product of (4.12) with x_0:

$$\langle x, x_0\rangle = 0 + \alpha\langle x_0, x_0\rangle = \alpha.$$

Therefore,

$$\forall x \in H \quad f(x) = \langle x, x_0\rangle f(x_0) = \langle x, y\rangle, \quad \text{where } y = f(x_0)x_0.$$

Let us prove the uniqueness of $y \in H$ such that $f(x) = \langle x, y\rangle$ for all $x \in H$. Suppose that there exist y_1 and y_2 such that for all $x \in H$

$$f(x) = \langle x, y_1\rangle \quad \text{and} \quad f(x) = \langle x, y_2\rangle.$$

Then

$$\forall x \in H \quad \langle x, y_1 - y_2\rangle = 0 \quad \Longleftrightarrow \quad (y_1 - y_2) \perp H,$$

which implies that $y_1 - y_2 = 0$.

Let us prove (4.11): In fact, by Cauchy-Schwartz-Bunjakowski inequality

$$|f(x)| = |\langle x, y\rangle| \leq \|x\|\|y\|,$$

but $f(y) = \|y\|^2$, thus we have (4.11). ☐

Thanks to $H = H^*$, we find that a Hilbert space is reflexive.

4.7 Operators defined by a sesquilinear/bilinear form

Theorem 4.7.1 *Let H be a Hilbert space. Let $a : H \times H \to \mathbb{C}$ be a continuous sesquilinear form. Then there exists a unique bounded operator $A : H \to H$ such that*

$$\forall (x, y) \in H, \quad a(x, y) = \langle x, Ay\rangle.$$

Proof. By the continuity hypothesis, for any fixed $y \in H$ the linear form $x \mapsto a(x, y)$ is continuous. We denote it by $f_y(x) = a(x, y)$ for a fixed $y \in H$. By Riesz representation theorem, there exists unique $z \in H$ such that $f_y(x) = \langle x, z\rangle$.

Therefore, for any $y \in H$, there exists unique $z \in H$ such that $a(x,y) = \langle x, z\rangle$. Define $A : H \to H$ by $Ay = z$. A is linear (the unicity in Riesz representation theorem and the sesquilinearity of $a(\cdot,\cdot)$). In addition,

$$\|Ay\|^2 = \langle Ay, Ay\rangle = a(Ay, y) \le C\|Ay\|\|y\|.$$

Thus

$$\|Ay\| \le C\|y\|,$$

from where we conclude that A is a bounded linear operator. □

Definition 4.7.1 *Let $A : H \to H$ be a bounded operator. Let $a : H \times H \to \mathbb{R}$ be the associated bilinear/sesquilinear form.*

We note $A \ge 0$ if a is positive. We note $A \ge B$ if $A - B \ge 0$.

Theorem 4.7.2 ***(Adjoint operator)*** *Let H be a Hilbert space. For all $A \in \mathcal{L}(H,H)$ there exists unique $A^* \in \mathcal{L}(H,H)$, which satisfies*

$$\forall x, y \in H \quad \langle Ax, y\rangle = \langle x, A^*y\rangle. \tag{4.13}$$

Moreover, $\|A\| = \|A^*\|$.

A^* is called the adjoint operator of A.

Proof. Let $A : H \to H$ be a bounded operator ($x \mapsto Ax$). From A, let us define $a : H \times H \to \mathbb{R}$ by $a(x,y) = \langle Ax, y\rangle$. It is a continuous bilinear form (since A is continuous and the inner product is continuous). By Theorem 4.7.1 there exists a unique bounded operator $A^* : H \to H$ such that

$$\forall (x,y) \in H, \quad a(x,y) = \langle x, A^*y\rangle.$$

Let us prove that $\|A\| = \|A^*\|$.

For a fixed y in H, $f_y(x) = \langle Ax, y\rangle$ is a linear continuous form on H. Thanks to the Riesz representation theorem,

$$\|A^*y\| = \|f_y\|,$$

and therefore, by definition of the norm of a linear functional

$$\|f_y\| = \sup_{x\neq 0} \frac{|f_y(x)|}{\|x\|},$$

we have

$$\|A^*\| = \sup_{y\neq 0} \frac{\|A^*y\|}{\|y\|} = \sup_{x,y\neq 0} \frac{|\langle x, A^*y\rangle|}{\|x\|\|y\|} = \sup_{x,y\neq 0} \frac{|\langle Ax, y\rangle|}{\|x\|\|y\|} = \|A\|. \square$$

Problem 4.7.1 *Prove that the mapping $A \mapsto A^*$ satisfies the following properties:*

1. *$(\lambda A)^* = \overline{\lambda} A^*$, (antilinear)*
2. *$(A + B)^* = A^* + B^*$,*

3. $(AB)^* = B^* A^*$,
4. $A^{**} = A$.

Definition 4.7.2 *Let H be a Hilbert space. Operator $A \in \mathcal{L}(H, H)$ is*

- ***auto-adjoint*** *(also hermitian for H over $\mathbb{C}$ and symmetric for H over $\mathbb{R}$) if $A^* = A$,*
- ***antihermitian*** *(antisymmetric for H over $\mathbb{R}$) if $A^* = -A$,*
- ***unitary*** *(orthogonal for H over $\mathbb{R}$) if $A^* A = I$ (by I we denote the identity operator).*

Proposition 4.7.1 *Let H be a Hilbert space and $A \in \mathcal{L}(H, H)$. The following assertions are equivalent:*

- *A is hermitian.*
- *The form $f_A(x, y) = \langle Ax, y\rangle$ is hermitian ($f_A(x, y) = \overline{f_A(y, x)}$).*
- *The quadratic form $\phi_A(x) = \langle Ax, x\rangle$ is real.*

In addition, we have

$$\|A\| = \sup_{x \neq 0} \frac{| < Ax, x > |}{\|x\|_H^2}.$$

Problem 4.7.2 *Prove Proposition 4.7.1*

Proposition 4.7.2 *Let H be a Hilbert space.*

1. *Mapping $A \mapsto iA$ is a bijection between all hermitian operators and antihermitian operators on $\mathcal{L}(H, H)$.*
2. *For all $A \in \mathcal{L}(H, H)$ there exist unique hermitian operators B and C in $\mathcal{L}(H, H)$ such that*
$$A = B + iC.$$

Proof.

1. Let $A \in \mathcal{L}(H, H)$ be a hermitian operator. We define $B = iA$ and thus $B^* = -iA^*$. As A is hermitian, then $A^* = A$, from where we obtain that $B^* = -B$ and hence B is antihermitian. Doing the proof in the inverse way, we obtain that if $B = iA$ is antihermitian, then A is hermitian.
2. We fixe $A \in \mathcal{L}(H, H)$. Let define $B = \frac{A+A^*}{2}$ and $C = \frac{A-A^*}{2i}$. We verify that B and C are hermitian operators. Moreover, $A^* = B - iC$ and then $A = B + iC$. □

Proposition 4.7.3 *Let H be a Hilbert space. For an operator $U \in \mathcal{L}(H, H)$ the following assertions are equivalent:*

- *U is an unitary operator.*
- *For all x, y in H $\langle Ux, Uy\rangle = \langle x, y\rangle$ and U is onto.*
- *U is an isometry (global) of H to H.*

Proof. 1) $\Rightarrow$ 2)

Since U is a unitary operator, then $U^* = U^{-1}$ and then

$$\forall (x,y) \in H \times H \quad \langle Ux, Uy\rangle = \langle x, U^*Uy\rangle = \langle x, U^{-1}Uy\rangle = \langle x, y\rangle.$$

<u>2) ⇒ 1)</u>

We have

$$\forall (x,y) \in H \times H \quad \langle Ux, Uy\rangle = \langle x, y\rangle.$$

Thus,

$$\forall (x,y) \in H \times H \quad \langle x, U^*y\rangle = \langle Ux, y\rangle = \langle Ux, UU^{-1}y\rangle = \langle x, U^{-1}y\rangle,$$

which implies that for all $y \in H$ $U^*y = U^{-1}y$ and finally $U^* = U^{-1}$.

$U^*U = I$, and then

<u>2) ⇒ 3)</u>

As

$$\|Ux\|^2 = \|x\|^2,$$

U is a isometry. Since there exists U^{-1}, then U is a bijection, and thus U is a global isometry of H to H.

<u>3) ⇒ 2)</u>

Since U is a global isometry of H to H, U is bijection ($\Rightarrow \exists U^{-1}$) and $\|Ux\| = \|x\|$. □

Problem 4.7.3 *Let H be a Hilbert space and $A \in \mathcal{L}(H,H)$. Prove that if $A = A^*$, then the operator $U = e^{iA}$ is a unitary operator on H ($i = \sqrt{-1}$).*

Proposition 4.7.4 *Let H be a Hilbert space and $A : H \to H$ a linear continuous operator ($A \in \mathcal{L}(H,H)$). If $A = A^*$, then $H = \mathsf{Ker}A \oplus \overline{\mathsf{Im}A}$.*

Attention : $\overline{\mathsf{Im}A}$ means the closure of $\mathsf{Im}A$.

For the proof of Proposition 4.7.4 see TD3.

Corollary 4.7.1 *The operator P of the orthogonal projection on a closed subspace S of a Hilbert space H is symmetric.*

See TD3 for the proof.

4.8 Continuity and coercivity of a sesquilinear form

Let us show the important fact:

Proposition 4.8.1 *Let H be a Hilbert space. Let $\langle \cdot, \cdot \rangle$ be the inner product on H and $\|\cdot\|$ be the norm corresponding to this inner product.*

Let $a(\cdot,\cdot)$ be a sesquilinear/bilinear form on H which is continuous and coercive.

Then $a(\cdot,\cdot)$ defines an inner product equivalent to $\langle \cdot, \cdot \rangle$ iff $a(\cdot,\cdot)$ is hermitian/symmetric.

Proof. By definition of an inner product, it is a definite positive hermitian sesquilinear form. Let $a(\cdot,\cdot)$ be a sesquilinear form on H such that

1. $a(\cdot,\cdot)$ is continuous: $\exists C > 0: \ \forall (x,y) \in H^2, \quad |a(x,y)| \le C\|x\|\|y\|$,
2. $a(\cdot,\cdot)$ is coercive: $\exists \alpha > 0: \ \forall x \in H \quad a(x,x) \ge \alpha\|x\|^2$.

From the coercivity of $a(\cdot,\cdot)$ follows that $a(\cdot,\cdot)$ is definite positive. If $a(\cdot,\cdot)$ is hermitian, then $a(\cdot,\cdot)$ is an inner product on H. Let us define the norm corresponding to the inner product $a(\cdot,\cdot)$:

$$\forall x \in H \quad \|x\|_a = \sqrt{a(x,x)}.$$

Let now prove that the norms $\|\cdot\|$ and $\|\cdot\|_a$ are equivalent in H. Thanks to the continuity and the coercivity of $a(\cdot,\cdot)$, we have

$$\alpha\|x\|^2 \le \|x\|_a^2 = a(x,x) \le C\|x\|^2. \quad \square$$

4.9 Lax Milgram Lemma

Remark 4.9.1 *Let us define the identical operator I or Id which maps all $x \in H$ to itself: $I(x) = x$.*

Theorem 4.9.1 *(**Bounded inverse**) Let H be a real Hilbert space. Let $\alpha > 0$ and $A : H \to H$ be a linear bounded operator such that (see Definition 4.7.1)*

$$A \ge \alpha I.$$

Then it holds

1. *A is bijective,*
2. *A^{-1} is bounded,*
3. *$\|A^{-1}\| \le \frac{1}{\alpha}$.*

Proof. <u>A is injective</u>:

Let $A : H \to H$ and x be in H. Since $A \ge \alpha I$, it means that

$$\forall x \in H \quad \langle x, (A - \alpha I)x\rangle \ge 0,$$

and thus, using the linearity and symmetry of the inner product,

$$\forall x \in H \quad \langle Ax, x\rangle \ge \alpha\langle x, x\rangle.$$

Thus, by Cauchy-Schwartz-Bunjakowski inequality,

$$\alpha\|x\|^2 \le \langle Ax, x\rangle \le \|Ax\|\|x\|.$$

We devide by $\|x\|$ to obtain

$$\alpha\|x\| \le \|Ax\| \quad (\text{ and } \|x\| \le \frac{\|Ax\|}{\alpha}).$$

If $Ax = 0$, then $\|Ax\| = 0$, from where, due to the last estimation, $0 \le \|x\| \le 0$, *i.e.*, $\|x\| = 0$, and then $x = 0$. Thus A is injective.

<u>A is surjective</u>:

Let $z \in \mathrm{Im}(A)^{\perp}$. Then, for any $y \in \mathrm{Im}(A)$, we have $\langle y, z\rangle = 0$. Since

$$Az \in \mathrm{Im}(A),$$

it follows that

$$\langle Az, z\rangle = 0.$$

Thus

$$\alpha\langle z, z\rangle \leq \langle Az, z\rangle = 0$$

and then $\|z\| = 0$. Hence $\mathrm{Im}(A)^{\perp} = \{0\}$. Consequently,

$$\overline{\mathrm{Im}(A)} = H.$$

Let us show that

$$\overline{\mathrm{Im}(A)} = \mathrm{Im}(A),$$

i.e., it is closed:

$$\text{if } Ax_n \to y \quad \Rightarrow \quad \exists x \in H : \quad y = Ax.$$

Indeed, if $Ax_n \to y$, then (Ax_n) is a Cauchy sequence. It implies that (x_n) is a Cauchy sequence in H as we have

$$\forall p \geq n \quad \alpha\|x_n - x_p\| < \|Ax_n - Ax_p\|.$$

Since H is complete, there exists $x \in H$ such that $x_n \to x$. Thanks to the continuity of A (a linear bounded operator is continuous), we have

$$Ax_n \to Ax.$$

By the unicity of the limit, we obtain that $Ax = y$.

Therefore, all limit points of $\mathrm{Im}(A)$ are in $\mathrm{Im}(A)$, hence $\mathrm{Im}(A)$ is closed and we conclude that $\mathrm{Im}(A) = H$.

<u>A^{-1} is bounded</u>: As A is bijective, there exists A^{-1}.

Let $y \in \mathrm{Im}(A) = H$ and note $x = A^{-1}y$.

We have

$$\|A^{-1}y\| = \|x\| \leq \frac{\|Ax\|}{\alpha} = \frac{1}{\alpha}\|y\|$$

Thus A^{-1} is bounded and its norm is bounded by $(1/\alpha)$. □

Theorem 4.9.2 ***(Lax Milgram Theorem)*** *Let H be a real Hilbert space. Let $a(\cdot, \cdot)$ be a continuous and coercive bilinear form. Let $f \in H^*$.*

- *The equation: for all $u \in H, a(x, u) = f(u)$ has one and only one solution $x \in H$.*
- *The application that associates f to x is linear and continuous from H^* to H.*

Proof.

By the Riesz representation theorem, there exists an unique z in H such that $\langle z, u\rangle = f(u)$, and moreover, the application $f \mapsto z$ is linear and continuous ($\|z\| = \|f\|$).

For all $u \in H$, $a(x,u) = f(u)$ is equivalent to

$$\forall u \in H, \quad a(x,u) = \langle z, u\rangle,$$

where z is defined from f by the Riesz representation theorem.

Since $a : H \times H \to \mathbb{R}$ is a continuous bilinear form, by Theorem 4.7.1, there exists an unique bounded operator $A : H \to H$ such that:

$$\forall (x,u) \in H \times H \quad a(x,u) = \langle Ax, u\rangle.$$

Therefore $Ax = z$.

Since $a(\cdot,\cdot)$ is coercive, there exists $\alpha > 0$ such that $A \geq \alpha I$. Thus A^{-1} is a linear bounded operator (see Theorem 4.9.1) and then $x = A^{-1}z$. Thus the application $f \mapsto x$ is linear and continuous:

$$\|x\| < \frac{1}{\alpha}\|z\| = \frac{1}{\alpha}\|f\|. \quad \square$$

Remark 4.9.2 *If we add in the statement of Lax-Milgram Theorem the assumption that $a(\cdot,\cdot)$ is symmetric, then, by Proposition 4.8.1, the bilinear form $a(\cdot,\cdot)$ defines an inner product on H. In this case, to prove the unique existence of the solution we can directely apply the Riesz representation theorem on H equipped with the inner product $a(\cdot,\cdot)$.*

Chapter 5

Weak and Weak* convergences

We have seen that a unit ball in a infinite dimensional normed space is not compact. It is due to to the definition of the convergence (or the topology). Let us change the type of the convergence, to see if it is possible to make it compact.

5.1 Weak convergence in a Banach space

Let us define weak convergence (for the definition of the weak topology, see Appendix A Section A.5.1):

Definition 5.1.1 *Let (x_n) be a sequence of elements of X. We say that (x_n) converges* ***weakly*** *to x, noted $x_n \rightharpoonup x$, if for all f in X^*, $f(x_n)$ converges to $f(x)$.*

The weak convergence define the weak topology, denoted by $\sigma(X, X^*)$.

Definition 5.1.2 *The convergence in the topology, defined by the norm in X, is called* ***strong*** *convergence:*

$$(x_n) \subset (X, \|\cdot\|_X), \quad \|x_n - x\|_X \to 0 \quad \Rightarrow \quad x_n \to x \quad \textit{strongly.}$$

There are fewer open sets in the weak topology. Hence, if a sequence strongly converges, then it weakly converges.

An open neighborhood of zero in strong topology on X can be given by

$$B_r(0) = \{x \in X | \ \|x\|_X < r\} \quad (r > 0).$$

What is an open neighborhood of zero in a weak topology on X?

Given any $\epsilon > 0$ and any finite set of continuous linear functionals

$$f_1, \dots, f_n \in X^*,$$

let us consider the set

$$U = U_{f_1,\dots,f_n;\epsilon} = \{x \in X | \ |f_i(x)| < \epsilon, \quad i = 1, \dots, n\}. \tag{5.1}$$

The set U is open in X and contains the point zero, *i.e.*, U is a neighborhood of zero.

The intersection of two such neighborhoods contains a set of the same type as U (5.1). Therefore, the system of all sets of the form (5.1) generates a topology, the weak topology on X (see [2] for the proof, it can be omitted).

Every subset of X which is open (respectively closed) in the weak topology is also open (respectively closed) in the strong topology of X, but the converse may not be true. As we will see in Theorem 5.1.2, it is true when X is a finite dimensional.

Let X be an infinite dimensional normed space. Then, in particular,

1. <u>the set $S = \{x \in X| \ \|x\|_X = 1\}$ is never closed in a weak topology $\sigma(X, X^*)$:</u>

   if we denote by $\overline{S}^{\sigma(X,X^*)}$ the closure of S in the topology $\sigma(X, X^*)$, then

   $$\overline{S}^{\sigma(X,X^*)} = \{x \in X| \ \|x\|_X \leq 1\}.$$

   The proof can be found in [2], based on the fact that in infinite dimensional space each neighborhood (in the topology $\sigma(X, X^*)$) U of a point x_0 contains a straight line passing by x_0.

2. <u>the set $B_1(0) = \{x \in X| \ \|x\|_X < 1\}$ is never open in a weak topology $\sigma(X, X^*)$.</u>

   Let us verify that the set of interior points of $B_1(0)$ in $\sigma(X, X^*)$ is empty. Suppose the inverse, that there exists $x_0 \in B_1(0)$ and a neighborhood V of x_0 for $\sigma(X, X^*)$ such that $V \subset B_1(0)$. Therefore, V contains a straight line passing by x_0. This is a contradiction with $V \subset B_1(0)$.

All closed sets for the weak topology $\sigma(X, X^*)$ are closed for the strong topology. For the convex sets the notions are equivalent:

Theorem 5.1.1 *Let X be a Banach space and let $A \subset X$ be a convex set. Then A is closed in $\sigma(X, X^*)$ (or weakly closed) iff A is closed in the strong topology on X (strongly closed).*

Let us now consider the properties of the weak convergence:

Proposition 5.1.1 *1. Weak limit is unique.*

2. If $x_n \to x$, $n \to \infty$ strongly in X, then $x_n \rightharpoonup x$, $n \to \infty$ weakly in X (the converse is false).

3. If (x_k) converges weakly to x, then (x_k) is (strongly) bounded:

$$\exists C > 0 : \quad \|x_n\|_X < C \quad \text{and} \quad \|x\| \leq \liminf \|x_n\|_X,$$

where $\liminf \|x_n\|_X = \lim_{n\to\infty}(\inf_{m\geq n} x_m)$.

4. $x_n \rightharpoonup x_0 \iff$

(a) $(\|x_n\|)$ is bounded,

(b) $\langle f, x_n\rangle \to \langle f, x_0\rangle \quad \forall f \in E, \quad \overline{E} = X^*$

5. If $x_n \rightharpoonup x_0$ in X and if $f_n \to f$ strongly in X^* (i.e. $\|f_n - f\|_{\mathcal{L}(X,\mathbb{R})} \to 0$), then

$$\langle f_n, x_n\rangle \to \langle f, x\rangle.$$

Proof.

1. Let (x_n) be a sequence in X such that
$$x_n \rightharpoonup x \quad \text{and } x_n \rightharpoonup \tilde{x} \quad n \to \infty.$$
Then for all $f \in X^*$, since f is continuous,
$$\langle f, x\rangle = \langle f, \tilde{x}\rangle.$$
Therefore, since f is linear,
$$\forall f \in X^* \quad \langle f, x - \tilde{x}\rangle = 0.$$
Thus, using Corollary 3.4.2 from the Hahn-Banach theorem, we obtain that $x = \tilde{x}$.

2. We use the estimate:
$$|\langle f, x_n\rangle - \langle f, x\rangle| = |\langle f, x_n - x\rangle| \le \|f\|_{\mathcal{L}(X,\mathbb{R})}\|x_n - x\|_X.$$
We will show that the converse is false: see Example 5.1.1 and 5.3.1.

3. For all $f \in X^*$ the sequence $(\langle f, x_n\rangle)$ is bounded. But $x_n \in X \subset X^{**}$ can be considered as an element of the space X^{**}, *i.e.* the linear functional. Thus, by the theorem of Banach-Steinhaus, the sequence $(\|x_n\|)$ is bounded.

   In addition,
$$|\langle f, x_n\rangle| \le \|f\|_{\mathcal{L}(X,\mathbb{R})}\|x_n\|_X = C(f),$$
thus for $n \to +\infty$ we have
$$|\langle f, x\rangle| \le \|f\|_{\mathcal{L}(X,\mathbb{R})} \liminf \|x_n\|_X.$$
Finally,
$$\|x\|_X = \sup_{\|f\|\le 1} |\langle f, x\rangle| \le \liminf \|x_n\|_X.$$

4. It is the corollary of the theorem of Banach-Steinhaus if we consider x_n as linear functionals on X^*.

5. It is the corollary of the following estimation:
$$\begin{aligned}&|\langle f_n, x_n\rangle - \langle f, x\rangle| \le |\langle f_n - f, x_n\rangle| + |\langle f, x_n - x\rangle| \\ &\le \|f_n - f\|\|x_n\| + |\langle f, x_n - x\rangle| = \|f_n - f\|\|x_n\| + |\langle f, x_n\rangle - \langle f, x\rangle|. \quad \square\end{aligned}$$

Definition 5.1.3 *A set $M \subset X$ of a normed space X is called* ***weakly bounded*** *if for all $f \in X^*$ the numerical set*
$$\{\langle f, x\rangle, \quad x \in M\}$$
is bounded.

If M is a bounded (strongly) in X, then M is weakly bounded.

Proposition 5.1.2 *If X is a Banach space and its subset M is weakly bounded, then M is bounded in X.*

Proof. Suppose M is not bounded:

$$\exists (x_n) \subset M \subset X : \quad \|x_n\| > n^2.$$

We consider the sequence $(\frac{x_n}{n})$:

$$\forall f \in X^* \quad |\langle f, \frac{x_n}{n}\rangle| \leq \frac{1}{n} \sup_{x\in M} |\langle f, x\rangle| \leq \frac{c}{n} \to 0 \; n \to +\infty.$$

Then

$$\frac{x_n}{n} \rightharpoonup 0 \quad \text{for} \quad n \to +\infty,$$

and, thanks to Proposition 5.1.1, $(\frac{x_n}{n})$ is bounded, *i.e.*

$$\exists C > 0 : \quad \forall n \in \mathbb{N} \quad \left\|\frac{x_n}{n}\right\| < C,$$

which gives the contradiction with our assumption. □

Theorem 5.1.2 *If* $\dim X = m < \infty$, *X is a normed space and $x_n \rightharpoonup x_0$ weakly in X, then $x_n \longrightarrow x_0$ strongly in X.*

Proof. Let $\{e_n\}_{i=1}^m$ be a basis of X. Then

$$x_n = \sum_{i=1}^m \alpha_i^{(n)} e_i, \quad x_0 = \sum_{i=1}^m \alpha_i^{(0)} e_i.$$

We define $f_i \in X^*$ such that

$$\langle f_i, e_j\rangle = \delta_{i,j} = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}.$$

Therefore,

$$\alpha_i^{(n)} = \langle f_i, x_n\rangle \to \langle f_i, x_0\rangle = \alpha_i^0 \quad n \to \infty,$$

and consequently,

$$x_n = (\alpha_1^{(n)}, \dots, \alpha_m^{(n)}) \to x_0 = (\alpha_1^0, \dots, \alpha_m^0) \quad n \to \infty,$$

what implies that $x_n \to x_0$ strongly in X. (As all norms are equivalent in X, it is sufficient to show that $x_n \to x_0$ by $\|\cdot\|_{\ell^\infty}$). □

Example 5.1.1 *Consider the space $C([a,b])$ of all functions continuous on $[a,b]$ equipped with the norm*

$$\|f\|_\infty = \max_{a\leq x\leq b} |f(x)|.$$

If $\|f_n - f\|_\infty \to 0$ in $C([a,b])$, it follows that the sequence (f_n) converges uniformly to f on $[a,b]$. Thus the strong convergence in $C([a,b])$ means the uniform convergence.

Let us now consider the weak convergence in $C([a,b])$. Let (f_n) be a sequence of functions in $C([a,b])$ converging weakly to a function $f \in C([a,b])$. Among the continuous linear functionals on $C([a,b])$, we have the functionals δ_{x_0}, $a < x_0 < b$, named the Dirac delta functions, which assign to each function $f(x) \in C([a,b])$ its value at some fixed point $x_0 \in [a,b]$.

Let us show that $\delta_{x_0} \in C^*([a,b])$. Indeed, by definition δ_{x_0} is linear mapping from $C([a,b])$

to $\mathbb{R}$*:*

$$\delta_{x_0}(\alpha f + \beta g) = \alpha f(x_0) + \beta g(x_0) = \alpha\delta_{x_0}(f) + \beta\delta_{x_0}(g) \quad \forall f, g \in C([a,b]),\ \forall \alpha, \beta \in \mathbb{R},$$

and in addition we have

$$|\delta_{x_0}(f)| = |f(x_0)| \leq \max_{a \leq x \leq b} |f(x)| = \|f\|_\infty,$$

where equality holds if $f(x) = const$*. Hence* δ_{x_0} *is bounded, thus it is continuous, with norm*

$$\|\delta_{x_0}\|_{\mathcal{L}(C([a,b]),\mathbb{R})} = 1.$$

From

$$\delta_{x_0}(f_n) \to \delta_{x_0}(f)$$

follows by the definition of the functional δ_{x_0} *that*

$$f_n(x_0) \to f(x_0).$$

Hence, if the sequence (f_n) *is weakly convergent in* $C([a,b])$*, then*

1. (f_n) *is uniformly bounded on* $[a,b]$*, i.e., there is a constant* $C \leq 0$ *such that*

$$|f_n(x)| \leq C \quad \forall n \in \mathbb{N} \quad \forall x \in [a,b],$$

2. (f_n) *is pointwise convergent on* $[a,b]$*, i.e.,* $(f_n(x))$ *is a convergent numerical sequence for every fixed* $x \in [a,b]$*.*

We can see that the strong convergence in $C([a,b])$ *implies the weak convergence, but not the converse.*

We give without proof the following result:

Theorem 5.1.3 *Let* X *be a Banach space.* X *is reflexive if and only if each bounded sequence in* X *contains a subsequence which converges weakly in* X*.*

When there are fewer open sets, thus it is more easy to be convergent, and thus to be compact too (the compact sets are very important for theorems of existence):

Theorem 5.1.4 *(**Kakutani**) Let* X *be a Banach space.* X *is reflexive iff the closed unit ball associated to the norm* $\overline{B_1(0)} = \{x \in X, \|x\| \leq 1\}$ *is compact in the weak topology* $\sigma(X, X^*)$ *.*

Remark 5.1.1 *As any Hilbert space* H *is reflexive, then the closed unit ball in a Hilbert space is weakly compact.*

Remark 5.1.2 *Moreover, for the infinite dimensional case, we saw that a closed unit ball* $\overline{B_1(0)}$ *is not compact in the strong topology. Theorem of Kakutani states that* $\overline{B_1(0)}$ *is compact in the weak topology iff* X *is a reflexive Banach space. Can we find a topology for which* $\overline{B_1(0)}$ *is compact even if* X *is not reflexive?*

5.2 Weak* convergence in a Banach space

Let X be a Banach space. Let X^* be its dual. Let X^{**} be its bidual (containing X). In X we can define

- the **strong topology**, defined by the norm $\|\cdot\|_X$ in X;
- the **weak topology**, defined by the weak convergence.

As we know, for all normed spaces X, its dual space X^* is a Banach space with the norm:

$$\|f\|_{\mathcal{L}(X,\mathbb{R})} = \sup_{x\in X, \|x\|_X \leq 1} |\langle f, x\rangle|.$$

Thus, X^* can be equipped with:

- the strong topology on X^* (defined by the convergence by the norm on X^*);
- the weak topology on X^* (defined by the weak convergence on X^*, that is, in notations of Appendix A, $\sigma(X^*, X^{**})$).

In fact, there are two ways of regarding the space X^* of continuous linear functionals on a given space X:

1. as an "original space" in its own right, with conjugate space X^{**},
2. as the space conjugate to the original space X.

These two points of view gives two different topologies (convergences):

- the weak topology on X^* (that is $\sigma(X^*, X^{**})$);
- the weak* topology on X^* defined by $\sigma(X^*, X)$ (the weak* convergence is denoted by $\overset{*}{\rightharpoonup}$).

Definition 5.2.1 *Let (f_n) be a sequence in X^*, the dual space to the normed space X. We say that the sequence of functionals (f_n) converges **weakly*** to $f \in X^*$, denoted by $f_n \overset{*}{\rightharpoonup} f$, if*

$$\langle f_n, x\rangle \to \langle f, x\rangle \quad \forall x \in X.$$

*The corresponding topology is called **weak* topology**, noted by $\sigma(X^*, X)$ (see also Definition A.5.2).*

Since $X \subset X^{**}$, the weak* topology $\sigma(X^*, X)$ is weaker than the weak topology $\sigma(X^*, X^{**})$, which is weaker than the strong topology, *i.e.* on X^* we always have

$$\boxed{\to} \quad \Rightarrow \quad \boxed{\rightharpoonup} \quad \Rightarrow \quad \boxed{\overset{*}{\rightharpoonup}},$$

but not converse.

Problem 5.2.1 *Show that for all finite sets $A \subset X$*

$$U_{A,\epsilon} = \{f \in X^* | \ |f(x)| < \epsilon \text{ for all finite } A\}$$

is an open neighborhood of zero in the weak* topology $\sigma(X^*, X)$.

Clearly, the weak convergence and the weak* convergence on X^* be the same if and only if X is reflexive. In particular, if X is a Hilbert space then $X^{**} = X$, therefore the weak* topology and the weak topology coincide.

Proposition 5.2.1 *Let (f_n) be a sequence in X^*, the dual space to the Banach space X. We have*

1. *The weak* limit is unique.*
2. $f_n \overset{*}{\rightharpoonup} f$ *in* $\sigma(X^*, X)$ *iff* $\langle f_n, x\rangle \to \langle f, x\rangle \quad \forall x \in X$
3. *If* $f_n \to f$, $n \to \infty$ *strongly in* X^*, *then* $f_n \rightharpoonup f$, $n \to \infty$ *weakly in* $\sigma(X^*, X^{**})$.
4. *If* $f_n \rightharpoonup f$, $n \to \infty$ *weakly in* $\sigma(X^*, X^{**})$, *then* $f_n \overset{*}{\rightharpoonup} f$ *in* $\sigma(X^*, X)$.
5. *If* $f_n \overset{*}{\rightharpoonup} f$ *in* $\sigma(X^*, X)$, *then* $\|f_n\|$ *is bounded:*
$$\exists C > 0 : \ \forall n \quad \|f_n\|_{X^*} < C \quad \text{and} \quad \|f\| \leq \liminf \|f_n\|_{X^*},$$
where $\liminf \|f_n\|_{X^*} = \lim_{n\to\infty}(\inf_{m\geq n} \|f_m\|_{X^*})$.
6. $f_n \overset{*}{\rightharpoonup} f \iff$
   (a) $(\|f_n\|)$ *is bounded,*
   (b) $\langle f_n, x\rangle \to \langle f, x\rangle \quad \forall x \in E, \quad \overline{E} = X$
7. *If* $f_n \overset{*}{\rightharpoonup} f$ *in* $\sigma(X^*, X)$ *and if* $x_n \to x$ *strongly in* X, *then*
$$\langle f_n, x_n\rangle \to \langle f, x\rangle.$$

Remark 5.2.1 *We recall that the notation $\langle f, x\rangle$ means the value of f at x: $f(x)$.*

The proof of Proposition 5.2.1 follows the proof of Proposition 5.1.1.

Let us prove Point 6). $\underline{\Rightarrow}$ It is obvious.

$\underline{\Leftarrow}$ If z is a linear combination of elements in E, then $\langle f_n, z\rangle \to \langle f, z\rangle$.

Let x now be an arbitrary element of X, and let (z_k) be a sequence of linear combinations of elements of E converging to x in X (such a sequence exists, since E is dense in X). Let us show that $\langle f_n, x\rangle \to \langle f, x\rangle$.

Let C be such that
$$\|f_n\| \leq C \quad \forall n \in \mathbb{N} \qquad \text{and } \|f\| \leq C.$$

Moreover, given any $\epsilon > 0$, choose k large enough so that
$$\|\langle f_n, z_k\rangle - \langle f, z_k\rangle\| < \epsilon$$
(this is possible, since $\langle f_n, z\rangle \to \langle f, z\rangle$ for all z in E).

Then
$$\begin{aligned}&|\langle f_n, x\rangle - \langle f, x\rangle| \leq |\langle f_n, x\rangle - \langle f_n, z_k\rangle| + |\langle f_n, z_k\rangle - \langle f, z_k\rangle| \\ &+|\langle f, z_k\rangle - \langle f, x\rangle| \leq \|f_n\|\|x - z_k\| + \epsilon + \|f\|\|z_k - x\| \leq \epsilon(1 + 2C).\end{aligned}$$

Therefore $f_n \overset{*}{\rightharpoonup} f$. $\square$

Let us finish by

Theorem 5.2.1 ***(Banach-Alaoglu-Bourbaki)*** *Let X be a normed space. The set $B_{X^*} = \{f \in X^* | \, \|f\| \leq 1\}$ is compact in the weak* topology $\sigma(X^*, X)$.*

Instead of this general result, let us prove

Theorem 5.2.2 *Let X be a separable normed linear space. Every bounded sequence (f_n) of linear bounded functionals, $f_n \in X^*$, contains a weakly* convergent subsequence.*

Proof. Since X is separable, there is a countable set of points

$$x_1, x_2, \ldots, x_n, \ldots$$

dense in X.

Suppose the sequence (f_n) of functionals in X^*, *i.e.,* continuous linear functionals on X, is bounded (in norm).

Then the numerical sequence

$$f_1(x_1), f_2(x_1), \ldots, f_n(x_1), \ldots$$

is bounded, and hence, by the Bolzano-Weierstrass theorem, (f_n) contains a subsequence

$$f_1^{(1)}, f_2^{(1)}, \ldots, f_n^{(1)}, \ldots$$

such that the numerical sequence

$$f_1^{(1)}(x_1), f_2^{(1)}(x_1), \ldots, f_n^{(1)}(x_1), \ldots$$

converges.

By the same token, the subsequence $(f_n^{(1)})$ in turn contains a subsequence

$$f_1^{(2)}, f_2^{(2)}, \ldots, f_n^{(2)}, \ldots$$

such that the sequence

$$f_1^{(2)}(x_2), f_2^{(2)}(x_2), \ldots, f_n^{(2)}(x_2), \ldots$$

converges. Continuing this construction, we get a system of subsequences $(f_n^{(k)})$, $k = 1, 2, \ldots$ such that

- $(f_n^{(k+1)})$ is a subsequence of $(f_n^{(k)})$ for all $k = 1, 2, \ldots$;
- $(f_n^{(k)})$ converges at the points $x_1, x_2, \ldots, x_k$.

Hence, taking the "diagonal sequence"

$$f_1^{(1)}, f_2^{(2)}, \ldots, f_n^{(n)}, \ldots,$$

we get a sequence of continuous linear functionals on X such that

$$f_1^{(1)}(x_n), f_2^{(2)}(x_n), \ldots, f_n^{(n)}(x_n), \ldots,$$

converges for all n. But then, by Proposition 5.2.1 point 6, the sequence

$$f_1^{(1)}(x), f_2^{(2)}(x), \ldots, f_n^{(n)}(x), \ldots,$$

converges for all $x \in X$. $\square$

Remark 5.2.2 *Let X be a separable normed linear space. Let B and B^* be the unit closed balls in X and X^* respectively. Then the topology induced in B^* by the weak* topology in X^* is metrizable by the metric*

$$d(f,g) = \sum_{n=1}^{\infty} 2^{-n} |\langle f-g, x_n\rangle|,$$

where $\{x_1, \ldots, x_n, \ldots\}$ is any fixed countable dense set in B.

As in metric space the sequentially compactness is equivalent to the compactness, then, thanks to Theorem 5.2.2, we can conclude that B^* is compact in $\sigma(X^*, X)$ if we prove that B^* is bounded in $\sigma(X^*, X)$.

Consequently, let us prove

Theorem 5.2.3 *Let X be a separable normed linear space. Every closed ball in the space X^* (closed by the strong topology in X^*) is compact in the weak* topology.*

Proof Let us prove actually that

Every closed ball in the space X^* (closed by the strong topology in X^*) is closed in the weak* topology.

In fact, since a shift in X^* carries every closed set (in the weak* topology) into another closed set, we need only prove the assertion for every ball of the form

$$B_r^* = \{f \in X^* | \|f\| \leq r\}.$$

Suppose $f_0 \notin B_r^*$. Then, by the definition of the norm of the functional f_0, there is an element $x \in X$ such that

$$\|x\| = 1 \quad \text{and } f_0(x) = \alpha > r.$$

But then the set

$$U = \left\{ f \in X^* | f(x) > \frac{\alpha + r}{2} \right\}$$

is a weak* neighborhood of f_0 containing no elements of B_r^*. Therefore, B_r^* is closed in the weak* topology.

By Theorem 5.2.2 and by the fact (without proof, see Remark 5.2.2) that any closed ball in X^* is a metric space for $\sigma(X^*, X)$, we conclude that B_r^* is compact in $\sigma(X^*, X)$. □.

5.3 Strong and weak convergence in a Hilbert space

Thanks to the Riesz Representation Theorem (see Chapter 4), all linear continuous functionals on H can be uniquely presented by the inner product in H. Therefore, a sequence (x_n) converges weakly to x in a Hilbert space H, if

$$\forall v \in H \quad \lim_{n\to\infty} \langle x_n, v\rangle = \langle x, v\rangle,$$

where $\langle \cdot, \cdot \rangle$ is the inner product in H.

In addition, if H is a Hilbert space, then H is reflexive, since H^* is isometric to H. Thus, the weak topology on H is equal to the weak* topology on H. Thanks to the Kakutani Theorem, every bounded set in H (in strong topology) is weakly compact: from a bounded sequence (x_n) in H one can extract a subsequence weakly converging in H.

However, the weak topology in H is still coarser than the strong topology:

Example 5.3.1 *($\rightharpoonup \not\Rightarrow \longrightarrow$) Let H be a Hilbert space and (e_k) be its orthonormal basis. Then*

$$\forall x \in H \quad \langle x, e_k\rangle \to 0 \quad k \to \infty$$

since $\langle x, e_k\rangle$ are Fourrier coefficients of x. Consequently, $e_k \rightharpoonup 0$. But if $n \neq m$ $\|e_n - e_m\|^2 = \langle e_n - e_m, e_n - e_m\rangle = 2$, thus (e_k) is not a Cauchy sequence in H and then (e_k) does not converge.

We know that the inner product is continuous by the strong convergence in H:

$$x_n \to x, \quad y_n \to y \quad \Rightarrow \quad \langle x_n, y_n\rangle \to \langle x, y\rangle.$$

If $x_n \rightharpoonup x, \quad y_n \rightharpoonup y$, it does not imply that $\langle x_n, y_n\rangle \to \langle x, y\rangle$. Let us give an example:

Example 5.3.2 *Let (e_n) be an orthonormal sequence in H and $x_n = y_n = e_n$. Then $e_n \rightharpoonup 0$ and*

$$\langle e_n, e_n\rangle = \|e_n\|^2 = 1 \nrightarrow 0 = \langle 0, 0\rangle.$$

Proposition 5.3.1 *Let H be a real Hilbert space. If $x_n \to x$ and $y_n \rightharpoonup y$ in H, then*

$$\langle x_n, y_n\rangle \to \langle x, y\rangle.$$

Proof. Since $y_n \rightharpoonup y$ in H, for all $n \in \mathbb{N}$ the norms $\|y_n\|$ are bounded. Let

$$M = \sup_n \|y_n\|.$$

Then, by the Cauchy-Schwartz inequality, we have

$$|\langle x_n, y_n\rangle - \langle x, y\rangle| \leq |\langle x_n - x, y_n\rangle| + |\langle x, y_n - y\rangle| \leq M\|x_n - x\| + |\langle x, y_n - y\rangle|.$$

Since $y_n \rightharpoonup y$ in H, then $|\langle x, y_n - y\rangle| \to 0$, and since $x_n \to x$, then $\|x_n - x\| \to 0$. Therefore, we conclude that $|\langle x_n, y_n\rangle - \langle x, y\rangle| \to 0$. $\square$

Let us give an example for the following property of weak convergence: if (x_n) converges weakly, then (x_n) is bounded:

$$\exists C > 0 : \quad \|x_n\|_H < C \quad \text{and} \quad \|x\| \leq \liminf \|x_n\|_H.$$

Example 5.3.3 *Let us consider $L^2([0,1])$ and*

$$x_n(t) = \sqrt{2}\sin(\pi n t).$$

The sequence (x_n) *is an orthonormal basis of* $L^2([0,1])$. *We have*

$$\|x_n\| = 1 \quad \Rightarrow \quad \lim_n \|x_n\| = 1.$$

For $\alpha(t) \in L^2([0,1])$ *we define, with notation* c_n *for Fourier coefficients,*

$$f(x_n) = \sqrt{2}\int_0^1 \alpha(t)\sin(\pi n t)dt = \sqrt{2}c_n$$

such that

$$f(x_n) \to 0,\ n \to \infty, \quad \textit{and} \quad x_n \rightharpoonup 0.$$

Thus, the limit $x = 0$ *and*

$$\|x\| = 0 < 1 = \lim_n \|x_n\|.$$

In a Hilbert space it holds

Proposition 5.3.2 *Let* H *be a real Hilbert space. Let* (x_n) *be a weakly converging sequence to* x. *Further assume that* $\|x_n\|$ *converges to* $\|x\|$. *Then* (x_n) *converges strongly to* x.

Proof. We have

$$\langle x_n - x, x_n - x\rangle = \langle x_n, x_n\rangle - \langle x_n, x\rangle - \langle x, x_n\rangle + \langle x, x\rangle.$$

For $n \to \infty$, we find that

$$\lim_{n\to\infty} \langle x_n - x, x_n - x\rangle = 0,$$

which means that $\|x_n - x\| \to 0$, *i.e.* $x_n \to x$ in H. $\square$

We can also obtain the strong convergence from the weak convergence in a Banach space in the following way:

Theorem 5.3.1 *Let* X *be a Banach space. Let* (x_n) *be a sequence of elements of* X *such that*

$$x_n \rightharpoonup x \quad \textit{in } X.$$

Then there exists a subsequence of linear combinations of the elements of (x_n), *denoted by*

$$\left(\sum_{k=1}^{k_n} c_k^{(n)} x_k\right),$$

which converges strongly to l *in* X.

Proof. Let's notice that theorem states that l belongs to a linear closed subspace $L = \overline{\mathrm{Span}((x_n)_{n\in\mathbb{N}^*})}$.

Suppose the converse, that $x \notin L$. Then, thanks to Corollary 3.4.3 of the Hahn-Banach Theorem from Chapter 3, there exists $f \in X^*$ such that

$$f(x) = 1, \quad \text{and } \forall n \quad f(x_n) = 0.$$

But it means that $f(x_n) \not\rightharpoonup f(x)$, which is in contradiction with the assumption that $x_n \rightharpoonup x$ in X. $\square$

Problem 5.3.1 *Give an example of a compact and a weakly compact set in a Hilbert space* H.

Chapter 6

Compact operators and spectral theory

6.1 Compact operators

Definition 6.1.1 *Let X and Y be Banach spaces. A linear operator $A : X \to Y$ is* ***compact****, if it maps all bounded sets of the space X to a relatively compact set of the space Y.*

Remark 6.1.1 *We recall that $M \subset Y$ is a relatively compact set of the space $(Y, \|\cdot\|)$ if its closure $\overline{M}$ is compact in $(Y, \|\cdot\|)$.*

Remark 6.1.2 *Let $B_1 = \{f \in X | \|f\|_X \leq 1\}$ be the closed unit ball in X. The operator A is compact iff AB_1 is a relatively compact set in Y.*

Indeed, if A is compact operator, then obviously AB_1 is a relatively compact set in Y.

Conversely, let AB_1 be a relatively compact set in Y. For all bounded sets M in X there exist $r > 0$, the radius of the ball including M:

$$M \subset B_r = rB_1.$$

Consequently, using the linearity of A,

$$AM \subset rAB_1.$$

Since AB_1 is a relatively compact set, thus rAB_1 too, what implies that AM is relatively compact (see Corollary B.1.1 and Theorem B.1.2).

Remark 6.1.3 *Let us denote by $K(X, Y)$ the set of all compact operators from X to Y (X and Y are Banach spaces).*

- *All compact operators from $(X, \|\cdot\|_X)$ to $(Y, \|\cdot\|_Y)$ are bounded operators:*

$$K(X, Y) \subset \mathcal{L}(X, Y).$$

Indeed, it is sufficient to notice that if AB_1 is relatively compact, then AB_1 is bounded.

- *If* $\dim X < \infty$ *or* $\dim Y < \infty$*, then the compactness of the operators is equivalent to the boundness:*
$$K(X,Y) = \mathcal{L}(X,Y).$$
Indeed, let $\dim X < \infty$*, then* AX *is finite dimensional. In addition,* $AB_1 \subset AX$ *and all bounded set in* AX *is relatively compact in* Y*. If* $\dim Y < \infty$*, we have* $AB_1 \subset Y$ *and there is no difference between the relatively compactness and the boundness in the final dimensional space.*

Theorem 6.1.1 *Let* X *and* Y *be Banach spaces. The set of all compact operators from* $(X, \|\cdot\|_X)$ *to* $(Y, \|\cdot\|_Y)$*, denoted by* $K(X,Y)$*, is a closed vector subspace of* $\mathcal{L}(X,Y)$*. In addition,* $K(X,X)$ *forms a closed ideal in* $\mathcal{L}(X,X)$*, i.e if* A *is a compact operator and* S *is a bounded operator, then the operators* AS *and* SA *are compact.*

Proof.

1. A is a compact operator $\Rightarrow$ $\forall \lambda \in \mathbb{C}$ λA is compact operator.

   It is obvious.

2. A and S are compact operators $\Rightarrow$ $A + S$ is compact operator.

   We have
$$(A+S)B_1 = \{Ax + Sx |\ x \in B_1\} \subset AB_1 + SB_1 = \{Ax + Sy |\ x, y \in B_1\}.$$
   Sets AB_1 and SB_1 are relatively compact in Y. In addition, the function $f(x,y) = x + y$ is continuous as a mapping $(x,y) \in Y^2 \mapsto x + y \in Y$, which implies that $AB_1 + SB_1 = f(AB_1, SB_1)$ is also relatively compact in Y (see Theorem B.2.1).

3. (A_n) is a sequence of compact operators such that $A_n \to A$ in $\mathcal{L}(X,Y)$ $\Rightarrow$ A is compact operator.

   We suppose that for all $n \in \mathbb{N}$ the sets $A_n B_1$ are relatively compact. We want to prove that AB_1 is a relatively compact set. Let us define for a fixed $x \in B_1$
$$y = Ax \in AB_1, \quad y_n = A_n x \in A_n B_1.$$
   Thus we have, using A, $A_n \in \mathcal{L}(X,Y)$, that for all $x \in B_1$
$$\|y - y_n\|_Y = \|(A - A_n)x\|_Y \leq \|A - A_n\|_{\mathcal{L}(X,Y)} \|x\|_X \leq \|A - A_n\|_{\mathcal{L}(X,Y)}.$$
   Therefore, uniformly on $x \in B_1$, we have
$$\forall \epsilon > 0 \quad \exists n_0(\epsilon) \in \mathbb{N}: \quad \forall n \geq n_0(\epsilon) \quad \|y - y_n\|_Y \leq \epsilon. \tag{6.1}$$
   Now, we use Theorem 2.3.4 of Chapter 2: A subset M of a complete metric space (E,d) is relatively compact if and only if it is totally bounded. Thus, M is relatively compact in a complete metric space (E,d) iff for all $\epsilon > 0$ there exists a finite ϵ-net of M, or equivalently,
$$\forall \epsilon > 0 \quad \exists m \in \mathbb{N} \text{ and } z_1, \ldots, z_m \in E: \quad M \subset \cup_{i=1}^m B_\epsilon(z_i),$$
   where $B_\epsilon(z_i) = \{x \in E |\ d(x, z_i) \leq \epsilon\}$.

Let us fixe $n \geq n_0(\epsilon)$ with $n_0(\epsilon)$ defined previously in the convergence of A_n to A. Let us prove that the existence of the ϵ-net for A_nB_1, which is relatively compact by the assumption, implies the existence of the 2ϵ-net for AB_1:

$$A_nB_1 \subset \cup_{i=1}^m B_\epsilon(z_i) \quad \Rightarrow \quad AB_1 \subset \cup_{i=1}^m B_{2\epsilon}(z_i).$$

Indeed, given $x \in B_1$ we define as previously $y = Ax \in AB_1$, and $y_n = A_nx \in A_nB_1$. For y_n there exists $i \in [1, \dots, m]$ such that

$$\|y_n - z_i\|_Y \leq \epsilon,$$

by the definition of the ϵ-net. Consequently, we also have

$$\|y - z_i\|_Y \leq \|y - y_n\|_Y + \|y_n - z_i\|_Y \leq \epsilon + \epsilon = 2\epsilon,$$

where $\|y - y_n\|_Y \leq \epsilon$ thanks to (6.1). Hence, $\cup_{i=1}^m B_{2\epsilon}(z_i)$ is the 2ϵ-net of AB_1 and consequently, AB_1 is relatively compact by Theorem **??**. Thus we conclude that A is compact.

4. $\underline{A \in K(X,X),\ S \in \mathcal{L}(X,X) \quad \Rightarrow \quad AS \text{ and } SA \text{ are compact operators.}}$

By AS and SA we understand $A \circ S$ and $S \circ A$ respectively. Since S maps bounded sets to bounded sets, we have that there exists $r > 0$ such that

$$(AS)B_1 = A(SB_1) \subset AB_r,$$

and since A is compact and B_r is bounded, thus AB_r is relatively compact. In other hand, $(SA)B_1 = S(AB_1)$, where AB_1 is relatively compact. Since $S \in \mathcal{L}(X,X)$, S is continuous and thus preserve the property of the relatively compactness (see Theorem B.2.1).

6.2 Adjoint operator on Banach spaces

Definition 6.2.1 *Let X and Y be normed vector spaces and $A \in \mathcal{L}(X,Y)$. The operator $A^* : Y^* \to X^*$, $A^* \in \mathcal{L}(Y^*, X^*)$ is called the **adjoint operator** to A if it is defined by*

$$\forall x \in X \quad \forall f \in Y^* \quad \langle f, Ax\rangle = \langle A^*f, x\rangle,$$

where the notation $\langle f, x\rangle$ means the value of f on x, $f(x)$.

Theorem 6.2.1 *Let A be a compact operator mapping a Banach space X into itself. Then the adjoint operator A^* is also compact.*

Proof. As $A : X \to X$, then $A^* : X^* \to X^*$. Let B_1^* be a closed unit ball in X^*. Let us prove that the set $A^*B_1^*$ is relatively compact in X^*. Now suppose we consider the elements of X^* as functionals not on the whole space X, but only on the compact set $\overline{AB_1}$, where, in our notations, $\overline{AB_1}$ is the closure of the image of the closed unit ball in X under the operator A. We define the set of functionals

$$\Phi \subset X^*, \quad \Phi = \{f \in (\overline{AB_1})^* | \ \|f\|_{X^*} \leq 1\}.$$

Thus Φ is uniformly bounded and uniformly equicontinuous (see Ascoli-Arzela's theorem, Theorem 2.3.7), since for all $f \in \Phi$ (thus $\|f\|_{X^*} \leq 1$)

$$\begin{aligned}
&\sup_{x\in\overline{AB_1}} |f(x)| = \sup_{x\in AB_1} |f(x)| = \sup_{x\in AB_1, x\neq 0} \left(\frac{|f(x)|}{\|x\|_X}\|x\|_X\right) \\
&\leq \sup_{x\in AB_1, x\neq 0} \left(\frac{|f(x)|}{\|x\|_X|}\right) \sup_{x\in AB_1} \|x\|_X \leq \sup_{x\in X, x\neq 0} \left(\frac{|f(x)|}{\|x\|_X}\right) \sup_{x\in B_1} \|Ax\|_X = \|f\|_{X^*}\|A\| \leq \|A\|,
\end{aligned}$$

and

$$|f(x_1) - f(x_2)| \leq \|f\|_{X^*}\|x_1 - x_2\|_X \leq \|x_1 - x_2\|_X.$$

Consequently, thanks to Ascoli-Arzela's theorem, the set Φ is relatively compact in $C(\overline{AB_1})$.

Let us prove that the set Φ with the metric induced by the usual metric of the space of continuous functions $C(\overline{AB_1})$, is isometric to the set $A^*B_1^*$, with the metric induced by the norm of the space X^*.

In fact, if $f_1,\ f_2 \in B_1^*$, then

$$\begin{aligned}
&\|A^*f_1 - A^*f_2\|_{X^*} = \sup_{x\in B_1} |\langle A^*f_1 - A^*f_2, x\rangle| = \sup_{x\in B_1} |\langle f_1 - f_2, Ax\rangle| \\
&= \sup_{z\in AB_1} |\langle f_1 - f_2, z\rangle| = \sup_{z\in\overline{AB_1}} |\langle f_1 - f_2, z\rangle| = d_{C(\overline{AB_1})}(f_1, f_2).
\end{aligned}$$

Being relatively compact, the set Φ is totally bounded, by Theorem **??**. Therefore the set $A^*B_1^*$ isometric to Φ is also totally bounded, and hence relatively compact, by the same theorem. $\square$

6.3 Spectral properties of compact operators in a Hilbert space

First we notice that

1. Let X be a normed vector space and $A \in \mathcal{L}(X, X)$, then the norm of A is defined by

$$\|A\| = \sup_{x\neq 0} \frac{\|Ax\|}{\|x\|}.$$

2. Let H be a Hilbert space and $A \in \mathcal{L}(H, H)$, then

$$\|A\| = \sup_{x,y\neq 0} \frac{|\langle Ax, y\rangle|}{\|x\|\|y\|}.$$

Theorem 6.3.1 *Let H be a Hilbert space. For all self-adjoint operator $A \in \mathcal{L}(H, H)$ we have*

$$\|A\| = \sup_{x\neq 0} \frac{|\langle Ax, x\rangle|}{\|x\|^2}. \tag{6.2}$$

Proof. Let
$$\alpha = \sup_{\|x\|=1} |\langle Ax, x\rangle|, \quad \beta = \sup_{\|x\|=\|y\|=1} |\langle Ax, y\rangle| = \|A\|.$$
We see that
$$\alpha \leq \beta$$
since
$$|\langle Ax, x\rangle| \leq \|Ax\|\|x\| \leq \|A\|\|x\|^2,$$
where we take the supremum over all x with $\|x\| = 1$.

Let us prove that $\beta \leq \alpha$.

We define
$$f(x, y) = \langle Ax, y\rangle, \quad \phi(x) = \langle Ax, x\rangle,$$
where ϕ is real and quadratic form. Since $A = A^*$, we verify that the real part of f can be found by
$$\operatorname{Re} f(x, y) = \frac{1}{4}[\phi(x + y) - \phi(x - y)].$$
(By information, the imaginary part is given by $\operatorname{Im} f(x, y) = \frac{1}{4}[-\phi(x + iy) + \phi(x - iy)]$.) Let us show that
$$\forall x \in H \quad \phi(x) = \langle Ax, x\rangle \leq \alpha\|x\|^2.$$
Indeed, for $\|x\| \leq 1$, thanks to the definition of α, it holds
$$|\langle Ax, x\rangle| \leq \alpha.$$
Thus, for $x \neq 0$ we have
$$\langle A\frac{x}{\|x\|}, \frac{x}{\|x\|}\rangle \leq \alpha,$$
which gives $\langle Ax, x\rangle \leq \alpha\|x\|^2$.

We fixe x and y and write
$$\langle Ax, y\rangle = \rho e^{i\mu},$$
where $i = \sqrt{-1}$ and ρ and μ are corresponding real numbers. Thus, since $\phi(x) \leq \alpha\|x\|^2$, we have
$$\begin{aligned} &|\langle Ax, y\rangle| = \rho = \langle Ax, ye^{-i\mu}\rangle = \frac{1}{4}[\phi(x + ye^{-i\mu}) - \phi(x - ye^{-i\mu})] \\ &\leq \frac{\alpha}{4}[\|x + ye^{-i\mu}\|^2 + \|x - ye^{-i\mu}\|^2]. \end{aligned}$$
We use the parallelogram law for the norm $\|\cdot\| = \sqrt{\langle\cdot,\cdot\rangle}$ in H:
$$\|x + y\|^2 + \|x - y\|^2 = 2(\|x\|^2 + \|y\|^2)$$
and we obtain
$$\begin{aligned} &|\langle Ax, y\rangle| \leq \frac{\alpha}{4}[\|x + ye^{-i\mu}\|^2 + \|x - ye^{-i\mu}\|^2] \\ &= \frac{\alpha}{2}(\|x\|^2 + \|y\|^2) = \alpha \quad \text{for } \|x\| = \|y\| = 1. \end{aligned}$$
Consequently, we proved that $\beta \leq \alpha$, hence $\alpha = \beta$. $\square$

Definition 6.3.1 *Let A be a linear operator mapping a topological linear space X into itself. Then a number λ is called an* ***eigenvalue*** *of A if the equation $Ax = \lambda x$ has at least one nonzero solution, and every such solution x is called an* ***eigenfunction*** *(or* ***eigenvector****) of A (corresponding to the eigenvalue λ).*

Theorem 6.3.2 *Let $H \neq \{0\}$ be a Hilbert space and $A : H \to H$ be a compact self-adjoint operator on H. Then at least one of the numbers $+\|A\|$ or $-\|A\|$ is an eigenvalue of A.*

Proof. We notice that

$$\|A\| = \sup_{\|x\|=1} |\langle Ax, x\rangle|, \quad |\langle Ax, x\rangle| \leq \|Ax\|\|x\| \leq \|A\|\|x\|^2.$$

Consequently, there exists a sequence (x_n) such that $\|x_n\| = 1$ for all n and

$$|\langle Ax_n, x_n\rangle| \to \|A\|.$$

In addition, it implies that $\|Ax_n\| \to \|A\|$. Therefore, there exists a subsequence (x_{n_k}) such that

$$\langle Ax_{n_k}, x_{n_k}\rangle \to +\|A\| \text{ or } -\|A\|.$$

To avoid the confusion, we denote the limit by α: $\langle Ax_{n_k}, x_{n_k}\rangle \to \alpha$. To simplify the notations, we also write x_k instead of x_{n_k}. Thus, we have

$$\|Ax_n - \alpha x_n\|^2 = \|Ax_n\|^2 - \alpha\langle Ax_n, x_n\rangle - \alpha\langle x_n, Ax_n\rangle + \alpha^2\|x_n\|^2 \to \alpha^2 - 2\alpha^2 + \alpha^2 = 0,$$

i.e. $z_n = Ax_n - \alpha x_n \to 0$ in H. We denote $Ax_n = y_n$. Since for all n $\|x_n\| = 1$, thus (x_n) is a bounded set in H. As A is a compact operator, then (Ax_n) is a relatively compact set in H. Moreover, there exists a subsequence (Ax_{n_j}) which converges in H. We denote by y its limit:

$$Ax_{n_j} = y_{n_j} \to y \quad j \to \infty.$$

In addition,

$$\alpha x_{n_j} = y_{n_j} - z_{n_j} \to y.$$

If $\alpha = 0$, it means that $A = 0$ and all vectors $x \in H$ are the eigenvectors of A. If $\alpha \neq 0$, we devide by α and obtain

$$x_{n_j} \to \frac{y}{\alpha} \overset{def}{=} x.$$

Thus we have that there is a sequence (x_n) with $\|x_n\| = 1$ for $n \in \mathbb{N}$ such that

$$x_n \to x \text{ in } H \quad \text{and} \quad y_n = Ax_n \to y.$$

Since A is continuous (and thus closed), it follows that $y = Ax$ and therefore, $Ax = \alpha x$ with $\|x\| = 1$. □

Example 6.3.1 ***(Importance of compactness)*** *Let $H = L^2([a, b])$ and $Ax(t) = tx(t)$ for $t \in [a, b] \subset \mathbb{R}$. This operator is self-adjoint but not compact. We can see that if*

$$(t - \lambda)x(t) = 0 \text{ a.e. in } [a, b]$$

and $t \neq \lambda$ then $x(t) = 0$ a.e. in $[a, b]$, thus $\|x\| = 0$, which implies that $x = 0$ in H. Consequently, A does not have eigenvectors in H.

Let us prove

Lemma 6.3.1 *Let H be a Hilbert space and $A : H \to H$, $A \in \mathcal{L}(H, H)$ be a self-adjoint operator in H. Then*

1. *all eigenvalues of A are real.*
2. *If x and y are eigenvectors of A corresponding to the eigenvalues λ and μ respectively, such that $\lambda \neq \mu$, then x is orthogonal to y in H.*
3. *If H_0 is a subspace of H and $AH_0 \subset H_0$, then $AH_0^\perp \subset H_0^\perp$.*

Proof.

1. Let $x \neq 0$ be a eigenvector of A corresponding to λ:
$$Ax = \lambda x.$$
It implies that
$$\langle Ax, x\rangle = \lambda\langle x, x\rangle.$$
As A is self-adjoint, its quadratic form $\langle Ax, x\rangle$ is real. In addition, $\langle x, x\rangle$ is real too and $\langle x, x\rangle \neq 0$. Therefore, $\lambda = \frac{\langle Ax,x\rangle}{\langle x,x\rangle}$ is a real number.
2. We have $Ax = \lambda x$ and $Ay = \mu y$. Thus
$$\langle Ax, y\rangle = \lambda\langle x, y\rangle, \quad \langle x, Ay\rangle = \langle x, \mu y\rangle.$$
Since μ is real, $\langle x, Ay\rangle = \mu\langle x, y\rangle$, and we also have $\langle Ax, y\rangle = \langle x, Ay\rangle$, since A is self-adjoint. Consequently, $(\lambda - \mu)\langle x, y\rangle = 0$ which implies, since $\lambda \neq \mu$, that $\langle x, y\rangle = 0$, which means that $x \perp y$ in H.
3. Let $x \in H_0$, $z \in H_0^\perp$, $y = Ax \in AH_0 \subset H_0$. Then
$$\langle y, z\rangle = 0, \quad \Longleftrightarrow \quad \langle x, Az\rangle = 0,$$
where $Az \in AH_0^\perp$. Hence, $AH_0^\perp \subset H_0^\perp$. $\square$

Theorem 6.3.3 ***(Hilbert-Schmidt)*** *Let $H \neq \{0\}$ be a Hilbert space and $A : H \to H$ be a compact self-adjoint operator in H. Then there is an orthonormal system $\phi_1, \phi_2, \ldots$ of eigenvectors of A which forms an orthonormal basis in H. In addition,*

1. *All nonzero eigenvalues $\lambda \neq 0$ of A have a finite multiplicity.*
2. *The set of different eigenvalues of A is finite or countable.*
3. *If the set of eigenvalues of A is countable, then the eigenvalues form a sequence (λ_n) which converges toward 0.*

Proof. Eigenvectors of A form an orthonormal basis in H:

We denote by
$$H_\lambda = \{x \in H |\ Ax = \lambda x\}$$
a closed subspace of H, which is nonzero iff λ is an eigenvalue of the operator A. Thanks to Lemma 6.3.1, for $\lambda \neq \mu$ we have $H_\lambda \perp H_\mu$, thus we can define a subspace of H
$$H_* = \oplus_\lambda H_\lambda.$$

We fixe an orthonormal basis S_λ in H_λ and define

$$S = \cup_\lambda S_\lambda.$$

Hence, $H_* = \overline{\mathrm{Span}(S)}$.

Let us notice, that, by the linearity of A, $A(\mathrm{span}(S)) \subset \mathrm{span}(S)$ what implies, using the continuity of A, that $AH_* \subset H_*$. Thanks to point 3 of Lemma 6.3.1, it follows that

$$AH_*^\perp \subset H_*^\perp.$$

Let us suppose that $H_*^\perp \neq \{0\}$. Then, by Theorem 6.3.2, there exist $x \neq 0$ in $H_*^\perp$ and $\alpha \neq 0$ in $\mathbb{R}$ such that $Ax = \alpha x$. This implies that $x \in H_\alpha \subset H_*$, thus $x \perp x$ and hence $x = 0$, what contradicts our assumption $x \neq 0$. Thus, $H_*^\perp = \{0\}$. Consequently,

$$H = H_* = \oplus_\lambda H_\lambda$$

and thus S is an orthonormal basis in H and all vectors of S are eigenvectors of A.

1. <u>Eigenvalues $\lambda \neq 0$ of A have a finite multiplicity</u>

   Let $\lambda \neq 0$ be a eigenvalue of A and $\dim H_\lambda = \infty$. Let $(e_n)_{n\in\mathbb{N}}$ be an orthonormal subsequence of the basis S_λ of H_λ. Thus for all m and n in $\mathbb{N}$ ($m \neq n$) we have

   $$\|Ae_n - Ae_m\|^2 = \|\lambda e_n - \lambda e_m\|^2 = 2\lambda^2.$$

   Thus, the distance $d(Ae_n, Ae_m) = \|Ae_n - Ae_m\| = \sqrt{2}|\lambda| \neq 0$. Therefore, from the sequence (Ae_n) it is not possible to extract a convergent subsequence, what implies that the set AB_1 is not relatively compact. This is the contradiction with the assumption that A is compact. Then the basis S_λ has a finite number of eigenvectors of A.

2. <u>The set of different eigenvalues of A is finite or countable.</u>

   Let
   $$S(\delta) = \cup_{|\lambda|\geq\delta} S_\lambda \quad (\delta > 0).$$
   Let us suppose that the set $S(\delta)$ is infinite. Then for any e_n and e_m from $S(\delta)$ we have

   $$\begin{aligned} &Ae_n = \lambda e_n, \quad Ae_m = \mu e_m, \quad e_n \perp e_m, \\ &\|Ae_n - Ae_m\|^2 = \|\lambda e_n - \mu e_m\|^2 = \lambda^2 + \mu^2 \geq 2\delta^2. \end{aligned}$$

   We repeat the argument of the proof of point 1 and obtain the contradiction with the compactness of A. Thus the set $S(\delta)$ is finite. For all neighborhood U of 0 in $\mathbb{R}$, in $\mathbb{R} \setminus U$ there exist a finite number of eigenvalues of A. Thus, for $\delta \to 0$, in the limit, the set of eigenvalues of A is countable or finite.

3. <u>If the set of eigenvalues of A is countable, then the eigenvalues form a sequence (λ_n) which converges toward 0.</u> Using the proof of the previous point, we see that there exists only one point which can be a limit point, which is 0. Moreover, if there exists an infinite number of eigenvalues of A (λ_n), it implies that $\lambda_n \to 0$. $\square$

Remark 6.3.1 *If $\lambda = 0$ is the eigenvalue of A, it implies that* $\mathrm{Ker}(A) \neq \{0\}$ *and* $\mathrm{Ker}(A)$ *is a closed linear subspace of H. For $A \equiv 0$ in H (A is compact and self-adjoint), we have* $H = H_0 = \mathrm{Ker}(A)$.

Example 6.3.2 *Let $H = L^2([a,b])$ and A be the Fredholm operator with the kernel $K(t,s) = \overline{K(s,t)}$ on $L^2([a,b]^2)$ ($\overline{K}$ is the complex conjugate of K). Then there exists an orthonormal basis in $L^2([a,b])$ of the eigenvectors of A (see TD3 and TD5 for $A = A^*$ and its compactness).*

Remark 6.3.2 *Let $H \neq \{0\}$ be a Hilbert space and $A : H \to H$ be a compact self-adjoint operator in H. Then, by Hilbert-Schmidt theorem, there is an orthonormal system $\phi_1, \phi_2, \dots$ of eigenvectors of A which forms an orthonormal basis in H. In particular, all $u \in H$ can be uniquely presented as*

$$u = \sum_{\lambda_i \neq 0} \langle \phi_i, u \rangle \phi_i + \psi,$$

where $\psi \in \operatorname{Ker} A$. If $\operatorname{Ker} A = \{0\}$, then

$$\forall u \in H \quad u = \sum_i \langle \phi_i, u \rangle \phi_i \quad \text{and } Au = \sum_i \lambda_i \langle \phi_i, u \rangle \phi_i.$$

6.4 Spectrum and resolvent of a linear operator.

Suppose X is finite-dimensional. Then all linear operators from X to X are bounded, and hence continuous. In addition, if a linear operator is injective, then it is also surjective and consequently bijective.

In an infinite dimensional space these two statements do not hold.

Definition 6.4.1 *Let X be a Banach space and $A \in \mathcal{L}(X,X)$.*

- *The operator function*
$$R_\lambda(A) = (A - \lambda I)^{-1} \quad (\lambda \in \mathbb{C})$$
*is called the **resolvent** of the operator A.*

- *The domain of the resolvent*
$$\rho(A) = \{\lambda \in \mathbb{C} |\ R_\lambda(A) \in \mathcal{L}(X,X)\}$$
*is called the **resolvent set** of the operator A. Elements of $\rho(A)$ are called the **regular points**.*

- *The set*
$$\sigma(A) = \mathbb{C} \setminus \rho(A)$$
*is called **the spectrum** of the operator A.*

Remark 6.4.1 *1. $\lambda \in \rho(A) \iff A - \lambda$ is a bijection of X on X.*

2. The set of all eigenvalues of A is a subset of its spectrum.

Theorem 6.4.1 *Let X be a Banach space and $A \in \mathcal{L}(X,X)$. Then*

- *the resolvent set $\rho(A)$ is an open set;*
- *the resolvent $R_\lambda(A)$ is analytical in $\rho(A)$;*
- *the spectrum $\sigma(A)$ is a compact set and it is contained in a ball of the raduis $\|A\|$.*

Proof.

1. Let $\lambda \neq 0$. We have

$$R_\lambda(A) = (-\lambda)^{-1}(1 - \lambda^{-1}A)^{-1} = (-\lambda)^{-1} \sum_{n=0}^{\infty} (\lambda^{-1}A)^n.$$

The series converges for $\|\lambda^{-1}A\| < 1$, *i.e.*

$$|\lambda| > \|A\|,$$

which implies that

$$\sigma(A) \subset B_{\|A\|}(0) = \{x \in X| \ \|x\|_X \leq \|A\|_{\mathcal{L}(X,X)}\}.$$

2. Let $\lambda_0 \in \rho(A)$. We have

$$A - \lambda = (A - \lambda_0) - \Delta\lambda, \text{ where } \Delta\lambda = \lambda - \lambda_0,$$
$$R_\lambda(A) = [(A - \lambda_0) - \Delta\lambda]^{-1} = R_{\lambda_0}(A)(1 - \Delta\lambda R_{\lambda_0}(A))^{-1} = R_{\lambda_0}(A) \sum_{n=0}^{\infty} (\Delta\lambda)^n R_{\lambda_0}(A)^n.$$

The series converges for $|\Delta\lambda| \|R_{\lambda_0}(A)\| < 1$, thus for $|\Delta\lambda| < \delta = \frac{1}{\|R_{\lambda_0}(A)\|}$. It means, that $\rho(A)$ is open and $R_\lambda(A)$ is analytical in $\rho(A)$. Moreover, it follows that $\sigma(A)$ is closed. As in addition, $\sigma(A)$ is bounded in $\mathbb{C}$, thus $\sigma(A)$ is a compact in $\mathbb{C}$. $\square$

Example 6.4.1 *Let H be a Hilbert space and $A \in \mathcal{L}(H, H)$ such that $A = A^*$. Then $\sigma(A) \subset [-\|A\|, \|A\|]$.*

Definition 6.4.2 *A real number*

$$r(A) = \max\{|\lambda| : \ \lambda \in \sigma(A)\}$$

*is called **spectral raduis** of the operator A.*

Remark 6.4.2 *We note that $r(A) \leq \|A\|$.*

Example 6.4.2 *Let $A : \mathbb{C}^2 \to \mathbb{C}^2$ be defined by the matrix*

$$A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}.$$

Then $r(A) = 0$ and $\|A\| = 1$ (by the Euclidian norm).

We give without proof the following theorem:

Theorem 6.4.2 *Let H be a Hilbert space and $A \in \mathcal{L}(H, H)$. Then*

$$r(A) = \lim_{n\to\infty} \|A^n\|^{\frac{1}{n}}.$$

Proposition 6.4.1 *Let H be a Hilbert space and $A \in \mathcal{L}(H, H)$. If $A = A^*$, then $r(A) = \|A\|$.*

Proof. Since A is self-adjoint, then $\|A\|^2 = \|A^2\|$.

Indeed,

$$\|Ax\|^2 = \langle A^*Ax, x\rangle = \langle A^2x, x\rangle;$$
$$\|A\|^2 = \sup_{\|x\|=1} \|Ax\|^2 = \sup_{\|x\|=1} |\langle A^*Ax, x\rangle| = \|A^*A\| = \|A^2\|.$$

Therefore, we have

$$\|A\| = \|A^2\|^{\frac{1}{2}}.$$

We replace A by A^2 or A^{2^k}, $k \in \mathbb{N}$ and find that for $n = 2^k$

$$\|A\| = \|A^2\|^{\frac{1}{2}} = \|A^2\|^{\frac{1}{4}} = \ldots = \|A^n\|^{\frac{1}{n}}.$$

Thus

$$r(A) = \lim_{n\to\infty} \|A^n\|^{\frac{1}{n}} = \lim_{n=2^k} \|A^n\|^{\frac{1}{n}} = \|A\|. \quad \square$$

Corollary 6.4.1 *A compact operator A mapping a Banach space X into itself cannot have a bounded inverse A^{-1} if X is infinite-dimensional.*

Proof. If A^{-1} were bounded, then, by Theorem 6.1.1, the identity operator $I = A^{-1}A$ would be compact. But the unit ball is not relatively compact in an infinite-dimensional Banach space.$\square$

Chapter 7

Distributions

7.1 Space $\mathcal{D}(\Omega)$ of "test functions"

Remark 7.1.1 *Let Ω be an open bounded set in $\mathbb{R}^n$. Let us notice that*

$$C(\Omega) = \{ \text{ all continuous functions } f : \Omega \to \mathbb{C}\},$$
$$C(\overline{\Omega}) = \{ \text{ all equicontinuous and bounded functions } f : \Omega \to \mathbb{C}\}.$$

More precisely, if $\overline{\Omega}$ is compact in $\mathbb{R}^n$, then $f \in C(\overline{\Omega})$ if

1. $\sup_{x\in\Omega} |f(x)| < \infty$,
2. $\forall \epsilon > 0\ \exists \delta = \delta(\epsilon) > 0 : \forall x, y \in \Omega,\ |x - y| < \delta \Rightarrow |f(x) - f(y)| < \epsilon.$

For instance, it is easy to see that for $\Omega =]0, 1[$, $n = 1$, the functions

$$f_1(x) = \frac{1}{x} : \text{ is not bounded and not equicontinuous } \Rightarrow \quad f_1 \in C(\Omega), \quad f_1 \notin C(\overline{\Omega})$$
$$f_2(x) = \sin\frac{1}{x} : \text{ is bounded but not equicontinuous } \Rightarrow \quad f_2 \in C(\Omega), \quad f_2 \notin C(\overline{\Omega}).$$

Definition 7.1.1 *A subset Ω of $\mathbb{R}^n$ is called* ***a domain*** *if Ω is open and connected (any two points of Ω can be joined by a path, actually, by a continuous line belonging to Ω. See Fig. 7.1.*

In what follows, we always suppose the Ω is a domain in $\mathbb{R}^n$.

Definition 7.1.2 *The intersection of the closure of Ω with the closure of its complement $\mathbb{R}^n \setminus \Omega$ is called* ***boundary*** *of Ω and denoted by $\partial\Omega = \overline{\Omega} \cap \overline{\mathbb{R}^n \setminus \Omega}$. A point $x \in \partial\Omega$ is called* ***boundary point*** *of Ω.*

Example 7.1.1
- *In $\mathbb{R}$, the domain $]0, 1[$ has two boundaries points 0 and 1. The set $\{0\} \cup \{1\}$ is the boundary of $]0, 1[$.*
- *In $\mathbb{R}^2$, let $\Omega = \{(x, y) \in \mathbb{R}^2 | x^2 + y^2 < 1\}$. Then $\partial\Omega = \{(x, y) \in \mathbb{R}^2 | x^2 + y^2 = 1\}$.*

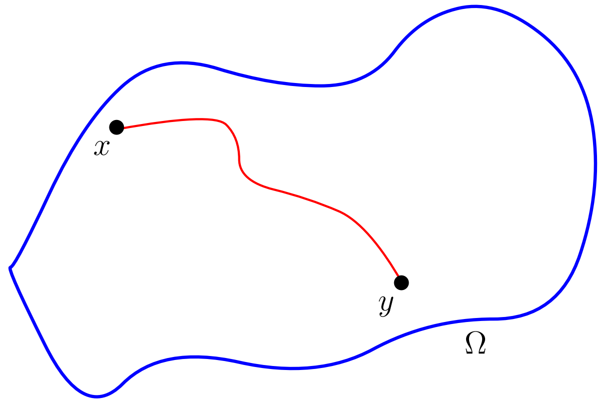

Figure 7.1 – Example of a connected set Ω: all points x and y in Ω can be joined by a continuous line (in red) belonging to Ω.

Definition 7.1.3 *Function $f : \mathbb{R}^n \to \mathbb{R}$ is said to be* ***compactly supported*** *in $\Omega \subset \mathbb{R}^n$ if the set, called its support,*

$$\operatorname{supp} f = \overline{\{x : f(x) \neq 0\}},$$

is a compact subset of Ω. (Actually, as we take the closure of the set $\{x : f(x) \neq 0\}$, $\operatorname{supp} f$ is compact, if it is bounded.)

Example 7.1.2 *Let us consider (see Fig 7.2) a function*

$$f(x) = \begin{cases} 2x, & 0 \leq x < \frac{1}{2} \\ 2(1-x), & \frac{1}{2} \leq x < 1 \\ 0, & \textit{elsewhere.} \end{cases}$$

Thus $\operatorname{supp} f = \overline{]0,1[} = [0,1]$.

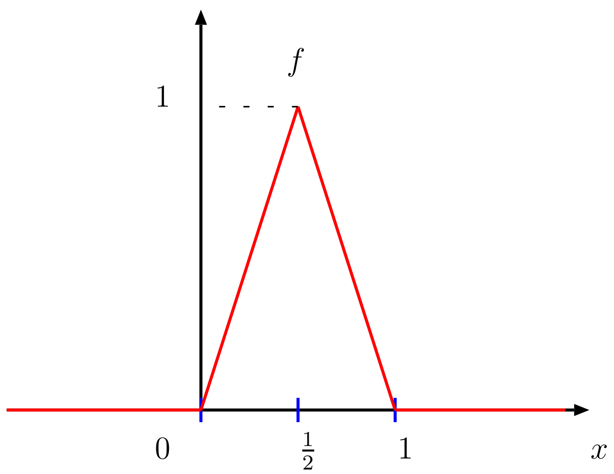

Figure 7.2 – Function f from Example 7.1.2.

Definition 7.1.4 *We define by*

1. *$C_0^l(\Omega)$ the set of all l times continuously differentiable functions compactly supported in Ω. (If $l = 0$, thus $C_0^0(\Omega)$ is the set of all continuous functions compactly supported*

in Ω. In what follows instead of $C_0^0(\Omega)$ we simply write $C_0(\Omega)$.)

2. $\mathcal{D}(\Omega) = C_0^\infty(\Omega)$ *the set of infinitely differentiable functions compactly supported in Ω.*

Definition 7.1.5 *A sequence (f_k) of elements of $\mathcal{D}(\Omega)$ converges toward $f \in \mathcal{D}(\Omega)$ iff*

1. *There exists $K \subsetneq \Omega$ compact such that*
$$\forall k \in \mathbb{N}, \quad \operatorname{supp} f_k \subset K,$$
2. *For all multi-index α $\partial^\alpha f_k(x)$ converges uniformly toward $\partial^\alpha f$ on K.*

The definition of the convergence is equivalent to the definition of the topology on $\mathcal{D}(\Omega)$ (see Example A.1.6).

Note that

- it follows immediately that $\operatorname{supp} f \subset K$.
- if in the space $C^l(\overline{\Omega})$ we introduce the norm
$$\|f\|_{C^l(\overline{\Omega})} = \max_{|\alpha| \le l} \sup_{x \in \Omega} |\partial^\alpha f(x)|,$$
then $(C^l(\overline{\Omega}), \|\cdot\|_{C^l(\overline{\Omega})})$ is a Banach space and its norm is the norm of the uniform convergence:
$$\forall \alpha \in N^n \quad \partial^\alpha f_k(x) \rightrightarrows \partial^\alpha f \text{ on } K \quad \Longleftrightarrow \quad \forall l \in \mathbb{N} \quad \|f_k - f\|_{C^l(K)} \to 0 \text{ for } k \to +\infty.$$

Remark 7.1.2 1. *The space $C^\infty(\overline{\Omega})$ contains infinitely continuous differentiable functions such that*
$$\forall l \in \mathbb{N} \quad \|f\|_{C^l(\overline{\Omega})} < \infty.$$
$C^\infty(\overline{\Omega})$, and in the same way $\mathcal{D}(\Omega)$, are not normable spaces (it is not possible to define a norm).

2. *The space $C_0^l(\Omega)$ is a closed subspace of $C^l(\overline{\Omega})$. In the same time, $C^l(\overline{\Omega}) \subset C^l(\Omega)$.*

A linear operator $A : \mathcal{D}(\Omega) \to \mathcal{D}(\Omega)$ is continuous on $\mathcal{D}(\Omega)$ iff
$$f_k \to f \text{ in } \mathcal{D}(\Omega) \quad \Rightarrow \quad Af_k \to Af \text{ in } \mathcal{D}(\Omega).$$

The elements of $\mathcal{D}(\Omega)$ are often called **testing functions.**

Problem 7.1.1 *Show that the following linear operators are continuous in $\mathcal{D}(\Omega)$:*

1. *For a multi-index $\alpha \in N^n$*
$$\partial^\alpha : \mathcal{D}(\Omega) \to \mathcal{D}(\Omega), \quad f(x) \mapsto \partial^\alpha f(x),$$
2. *For a fixed function $\eta \in C^\infty(\Omega)$*
$$A_\eta : \mathcal{D}(\Omega) \to \mathcal{D}(\Omega), \quad f(x) \mapsto \eta(x) f(x),$$
3. *For $\Omega = \mathbb{R}^n$ and a fixed $g \in \mathcal{D}(\mathbb{R}^n)$*
$$S_g : \mathcal{D}(\Omega) \to \mathcal{D}(\Omega), \quad f(x) \mapsto S_g(f)(x) = (f * g)(x) = \int_{\mathbb{R}^n} f(x-y) g(y) dy.$$

Example 7.1.3 *Functions, called mollifiers and which we consider in next Subsection, gives an example of a non trivial function from $\mathcal{D}(\Omega)$. For instance, the familly of functions*

$$w_h(x) = \frac{1}{h^n}\eta\left(\frac{x}{h}\right), \quad \text{where } \eta(x) = \begin{cases} ce^{-\frac{1}{1-|x|^2}}, & |x| < 1 \\ 0, & |x| \geq 1 \end{cases}$$

with a constant c such that $\int_{\mathbb{R}^n} w_h(x)dx = 1$, belongs to $\mathcal{D}(]-1,1[)$ for all $h > 0$ (see Fig. 7.3).

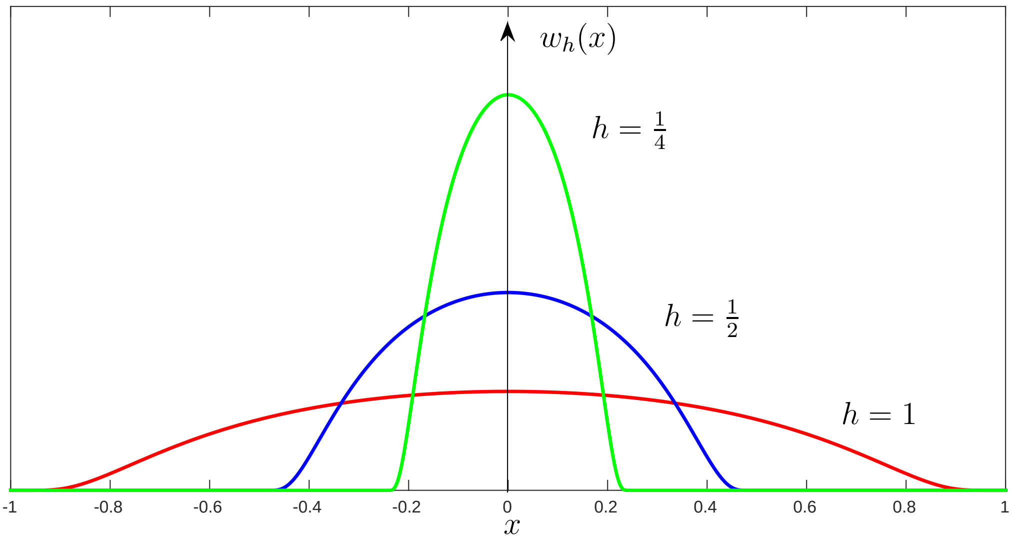

Figure 7.3 – Example of non trivial elements of $\mathcal{D}(]-1,1[)$: $w_h(x) = \frac{1}{h}\eta\left(\frac{x}{h}\right)$, where $\eta(x) = e^{-\frac{1}{1-|x|^2}}$ if $|x| < 1$ and $\eta(x) = 0$ if $|x| \geq 1$.

In addition, we will prove (see Theorem 7.1.1) that $\mathcal{D}(\Omega)$ is dense in $L^p(\Omega)$.

7.1.1 Mollifiers

Let $\Omega \subsetneq \mathbb{R}^n$ be a bounded domain (see Definition 7.1.1). For $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$, we write

$$|x| = \sqrt{\sum_{i=1}^{n} |x_i|^2}.$$

Let $\varphi \in L^p(\Omega)$. We extend φ by 0 for $x \notin \Omega$:

$$\tilde{\varphi}(x) = \begin{cases} \varphi(x), & x \in \Omega \\ 0, & x \notin \Omega \end{cases}$$

Then $\tilde{\varphi} \in L^p(\mathbb{R}^n)$ and the boundary $\partial\Omega$ contains a part of discontinuities of $\tilde{\varphi}$. Let us mollify (smooth) these discontinuities.

Definition 7.1.6 *Let $h > 0$ and $\Omega \subset \mathbb{R}^n$ be a domain. For $\varphi \in L^p(\Omega)$ we define a **mollifier***

$$\varphi_h(x) = \int_{B_h(x)} w_h(x-y)\tilde{\varphi}(y)dy, \quad x \in \mathbb{R}^n, \tag{7.1}$$

*where $w_h(x)$ is a function, called the **kernel** of φ_h, such that*

1. $w_h \in C_0^\infty(\mathbb{R}^n)$,
2. $w_h(x) > 0$ *if* $|x| < h$ *and* $w_h(x) = 0$ *if* $|x| \geq h$. *This statement is equivalent to*
$$\operatorname{supp} w_h(x) = \overline{B_h(0)}.$$
3. $\int_{\mathbb{R}^n} w_h(x)dx = 1$
4. $w_h(-x) = w_h(x)$.

Remark 7.1.3 *In this Section we write $\int_\Omega f(x,y)dx$ understanding the Lebesgue integral of f with respect to $x \in \Omega$. When the function f depends only on one variable x, we will write $\int_\Omega f d\mu$ for the Lebesgue integral of f over Ω.*

Remark 7.1.4 *We notice that in our notations*
$$\overline{B_h(x)} = \{y \in \mathbb{R}^n |\, |x-y| \leq h\}.$$
Since for a fixed x
$$w_h(x-y)\varphi(y) \in L_p(\overline{B_h(x)}) \subset L_1(\overline{B_h(x)}),$$
we conclude that the definition of φ_h is well defined.

As $w_h(x-y) = 0$ for $|x-y| \geq h$, then for a fixed x
$$w_h(x-y)\tilde{\varphi}(y) \in L^p(\mathbb{R}^n).$$

We summarize the properties of the mollifiers in the following proposition:

Proposition 7.1.1 *Let $\varphi \in L^p(\Omega)$, $p \geq 1$. Then we have:*

1. $\forall h > 0 \ \varphi_h \in C^\infty(\mathbb{R}^n)$
2. *If $d(x,\Omega) > h$, then $\varphi_h(x) = 0$.*
3. *If Ω is bounded then for all $h > 0$ $\varphi_h \in C_0^\infty(\mathbb{R}^n)$.*
4. $\|\varphi_h\|_{L^p(\Omega)} \leq \|\varphi\|_{L^p(\Omega)}$
5. *$\varphi_h \to \varphi$ for $h \to 0$ in $L^p(\Omega)$.*

Proof. Let us prove points 2), 4) and 5). The point 3) is a direct corollary of point 1 and point 2. For point 1) see for example [3].

1. Point 2): If $d(x,\Omega) > h$, then $\varphi_h(x) = 0$.

   If $d(x,\Omega) > h$, then $\Omega \cap \overline{B_h(x)} = \varnothing$ (see Fig. 7.4). We suppose that φ is defined in $\mathbb{R}^n$ using the extention of φ by 0 for $x \notin \Omega$. Thus, by definition of the kernel w_h, we have
$$\forall y \in \mathbb{R}^n \quad w_h(x-y)\varphi(y) = \begin{cases} 0, & |x-y| \geq h \quad (\text{as } w_h(x-y) = 0), \\ 0, & |x-y| < h \quad (\text{as } \varphi(y) = 0). \end{cases}$$
   Hence,
$$\varphi_h = \int_{\mathbb{R}^n} w_h(x-y)\varphi(y)dy = 0.$$

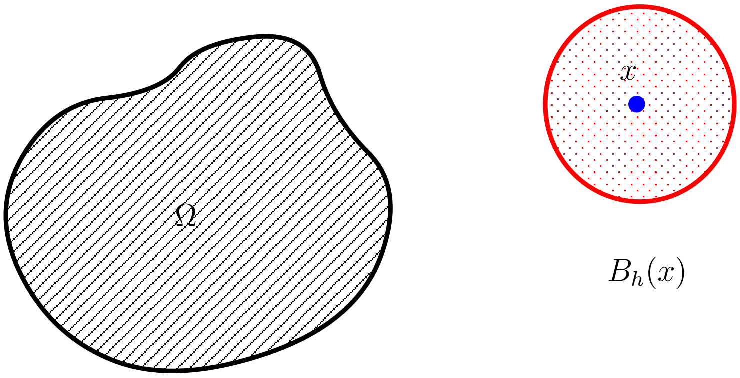

Figure 7.4 – If $d(x,\Omega) > h$, then $\Omega \cap \overline{B_h(x)} = \varnothing$.

2. Point 4): $\|\varphi_h\|_{L^p(\Omega)} \leq \|\varphi\|_{L^p(\Omega)}$

We separate two cases: $p = 1$ and $p > 1$.

Let $p = 1$.

We have

$$\varphi_h(x) = \int_\Omega w_h(x-y)\varphi(y)dy.$$

Thus we find

$$|\varphi_h(x)| \leq \int_\Omega w_h(x-y)|\varphi(y)|dy,$$

using that $w_h(x-y) \geq 0$ for all x and y. Then

$$\|\varphi_h\|_{L^1(\Omega)} = \int_\Omega |\varphi_h(x)|dx \leq \int_\Omega \left(\int_\Omega w_h(x-y)|\varphi(y)|dy\right) dx,$$

and by the Fubini Theorem,

$$\begin{aligned}\|\varphi_h\|_{L^1(\Omega)} &\leq \int_\Omega \left(\int_\Omega w_h(x-y)|\varphi(y)|dy\right) dx = \int_\Omega |\varphi(y)| \left(\int_\Omega w_h(x-y)dx\right) dy \\ &\leq \int_\Omega |\varphi(y)| \left(\int_{\mathbb{R}^n} w_h(x-y)dx\right) dy = \int_\Omega |\varphi(y)|dy = \|\varphi\|_{L^1(\Omega)}.\end{aligned}$$

Here we have used that $\int_{\mathbb{R}^n} w_h(x-y)dx = 1$, by definition of w_h. Thus we have

$$\|\varphi_h\|_{L^1(\Omega)} \leq \|\varphi\|_{L^1(\Omega)}.$$

Let $p > 1$.

Since $\frac{1}{p} + \frac{1}{p'} = 1$, we have

$$|\varphi_h(x)| \leq \int_\Omega w_h(x-y)|\varphi(y)|dy = \int_\Omega w_h^{\frac{1}{p}}(x-y)w_h^{\frac{1}{p'}}(x-y)|\varphi(y)|dy.$$

We apply the Hölder inequality and use the fact that $\int_{\mathbb{R}^n} w_h(x-y)dx = 1$:

$$\begin{aligned}|\varphi_h(x)| &\leq \int_\Omega w_h^{\frac{1}{p'}}(x-y)w_h^{\frac{1}{p}}(x-y)|\varphi(y)|dy \\ &\leq \left[\int_\Omega \left(w_h^{\frac{1}{p'}}(x-y)\right)^{p'} dy\right]^{\frac{1}{p'}} \left[\int_\Omega \left(w_h^{\frac{1}{p}}(x-y)|\varphi(y)|\right)^{p} dy\right]^{\frac{1}{p}} \\ &\leq \left[\int_{\mathbb{R}^n} w_h(x-y)dy\right]^{\frac{1}{p'}} \left[\int_\Omega w_h(x-y)|\varphi(y)|^p dy\right]^{\frac{1}{p}} \leq \left[\int_\Omega w_h(x-y)|\varphi(y)|^p dy\right]^{\frac{1}{p}}.\end{aligned}$$

Thus,

$$|\varphi_h(x)|^p \le \int_\Omega w_h(x-y)|\varphi(y)|^p dy.$$

As for the case $p=1$, we integrate the last inequality over $y \in \Omega$ and, applying the Fubini Theorem, we find

$$\|\varphi_h\|^p_{L^p(\Omega)} = \int_\Omega |\varphi_h(x)|^p dx \le \int_\Omega |\varphi_h(y)|^p dy = \|\varphi\|^p_{L^p(\Omega)}.$$

3. Point 5): $\varphi_h \to \varphi$ for $h \to 0$ in $L^p(\Omega)$.

   We separate two cases: $p=1$ and $p>1$.

   Let $p=1$.

   Let us consider $\varphi_h(x)-\varphi(x)$. Since $\int_{B_h(x)} w_h(x-y)dy = 1$, we have

$$\begin{aligned}\varphi_h(x)-\varphi(x) &= \int_{B_h(x)} w_h(x-y)\varphi(y)dy - \varphi(x)\int_{B_h(x)} w_h(x-y)dy\\ &= \int_{B_h(x)} w_h(x-y)(\varphi(y)-\varphi(x))dy.\end{aligned}$$

   We perform the change of variables $x-y=z \in \mathbb{R}^n$ with $dy=(-1)^n dz$ ($dz = dz_1 \cdot \ldots \cdot dz_n$):

$$\varphi_h(x)-\varphi(x) = \int_{B_h(x)} w_h(x-y)(\varphi(y)-\varphi(x))dy = \int_{|z|<h} w_h(z)(\varphi(x-z)-\varphi(x))dz.$$

   Thus

$$\int_\Omega |\varphi_h(x)-\varphi(x)|dx \le \int_\Omega \int_{|z|<h} w_h(z)|\varphi(x-z)-\varphi(x)|dzdx,$$

   and by the Fubini Theorem

$$\int_\Omega |\varphi_h(x)-\varphi(x)|dx \le \int_{|z|<h} w_h(z)\left(\int_\Omega |\varphi(x-z)-\varphi(x)|dx\right)dz.$$

   By the continuity of the Lebesgue integral, we have (the result is assumed without proof)

$$\forall \epsilon>0 \quad \exists \delta(\epsilon)>0: \quad \int_\Omega |\varphi(x-z)-\varphi(x)|dx < \epsilon \quad \text{for } |z|<\delta(\epsilon).$$

   Let us take $h<\delta(\epsilon)$. Then

$$\|\varphi_h-\varphi\|_{L^1(\Omega)} = \int_\Omega |\varphi_h(x)-\varphi(x)|dx < \epsilon \int_{|z|<h} w_h(z)dz = \epsilon.$$

   Let $p>1$.

   The proof follows to the case of $p=1$ with the differences detailed in the point 4, case $p>1$. □

7.1.2 Density of $\mathcal{D}(\Omega)$ in $L^p(\Omega)$.

From Proposition 7.1.1 it follows

Theorem 7.1.1 *Let $p \in [1, \infty[$. $\mathcal{D}(\Omega)$ is dense in $L^p(\Omega)$.*

Proof. Let $\varphi \in L^p(\Omega)$. We define (see Fig. 7.5)

$$\Omega^{(r)} = \{x \in \Omega | d(x, \partial\Omega) > r, \quad |x| < \frac{1}{r}\}.$$

Let us prove that

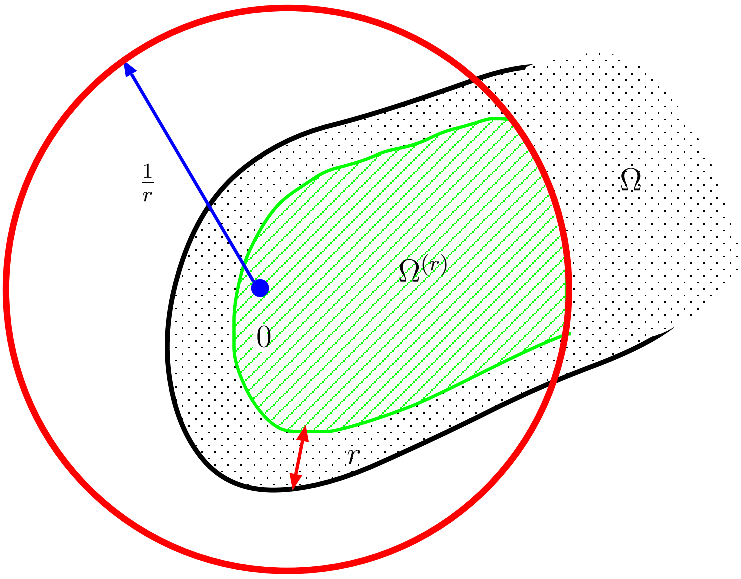

Figure 7.5 – Subdomain $\Omega^{(r)}$.

$$\forall \epsilon > 0 \quad \exists r > 0 : \quad \int_{\Omega \setminus \Omega^{(r)}} |\varphi(x)|^p dx < \left(\frac{\epsilon}{2}\right)^p. \tag{7.2}$$

Firstly, we notice that (see Fig. 7.5)

1. if $r_1 < r_2$ then $\Omega_{r_2} \subset \Omega_{r_1}$ (obviously).
2. if $r = \frac{1}{m}$, $m \in \mathbb{N}^*$, then
$$\cup_{m=1}^{\infty} \Omega^{\left(\frac{1}{m}\right)} = \Omega.$$

Indeed, we have
$$\forall m \in \mathbb{N}^* \quad \Omega^{\left(\frac{1}{m}\right)} \subset \Omega \quad \Rightarrow \quad \cup_{m=1}^{\infty} \Omega^{\left(\frac{1}{m}\right)} \subset \Omega.$$
But Ω is open (in $\mathbb{R}^n$) by Definition 7.1.1. Then for all $x \in \Omega$ (Ω bounded or not) we have

$$|x| < \infty \quad \text{and} \quad d(x, \partial\Omega) > 0.$$

It implies that
$$\exists m = m_x \in \mathbb{N}^* : \quad |x| < m, \quad d(x, \partial\Omega) > \frac{1}{m},$$

thus $x \in \Omega^{\left(\frac{1}{m_x}\right)}$ and then

$$x \in \cup_{m=1}^{\infty} \Omega^{\left(\frac{1}{m}\right)}, \quad i.e., \quad \Omega \subset \cup_{m=1}^{\infty} \Omega^{\left(\frac{1}{m}\right)}.$$

Consequently, we find that $\Omega = \cup_{m=1}^{\infty} \Omega^{\left(\frac{1}{m}\right)}$.

From 1) and 2) it follows that

$$\int_{\Omega} |\varphi(x)|^p dx = \lim_{m \to \infty} \int_{\Omega^{\left(\frac{1}{m}\right)}} |\varphi(x)|^p dx, \quad i.e., \quad \forall \varphi \in L^p(\Omega) \quad \int_{\Omega \setminus \Omega^{\left(\frac{1}{m}\right)}} |\varphi(x)|^p dx \to 0 \text{ for } m \to \infty.$$

Thus we obtain (8.7).

For

$$\tilde{\varphi}(x) = \begin{cases} \varphi, & x \in \Omega^{(r)}, \\ 0, & x \in \Omega \setminus \Omega^{(r)} \end{cases}$$

we construct its mollifier $\tilde{\varphi}_h$. Let us take h such that $h < r$. If x is such that $d(x, \partial\Omega) < r - h$, then $\Omega^{(r)} \cap \overline{B_h(x)} = \varnothing$ and consequantly, by Proposition 7.1.1, $\tilde{\varphi}_h = 0$. Thus, $\tilde{\varphi}_h \in C_0^{\infty}(\Omega)$. Let us show that

$$\forall \epsilon > 0 \quad \exists h > 0 : \quad \|\varphi - \tilde{\varphi}_h\|_{L^p(\Omega)} \leq \epsilon.$$

Indeed,

$$\|\varphi - \tilde{\varphi}_h\|_{L^p(\Omega)} \leq \|\varphi - \tilde{\varphi}\|_{L^p(\Omega)} + \|\tilde{\varphi} - \tilde{\varphi}_h\|_{L^p(\Omega)}.$$

We have, thanks to (8.7),

$$\|\varphi - \tilde{\varphi}\|_{L^p(\Omega)} = \left(\int_{\Omega} |\varphi(x) - \tilde{\varphi}(x)|^p dx\right)^{\frac{1}{p}} = \left(\int_{\Omega \setminus \Omega^{(r)}} |\varphi(x)|^p dx\right)^{\frac{1}{p}} < \frac{\epsilon}{2}$$

and in addition

$$\|\tilde{\varphi} - \tilde{\varphi}_h\|_{L^p(\Omega)} \to 0 \quad h \to 0.$$

We conclude that for h small enough

$$\|\varphi - \tilde{\varphi}_h\|_{L^p(\Omega)} \leq \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon. \quad \square$$

7.1.3 Completeness of $\mathcal{D}(\Omega)$

Firstly, we note that

Lemma 7.1.1 *For all domain $\Omega \subset \mathbb{R}^n$ and all $\epsilon > 0$ there exists a function $\eta \in C^{\infty}(\mathbb{R}^n)$ such that (see Fig. 7.6 for $\Omega =]a, b[$ $(a < b)$)*

1. $\forall x \in \mathbb{R}^n \ 0 \leq \eta(x) \leq 1$;
2. $\forall x \in \Omega_{\epsilon} \ \eta(x) = 1$;
3. $\forall x \notin \Omega_{3\epsilon} \ \eta(x) = 0$,

where

$$\Omega_{\epsilon} = \Omega \cup \{x \in \mathbb{R}^n \setminus \Omega | \ d(\partial\Omega, x) < \epsilon\}.$$

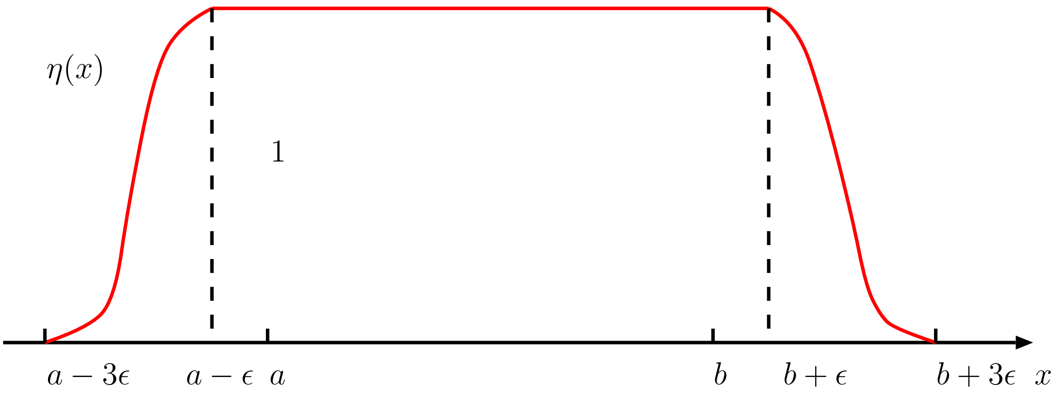

Figure 7.6 – Example of $\eta \in \mathcal{D}(]a,b[)$ from Lemma 7.1.1.

Proof. Let $\mathbb{1}_{\Omega_{2\epsilon}}$ be the characteristic function of the set $\Omega_{2\epsilon}$:

$$\mathbb{1}_{\Omega_{2\epsilon}} = \begin{cases} 1, & x \in \Omega_{2\epsilon} \\ 0, & x \neq \Omega_{2\epsilon}. \end{cases}$$

Let us show that the function

$$\eta(x) = \int_{\mathbb{R}^n} \mathbb{1}_{\Omega_{2\epsilon}}(y) w_\epsilon(x-y) dy$$

satisfies 1)-3). Here $w_\epsilon(x) \in \mathcal{D}(\mathbb{R}^n)$ is the kernel defined in Definition 7.1.6 such that

- $\forall x \in \mathbb{R}^n \quad w_\epsilon(x) \geq 0$,
- $\operatorname{supp} w_\epsilon(x) = \overline{B_\epsilon(0)}$,
- $\int_{\mathbb{R}^n} w_\epsilon(x)dx = 1$.

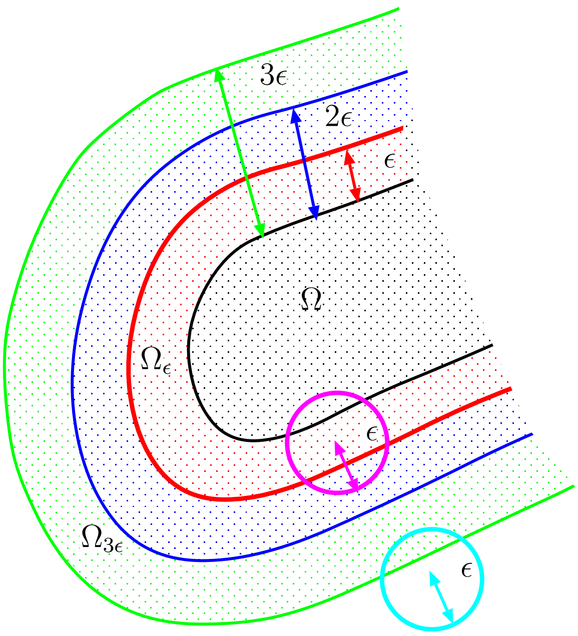

Figure 7.7 – Example of supports of $w_\epsilon(x-y)$ for $x \in \Omega_\epsilon$ and $x \notin \Omega_{3\epsilon}$ for the proof of Lemma 7.1.1.

Consequently, (see Fig. 7.7) we have

$$\eta(x) = \int_{\Omega_{2\epsilon}} w_\epsilon(x-y)dy \in C^\infty(\mathbb{R}^n);$$
$$0 \le \eta(x) \le \int_{\mathbb{R}^n} w_\epsilon(x-y)dy = \int_{\mathbb{R}^n} w_\epsilon(z)dz = 1;$$
$$\eta(x) = \int_{B_\epsilon(x)} \mathbb{1}_{\Omega_{2\epsilon}}(y) w_\epsilon(x-y)dy =$$
$$= \begin{cases} \int_{B_\epsilon(x)} w_\epsilon(x-y)dy = \int_{\mathbb{R}^n} w_\epsilon(z)dz = 1, & x \in \Omega_\epsilon; \\ 0, & x \notin \Omega_{3\epsilon}. \end{cases} \qquad \square$$

In addition, we know (see Theorem 7.1.1) that $\mathcal{D}(\Omega)$ is dense in $L^p(\Omega)$. We have defined what a Cauchy sequence is in the metric spaces, but it can also be defined for the topological spaces which are not metrizable.

In particular, let us define what a Cauchy sequence is in $\mathcal{D}(\Omega)$:

Definition 7.1.7 *Let $\varphi_k \in \mathcal{D}(\Omega)$ for $k \in \mathbb{N}$. We say that (φ_k) is* ***a Cauchy sequence*** *in $\mathcal{D}(\Omega)$ if there is some compact set $K \subsetneq \Omega$ such that for all $k \in \mathbb{N}$ $\operatorname{supp}\varphi_k \subset K$ and such that*

$$\forall \alpha \in \mathbb{N}^n \quad \forall l \in \mathbb{N} \quad \|\partial^\alpha(\varphi_k - \varphi_m)\|_{C^l(K)} \to 0 \text{ as } k, m \to \infty.$$

Thus we can prove:

Theorem 7.1.2 *Let Ω be a domain in $\mathbb{R}^n$. $\mathcal{D}(\Omega)$ is complete.*

Proof. Let (φ_k) be a Cauchy sequence in $\mathcal{D}(\Omega)$ (see Definition 7.1.7). Since

$$\forall\, l \in \mathbb{N} \quad (C^l(K), \|\cdot\|_{C^l(K)}) \text{ is a Banach space,}$$

it follows that there exists a function $\varphi \in C^\infty(K)$ such that

$$\forall \alpha \in \mathbb{N}^n \quad \forall l \in \mathbb{N} \quad \|\partial^\alpha(\varphi_k - \varphi)\|_{C^l(K)} \to 0 \text{ as } k \to \infty.$$

But if for all $k \in \mathbb{N}$ $\operatorname{supp}\varphi_k \subset K$, then $\operatorname{supp}\varphi \subset K$ also. Hence $\varphi \in C_0^\infty(\Omega) = \mathcal{D}(\Omega)$ and $\varphi_k \to \varphi$ in $\mathcal{D}(\Omega)$. $\square$

Problem 7.1.2 *Prove that $\mathcal{D}(\Omega)$ is dense in $C^\infty(\Omega)$:*

$$\forall f \in C^\infty(\Omega) \quad \exists (\varphi_k),\ \varphi_k \in \mathcal{D}(\Omega) \quad \text{such that } \varphi_k \to f\ (k \to \infty) \text{ in } C^\infty(\Omega).$$

7.1.4 Lemma of du Bois-Reymond

Lemma 7.1.2 *Let $f \in L^1_{loc}(\Omega)$ and*

$$\forall x \in \Omega \quad \exists \text{ open neigborhood } U_x : \quad f(x) = 0 \text{ a.e. in } U_x.$$

Then $f(x) = 0$ a.e. in Ω.

Proof. Let K be a compact in Ω.

Let us show that $\underline{f(x) = 0 \text{ a.e. in } K.}$

Since K is compact, the infinite (open) cover $\cup_{x\in K} U_x \supset K$ contains a finite subcover

$$K \subset \cup_{i=1}^N U_{x^i},$$

where U_{x^i} are open neigborhoods of points x^i such that $f(x) = 0$ a.e. in U_{x^i}. Thus,

$$\{x \in K | f(x) \neq 0\} \subset \cup_{i=1}^N \{x \in U_{x^i} | f(x) \neq 0\}.$$

Since a finite union of sets of zero measure is a set of zero measure, and a subset of a set of zero measure is also a set of zero measure, we conclude that $\mu(\{x \in K | f(x) \neq 0\}) = 0$.

Let's now prove that $\underline{f(x) = 0 \text{ a.e. in } \Omega.}$

We define for $r \in \mathbb{N}$

$$K^r = \{x \in \Omega |\ |x| \leq r \text{ and } d(x, \partial\Omega) \geq \frac{1}{r}\}.$$

Then K^r is a compact subset of Ω and

$$K^1 \subsetneq K^2 \subsetneq \ldots \subsetneq K^r \subsetneq \ldots \subsetneq \Omega.$$

Thus $\Omega = \cup_{r=1}^\infty K^r$ and consequently

$$\{x \in \Omega |\ f(x) \neq 0\} \subset \cup_{r=1}^\infty \{x \in K^r | f(x) \neq 0\}.$$

Therefore,

$$\mu(\{x \in \Omega |\ f(x) \neq 0\}) \leq \mu(\cup_{r=1}^\infty \{x \in K^r | f(x) \neq 0\}).$$

As $\forall r \in \mathbb{N} \quad \mu(\{x \in K^r | f(x) \neq 0\}) = 0$, then

$$\mu(\cup_{r=1}^\infty \{x \in K^r | f(x) \neq 0\}) \leq \cup_{r=1}^\infty \mu(\{x \in K^r | f(x) \neq 0\}) = 0,$$

from where $\mu(\{x \in \Omega |\ f(x) \neq 0\}) = 0$. $\square$

Finally, let us prove using Lemma 7.1.2 the following result:

Lemma 7.1.3 ***(du Bois-Reymond)***

Let f be a function of $L^1_{loc}(\Omega)$ such that

$$\forall u \in C_0(\Omega), \quad \int_\Omega f u d\mu = 0,$$

then $f = 0$ a.e. on Ω.

Proof. We want to show that

$$\forall x^0 \in \Omega \quad \exists B_r(x^0) : \quad f(x) = 0 \text{ a. e. in } B_r(x^0).$$

If it is true, then we apply Lemma 7.1.2, which finishes the proof.

Indeed, since Ω is a domain of $\mathbb{R}^n$, then Ω is open:

$$\forall x^0 \in \Omega \quad \exists \text{ open neigborhood } U_{x^0} : \quad U_{x^0} \subset \Omega.$$

In addition, since U_{x^0} is open, we can also say that

$$\forall x^0 \in \Omega \quad \exists B_r(x^0) : B_r(x^0) \subset \Omega.$$

Thus, if $\forall x_0 \in \Omega\ f(x) = 0$ a. e. in $B_r(x_0)$, then, by Lemma 7.1.2 with $U_{x_0} = B_r(x_0)$, we conclude that $f(x) = 0$ a. e. in Ω.

Let's take r such that $\overline{B_{2r}(x^0)} \subset \Omega$. We define

$$f_1(x) = \begin{cases} f(x), & x \in \overline{B_{2r}(x^0)} \\ 0, & x \notin B_{2r}(x^0) \end{cases}.$$

Thus, $f_1 \in L^1(\Omega)$. Let $x \in B_r(x^0)$. We take $h < r$ such that $B_h(x) \subset B_{2r}(x^0)$ (see Fig. 7.8). Let $f_{1,h}$ be the mollifier of f_1. Thus

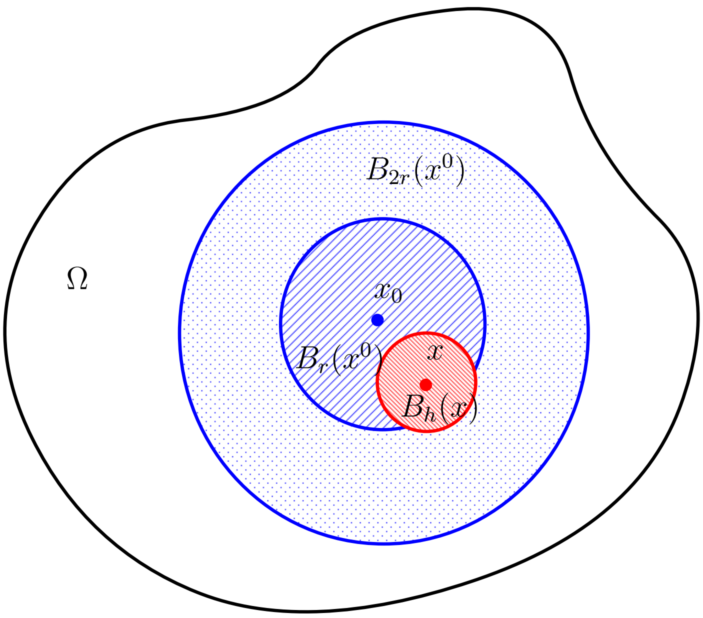

Figure 7.8 – Construction of the proof.

$$f_{1,h}(x) = \int_{B_h(x)} w_h(x-y) f_1(y) dy = \int_{B_h(x)} w_h(x-y) f(y) dy.$$

Since

$$\forall x \in B_r(x^0) \quad w_h(x-y) \in C_0^\infty(\Omega)$$

as a function of y, and

$$\operatorname{supp}(w_h(x-y)) = \overline{B_h(x)} \subset \overline{B_{2r}(x^0)},$$

then $f_{1,h}(x) = 0$ in $\overline{B_r(x^0)}$.

Therefore, we have

$$\int_{B_r(x^0)} |f(x)| dx = \int_{B_r(x^0)} |f_1(x)| dx = \int_{B_r(x^0)} |f_1(x) - f_{1,h}(x)| dx.$$

By point 5) of Proposition 7.1.1, that

$$\int_{B_r(x^0)} |f_1(x) - f_{1,h}(x)|dx \to 0 \quad \text{for } h \to 0.$$

But $\int_{B_r(x^0)} |f(x)|dx$ does not depends on h. Consequently, $\int_{B_r(x^0)} |f(x)|dx = 0$, from where $f = 0$ a.e. in $B_r(x^0)$. $\square$

7.2 Dual space of $\mathcal{D}(\Omega)$. Distributions

Definition 7.2.1 *Let $\mathcal{D}'(\Omega)$ be the dual space of $\mathcal{D}(\Omega)$: space of linear continuous functionals on $\mathcal{D}(\Omega)$. Elements $\mathcal{D}'(\Omega)$ of are called* ***distributions****.*

Example 7.2.1 ***(Regular distributions)*** *For f in $L^1_{loc}(\Omega)$ define $T_f : \mathcal{D}(\Omega) \to \mathbb{R}$ by*

$$T_f(\varphi) = \langle T_f, \varphi\rangle = \int_\Omega f(x)\varphi(x)dx.$$

T_f is linear and continuous: $T_f \in \mathcal{D}'(\Omega)$. Indeed, as $\operatorname{supp}\varphi$ is a compact in Ω, we have for all $l \in \mathbb{N}$

$$|\langle T_f, \varphi\rangle| = \left|\int_\Omega f(x)\varphi(x)dx\right| \le \int_{\operatorname{supp}\varphi} |f(x)|\,|\varphi(x)|dx \le \|\varphi\|_{C^l(\operatorname{supp}\varphi)} \int_{\operatorname{supp}\varphi} |f(x)|dx.$$

As $f \in L^1_{loc}(\Omega)$ and $\operatorname{supp}\varphi$ is a compact in Ω, we have

$$\int_{\operatorname{supp}\varphi} |f(x)|dx < \infty.$$

Thus, it follows that if $\varphi_k \to \varphi$ in $\mathcal{D}(\Omega)$ then $T_f(\varphi_k) \to T_f(\varphi)$, i.e., T_f is continuous.

It is convenient to identify $f \in L^1_{loc}(\Omega)$ with $T_f \in \mathcal{D}'(\Omega)$ and to consider the function f as being a distribution (namely that given by T_f). So, elements of $L^1_{loc}(\Omega)$ "are" distributions. They are called ***regular distributions****.*

Remark 7.2.1 *If $f_1,\ f_2 \in L^1_{loc}(\Omega)$ and $f_1 = f_2$ a.e. in Ω, then*

$$\forall\varphi \in \mathcal{D}(\Omega) \quad \langle f_1, \varphi\rangle = \langle f_2, \varphi\rangle$$

i.e.,

$$f_1 = f_2 \text{ in } \mathcal{D}'(\Omega).$$

The converse is also true thanks to du Bois-Reymond Lemma 7.1.3:

$$\text{if } f_1,\ f_2 \in L^1_{loc}(\Omega) \text{ and } f_1 = f_2 \text{ in } \mathcal{D}'(\Omega) \quad \Rightarrow \quad f_1 = f_2 \quad \text{a.e. in } \Omega.$$

Example 7.2.2 ***(Dirac)*** *Define $T : \mathcal{D}(\Omega) \to \mathbb{R}$ by*

$$T(\varphi) = \langle T, \varphi\rangle = \varphi(0).$$

T is linear and continuous: $T \in \mathcal{D}'(\Omega)$.

*T is called a **Dirac** (sometimes a "Dirac delta function"). This distribution is usually denoted by δ. It is not in $L^1_{loc}(\Omega)$ (prove it!).*

Example 7.2.3 ***(Single layer)*** *Let S be a piecewise regular (for instance, C^1) surface in $\mathbb{R}^n$ and μ be a continuous function defined on S. The distribution $\mu\delta_S \in \mathcal{D}'(\mathbb{R}^n)$, called a single layer, is defined by the formula:*

$$\forall \varphi \in \mathcal{D}(\mathbb{R}^n) \quad \langle \mu\delta_S, \varphi\rangle = \int_{\mathbb{R}^n} \mathbb{1}_S(x)\mu(x)\varphi(x)\mathrm{d}x,$$

where $\mathbb{1}_S(x) = 0$ if $x \notin S$ and $= 1$ if $x \in S$. We see that $\mu\delta_S$ is a generalization of the Dirac distribution for surfaces.

Problem 7.2.1 *An other example of a singular distribution is $\mathcal{P}\frac{1}{x}$, called the principal part of the inegral of $\frac{1}{x}$:*

$$\forall \varphi \in \mathcal{D}(\mathbb{R}) \quad \langle \mathcal{P}\frac{1}{x}, \varphi\rangle = \lim_{\epsilon\to+0}\left(\int_{-\infty}^{-\epsilon} + \int_{\epsilon}^{\infty}\right)\frac{\varphi(x)}{x}\mathrm{d}x.$$

Prove that $\mathcal{P}\frac{1}{x} \in \mathcal{D}'(\mathbb{R})$.

To describe the space $\mathcal{D}'(\Omega)$, we give (without the proof) the following Theorem which gives a criteria when a linear form on $D(\Omega)$ is continuous:

Theorem 7.2.1 *Let $T : D(\Omega) \to \mathbb{R}$ be a linear form. Then*

$$T \in \mathcal{D}'(\Omega) \quad \iff \quad \begin{array}{l} \forall \text{ bounded } \Omega' \subsetneq \Omega \quad \exists M = M(\Omega') \in]0,\infty[\text{ and } m = m(\Omega') \in \mathbb{N} \text{ such that} \\ \forall \varphi \in \mathcal{D}(\Omega') \quad |\langle T, \varphi\rangle| \le M\|\varphi\|_{C^m(\overline{\Omega'})}. \end{array}$$

Definition 7.2.2 *Let $T \in \mathcal{D}'(\Omega)$ be a distribution and $g \in C^\infty(\Omega)$. We define the operator*

$$gT : T \in \mathcal{D}'(\Omega) \mapsto g \cdot T \in \mathcal{D}'(\Omega)$$

of the product $g \cdot T$ by

$$\forall \varphi \in \mathcal{D}(\Omega) \quad \langle gT, \varphi\rangle = \langle T, g\varphi\rangle.$$

(Since $g \in C^\infty(\Omega)$ and $\varphi \in \mathcal{D}(\Omega)$ then $g\varphi \in \mathcal{D}(\Omega)$.)

Definition 7.2.3 *Let $T \in \mathcal{D}'(\Omega)$ be a distribution.*

*We define $T' : \mathcal{D}(\Omega) \to \mathbb{R}$, **the derivative of the distribution** T, by*

$$T'(\varphi) = -T(\varphi').$$

(Since $\varphi \in \mathcal{D}(\Omega)$, then $\varphi' \in \mathcal{D}(\Omega)$, and since $\mathcal{D}'(\Omega)$ is a linear vector space, $-T \in \mathcal{D}'(\Omega)$. Consequently, $T' \in \mathcal{D}'(\Omega)$, i.e. a linear continuous functional on $\mathcal{D}(\Omega)$.)

Therefore, we define the operator $D : \mathcal{D}'(\Omega) \to \mathcal{D}'(\Omega)$ by

$$\forall \varphi \in \mathcal{D}(\Omega), \quad \langle T', \varphi\rangle = -\langle T, \varphi'\rangle.$$

This is consistent with the derivative of functions thanks to the integration by parts: since φ has a compact support in Ω, then $\varphi(x) = 0$ for all $x \in \partial\Omega$ and the boundary term in the formula of the integration by parts is equal to zero.

Example 7.2.4 *We consider $\Omega = \mathbb{R}$. Let us consider the derivative of the Heaviside function:*

$$\theta(x) = \begin{cases} 0, & x < 0, \\ 1, & x \geq 0. \end{cases}$$

Clearly, $\theta \in L^1_{loc}(\mathbb{R})$ and thus $\theta \in \mathcal{D}'(\mathbb{R})$ with

$$\forall \varphi \in \mathcal{D}(\mathbb{R}) \quad \langle \theta, \varphi \rangle = \int_{\mathbb{R}} \theta(x)\varphi(x)dx.$$

Let us calculate θ':

$$\forall \varphi \in \mathcal{D}(\mathbb{R}) \quad \langle \theta', \varphi \rangle = -\langle \theta, \varphi' \rangle = -\int_{\mathbb{R}} \theta(x)\varphi'(x)dx = -\int_0^{+\infty} \varphi'(x)dx = -\varphi(x)|_0^{+\infty}$$
$$= -0 + \varphi(0) = \langle \delta, \varphi \rangle.$$

Finally,

$$\forall \varphi \in \mathcal{D}(\mathbb{R}) \quad \langle \theta', \varphi \rangle = \langle \delta, \varphi \rangle,$$

which means that $\theta' = \delta$ in the sense of distributions.

More generally,

Definition 7.2.4 *Let Ω be a domain in $\mathbb{R}^n$. Let $T \in \mathcal{D}'(\Omega)$ be a distribution and let $\alpha \in \mathbb{N}^n$ be a multi-index.*

We define the derivative operator $D^\alpha : \mathcal{D}'(\Omega) \to \mathcal{D}'(\Omega)$ by

$$\forall \varphi \in \mathcal{D}(\Omega) \quad \langle D^\alpha T, \varphi \rangle = (-1)^{|\alpha|} \langle T, \partial^\alpha \varphi \rangle.$$

Remark 7.2.2 *If $f \in C^\infty(\Omega)$ then all its derivatives $\partial^\alpha f$ (in the usual, or classical sense) are equal to the corresponding derivatives of f in the sense of distributions:*

- *As $f \in C^\infty(\Omega)$ then $f \in L^1_{loc}(\Omega)$ and thus $f \in \mathcal{D}'(\Omega)$. We summarize:*

$$C^\infty(\Omega) \subsetneq \mathcal{D}'(\Omega).$$

- *Therefore, by integration by parts,*

$$\forall \varphi \in \mathcal{D}(\Omega) \quad \langle D^\alpha f, \varphi \rangle = (-1)^{|\alpha|} \int_\Omega f(x)\partial^\alpha \varphi(x)dx = \int_\Omega \partial^\alpha f(x)\varphi(x)dx = \langle \partial^\alpha f, \varphi \rangle.$$

Example 7.2.5 *Let $T \in \mathcal{D}'(\mathbb{R}^3)$ be a distribution and let $\alpha = (1, 4, 5)$. By Definition 7.2.4, the operator $D^{(1,4,5)} : \mathcal{D}(\mathbb{R}^3) \to \mathbb{R}$ is defined by*

$$\forall \varphi \in \mathcal{D}(\mathbb{R}^3), \quad \langle D^{(1,4,5)} T, \varphi \rangle = \langle T, \partial^{(1,4,5)} \varphi \rangle$$

Corollary 7.2.1 *1. All distributions are infinitely derivable.*

2. All convergent series in $\mathcal{D}'(\Omega)$ are infinitely derivable term by term:

$$\text{if } (u_k)_{k \in \mathbb{N}} \subsetneq \mathcal{D}'(\Omega) \text{ such that } \sum_{k \in \mathbb{N}} u_k = S \in \mathcal{D}'(\Omega), \quad \text{then}$$

$$\forall \alpha \in \mathbb{N}^n \quad \sum_{k \in \mathbb{N}} D^\alpha u_k = D^\alpha S \text{ in } \mathcal{D}'(\Omega).$$

Proof. The first statement follows from the definition of a distribution: since $\varphi \in \mathcal{D}(\Omega)$, then for all $\alpha \in \mathbb{N}^n$ $\partial^\alpha \varphi \in \mathcal{D}(\Omega)$, therefore, for all $T \in \mathcal{D}'(\Omega)$ the distribution $D^\alpha T$ is correctly defined for all $\alpha \in \mathbb{N}^n$.

The second point follows from

$$S = \lim_{m\to\infty} \sum_{k=1}^m u_k \quad \text{in } \mathcal{D}'(\Omega),\ i.e. \quad \forall \varphi \in \mathcal{D}(\Omega) \quad \langle S, \varphi\rangle = \lim_{m\to\infty} \langle \sum_{k=1}^m u_k, \varphi\rangle,$$

$$\Rightarrow \forall \varphi \in \mathcal{D}(\Omega) \quad \forall \alpha \in \mathbb{N}^n \quad \langle D^\alpha S, \varphi\rangle = (-1)^{|\alpha|}\langle S, \partial^\alpha \varphi\rangle = (-1)^{|\alpha|} \lim_{m\to\infty} \langle \sum_{k=1}^m u_k, \partial^\alpha \varphi\rangle$$

$$= \lim_{m\to\infty} \langle D^\alpha (\sum_{k=1}^m u_k), \varphi\rangle = \lim_{m\to\infty} \langle \sum_{k=1}^m D^\alpha u_k, \varphi\rangle = \langle \sum_{k\in\mathbb{N}} D^\alpha u_k, \varphi\rangle. \Box$$

Remark 7.2.3 *For all $\Omega \subset \mathbb{R}^n$, it holds (prove it!)*

$$L^p(\Omega) \subset L^1_{loc}(\Omega).$$

Therefore functions in $L^p(\Omega)$ are regular distributions.

Thus, we can now differentiate any function in L^p! But the derivative may (or may not) be a function in L^p. When it is the case, we have a Sobolev space, we will see this in the next section.

7.2.1 Direct product of distributions

Let $f(x) \in L^1_{loc}(\mathbb{R}^n)$ and $g(y) \in L^1_{loc}(\mathbb{R}^m)$. Then the function $f(x)g(y) \in L^1_{loc}(\mathbb{R}^{n+m})$. Thus, fg define a regular distribution $T_{fg} = f \times g$ by the formula

$$\forall \varphi \in \mathcal{D}(\mathbb{R}^{n+m}) \quad \langle f \times g, \varphi\rangle = \int_{\mathbb{R}^{n+m}} f(x)g(y)\varphi(x,y)\mathrm{d}x\mathrm{d}y = \int_{\mathbb{R}^n} f(x) \int_{\mathbb{R}^m} g(y)\varphi(x,y)\mathrm{d}y\mathrm{d}x$$
$$= \langle T_f, \langle T_g, \varphi\rangle\rangle, \quad (7.3)$$

where we have used the Fubini Theorem (see [4]).

Definition 7.2.5 ***(Direct product in $\mathcal{D}'$)*** *The direct product $f \times g$ of two distributions $f \in \mathcal{D}'(\mathbb{R}^n)$ and $g \in \mathcal{D}'(\mathbb{R}^m)$ is called the distribution from $\mathcal{D}'(\mathbb{R}^{n+m})$ defined by the formula [see (7.3)]:*

$$\forall \varphi \in \mathcal{D}(\mathbb{R}^{n+m}) \quad \langle f \times g, \varphi\rangle = \langle f, \langle g, \varphi\rangle\rangle. \quad (7.4)$$

Let us verify that the definition is correct, *i.e.* the right hand side of (7.4) defines a linear continuous functional on $\mathcal{D}(\mathbb{R}^{n+m})$.

We use the following lemma (for the proof see [10])

Lemma 7.2.1 *For all $g \in \mathcal{D}'(\mathbb{R}^m)$ and $\varphi \in \mathcal{D}(\mathbb{R}^{n+m})$ the function*

$$\psi(x) = \langle g(y), \varphi(x,y)\rangle$$

belongs to $\mathcal{D}(\mathbb{R}^n)$, and for all $\alpha \in \mathbb{N}^n$

$$D^\alpha \psi(x) = \langle g(y), \partial_x^\alpha \varphi(x,y)\rangle. \quad (7.5)$$

In addition, if $\varphi_k \to 0$ for $k \to +\infty$ in $\mathcal{D}(\mathbb{R}^{n+m})$, then

$$\psi_k(x) = \langle g(y), \varphi_k(x,y)\rangle \to 0 \quad \text{for } k \to +\infty \text{ in } \mathcal{D}(\mathbb{R}^n).$$

Lemma 7.2.1 ensures that the right hand side of (7.4) define a functional on $\mathcal{D}(\mathbb{R}^{n+m})$. From the linearity of the functionals f and g it follows the linearity of $f \times g$. Let us prove that the linear functional $f \times g$ is continuous on $\mathcal{D}(\mathbb{R}^{n+m})$. Suppose $\varphi_k \to 0$, $k \to \infty$ in $\mathcal{D}(\mathbb{R}^{n+m})$. Then, by Lemma 7.2.1,

$$\langle g, \varphi_k\rangle \to 0, \quad k \to +\infty \quad \text{in } \mathcal{D}(\mathbb{R}^n).$$

Therefore, since f is continuous on $\mathcal{D}(\mathbb{R}^n)$, we obtain that

$$\langle f, \langle g, \varphi_k\rangle\rangle \to 0, \quad k \to +\infty,$$

which means that $f \times g$ is continuous. Hence, $f \times g \in \mathcal{D}'(\mathbb{R}^{n+m})$.

We give the main properties of the direct product (for the proof see [10]):

Proposition 7.2.1 *Let $f \in \mathcal{D}'(\mathbb{R}^n)$ and $g \in \mathcal{D}'(\mathbb{R}^m)$.*

1. *The direct product in $\mathcal{D}'$ is commutative:*

$$f \times g = g \times f \in \mathcal{D}'(\mathbb{R}^{n+m}).$$

2. *The direct product in $\mathcal{D}'$ is associative: if $f \in \mathcal{D}'(\mathbb{R}^n)$, $g \in \mathcal{D}'(\mathbb{R}^m)$ and $h \in \mathcal{D}'(\mathbb{R}^\ell)$, then*

$$f \times [g \times h] = [f \times g] \times h \in \mathcal{D}'(\mathbb{R}^{m+n+\ell}).$$

3. *The direct product $f \times g$ is linear and continuous with respect to f, as the mapping*

$$f \in \mathcal{D}'(\mathbb{R}^n) \mapsto f \times g \in \mathcal{D}'(\mathbb{R}^{n+m}),$$

and with respect to g, as the mapping

$$g \in \mathcal{D}'(\mathbb{R}^m) \mapsto f \times g \in \mathcal{D}'(\mathbb{R}^{n+m}).$$

For instance,

$$[\lambda f + \mu f_1] \times g = \lambda[f \times g] + \mu[f_1 \times g] \quad \forall f, f_1 \in \mathcal{D}'(\mathbb{R}^n),\ g \in \mathcal{D}'(\mathbb{R}^m), \lambda, \mu \in \mathbb{R},$$

and if $f_k \to 0$ in $\mathcal{D}'(\mathbb{R}^n)$, then $f_k \times g \to 0$, $k \to +\infty$ in $\mathcal{D}'(\mathbb{R}^{n+m})$.

4. *Derivation of the direct product:*

$$D_x^\alpha[f(x) \times g(y)] = D^\alpha f(x) \times g(y).$$

5. *Multiplication by a C^∞-function: if $a \in C^\infty(\mathbb{R}^n)$ then*

$$a(x)[f(x) \times g(y)] = a(x)f(x) \times g(y).$$

6. *Translation on a vector:*

$$(f \times g)(x+h, y) = f(x+h) \times g(y).$$

7. *Distribution $f(x) \times 1(y)$ is independing on y: for all $\varphi \in \mathcal{D}(\mathbb{R}^{n+m})$ it holds*

$$\langle f(x) \times 1(y), \varphi\rangle = \langle f, \int_{\mathbb{R}^m} \varphi(x,y)\mathrm{d}y\rangle = \langle 1(y) \times f(x), \varphi\rangle = \int_{\mathbb{R}^m} \langle f(x), \varphi(x,y)\rangle \mathrm{d}y.$$

Therefore, for all $f \in \mathcal{D}'(\mathbb{R}^n)$ and $\varphi \in \mathcal{D}(\mathbb{R}^{n+m})$ it holds

$$\langle f, \int_{\mathbb{R}^m} \varphi(x,y)\mathrm{d}y\rangle = \int_{\mathbb{R}^m} \langle f(x), \varphi(x,y)\rangle \mathrm{d}y.$$

7.2.2 Convolution in $\mathcal{D}'(\Omega)$

Let f and g be two functions in $L^1_{loc}(\mathbb{R}^n)$ and

$$h(x) = \int_{\mathbb{R}^n} |g(y)f(x-y)|\mathrm{d}y \in L^1_{loc}(\mathbb{R}^n).$$

As it was defined in [4], the convolution $f * g$ is a function

$$(f * g)(x) = \int_{\mathbb{R}^n} f(y)g(x-y)\mathrm{d}y = \int_{\mathbb{R}^n} g(y)f(x-y)\mathrm{d}y = (g * f)(x).$$

Note that $f * g$ and $|f| * |g| = h$ exist in the same time and, since

$$\text{for almost all } x \quad |(f*g)(x)| \le h(x) \quad \Rightarrow \quad f * g \in L^1_{loc}(\mathbb{R}^n),$$

then $f * g$ is a regular distribution. Moreover, for all $\varphi \in \mathcal{D}(\mathbb{R}^n)$, using Fubini Theorem from [4],

$$\begin{aligned}\langle f*g, \varphi\rangle &= \int_{\mathbb{R}^n} (f*g)(\xi)\varphi(\xi)\mathrm{d}\xi = \int_{\mathbb{R}^n} \left[\int_{\mathbb{R}^n} g(y)f(\xi-y)\mathrm{d}y\right]\varphi(\xi)\mathrm{d}\xi \\ &= \int_{\mathbb{R}^n} g(y)\left[\int_{\mathbb{R}^n} f(\xi-y)\varphi(\xi)\mathrm{d}\xi\right]\mathrm{d}y = \int_{\mathbb{R}^n} g(y)\left[\int_{\mathbb{R}^n} f(x)\varphi(x+y)\mathrm{d}x\right]\mathrm{d}y,\end{aligned}$$

i.e.

$$\forall \varphi \in \mathcal{D}(\mathbb{R}^n) \quad \langle f*g, \varphi\rangle = \int_{\mathbb{R}^{2n}} f(x)g(y)\varphi(x+y)\mathrm{d}x\mathrm{d}y. \tag{7.6}$$

Here, the integration is over $\mathbb{R}^{2n}$ while the testing function $\varphi \in \mathcal{D}(\mathbb{R}^n)$.

Let the sequence (η_k) of functions from $\mathcal{D}(\mathbb{R}^n)$ converges in $\mathbb{R}^n$ toward to 1. For example, let $\eta \in \mathcal{D}(\mathbb{R}^n)$ such that $\eta(x) = 1$ for all x in the unit ball of $\mathbb{R}^n$, then we can define $\eta_k(x) = \eta\left(\frac{x}{k}\right)$.

We admit (see [10]) that equation (7.6) can be written in the form

$$\forall \varphi \in \mathcal{D}(\mathbb{R}^n) \quad \langle f*g, \varphi\rangle = \lim_{k\to\infty} \langle f \times g, \eta_k(x,y)\varphi(x+y)\rangle, \tag{7.7}$$

where (η_k) is any sequence converging toward 1 in $\mathbb{R}^{2n}$. We note that for all k the function $\eta_k(x,y)\varphi(x+y)$ belongs to $\mathcal{D}(\mathbb{R}^{2n})$. Let us take f and g from $\mathcal{D}'(\mathbb{R}^n)$ in a such way that

that their direct product $f \times g$ allows an extention $\langle f(x) \times g(y), \varphi(x+y)\rangle$ on the functions of the form $\varphi(x+y)$ (for all $\varphi \in \mathcal{D}(\mathbb{R}^n)$) in the following sense:

for any sequence (η_k) from $\mathcal{D}(\mathbb{R}^n)$, converging in $\mathbb{R}^n$ toward to 1, there exists a limite (independing on the choice of (η_k)) of the numerical sequence

$$\lim_{k\to\infty} \langle f(x) \times g(y), \eta_k(x,y)\varphi(x+y)\rangle = \langle f(x) \times g(y), \varphi(x+y)\rangle.$$

Definition 7.2.6 *Convolution of two distributions f and g from $\mathcal{D}'(\mathbb{R}^n)$ is called the distribution $f * g \in \mathcal{D}'(\mathbb{R}^n)$ given by the formula*

$$\forall \varphi \in \mathcal{D}(\mathbb{R}^n) \quad \langle f * g, \varphi\rangle = \langle f(x) \times g(y), \varphi(x+y)\rangle = \lim_{k\to\infty} \langle f(x) \times g(y), \eta_k(x,y)\varphi(x+y)\rangle, \tag{7.8}$$

where (η_k) is any sequence converging toward 1 in $\mathbb{R}^{2n}$.

Remark 7.2.4 *As $\varphi(x+y)$ does not belong to $\mathcal{D}(\mathbb{R}^{2n})$ (it has not a compact support in $\mathbb{R}^{2n}$), the right hand part of equation (7.8) can be not exist. Hence the existence of the convolution depends on the choice of the distributions f and g.*

Example 7.2.6 *For all distribution f there exists its convolution with the Dirac distribution δ and*

$$f * \delta = \delta * f = f.$$

Let us give the main properties of the convolution (for the proof see [10]):

Proposition 7.2.2 *Let $f, g \in \mathcal{D}'(\mathbb{R}^n)$.*

1. *The mappings*

   $$f \in \mathcal{D}'(\mathbb{R}^n) \mapsto f * g \in \mathcal{D}'(\mathbb{R}^n) \text{ and } g \in \mathcal{D}'(\mathbb{R}^n) \mapsto f * g \in \mathcal{D}'(\mathbb{R}^n)$$

   are linear but not continuous. (For example, $\delta(x-k) \to 0$, $k \to \infty$ in $\mathcal{D}'(\mathbb{R})$, but $1 * \delta(x-k) = 1 \nrightarrow 0$, $k \to \infty$ in $\mathcal{D}'(\mathbb{R})$.)

2. *If the convolution $f * g$ exists, then there exists the convolution $g * f$ and*

   $$f * g = g * f.$$

3. *If the convolution $f * g$ exists, then there exist the convolutions $D^\alpha f * g$ and $f * D^\alpha g$. In addition*

   $$D^\alpha f * g = D^\alpha (f * g) = f * D^\alpha g.$$

   But from the existence of the convolutions $D^\alpha f * g$ and $f * D^\alpha g$ does not follows the existence of the convolution $f * g$:

   $$\theta' * 1 = \delta * 1 = 1, \quad \text{but} \quad \theta * 1' = \theta * 0 = 0,$$

   where θ is the Heaviside function. Hence, the convolution is not associative:

   $$(\theta * \delta') * 1 = \theta' * 1 = 1, \quad \text{but} \quad \theta * (\delta' * 1) = \theta * 0 = 0.$$

   If there exist the convolutions $f * g * h$, $f * g$, $g * h$ and $f * h$ then there exist the convolutions $(f * g) * h$, $f * (g * h)$ and $(f * h) * g$. In this case, it holds

   $$f * g * h = (f * g) * h = f * (g * h) = (f * h) * g.$$

4. *If the convolution $f * g$ exists, then there exists the convolution $f(x+h) * g(x)$ and*

$$\forall h \in \mathbb{R}^n \quad f(x+h) * g(x) = (f * g)(x+h),$$

i.e. the translation and the convolution commute.

Chapter 8

Sobolev spaces

8.1 Weak derivatives

Let Ω be a domain. As we know

$$L^p(\Omega) \subset L^1_{loc}(\Omega) = \{\text{regular distributions}\}.$$

Thus

$$\forall f \in L^p(\Omega) \quad \forall \alpha \in \mathbb{N}^n \quad \exists D^\alpha f \in \mathcal{D}'(\Omega).$$

When can we say that for $f \in L^p(\Omega)$ all its derivatives of the order $|\alpha| \leq m$ $D^\alpha f \in L^p(\Omega)$?

Definition 8.1.1 *Let $f \in L^1_{loc}(\Omega)$. If there exists $w \in L^1_{loc}(\Omega)$ such that*

$$\forall \varphi \in \mathcal{D}(\Omega) \quad \int_\Omega w(x)\phi(x)\mathrm{d}x = (-1)^{|\alpha|} \int_\Omega f(x)\partial^\alpha \phi(x)\mathrm{d}x,$$

then $w(x) = \partial^\alpha f(x)$ is called the weak derivative of u (of the order α).

Remark 8.1.1 *The weak derivative is more restrictive than the derivative in the sense of distributions. The advantage of the definition of the weak derivative, that it maps the regular distributions to the regular distributions.*

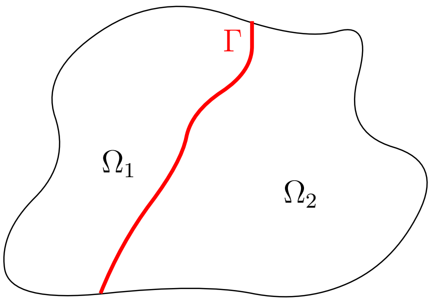

Figure 8.1 – Domain Ω devided in three parts: $\Omega = \Omega_1 \cup \Gamma \cup \Omega_2$.

We notice from Definition 8.1.1 the following important properties of the weak derivative:

Proposition 8.1.1 *1. Let $\partial^\alpha u \in C(\Omega)$ be a "classical" derivative of u, defined for all $x \in \Omega$. Then $\partial^\alpha u$ is the weak derivative of u of the order α on Ω.*

2. *If $\partial^\alpha f \in L^1_{loc}(\Omega)$, then for all subdomain $\tilde{\Omega} \subset \Omega$ $\partial^\alpha f \in L^1_{loc}(\tilde{\Omega})$.*
3. *The weak derivative is uniquely defined up to the equivalence (up to the equivalence classes).*
4. *From the existence of the weak derivative of the order α does not follows the existence the weak derivatives of smaller orders.*
5. *Let Ω b devided in tree parts: two subdomains Ω_1 and Ω_2 and their commun boundary set Γ in such a way that, $\Omega_1 \cap \Omega_2 = \varnothing$, and $\Omega = \Omega_1 \cup \Gamma \cup \Omega_2$ (see Fig. 8.1). From the existence of the weak derivative $\partial^\alpha u$ in each Ω_i $(i = 1, 2)$ does not implies the existence of the weak derivative $\partial^\alpha u$ in all Ω.*

8.2 $W^{m,p}$ Spaces

8.2.1 Definition and main properties

Let Ω be a domain in $\mathbb{R}^n$ (not necessarily bounded).

Definition 8.2.1 *Let $m \in \mathbb{N}$ and $p \in [1, \infty]$. The Sobolev Space $W^{m,p}(\Omega)$ is defined by*

$$W^{m,p}(\Omega) = \{f \in L^p(\Omega) |\ D^\alpha f \in L^p(\Omega) \text{ for any } |\alpha| \leq m\}$$

(derivative in the distributional sense) with a norm

$$\|f\|_{m,p} = \|f\|_{W^{m,p}(\Omega)} = (\sum_{0\leq|\alpha|\leq m} \int_\Omega |D^\alpha f|^p)^{\frac{1}{p}} = (\sum_{0\leq|\alpha|\leq m} \|D^\alpha f\|^p_{L^p(\Omega)})^{\frac{1}{p}} \quad \text{for } p < \infty \tag{8.1}$$

$$\|f\|_{m,\infty} = \|f\|_{W^{m,\infty}(\Omega)} = \max_{0\leq|\alpha|\leq m} \|D^\alpha f\|_{L^\infty(\Omega)} \quad \text{for } p = \infty. \tag{8.2}$$

Problem 8.2.1 *The norm (8.1) is equivalent in $W^{m,p}(\Omega)$ to the norm*

$$N(f) = \sum_{0\leq|\alpha|\leq m} \|D^\alpha f\|_{L^p(\Omega)}. \tag{8.3}$$

Therefore, sometimes the norm in $W^{m,p}(\Omega)$ is directely defined by (8.3). We notice that for $p = 1$ the norms (8.1) and (8.3) are identically the same and for $p = \infty$ the norms (8.2) and (8.3) are also equivalent.

Remark 8.2.1 *We notice that*

$$W^{0,p}(\Omega) = L^p(\Omega).$$

Theorem 8.2.1 *Let $m \in \mathbb{N}$ and $p \in [1, \infty]$. The Sobolev space $W^{m,p}(\Omega)$ with the norm (8.3) (and thus with (8.1)-(8.2)) is a Banach space.*

Proof. Let (u_i) be a Cauchy sequence in $W^{m,p}(\Omega)$:

$$\forall \epsilon > 0 \quad \exists N \in \mathbb{N} : \quad \forall m, k \geq N \quad \|u_m - u_k\|_{W^{m,p}(\Omega)} < \epsilon.$$

The inequality $\|u_m - u_k\|_{W^{m,p}(\Omega)} < \epsilon$ means that

$$\sum_{0\leq|\alpha|\leq m} \|D^\alpha u_m - D^\alpha u_k\|_{L^p(\Omega)} < \epsilon.$$

Then for $0 \le |\alpha| \le m$ the sequence $(D^\alpha u_i)$ is a Cauchy sequence in $L^p(\Omega)$. Since $L^p(\Omega)$ is complete, there exist functions $v \in L^p(\Omega)$ and $v_\alpha \in L^p(\Omega)$ for $1 \le |\alpha| \le m$ such that

$$u_i \to v \text{ in } L^p(\Omega) \quad \text{and} \quad D^\alpha u_i \to v_\alpha \text{ in } L^p(\Omega) \quad (1 \le |\alpha| \le m).$$

To finish the proof, we need to show that

$$D^\alpha v = v_\alpha \text{ for all } 1 \le |\alpha| \le m,$$

from where we could conclude that $v \in W^{m,p}(\Omega)$ and thus (u_i) converge to $v \in W^{m,p}(\Omega)$ in $W^{m,p}(\Omega)$.

As $L^p(\Omega) \subset L^1_{loc}(\Omega)$ and so u_i determines a regular distribution T_{u_i} (or itself can be considered as a regular distribution):

$$\forall \varphi \in \mathcal{D}(\Omega) \quad < u_i, \varphi >= \int_\Omega u_i(x)\varphi(x)dx.$$

For any $\varphi \in \mathcal{D}(\Omega)$ we have

$$| < u_i, \varphi > - < v, \varphi > | \le \int_\Omega |u_i(x) - v(x)||\varphi(x)|dx \le \|\varphi\|_{L^{p'}(\Omega)} \|u_i - v\|_{L^p(\Omega)},$$

where we have applied the Hölder inequality. Hence $< u_i, \varphi > \to < v, \varphi >$ for every $\varphi \in \mathcal{D}(\Omega)$ as $i \to \infty$.

Similarly,

$$\forall \varphi \in \mathcal{D}(\Omega) \quad < D^\alpha u_i, \varphi > \to < v_\alpha, \varphi > \quad i \to \infty.$$

It follows that

$$\forall \varphi \in \mathcal{D}(\Omega) \quad < v_\alpha, \varphi >= \lim_{i\to\infty} < D^\alpha u_i, \varphi >= \lim_{i\to\infty} (-1)^{|\alpha|} < u_i, \partial^\alpha \varphi >= (-1)^{|\alpha|} < v, \partial^\alpha \varphi > .$$

Thus $D^\alpha v = v_\alpha$ for all $1 \le |\alpha| \le m$, from where $v \in W^{m,p}(\Omega)$ and $\|u_i - v\|_{W^{m,p}(\Omega)} \to 0$ for $i \to \infty$. $\square$

Remark 8.2.2 *If classical partial derivatives exist and are continuous then they coincide with the distributional partial derivative. Thus the set*

$$S = \{f \in C^m(\Omega) | \|f\|_{W^{m,p}(\Omega)} < \infty\}$$

is contained in $W^{m,p}(\Omega)$. Since $W^{m,p}(\Omega)$ is complete, it is possible to prove that

$$\overline{S}^{\|\cdot\|_{W^{m,p}(\Omega)}} = W^{m,p}(\Omega)$$

(it is the theorem of Meyers and Serrin, see [1] p.52). It means that $C^m(\Omega)$ is dense in $W^{m,p}(\Omega)$.

We give without the proof the following Proposition (see [1] for more details):

Proposition 8.2.1 *The function*

$$|f|_{m,p} = |f|_{W^{m,p}(\Omega)} = (\sum_{|\alpha|=m} \int_\Omega |D^\alpha f|^p)^{\frac{1}{p}}$$

is a semi-norm on the Sobolev space $W^{m,p}(\Omega)$. The function

$$N(f) = \|f\|_{L^p(\Omega)} + \sum_{|\alpha|=m} \|D^\alpha f\|_{L^p(\Omega)}$$

is a norm in $W^{m,p}(\Omega)$. For bounded and regular (see Definition 8.2.3) $\partial\Omega$ this norm is equivalent to the norm $\|\cdot\|_{W^{m,p}(\Omega)}$ (see (8.3) or (8.1) and (8.2)).

Example 8.2.1 *Let us consider*

$$W^{2,3}(]0,1[) = \{f \in L^3(]0,1[) |\ f' \in L^3(]0,1[), f'' \in L^3(]0,1[)\}$$

with the norm

$$\|f\|_{W^{2,3}(]0,1[)} = (\int_{]0,1[} |f|^3 + \int_{]0,1[} |f'|^3 + \int_{]0,1[} |f''|^3)^{\frac{1}{3}}.$$

Here $x \mapsto x^{\frac{3}{2}}$ *does not belong to* $W^{2,3}(]0,1[)$*, but* $x \mapsto x^{\frac{9}{5}}$ *belongs to* $W^{2,3}(]0,1[)$*.*

Example 8.2.2 *The Sobolev space*

$$W^{1,2}(\mathbb{R}) = \{f \in L^2(\mathbb{R}) |\ f' \in L^2(\mathbb{R})\},$$

endowed with the norm

$$\|f\|_{W^{1,2}(\mathbb{R})} = (\int_{\mathbb{R}} |f|^2 + \int_{\mathbb{R}} |f'|^2)^{\frac{1}{2}}$$

is a Hilbert space with the inner product

$$\langle f, g\rangle = \int_{\mathbb{R}} fg + \int_{\mathbb{R}} f'g'.$$

Theorem 8.2.2 $W^{m,p}(\Omega)$ *is separable if* $1 \le p < \infty$*, and is reflexive and uniformly convex if* $1 < p < \infty$*.*

Proof. (See [1] for more details.)

Let us present $W^{m,p}(\Omega)$ as a closed subspace of a Cartesian product of spaces $L^p(\Omega)$. Let N be the number of multi-indices α satisfying $0 \le |\alpha| \le m$.

For $1 \le p \le \infty$ let

$$L^p_N = \prod_{j=1}^{N} L^p(\Omega).$$

We define the norm of $u = (u_1, \ldots, u_N)$ in L^p_N by

$$\|u\|_{L^p_N} = \begin{cases} \left(\sum_{j=1}^N \|u_j\|^p_{L^p}\right)^{\frac{1}{p}}, & \text{if } 1 \le p < \infty, \\ \max_{1\le j\le N} \|u_j\|_{L^\infty}, & \text{if } p = \infty. \end{cases}$$

Using the properties of the L^p spaces, we obtain that L^p_N is a Banach space that is separable if $1 \le p < \infty$ and reflexive and uniformly convex if $1 < p < \infty$.

Let us suppose that the N multi-indices α satisfying $0 \le |\alpha| \le m$ are linearly ordered in some convenient fashion so that to each $u \in W^{m,p}(\Omega)$ we may associate the well-defined vector $Pu \in L^p_N$ given by

$$Pu = (D^\alpha u)_{0\le|\alpha|\le m}.$$

Since

$$\|Pu\|_{L^p_N} = \|u\|_{W^{m,p}(\Omega)},$$

P is an isomorphism of $W^{m,p}(\Omega)$ onto a subspace $W \subset L^p_N$.

Since $W^{m,p}(\Omega)$ is complete, W is a closed subspace of L^p_N. Thus, W is separable if $1 \le p < \infty$ and is reflexive and uniformly convex if $1 < p < \infty$.

The same conclusions must therefore hold for $W^{m,p}(\Omega) = P^{-1}(W)$. $\square$

Let us also give the following properties of Sobolev spaces $W^{m,p}$:

Proposition 8.2.2 *1. $\forall \Omega' \subsetneq \Omega$ if $f \in W^{m,p}(\Omega)$, then $f \in W^{m,p}(\Omega')$.*

2. If $f \in W^{m,p}(\Omega)$ and $g \in C^m(\overline{\Omega})$, then $gf \in W^{m,p}(\Omega)$.

8.2.2 Sobolev spaces $W^{m,p}(\Omega)$ for bounded domains with a regular boundary $\partial\Omega$

Definition 8.2.2 *Let Ω be a bounded domain of $\mathbb{R}^n$. Its boundary $\partial\Omega$ is* ***locally Lipschitz*** *if*

$\forall x \in \partial\Omega \quad \exists$ *neighborhood U_x : $\partial\Omega \cap U_x$ is the graph of a Lipschitz continuous function f.*

Remark 8.2.3 *Here by a Lipschitz continuous function we understand $f : \mathbb{R}^{n-1} \to \mathbb{R}$ which is continuous and there exists $M > 0$ such that*

$$\forall x, y \in \mathbb{R}^{n-1} \quad |f(x) - f(y)| \leq M\|x - y\|_{\mathbb{R}^{n-1}}.$$

Definition 8.2.3 *The boundary $\partial\Omega$ is called* ***a regular boundary of the class*** *C^k ($k \geq 1$, see Fig. 8.2):*

$$\forall x_0 \in \partial\Omega \quad \exists f(x) \in C^k(U_{x_0}) \text{ such that } \{x \in U_{x_0} |\, f(x) = 0\} = \partial\Omega \cap \overline{U}_{x_0}, \quad df|_{U_{x_0}} \neq 0.$$

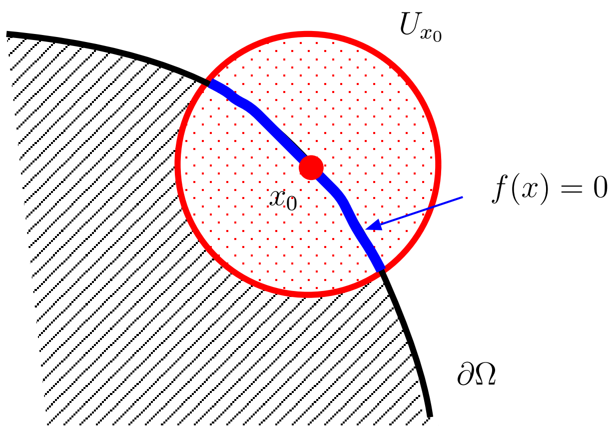

Figure 8.2 – Example of a regular boundary from Definition 8.2.3.

If Ω is a bounded domain with a regular boundary (at least locally Lipschitz), then we have following results:

Theorem 8.2.3 *Let Ω be a bounded domain with a locally Lipschitz boundary. Then it holds*

1. *$C^\infty(\overline{\Omega})$ is dense in $W^{m,p}(\Omega)$.*

2. *We can extend $f \in W^{1,p}(\Omega)$ from Ω to a bigger domain Ω' ($\Omega \subsetneq \Omega'$) without loss of regularity: $\tilde{f} \in W^{1,p}(\Omega')$ and $\tilde{f} = f$ on Ω. In addition, it holds*

$$\|\tilde{f}\|_{W^{1,p}(\Omega')} \leq C(p, \Omega, \Omega')\|f\|_{W^{1,p}(\Omega)},$$

where the constant $C(p, \Omega, \Omega')$ does not depend on f.

8.3 H^m Spaces

8.3.1 Definition and properties

Definition 8.3.1 *Let Ω be a domain in $\mathbb{R}^n$ and $m \in \mathbb{N}$. Define $H^m(\Omega) = W^{m,2}(\Omega)$. For f and g in $H^m(\Omega)$, we define a inner product*

$$\langle f, g\rangle_{H^m(\Omega)} = \sum_{0\le|\alpha|\le m} \int_\Omega D^\alpha f D^\alpha g = \langle f, g\rangle_{L^2(\Omega)} + \sum_{|\alpha|=1}^{m} \langle D^\alpha f, D^\alpha g\rangle_{L^2(\Omega)}, \tag{8.4}$$

which is associated with the norm

$$\|f\|_m = \|f\|_{H^m(\Omega)} = (\sum_{|\alpha|\le m} \int_\Omega |D^\alpha f|^2)^{\frac{1}{2}} = (\sum_{|\alpha|\le m} \|D^\alpha f\|^2_{L^2(\Omega)})^{\frac{1}{2}}.$$

Corollary 8.3.1 *$H^m(\Omega)$ with the inner product (8.4) is a separable Hilbert space.*

Proposition 8.3.1 *1. Canonic injection of $H^m(\Omega)$ in $L^2(\Omega)$ is continuous:*

$$\forall v \in H^m(\Omega) \quad \|v\|_{L^2(\Omega)} \le \|v\|_{H^m(\Omega)}.$$

2. $H^m(\Omega)$ is dense in $L^2(\Omega)$.

Proof.

1. It is a direct corollary of the definition of the norms in H^m and in L^2.
2. Firstly, we notice that
$$\mathcal{D}(\Omega) \subsetneq H^m(\Omega) \subsetneq L^2(\Omega).$$
Secondly, $\mathcal{D}(\Omega)$ is dense in $L^2(\Omega)$ by the norm $\|\cdot\|_{L^2}$ (see Theorem 7.1.1). Thus, $H^m(\Omega)$ is dense in $L^2(\Omega)$. □

8.3.2 Dual space $(H^m(\Omega))^*$

As $H^m(\Omega)$ is a normed (Banach) space, then its dual space is $(H^m(\Omega))^* = \mathcal{L}(H^m(\Omega), \mathbb{R})$. Since $H^m(\Omega)$ is a Hilbert space, the Riesz representation theorem yields

$$(H^m(\Omega))^* = H^m(\Omega),$$

where elements of $(H^m(\Omega))^*$ are linear continuous forms defined by the inner product of $H^m(\Omega)$. It means that to $f \in H^m(\Omega)$ corresponds the following element of $(H^m(\Omega))^*$:

$$u \in H^m(\Omega) \mapsto \langle f, u\rangle_{H^m(\Omega)} = \sum_{|\alpha|\le m} \int_\Omega D^\alpha f \, D^\alpha u.$$

By definition of H^m, we see that

$$H^m(\Omega) \subsetneq L^2(\Omega),$$

where $L^2(\Omega)$ is also a Hilbert space for $\langle f, g\rangle_{L^2(\Omega)} = \int_\Omega fg$. To avoid this kind of inclusions

$$H^m(\Omega) = (H^m(\Omega))^* \subsetneq L^2(\Omega)^* = L^2(\Omega),$$

which are not true, we don't identify the dual space of H^m with H^m, but identify a subspace of $(H^m)^*$ in such a way that holds

$$H^m \subsetneq L^2 = (L^2)^* \subsetneq (H^m)^*.$$

Definition 8.3.2 *By tradition, $L^2(\Omega)^*$ is always identified with $L^2(\Omega)$ by Riesz Theorem, but H^m is not identified with $(H^m)^*$. By the mapping $\Phi(f) : L^2(\Omega) \to (H^m(\Omega))^*$ defined for all $u \in H^m(\Omega)$ by*

$$f \in L^2(\Omega) \mapsto [\Phi(f)](u) = \langle f, u\rangle_{L^2(\Omega)} \in (H^m(\Omega))^*,$$

we indentify a subspace of $(H^m)^*$ with L^2:

$$H^m \subsetneq L^2 = (L^2)^* \subsetneq (H^m)^*.$$

For instance, for a fixed $f \in L^2(\Omega)$ the application

$$T: \ u \in H^m(\Omega) \to \langle f, u\rangle_{L^2(\Omega)} \in \mathbb{R}$$

is a linear continuous form on $L^2(\Omega)$ and also on $H^m(\Omega)$. Thus $T \in (H^m(\Omega))^*$. We define $\Phi(f) : L^2(\Omega) \to (H^m(\Omega))^*$ by

$$f \in L^2(\Omega) \mapsto \Phi(f) = [u \in H^m(\Omega) \to \langle f, u\rangle \in \mathbb{R}] = T \in (H^m(\Omega))^*.$$

Let us prove the following properties of $\Phi(f)$:

Lemma 8.3.1

1. *$\|\Phi(f)\|_{(H^m)^*} \le \|f\|_{L^2}$*
2. *$\Phi(f)$ is injective*
3. *$\Phi(L^2)$ is dense in $(H^m)^*$*

Proof.

1. $\underline{\|\Phi(f)\|_{(H^m)^*} \le \|f\|_{L^2}}$:

   We have for all $u \in H^m$

   $$|\langle f, u\rangle_{L^2}| \le \|f\|_{L^2}\|u\|_{L^2} \le \|f\|_{L^2}\|u\|_{H^m},$$

   from where

   $$\|\Phi(f)\|_{(H^m)^*} = \sup_{u\neq 0} \frac{|\langle f, u\rangle_{L^2}|}{\|u\|_{H^m}} \le \|f\|_{L^2}.$$

2. $\underline{\Phi(f) \text{ is injective}}$:

   By definition of $\Phi(f)$,

   $$\forall u \in H^m, \quad \forall f \in L^2 \quad \Phi(f) = \langle f, u\rangle_{L^2}.$$

   Therefore,

   $$\mathsf{Ker}\Phi(f) = \{f \in L^2 | \ \forall u \in H^m \quad \langle f, u\rangle_{L^2} = 0\}.$$

   As H^m is dense in L^2 (see Proposition 8.3.1), it follows that if $f \in \mathsf{Ker}\Phi(f)$ then

   $$\forall u \in L^2 \quad \langle f, u\rangle_{L^2} = 0 \quad \Longleftrightarrow \quad f = 0.$$

3. $\Phi(L^2)$ is dense in $(H^m)^*$:

   Let us apply Corollary 3.4.4 of Chapter 3: By Theorem 3.2.1, $(H^m)^*$ is a Banach space. Clearly, $\Phi(L^2) = \mathsf{Im}(\Phi(f))$ is a subspace of $(H^m)^*$. We also notice that H^m is dense in L^2 and $H^m \subset (H^m)^{**} \subset L^2$.

   If $\Phi(L^2)$ is not dense in $(H^m)^*$, then, by Corollary 3.4.4,

$$\exists u \in H^m \quad u \neq 0 \text{ such that } \forall f \in L^2 \quad \langle u, f\rangle_{L^2} = 0,$$

   which is not possible $(u = 0)$. □

Therefore, using $\Phi(f)$, we prove the inclusion of $L^2 = (L^2)^*$ in $(H^m)^*$:

$$H^m \subsetneq L^2 = (L^2)^* \subsetneq (H^m)^*.$$

L^2 is called **the pivot space**. From now on, we will make this choice.

8.3.3 Trace operator

In what follows, we suppose that Ω is a bounded domain of $\mathbb{R}^n$ and its boundary $\partial\Omega$ is locally Lipschitz, writing that $\partial\Omega$ is regular or sufficiently smooth.

We note (see [3] p. 252 for the proof)

Proposition 8.3.2 *For a regular $\partial\Omega$ of a bounded domain Ω, $C^\infty(\overline{\Omega})$ is dense in $H^m(\Omega)$.*

For a regular $\partial\Omega$ of the class C^1 of a bounded domain Ω, it is possible to define a boundary measure and define the space $L^2(\partial\Omega)$:

$$L^2(\partial\Omega) = \{v \text{ measurable on } \partial\Omega | \int_{\partial\Omega} |v(x)|^2 dx < \infty\}.$$

For $u \in C^\infty(\overline{\Omega})$ we have a continuous mapping

$$T: \ u \in C^\infty(\overline{\Omega}) \mapsto u|_{\partial\Omega} \in L^2(\partial\Omega),$$

such that

$$\|u(x)|_{\partial\Omega}\|_{L^2(\partial\Omega)} \leq C\|u\|_{H^1(\Omega)}. \tag{8.5}$$

Here by $u|_{\partial\Omega}$ we understand the trace, or the restriction of $u(x)$ for $x \in \partial\Omega$.

Proposition 8.3.3 *Using (8.5), the operator T can be uniquely extended to a linear continuous operator*

$$\mathsf{tr} : H^1(\Omega) \to L^2(\partial\Omega),$$

named the trace operator.

Proof. Construction of tr:

Thanks to Proposition 8.3.2, we have that

$$\forall u \in H^1(\Omega) \quad \exists (u_m) \subsetneq C^\infty(\overline{\Omega}) \text{ such that } u_m \to u \text{ for } m \to \infty \text{ in } H^1(\Omega).$$

As H^1 is complete, then (u_m) is a Cauchy sequence in H^1:

$$\forall \epsilon > 0 \quad \exists N \in \mathbb{N} \text{ such that } \forall m, k \geq N \quad \|u_m - u_k\|_{H^1(\Omega)} < \frac{\epsilon}{C(\Omega)}.$$

Therefore, using (8.5), we find

$$\forall \epsilon > 0 \quad \exists N \in \mathbb{N} \text{ such that } \forall m, k \geq N \quad \|u_m|_{\partial\Omega} - u_k|_{\partial\Omega}\|_{L^2(\partial\Omega)} \leq C(\Omega)\|u_m - u_k\|_{H^1(\Omega)} < \epsilon.$$

Thus $(u_m|_{\partial\Omega})$ is a Cauchy sequence in $L^2(\partial\Omega)$. The space $L^2(\partial\Omega)$ is complete, consequently

$$\exists v \in L^2(\partial\Omega) \text{ such that } u_m|_{\partial\Omega} \to v \text{ for } m \to \infty \text{ in } L^2(\partial\Omega).$$

Uniqueness of v:

Let $(\tilde{u}_k)$ be a sequence in $C^\infty(\overline{\Omega})$ such that

$$(\tilde{u}_k) \neq (u_m) \quad \text{and} \quad \begin{cases} \tilde{u}_k \to u \text{ in } H^1(\Omega), \\ \tilde{u}_k|_{\partial\Omega} \to \tilde{v} \text{ in } L^2(\partial\Omega). \end{cases}$$

Then, using (8.5) and the triangle inequality, we obtain

$$\begin{aligned} &\|v - \tilde{v}\|_{L^2(\partial\Omega)} = \|v - u_m|_{\partial\Omega} + u_m|_{\partial\Omega} - u|_{\partial\Omega} + u|_{\partial\Omega} - \tilde{u}_k|_{\partial\Omega} + \tilde{u}_k|_{\partial\Omega} - \tilde{v}\|_{L^2(\partial\Omega)} \\ &\leq \|v - u_m|_{\partial\Omega}\|_{L^2(\partial\Omega)} + \|u_m|_{\partial\Omega} - u|_{\partial\Omega}\|_{L^2(\partial\Omega)} + \|u|_{\partial\Omega} - \tilde{u}_k|_{\partial\Omega}\|_{L^2(\partial\Omega)} + \|\tilde{u}_k|_{\partial\Omega} - \tilde{v}\|_{L^2(\partial\Omega)} \\ &\leq \|v - u_m|_{\partial\Omega}\|_{L^2(\partial\Omega)} + \|\tilde{u}_k|_{\partial\Omega} - \tilde{v}\|_{L^2(\partial\Omega)} + C(\Omega)(\|u_m - u\|_{H^1(\Omega)} + \|u - \tilde{u}_k\|_{H^1(\Omega)}) \to 0 \quad \text{for } m, k \to \infty. \end{aligned}$$

Hence, $\|v - \tilde{v}\|_{L^2(\partial\Omega)} = 0$ and thus $v = \tilde{v}$. We conclude that for all $u \in H^1(\Omega)$ our construction allows to define an unique element of $L^2(\partial\Omega)$, which we call the trace of u. □

Attention : $\mathrm{tr}(H^1(\Omega)) \subsetneq L^2(\partial\Omega)$. It is denoted by $H^{\frac{1}{2}}(\partial\Omega) = \mathrm{tr}(H^1(\Omega))$.

8.3.4 Spaces $H_0^m(\Omega)$

Remark 8.3.1 *In what following we write ∂^α for functions in Sobolev spaces understanding the derivative in the sense of distributions.*

Definition 8.3.3 *Define $H_0^m(\Omega)$ as the closure of $\mathcal{D}(\Omega)$ in $H^m(\Omega)$.*

As we know, $\mathcal{D}(\Omega)$ is dense in $L^2(\Omega)$, but $\mathcal{D}(\Omega)$ is not dense in $H^m(\Omega)$. Thanks to Definition 8.3.3, $\mathcal{D}(\Omega)$ is dense in $H_0^m(\Omega)$!

We admit this characterization of $H_0^m(\Omega)$:

Theorem 8.3.1 *For all domains Ω in $\mathbb{R}^n$ (bounded or not)*

1. *$u \in H_0^m(\Omega) \Leftrightarrow \exists (u_m) \subsetneq \mathcal{D}(\Omega)$ such that $u_m \to u$ in $H^1(\Omega)$,*
2. *$H_0^m(\Omega)$ is a Hilbert space with the inner product of $H^m(\Omega)$,*
3. *$H_0^m(\Omega) \subsetneq H^m(\Omega)$ (if $\Omega = \mathbb{R}^n$ then $H_0^m(\Omega) = H^m(\Omega)$),*
4. *If $u \in H^m(\Omega)$ and $v \in \mathcal{D}(\Omega)$, then $vu \in H_0^m(\Omega)$,*

5. $H_0^1(\Omega) = \{f \in H^1(\Omega) |\ f = 0 \ \text{on}\ \partial\Omega\}$,
6. $H_0^m(\Omega) = \{f \in H^m(\Omega) |\ \text{for}\ |\alpha| < m, D^\alpha f = 0\ \text{on}\ \partial\Omega\}$.

Theorem 8.3.2 ***(Poincaré inequality)*** *Let Ω be a bounded domain of $\mathbb{R}^n$.*

Then there exists a constant $C = C(\Omega) > 0$ depending only on Ω, such that for all u in $H_0^1(\Omega)$,

$$\|u\|_{L_2(\Omega)} \leq C(\Omega)\|\nabla u\|_{L_2(\Omega)}, \tag{8.6}$$

where

$$\|\nabla u\|_{L_2(\Omega)} = \sum_{k=1}^{n} \left\| \frac{\partial u}{\partial x_k} \right\|_{L_2(\Omega)}.$$

Proof. Since $\mathcal{D}(\Omega)$ is dense in $H_0^1(\Omega)$, for all $u \in H_0^1(\Omega)$ there exists a sequence of functions $v_k \in \mathcal{D}(\Omega)$ which converges toward u in $H_0^1(\Omega)$:

$$\|v_k - u\|_{H^1(\Omega)} \to 0 \quad \text{for } k \to \infty.$$

Firstly we show (8.6) for $v_k \in \mathcal{D}(\Omega)$. After it, since the convergence by the norm of $H^1(\Omega)$ implies the convergence of $\|v_k - u\|_{L_2(\Omega)}$ and $\|\nabla(v_k - u)\|_{L_2(\Omega)}$ toward 0, we obtain (8.6) for $u \in H_0^1(\Omega)$ taking the limit for $k \to \infty$ by the norm of H^1.

Let Π be a parallelepiped such that $\Omega \subsetneq \Pi$ and, without loss a generality, we suppose that

$$\Pi = \{x \in \mathbb{R}^n | 0 < x_i < d_i,\ i = 1, \ldots, n\} \quad \text{and } d_1 = \min_{i=1,\ldots,n} d_i.$$

As for all k $\operatorname{supp} v_k \subsetneq \Omega$, we extend v_k from Ω to Π by 0.

For all $v_k \in \mathcal{D}(\Omega)$ we have

$$v_k(x_1, \ldots, x_n) = \int_0^{x_1} \frac{\partial v_k}{\partial y_1}(y_1, x_2, \ldots, x_n) \mathrm{d}y_1. \tag{8.7}$$

Introducing the following notations $x' = (x_2, \ldots, x_n)$,

$$\Pi' = \{x' \in \mathbb{R}^{n-1} | 0 < x_i < d_i \quad i = 2, \ldots, n\},$$

we take the square of (8.7) and integrate it over Π:

$$\int_\Pi v_k^2 \mathrm{d}x = \int_0^{d_1} \mathrm{d}x_1 \int_{\Pi'} \left(\int_0^{x_1} \frac{\partial v_k}{\partial y_1}(y_1, x') \mathrm{d}y_1 \right)^2 \mathrm{d}x'. \tag{8.8}$$

Using the $1D$ Cauchy-Schwartz inequality

$$\left(\int_0^{x_1} 1 \cdot \frac{\partial v_k}{\partial y_1}(y_1, x') \mathrm{d}y_1 \right)^2 \leq \int_0^{x_1} 1^2 \mathrm{d}y_1 \int_0^{x_1} \left(\frac{\partial v_k}{\partial y_1} \right)^2 \mathrm{d}y_1 \leq x_1 \int_0^{d_1} \left(\frac{\partial v_k}{\partial y_1} \right)^2 \mathrm{d}y_1,$$

we obtain

$$\int_\Pi v_k^2 \mathrm{d}x \leq \int_0^{d_1} \mathrm{d}x_1 \int_{\Pi'} \left[x_1 \int_0^{d_1} \left(\frac{\partial v_k}{\partial y_1} \right)^2 \mathrm{d}y_1 \right] \mathrm{d}x' = \frac{d_1^2}{2} \int_\Pi \left(\frac{\partial v_k}{\partial y_1} \right)^2 \mathrm{d}x.$$

Since $v_k \equiv 0$ out of Ω, we conclude that

$$\int_\Omega v_k^2 \mathrm{d}x \leq \frac{d_1^2}{2} \int_\Omega \left(\frac{\partial v_k}{\partial y_1} \right)^2 \mathrm{d}x,$$

from where taking the limit by the norm of H^1, we find (8.6).

Remark 8.3.2 *The Poincaré inequality holds true for domains bounded at least in one direction. But the result is not true in $H^1(\Omega)$. $u = 1$ everywhere on Ω would not work!*

Theorem 8.3.3 *When the open set Ω is bounded, the semi-norm $|\cdot|_m$ is a norm on H_0^m which is equivalent to $\|\cdot\|_m$.*

Proof. Instead of the general result, let us prove the equivalence of the norm

$$\|u\|_{H_0^1(\Omega)} = \sqrt{\int_\Omega (u^2 + |\nabla u|^2)\mathrm{d}x} \tag{8.9}$$

and

$$|u|_1 = \sqrt{\int_\Omega |\nabla u|^2 \mathrm{d}x} = \|\nabla u\|_{L_2(\Omega)}. \tag{8.10}$$

We need to show that there exist constants $C_1 > 0$ and $C_2 > 0$, such that

$$C_2 |u|_1 \leq \|u\|_{H_0^1(\Omega)} \leq C_1 |u|_1.$$

Thanks to the Poincaré inequality and the positivity of the norm, we find that indeed

$$\|\nabla u\|^2_{L_2(\Omega)} \leq \|u\|^2_{L_2(\Omega)} + \|\nabla u\|^2_{L_2(\Omega)} \leq (C^2(\Omega) + 1)\|\nabla u\|^2_{L_2(\Omega)},$$

from where it is sufficient to take $C_2 = 1$ and $C_1 = \sqrt{C(\Omega)^2 + 1}$. $\square$

8.3.5 Green's formula and integration by parts

Definition 8.3.4
- *If $\partial\Omega$ is C^1, then along $\partial\Omega$ is defined the outward pointing unit normal vector field $\nu = (\nu_1, \ldots, \nu_n)$. The unit normal at any point $x_0 \in \partial\Omega$ is $\nu(x_0)$.*
- *We denote by the normal derivative of u*

$$\frac{\partial u}{\partial \nu} = \langle \nu, \nabla u \rangle_{\mathbb{R}^n}.$$

Theorem 8.3.4 *Let Ω be sufficiently smooth. Let ν be the unit outward normal to Ω. Let $\{e_i\}$ be the canonical basis of $\mathbb{R}^n$. For any u and v in $H^1(\Omega)$ we have*

$$\int_\Omega u(\partial_i v) = -\int_\Omega (\partial_i u)v + \int_{\partial\Omega} uv\cos(\nu, e_i)$$

In particular, there is useful formula which can be considered as a classical integration by parts (see [3, 8]) for sufficient regular u and v:

$$\int_\Omega \Delta uv = \int_{\partial\Omega} \frac{\partial u}{\partial \nu} v - \int_\Omega \nabla u \nabla v.$$

8.4 Compact embedding of $H^1(\Omega)$ to $L^2(\Omega)$

Let us define the inclusion operator

Definition 8.4.1 *The operator $i(u) = u$ which associates $u \in H^1(\Omega)$ with the same u as an element of $L^2(\Omega)$ is called the **operator of inclusion** of $H^1(\Omega)$ in $L^2(\Omega)$.*

Instead to prove Theorem E.1.1, we prove the particular case

Theorem 8.4.1 *Let Ω be a bounded domain in $\mathbb{R}^n$ with locally Lipschitz boundary $\partial\Omega$. Then the inclusion operator $i : H^1(\Omega) \to L^2(\Omega)$ is linear, bounded and compact.*

Proof. Continuity

We have

$$\|i(u)\|_{L^2(\Omega)} \leq \|u\|_{H^1(\Omega)}.$$

Compactness

Let (u_m) is a bounded sequence in $H^1(\Omega)$:

$$\forall m \quad \|u_m\|_{H^1(\Omega)} \leq C.$$

Let define a parallelepiped Π in the way as

$$\Omega \subsetneq \Pi, \quad \Pi = \{x | 0 < x_i < d_i\}.$$

Let

$$\Pi = \sqcup_{i=1}^{M} \Pi_i, \text{ where } \Pi_i = \otimes_{k=1}^{n} [a_i, a_i + \frac{d_k}{N}].$$

By Theorem 8.2.3 point 2, we can extend u_m from Ω to a parallelepiped Π, containing Ω, such that the extensions $\tilde{u}_m$

$$\tilde{u}_m \in H^1(\Pi), \quad \tilde{u}_m|_\Omega = u_m, \quad \|u_m\|_{H^1(\Omega)} \leq \|\tilde{u}_m\|_{H^1(\Pi)}$$

and in addition there exists a constant $C(\Omega, \Pi)$ independing on u, such that

$$\|\tilde{u}_m\|_{H^1(\Pi)} \leq C(\Omega, \Pi) \|u_m\|_{H^1(\Omega)}.$$

Thus, the sequence $(\tilde{u}_m)$ is also a bounded sequence in $H^1(\Pi)$.

For this Π it is possible to prove the following inequality for all $u \in H^1(\Pi)$:

$$\int_\Pi u^2 dx \leq \sum_{i=1}^{N^n} \frac{1}{|\Pi_i|} \left(\int_{\Pi_i} u dx \right)^2 + \frac{n}{2} \int_\Pi \sum_{k=1}^{n} \left(\frac{d_k}{N} \right)^2 \left(\frac{\partial u}{\partial x_k} \right)^2 dx. \tag{8.11}$$

We have supposed that (u_m) is a bounded sequence in $H^1(\Omega)$. Therefore, as we have seen, $(\tilde{u}_m)$ is a bounded sequence in $H^1(\Pi)$. Thus, since i is a continuous operator from $H^1(\Pi)$ to $L^2(\Pi)$, the sequence $(i(\tilde{u}_m))$ is also bounded in $L^2(\Pi)$.

On the other hand, $L^2(\Pi)$ is a Hilbert space, thus weak* topology on it is equial to the weak topology. Moreover, as L^2 is separable, from Theorem 5.2.3, it follows that all closed bounded sets in $L^2(\Pi)$ are weakly sequentially compact (or compact in the weak topology since here the weak topology is metrizable).

Consequently, the sequence $(i(\tilde{u}_m))$ is weakly sequentially compact in $L^2(\Pi)$.

To simplify the notations, we will simply write u_m for $i(\tilde{u}_m) \in L^2(\Pi)$.

Since (u_m) is weakly sequentially compact in $L^2(\Pi)$, we have

$$\exists (u_{m_k}) \subset (u_m) \ : \quad \exists u \in L^2(\Pi) \quad u_{m_k} \rightharpoonup u.$$

Here u is an element of L^2, not necessarily in H^1.

As $(L^2(\Pi))^* = L^2(\Pi)$, by the Riesz theorem,

$$u_{m_k} \rightharpoonup u \in L^2(\Pi) \quad \Leftrightarrow \quad \forall v \in L^2(\Pi) \quad \int_\Pi (u_{m_k} - u) v dx \to 0.$$

We also notice that $u_{m_k} \rightharpoonup u \in L^2(\Pi)$ implies that (u_{m_k}) is a Cauchy sequence in weak topology on L^2. In addition, if we take $v = \mathbb{1}_\Pi$, then

$$\int_\Pi (u_{m_k} - u_{m_j}) dx \to 0 \quad \text{for } k, j \to +\infty.$$

Thus, using (8.11), for two members of the subsequence u_{m_k} with sufficiently large ranks p and q, we have

$$\begin{aligned}
&\|u_p - u_q\|^2_{L^2(\Omega)} \le \|u_p - u_q\|^2_{L^2(\Pi)} \\
&\qquad \le \sum_{i=1}^{N^n} \frac{1}{|\Pi_i|} \left(\int_{\Pi_i} (u_p - u_q) dx \right)^2 + \frac{n}{2N^2} \sum_{k=1}^n d_k^2 \left\| \frac{\partial u_p}{\partial x_k} - \frac{\partial u_q}{\partial x_k} \right\|^2_{L^2(\Pi)} < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon.
\end{aligned}$$

Here we have chosen N such that

$$\frac{n}{2N^2} \sum_{k=1}^n d_k^2 \left\| \frac{\partial u_p}{\partial x_k} - \frac{\partial u_q}{\partial x_k} \right\|^2_{L^2(\Pi)} < \frac{\epsilon}{2}.$$

Consequently, (u_{m_k}) (*i.e.*, $(i(u_{m_k}))$) is a Cauchy sequence in $L^2(\Omega)$, and thus converges strongly in $L^2(\Omega)$. $\square$

Appendix A

Topology

A.1 Open sets and topology

A.1.1 Definition and examples

Definition A.1.1 *Let X be a set and $\mathcal{T}$ be a family of subsets of X. $\mathcal{T}$ is called a **topology** on X if:*

1. *The empty set $\varnothing$ and X are elements of $\mathcal{T}$.*
2. *Arbitrary (finite or infinite) unions $\cup_\alpha U_\alpha$ of elements of $\mathcal{T}$ belong to $\mathcal{T}$ (or equivalently, $\mathcal{T}$ is **stable** by arbitrary unions).*
3. *Any finite intersection $\cap_\alpha U_\alpha$ of elements of $\mathcal{T}$ is in $\mathcal{T}$ (or equivalently, $\mathcal{T}$ is **stable** by finite intersection).*

Definition A.1.2 *Set X with a given topology defined on it, i.e. the pair $(X, \mathcal{T})$, is a **topological space**.*

Definition A.1.3 *Elements of $\mathcal{T}$ are called **open sets**.*

To specify a topological space, it means to define a set X and a topology in X, *i.e.*, to indicate which subsets of X are considered as open sets. Clearly, we can define on X various different topologies and therefore obtain different topological spaces constructed on the same set X.

Example A.1.1 *If $X = \{1, 2, 3, 4, 5\}$, then*

$$\mathcal{T}_1 = \{\varnothing, \{1,2\}, \{3,4\}, \{1,2,3,4\}, X\} \quad \textit{and} \quad \mathcal{T}_2 = \{\varnothing, \{1,2\}, \{2,3\}, \{2\}, \{1,2,3\}, X\}$$

are different topologies on X, as the three properties in Definition (A.1.1) are satisfied and $\mathcal{T}_1 \neq \mathcal{T}_2$.

Any set X can always be considered as a topological space:

Example A.1.2 *For any set X it is always possible to define:*

- *the **trivial topology** $\mathcal{T}_t = \{\varnothing, X\}$;*
- *the **discrete topology** $\mathcal{T}_d = \{$all subsets of $X\}$ (a usual notation for $\mathcal{T}_d$ is 2^X).*

We notice that in the case of the discrete topology every subset of $(X, \mathcal{T}_d)$ is open.

Let us note the procedure for the construction of a topology $\mathcal{T}$ on a given family $\mathcal{F}$ of sets in X (while adding the fewest possible sets):

1. Add $\varnothing$ and the whole space X to $\mathcal{T}$.
2. Add to $\mathcal{T}$ all finite intersections of elements of $\mathcal{F}$. Thus $\mathcal{T}$ is a family of subsets of X stable by any finite intersection.
3. Add to $\mathcal{T}$ all unions of elements of $\mathcal{T}$ constructed in the step 2. $\mathcal{T}$ is now stable by unions. It can be proved that the constructed $\mathcal{T}$ is stable by finite intersections and consequently, $\mathcal{T}$ is a topology in X.

Remark A.1.1 *Steps 2 and 3 of the construction of a topology $\mathcal{T}$ cannot be permuted: if we take, firstly, all unions in X, and after it all finite intersections, we obtain a family of sets stable by all finite intersections, but not by any unions. To remedy this fact, we should take again all unions of elements in $\mathcal{T}$.*

Definition A.1.4 ***(Usual topology on $\mathbb{R}$)*** *Let $X = \mathbb{R}$. The **usual topology on** $\mathbb{R}$ called the topology $\mathcal{T}$ defined by*

$$O \in T \text{ iff } \forall x \in O \ \exists \epsilon > 0 :]x - \epsilon, x + \epsilon[\subset O.$$

Definition A.1.5 *Complements of open sets are called **closed** sets: for any $U \in T$, which is open, its complement $F = X \setminus U$ is closed.*

The elements of a topological space $(X, \mathcal{T})$, *i.e.* $x \in X$, are called the **points** of $(X, \mathcal{T})$.

The elements of a topology $\mathcal{T}$, defined on a set X, are subsets of X.

Let us recall the following **de Morgan's laws** (also sometimes called "duality principle") from set theory: the complement of a union equals the intersection of the complements, and the complement of an intersection equals the union of the complements, *i.e.*

$$X \setminus \cup_\alpha U_\alpha = \cap_\alpha (X \setminus U_\alpha), \tag{A.1}$$
$$X \setminus \cap_\alpha U_\alpha = \cup_\alpha (X \setminus U_\alpha). \tag{A.2}$$

Problem A.1.1 *Prove relations (A.1) and (A.2).*

According to de Morgan's laws, it follows from Definition A.1.1 that:

1. The space $(X, \mathcal{T})$ itself and the empty set $\varnothing$ are closed;
2. Arbitrary (finite or infinite) intersections $\cap_\alpha F_\alpha$ and finite unions $\cup_\alpha F_\alpha$ of closed sets of $(X, \mathcal{T})$ are closed.

We introduce now the concepts of neighborhood, contact point, limit point and closure of a set:

Definition A.1.6
1. *The set U is called a **neighborhood** of a point x of the topological space $(X, \mathcal{T})$ if there exists an open set $V \in T$ such that $x \in V$ and $V \subset U$. The set of neighborhoods of x is noted $\mathcal{V}(x)$.*
2. *A point $x \in (X, \mathcal{T})$ is called a **contact point** of a set $A \subset (X, \mathcal{T})$ if every neighborhood of x contains at least one point of A;*

3. *A point $x \in (X, \mathcal{T})$ is called a* ***limit point*** *of a set $A \subset (X, \mathcal{T})$ if every neighborhood of x contains at least one point of A different of x:*

$$V \cap (A \setminus \{x\}) \neq \varnothing \quad \textit{for all neighborhoods } V \textit{ of } x;$$

4. *The set of all contact points of a set $A \subset (X, \mathcal{T})$ is called the* ***closure*** *of A, denoted by $\overline{A}$.*

Problem A.1.2 *Given a topological space $(X, \mathcal{T})$, prove that a set $A \subset X$ is open if and only if every point $x \in A$ has a open neighborhood contained in A.*

Problem A.1.3 *Let $A \subset (X, \mathcal{T})$. Then $\overline{A} = A$ iff A is closed.*

From Problem A.1.3 it follows that the closure of A is the minimal closed set containing A.

Definition A.1.7 *The largest open set contained in a given set A is called the* ***interior*** *of A.*

Example A.1.3 *Every closed interval $[a, b]$ on the real line is a closed set for the usual topology on $\mathbb{R}$. Indeed, all points of $[a, b]$ are limit and, thus, contact points. Therefore, $\overline{[a, b]} = [a, b]$ and then it is a closed set. For the open interval, $]a, b[$, the points a and b are not in $]a, b[$, but they are still contact and limit points. Consequently, $\overline{]a, b[} = [a, b]$.*

Moreover, we have the following theorem:

Theorem A.1.1 *Let A be a subset of a topological space $(X, \mathcal{T})$. Then*

1. *$A \subset \overline{A}$ ($\overline{A}$ is the smallest closed set containing A),*
2. *$\overline{\overline{A}} = \overline{A}$,*
3. *if $B \subset A$, then $\overline{B} \subset \overline{A}$,*
4. *for all A, B in X, $\overline{A \cup B} = \overline{A} \cup \overline{B}$.*

Proof. Property 1) holds, since every point of A is a contact point of A.
Let's prove property 2). Thanks to Problem A.1.3, $\overline{A}$ is a closed set and therefore, $\overline{\overline{A}} = \overline{A}$.
Property 3) is obvious.
To prove property 4), let $x \in \overline{A \cup B}$ and suppose $x \notin \overline{A} \cup \overline{B}$. Then $x \notin \overline{A}$ and $x \notin \overline{B}$. But then there exist open neighborhoods V_A and V_B of x such that V_A contains no points of A while V_B contains no points of B. It follows that the set $V = V_A \cap V_B$ is the open neighborhood of x which contains no points of either A or B, and hence no points of $A \cup B$, contrary to the assumption that $x \in \overline{A \cup B}$. Therefore $x \in \overline{A} \cup \overline{B}$, and consequently

$$\overline{A \cup B} \subset \overline{A} \cup \overline{B},$$

since x is an arbitrary point of $\overline{A \cup B}$. On the other hand, since $A \subset A \cup B$ and $B \subset A \cup B$, using 3) we obtain that

$$\overline{A} \cup \overline{B} \subset \overline{A \cup B}.$$

As $\overline{A \cup B} \subset \overline{A} \cup \overline{B} \subset \overline{A \cup B}$, we conclude that $\overline{A \cup B} = \overline{A} \cup \overline{B}$. □

Example A.1.4 1. *For the discrete topology $(X, \mathcal{T}_d)$ introduced in Example A.1.2, every set $A \subset (X, \mathcal{T}_d)$ is both open and closed and coincides with its own closure.*

2. *If the topology on X is trivial, the closure of every nonempty set is the whole space X. Therefore, $(X, \mathcal{T}_t)$ can be called "space of coalesced points".*

Example A.1.5 *Let X be the set $\{a,b\}$, consisting of just two points a and b, and let the open sets in X be X itself, the empty set and the single-element set $\{b\}$:*

$$\mathcal{T} = \{\varnothing, \{b\}, X\}.$$

Then the three properties in Definition A.1.1 are satisfied, and $(X, \mathcal{T})$ is a topological space. The closed sets in this space are X itself, the empty set and the set $\{a\}$. Note that the closure of $\{b\}$ is the whole space X.

Example A.1.6 ***(Topology of $\mathcal{D}(\Omega)$, see Chapter 7)***

*Let Ω be an open domain of $\mathbb{R}^n$ (open in the usual topology of $\mathbb{R}^n$). V is an **open set** in $\mathcal{D}(\Omega)$ iff it can be presented as an (finite or countable) union $\cup_{j\in J} U_j$ of sets U_j of the following type:*

$U_j \subset \mathcal{D}(\Omega)$ such that

for all $K \subsetneq \Omega$ compact (in $\mathbb{R}^n$!), and $f \in U_j$ with $\operatorname{supp} f \subset K$ (thus $f \in \mathcal{D}(\Omega)$!), there exist $\epsilon > 0$ and a multi-index α, such that

$$\{g \in \mathcal{D}(\Omega) |\operatorname{supp} g \subset K \textit{ and } \forall x \in K, |\partial^\alpha f(x) - \partial^\alpha g(x)| < \epsilon\} \subset U_j.$$

A.1.2 Comparison of topologies

Definition A.1.8 *Let $\mathcal{T}_1$ and $\mathcal{T}_2$ be two topologies defined in the same set X. Then the topology $\mathcal{T}_1$ is **stronger** than the topology $\mathcal{T}_2$(or equivalently, $\mathcal{T}_2$ is **weaker** than $\mathcal{T}_1$) if $\mathcal{T}_2 \subset \mathcal{T}_1$, i.e., if every set of the system $\mathcal{T}_2$ is a set of the system $\mathcal{T}_1$.*

Let τ be the set of all topologies in X. Then for all $T \in \tau$

$$\mathcal{T}_t \subset T \subset \mathcal{T}_d,$$

where $\mathcal{T}_t$ is the trivial topology in X and $\mathcal{T}_d$ is the discrete topology in X. In other words, $\mathcal{T}_d$ is the maximal element of τ (the strongest topology in X) and $\mathcal{T}_t$ is the minimal element of τ (the weakest topology in X).

Theorem A.1.2 *Let $\{\mathcal{T}_\alpha\}$ be any set of topologies in X. Then the intersection $\mathcal{T} = \cap_\alpha \mathcal{T}_\alpha$ is also a topology in X.*

Proof. We need to verify Definition A.1.1 for $\mathcal{T}$. Clearly $\cap_\alpha \mathcal{T}_\alpha$ contains X and $\varnothing$. Moreover, since every $\mathcal{T}_\alpha$ is stable by the operations of taking arbitrary unions and finite intersections, the same is true for $\cap_\alpha \mathcal{T}_\alpha$. □

Corollary A.1.1 *Let $\mathcal{A}$ be any system of subsets of a set X. Then there exists a minimal topology in X containing $\mathcal{A}$, i.e., a topology $T(\mathcal{A})$ containing $\mathcal{A}$ and contained in every topology containing $\mathcal{A}$.*

Proof. A topology containing $\mathcal{A}$ always exists, e.g., the discrete topology in which every subset of X is open. The intersection of all topologies containing $\mathcal{A}$ is the desired minimal topology $T(\mathcal{A})$, often called the topology **generated** by the system $\mathcal{A}$. □

Definition A.1.9 *Let $\mathcal{A}$ be a system of subsets of X and A a fixed subset of X. Then the system $\mathcal{A}_A$ consisting of all subsets of X of the form $A \cap B$, $B \in \mathcal{A}$ is called **the trace** of the system $\mathcal{A}$ on the set A.*

It is easy to see that the trace (on A) of a topology $\mathcal{T}$ (defined in X) is a topology $\mathcal{T}_A$ in A. (Such a topology is often called a **relative** or **induced** topology.) In this sense, every subset A of a given topological space $(X, \mathcal{T})$ generates a new topological space $(A, \mathcal{T}_A)$, called a **subspace** of the original topological space $(X, \mathcal{T})$. Let us notice that in general, if $U \subset A$ is open in $(A, \mathcal{T}_A)$, then U is not necessary a open set in $(X, \mathcal{T})$:

Example A.1.7 *The set* $]0, 1] \subset \mathbb{R}$ *is not open in the usual topology of* $\mathbb{R}$*, but it is open in the induced topology when considered as a subset of* $A = [-1, 1]$*.*

In the same time, in the particular case, when $A \subset X$ is itself a open set in $(X, \mathcal{T})$, we can verify that $U \subset A$ is a open set of $(A, \mathcal{T}_A)$ iff U is a open set in $(X, \mathcal{T})$.

Let us also notice that if $\mathcal{T}_1$ and $\mathcal{T}_2$ are different topologies in X, they can generate the same relative topology $\mathcal{T}_A$ in A.

Definition A.1.10 *Let* $(X, \mathcal{T}_X)$ *and* $(Y, \mathcal{T}_Y)$ *be two topological spaces. The* ***product topology*** *in* $X \times Y$*, denoted* $\mathcal{T} = \mathcal{T}_X \otimes \mathcal{T}_Y$*, is defined by calling* $U \subset X \times Y$ *open if*

$$\forall (x, y) \in U \ \exists V \in \mathcal{T}_X \text{ and } W \in \mathcal{T}_Y \text{ such that } x \in V,\ y \in W \text{ and } V \times W \subset U.$$

Example A.1.8 $\mathbb{R}^2 = \mathbb{R} \times \mathbb{R}$ *is a topological space with the usual product topology constructed on the two usual topologies in* $\mathbb{R}$*.*

A.1.3 Dense subsets and connected topological spaces

Definition A.1.11 *Let* A *and* B *be subsets of a topological space* $(X, \mathcal{T})$*. Let* $A \subset B$*. The set* A *is* ***dense*** *in* B *if* $\overline{A} \supset B$*. Here we take* $\overline{A}$ *in topology on* X*:*

$$\forall x \in B \quad \forall \text{ open neighborhoods } U \text{ of } x \quad (U \in \mathcal{T}) \quad U \cap A \neq \varnothing.$$

Remark A.1.2 *We can also say for* $A \subset B$ *that the set* A *is dense in* B *if* $\overline{A} = B$ *for the induced topology on* B *(we have* $=$ *, because, in this case by definition of the induced topology* $\overline{A}$ *cannot be bigger than* B*.) I.e.*

$$\forall x \in B \quad \forall \text{ open neighborhoods } U \cap B \text{ of } x \quad (U \in \mathcal{T}) \quad (U \cap B) \cap A \neq \varnothing.$$

Definition A.1.12 *Let* A *be a subset of a topological space* $(X, \mathcal{T})$*. The set* A *is said to be* ***dense in*** X *if* $\overline{A} = X$*. A set* A *is said to be* ***nowhere dense*** *if it is dense in no (open) set at all.*

Example A.1.9 *The set of all rational numbers* $\mathbb{Q}$ *is dense in* $\mathbb{R}$*.*

Given any topological space $(X, \mathcal{T})$, the empty set $\varnothing$ and the space X itself are both open and closed, by definition.

Definition A.1.13 *A topological space* $(X, \mathcal{T})$ *is said to be connected if it has no subsets other than* $\varnothing$ *and* X *which are both open and closed.*

Example A.1.10 *The real line* $\mathbb{R}$ *is connected, but not the set* $\mathbb{R} \setminus \{x_0\}$ *obtained from* $\mathbb{R}$ *by deleting the point* x_0*.*

Example A.1.11 *Let (X, d) be a metric space. Then, the distance d define a topology on X by*

$$\mathcal{T}_d = \{O \subset E | \; \forall x \in O \; \exists r > 0, B_r(x) \subset O\}, \tag{A.3}$$

where by $B_r(x)$ is denoted the open ball of radius r centered in x.

A.2 Convergence and continuity

A.2.1 Continuous mappings. Homeomorphism

Definition A.2.1 *Let f be a mapping of one topological space $(X, \mathcal{T}_X)$ into another topological space $(Y, \mathcal{T}_Y)$, so that f associates an element $y = f(x) \in Y$ with each element $x \in X$. Then f is said to be* ***continuous at the point*** *$x_0 \in X$ if, given any neighborhood W_{y_0} of the point $y_0 = f(x_0)$, there is a neighborhood U_{x_0} of the point x_0 such that $f(U_{x_0}) \subset W_{y_0}$.*

Definition A.2.2 *The mapping $f : (X, \mathcal{T}_X) \to (Y, \mathcal{T}_Y)$ is said to be* ***continuous on*** *X if it is continuous at every point of X.*

In particular, a continuous mapping of a topological space $(X, \mathcal{T}_X)$ into the real line $(\mathbb{R}, \mathcal{T}_{\mathbb{R}})$ is called a continuous real mapping on X. Here by $\mathcal{T}_{\mathbb{R}}$ we denote the usual topology on $\mathbb{R}$ (see Definition A.1.4).

The notion of continuity of a mapping f of one topological space into another is easily stated in terms of open sets, i.e., in terms of the topologies of the two spaces:

Theorem A.2.1 *A mapping f of a topological space $(X, \mathcal{T}_X)$ into a topological space $(Y, \mathcal{T}_Y)$ is continuous if and only if the inverse image $f^{-1}(W)$ of every open set $W \in \mathcal{T}_Y$ is open, i.e. $f^{-1}(W) \in \mathcal{T}_X$.*

Proof. Suppose f is continuous on $(X, \mathcal{T}_X)$, and let W be any open subset of $\mathcal{T}_Y$. Choose any point $x \in f^{-1}(W)$, and let $y = f(x)$ (see Fig. A.1). Then W is an open neighborhood of the point y. Hence, by the continuity of f, there is a neighborhood U_X of x such that $f(U_X) \subset W$,

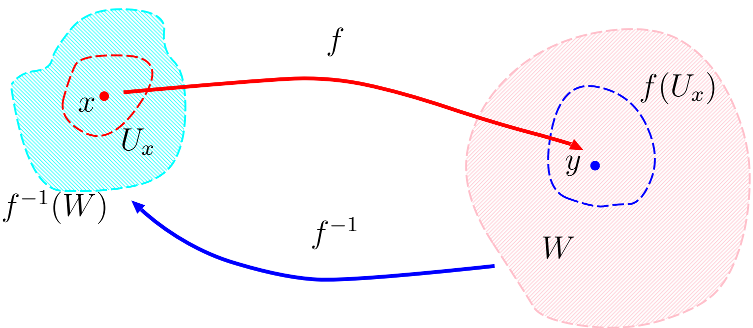

Figure A.1 – Illustration of the proof of Theorem A.2.1: for any open W in Y, any point $x \in f^{-1}(W)$ is associated with $y = f(x)$; U_X is a neighborhood of x such that $f(U_X) \subset W$.

i.e., $U_X \subset f^{-1}(W)$. In other words, every point $x \in f^{-1}(W)$ has a neighborhood contained in $f^{-1}(W)$. Consequently $f^{-1}(W)$ is open.

Conversely, suppose $f^{-1}(W)$ is open whenever $W \subset Y$ is open. Given any point $x \in X$, let U_Y be any open neighborhood of the point $y = f(x)$.

Then clearly $x \in f^{-1}(U_Y)$, and moreover $f^{-1}(U_Y)$ is open, by hypothesis. Therefore $U_X = f^{-1}(U_Y)$ is a neighborhood of x such that $f(U_X) \subset U_Y$. In other words, f is continuous at x and hence on X, since x is an arbitrary point of X. □

Naturally, Theorem A.2.1 has the following dual form:

Theorem A.2.2 *A mapping f of a topological space $(X, \mathcal{T}_X)$ into a topological space $(Y, \mathcal{T}_Y)$ is continuous if and only if the inverse image $f^{-1}(W)$ of every closed set $W \in Y$ is closed in X.*

Proof. Use the fact that the inverse image of a complement is the complement of the inverse image.□

Let us recall the following Lemma:

Lemma A.2.1 *1. The inverse image of a union (or intersection) of sets equals the union (or intersection) of the inverse images of the sets:*

$$f^{-1}(A \cup B) = f^{-1}(A) \cup f^{-1}(B), \qquad f^{-1}(A \cap B) = f^{-1}(A) \cap f^{-1}(B) \quad \forall A, B \subset Y.$$

2. The inverse image of the complement of a set is the complement of the inverse image of the set:

$$\forall U \subset Y \quad f^{-1}(Y \setminus U) = X \setminus f^{-1}(U).$$

Remark A.2.1 *Suppose $f : (X, \mathcal{T}_X) \to (Y, \mathcal{T}_Y)$ is a mapping of a topological space $(X, \mathcal{T}_X)$ into a topological space $(Y, \mathcal{T}_Y)$. Let $f^{-1}(\mathcal{T}_Y)$ be the inverse image of the topology $\mathcal{T}_Y$:*

$$f^{-1}(\mathcal{T}_Y) = \{ \text{ system of all sets } f^{-1}(U) \mid U \in \mathcal{T}_Y \}.$$

Thanks to the point 1 of Lemma A.2.1, we obtain that $f^{-1}(\mathcal{T}_Y)$ is a topology in X.

Problem A.2.1 *Prove the following theorem:*

Theorem A.2.3 *Let $f : (X, \mathcal{T}_X) \to (Y, \mathcal{T}_Y)$ be a mapping of a topological space $(X, \mathcal{T}_X)$ into a topological space $(Y, \mathcal{T}_Y)$. The mapping f is continuous if and only if the topology $\mathcal{T}_X$ is stronger than the topology $f^{-1}(\mathcal{T}_Y)$.*

Thanks to point 2 of Lemma A.2.1, we obtain the "dual" form of Theorem A.2.1:

Theorem A.2.4 *A mapping f of a topological space $(X, \mathcal{T}_X)$ into a topological space $(Y, \mathcal{T}_Y)$ is continuous if and only if the inverse image $f^{-1}(W)$ of every closed set $W \subset Y$ is closed in X.*

It is important to notice that the image (as opposed to the inverse image) of an open set under a continuous mapping need not be open. Similarly, the image of a closed set under a continuous mapping need not be closed.

Problem A.2.2 *Give an example of a continuous mapping $f : X \to Y$ which maps a closed set of X in an open set of Y.*

As a direct Corollary of Theorem A.2.1, we have the theorem on continuity of composite mapping:

Theorem A.2.5 *Given topological spaces $(X, \mathcal{T}_X)$, $(Y, \mathcal{T}_Y)$ and $(Z, \mathcal{T}_Z)$, suppose f is a continuous mapping of $(X, \mathcal{T}_X)$ into $(Y, \mathcal{T}_Y)$ and g is a continuous mapping of $(Y, \mathcal{T}_Y)$ into $(Z, \mathcal{T}_Z)$. Then the mapping $g \circ f$: $x \in X \mapsto g(f(x)) \in Z$ is continuous.*

Definition A.2.3 *Given two topological spaces $(X, \mathcal{T}_X)$ and $(Y, \mathcal{T}_Y)$, let f be a bijection of $(X, \mathcal{T}_X)$ onto $(Y, \mathcal{T}_Y)$, and suppose f and f^{-1} are both continuous. Then f is called a* ***homeomorphic mapping*** *or simply a* ***homeomorphism*** *(between X and Y). Two spaces $(X, \mathcal{T}_X)$ and $(Y, \mathcal{T}_Y)$ are said to be homeomorphic if there exists a homeomorphism between them. We note the homeomorphic spaces by $(X, \mathcal{T}_X) \approx (Y, \mathcal{T}_Y)$.*

Homeomorphic spaces have the same topological properties, and from the topological point of view are merely two "representatives" of one and the same space. In fact, if f is a homeomorphic mapping of $(X, \mathcal{T}_X)$ onto $(Y, \mathcal{T}_Y)$, then $\mathcal{T}_X = f^{-1}(\mathcal{T}_Y)$ and $\mathcal{T}_Y = f(\mathcal{T}_X)$.

Example A.2.1 *The interval* $]-\frac{\pi}{2}, \frac{\pi}{2}[$ *equipped with the usual topology on* $\mathbb{R}$ *is homeomorphic to* $\mathbb{R}$ *also equipped with the usual topology, as* $f(x) = \tan(x)$ *is a homeomorphic mapping of* $]-\frac{\pi}{2}, \frac{\pi}{2}[$ *into* $\mathbb{R}$.

Remark A.2.2 *The relation* $\approx$ *of being homeomorphic is* ***reflexive****, i.e. for any topological space* $(X, \mathcal{T}_X)$*,*

$$(X, \mathcal{T}_X) \approx (X, \mathcal{T}_X),$$

symmetric*, i.e. for all topological spaces* $(X, \mathcal{T}_X)$ *and* $(Y, \mathcal{T}_Y)$*,*

$$(X, \mathcal{T}_X) \approx (Y, \mathcal{T}_Y) \Longleftrightarrow (Y, \mathcal{T}_Y) \approx (X, \mathcal{T}_X),$$

and ***transitive****, i.e. for all topological spaces* $(X, \mathcal{T}_X)$*,* $(Y, \mathcal{T}_Y)$ *and* $(Z, \mathcal{T}_Z)$*,*

$$(X, \mathcal{T}_X) \approx (Y, \mathcal{T}_Y) \text{ and } (Y, \mathcal{T}_Y) \approx (Z, \mathcal{T}_Z) \Rightarrow (X, \mathcal{T}_X) \approx (Z, \mathcal{T}_Z),$$

and hence is called an ***equivalence relation****. Therefore any given family of topological spaces can be partitioned into disjoint classes of homeomorphic spaces.*

A.2.2 Converging sequences in $(X, \mathcal{T})$

Let $\mathcal{V}(x)$ denote the set of all neighborhoods of $x \in X$ and $\mathcal{O}(x)$ denote the set of all **open** neighborhoods of $x \in X$.

Definition A.2.4 *Let* $(X, \mathcal{T})$ *be a topological space and* (x_n) *be a sequence of elements of* X*. We say that* (x_n) *converges to* x*, if*

$$\forall U \in \mathcal{V}(x), \exists N \in \mathbb{N} \text{ such that } \forall n \geq N \Rightarrow x_n \in U.$$

We note that

- (x_n) may converge to several elements of X;
- If the topology on X is stronger (larger/finer), it is "harder" for (x_n) to converge;
- If X is equipped with the discrete topology, only sequences that become constant converge.

Remark A.2.3 *Thanks to the definition of the neighborhood* U*, as any set containing an open set* V *such that* $x \in V$ *(and thus* V *is an open neighborhood of* x*!), we can consider the equivalent form of Definition A.2.4 choosing only open neighborhoods of* l*:*

Definition A.2.5 *Let* $(X, \mathcal{T})$ *be a topological space and* (x_n) *be a sequence of elements of* X*. We say that* (x_n) *converges to* x *if*

$$\forall V \in \mathcal{O}(x), \exists N \in \mathbb{N} \text{ such that } \forall n \geq N \Rightarrow x_n \in V.$$

For the sake of clarity, in what follows we will use Definition A.2.5.

Proposition A.2.1 *Let* $(X, \mathcal{T}_X)$ *and* $(Y, \mathcal{T}_Y)$ *be two topological spaces. Let* $f : X \to Y$ *be a continuous mapping and* (x_n) *be a sequence in* X *converging to* x*. Define* $y_n = f(x_n)$*. Then* (y_n) *converges to* $f(x)$ *in* Y*.*

Proof. Let U be any open neighborhood of $f(x)$. Since f is continuous, $f^{-1}(U)$ is open (thanks to the definition of continuity). Since $f(x) \in U$, we have $l \in f^{-1}(U)$ and consequently $f^{-1}(U)$ is a open neighborhood of x. Since (x_n) converges to l in X,

$$\exists N \in \mathbb{N} \text{ such that } \forall n \geq N \; x_n \in f^{-1}(U),$$

implies $y_n = f(x_n) \in U$ for all $n \geq N$. As U is an arbitrary chosen open neighborhood of $f(x)$, we find that

$$\forall U \in \mathcal{O}(f(x)), \exists N \in \mathbb{N} \text{ such that } \forall n \geq N \Rightarrow y_n \in U.$$

So (y_n) converges to $f(x)$. □

Example A.2.2 *Let us consider the topological space $(X, \mathcal{T}_1)$ from Example A.1.1. We denote by $[x]$ the integer part of x. The sequence $(x_n)_{n\geq 1}$ is defined as*

$$x_n = \left[\frac{4}{n}\right] + \frac{(-1)^n + 3}{2},$$

from where we find

$$x_1 = 5,\; x_2 = 4,\; x_3 = 2,\; x_4 = 3,\; x_5 = 1,\; x_6 = 2,\; x_7 = 1,\; x_8 = 2, \ldots$$

We notice that for $n \geq 2$ $x_{2n+1} = 1$ and for $n \geq 3$ $x_{2n} = 2$. It means that for $n \geq 5$ $x_n \in \{1, 2\}$. The set $\{1, 2\}$ is the smallest open neighborhood containing 1 or 2 in $(X, \mathcal{T}_1)$:

$$1, 2 \in \{1, 2\} \subsetneq \{1, 2, 3, 4\} \subsetneq X.$$

Therefore, (x_n) converge to 1 and to 2 (the limit x is not unique) in $(X, \mathcal{T}_1)$.

Example A.2.3 *Let us consider $\mathbb{R}$ with the usual topology. The sequence $(\frac{1}{n})_{n\geq 1}$ converge to 0 in $\mathbb{R}$: for any open neighborhood of zero $]-\frac{1}{m}, \frac{1}{m}[$ with $m \in \mathbb{R} \setminus \{0\}$ there exists $n_0 \in \mathbb{N}$ such that for all $n > n_0$ we have $\frac{1}{n} \in]-\frac{1}{m}, \frac{1}{m}[$.*

A.3 Initial topology

We can equip a set X with a topology that makes every mapping f_i on X continuous. If all else fails, the discrete topology will work!

Definition A.3.1 *Let $f_i : X \to Y_i$ be given mappings defined on a set X ($i \in I$). We call **initial topology**, noted $\sigma(X, \{f_i, i \in I\})$, the coarsest topology in X that makes every f_i continuous, as mappings from the topological space $(X, \sigma(X, \{f_i, i \in I\}))$ to the topological space $(Y_i, \mathcal{T}_{Y_i})$.*

Example A.3.1 *Let $X = \mathbb{R}$, $Y = \mathbb{R}$ and f be defined by*

$$f(x) = \begin{cases} 0 & \text{if } x \leq 0 \\ 1 & \text{if } x > 0 \end{cases}.$$

Equip Y with the usual topology. The initial topology on X for f is

$$\sigma(\mathbb{R}, \{f\}) = \{\varnothing,\;]-\infty, 0],\;]0, +\infty[,\; \mathbb{R}\}.$$

Remark A.3.1 *Let $f_i : X \to Y_i$. We notice that the initial topology $\sigma(X, \{f_i, i \in I\})$ is constructed on the family of sets $f_i^{-1}(V_i)$ for open V_i in Y_i (see Section A.1.1). Moreover, the family of sets $\cap_{finite} f_i^{-1}(V_i)$ is a **base** of the initial topology (see [2] or [7] p.80).*

Let us formulate the following important result for the initial topology (see for instance, [2] for the proof):

Proposition A.3.1 *Let $f_i : X \to Y_i$ be given mappings defined on a set X ($i \in I$ finite or not) with image in sets Y_i equipped with topologies $\mathcal{T}_{Y_i}$ and let (x_n) be a sequence of X. In the topological space $(X, \sigma(X, \{f_i, i \in I\}))$, equipped with the initial topology, for $n \to \infty$*

$$x_n \to x \quad \text{if and only if} \quad \forall i \in I,\ f_i(x_n) \to f_i(x) \quad [\text{in the topological spaces } (Y_i, \mathcal{T}_{Y_i})].$$

Two natural examples of the initial topology is the weak and weak* topologies on a Banach space.

A.4 What a topology "sees" and does not "see". Separation of topological spaces

Let $(X, \mathcal{T}_X)$ be a topological space. We would like to address these questions:

- If x is a limit point of $A \subset X$. Is there a sequence (x_n) of points of A converging to x?
- Let $x_n \to x$ in $(X, \mathcal{T}_X)$. Is the limit unique?
- Let x and y be two different points of $(X, \mathcal{T}_X)$. Can we "separate" them?

To do so, we need to specify the separation properties of the topological space $(X, \mathcal{T}_X)$.

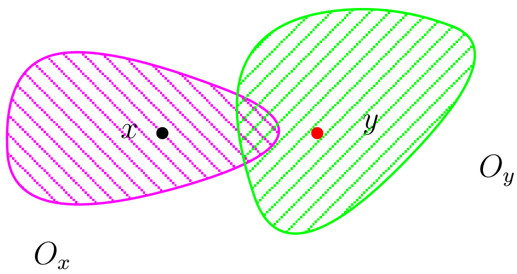

Figure A.2 – First axiom of separation.

Definition A.4.1 *A topological space $(X, \mathcal{T}_X)$ said to satisfy the **First axiom of separation** (or to be $\mathcal{T}_1$ **space**) if for all two distinct points x and y in $(X, \mathcal{T}_X)$ there exists an open neighborhood O_x of the point x such that $y \notin O_x$ and there exists a open neighborhood O_y of the point y such that $x \notin O_y$. (see Fig.A.2)*

Example A.4.1 *The topological space $(X, \mathcal{T}_X)$ constructed in Example A.1.5 is not a $\mathcal{T}_1$-space.*

In a $\mathcal{T}_1$-space singleton point is a closed set. Indeed, if $x \neq y$, then there exists an open neighborhood O_y of the point y such that $x \notin O_y$, *i.e.* $y \notin \overline{x}$. Thus, $\overline{x} = x$. Consequently, in a $\mathcal{T}_1$-space any finite union of points is a closed set.

In the topological spaces which are not $\mathcal{T}_1$-spaces, even sets composed only of a finite number of points can possess limit points. In the topological space $(X, \mathcal{T}_X)$ constructed in Example A.1.5, the point $\{a\}$ is the limit point for the set $W = \{b\}$.

But in a $\mathcal{T}_1$-space, it holds

Lemma A.4.1 *Point x is a limit point of the set W in a $\mathcal{T}_1$-space if and only if all open neighborhoods U of x contains infinite number of points of W.*

Proof. If any open neighborhood U of x contains an infinite number of points from W it is obvious (see Definition A.1.6) that in this case x is a limit point of W. Let us prove the converse. Let x be a limit point of W. Suppose that there exists an open neighborhood U of x such that U contains only a finite number of points $\{x_1, \ldots, x_n\}$ of W (where $x_i \neq x$ for all i in the case when $x \in W$). As in a $\mathcal{T}_1$-space any finite union of points is a closed set, $\{x_1, \ldots, x_n\}$ is closed. Therefore, $O = U \setminus \{x_1, \ldots, x_n\}$ is open and, thus, O is an open neighborhood of x such that $O \cap (W \setminus \{x\}) = \varnothing$, which contradicts the definition of a limit point. $\square$ Let us introduce Hausdorff spaces.

Definition A.4.2 *A topological space $(X, \mathcal{T}_X)$ is a **Hausdorff space** (or a $\mathcal{T}_2$**-space** or a **separated space**) if all two distinct points in X have two disjoint neighborhoods. (see Fig.A.3)*

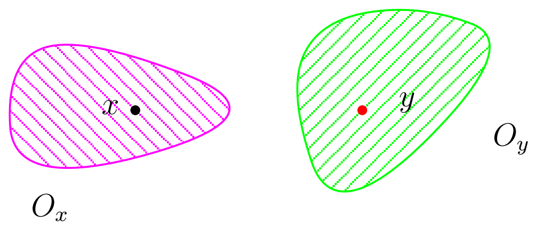

Figure A.3 – Hausdorff space.

The main advantage of a Hausdorff space is that the limit of a sequence is unique.

All Hausdorff spaces are $\mathcal{T}_1$-space, but not converse.

Example A.4.2 ***($\mathcal{T}_1$-space, but not $\mathcal{T}_2$-space)*** *Let us consider the interval $[0,1]$ and all sets*

$$A_{m,a_1,\ldots,a_m} = [0,1] \setminus (\cup_{i=1}^{m}\{a_i\}), \quad \textit{where } a_i \in [0,1] \textit{ and } 0 \leq m \leq +\infty.$$

If $X = [0,1]$ with the topology $\mathcal{T}$ composed by $\varnothing$ and all sets $A_{m,a_1,\ldots,a_m}$, then $(X, \mathcal{T})$ is $\mathcal{T}_1$-space, but not a Hausdorff space.

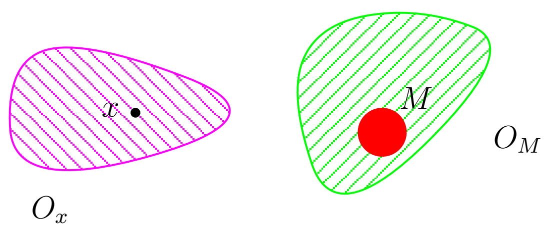

Figure A.4 – Third axiom of separation

Definition A.4.3 *A topological space $(X, \mathcal{T}_X)$ is said to satisfy the **Third axiom of separation** (or to be $\mathcal{T}_3$ **space**) if for all points x and closed sets M in $(X, \mathcal{T}_X)$ not containing x, there exist two disjoint neighborhoods O_x of the point x and O_M of the set M. (see Fig.A.4)*

We note that the open neighborhood of a set M in the topological space $(X, \mathcal{T})$ is called any open set U containing M.

Problem A.4.1 *Show that third axiom of separation can be also formulated in the following equivalent form:*

All open neighborhood U of a point $x \in (X, \mathcal{T})$ contains the closure of a smaller neighborhood O of x: $x \in \overline{O} \subset U$.

Definition A.4.4 *Topological spaces which satisfy the axioms $\mathcal{T}_1$ and $\mathcal{T}_3$ are called **regular**.*

Obviously, each regular space is a Hausdorff space. But not converse:

Example A.4.3 ***(Hausdorff spaces which are not regular)*** *Let $X = [0, 1]$. Let all points of X different to 0 have the usual neighborhoods of the usual topology. Define the neighborhoods of zero as all semi-intervals $[0, \alpha[$ without points $\frac{1}{n}$, $n \in \mathbb{N}^*$. This is a Hausdorff space, but the point 0 and the closed set $\{\frac{1}{n}\}_{n \in \mathbb{N}^*}$ are not separated by disjoint neighborhoods, i.e. axiom $\mathcal{T}_3$ is false.*

Definition A.4.5 *A topological $\mathcal{T}_1$-space $(X, \mathcal{T}_X)$ is said to satisfy the **Fourth axiom of separation** (or to be a $\mathcal{T}_4$ **space** or a **normal** space) if for all two disjoint closed sets M and P in $(X, \mathcal{T}_X)$, there exist two disjoint neighborhoods. (see Fig.A.5)*

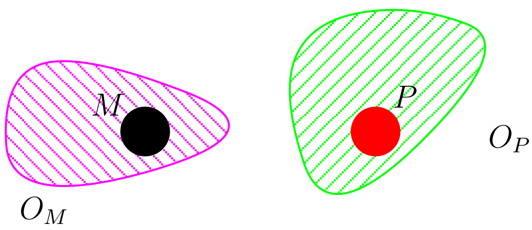

Figure A.5 – Normal space.

As every metric space is a topological space, we can also use it as the example of $\mathcal{T}_1$-spaces:

Theorem A.4.1 *Every metric space (E, d) is a normal space, and thus, a Hausdorff space, and thus, a $\mathcal{T}_1$-space.*

Proof. Let X and Y be any two disjoint closed subsets of (E, d). Every point $x \in X$ has an open neighborhood O_x disjoint from Y, and hence is at a positive distance r_x from Y (recall Problem **??**). Similarly, every point $y \in Y$ is at a positive distance r_y from X. Consider the open sets

$$U = \cup_{x \in X} B_{\frac{r_x}{2}}(x), \quad V = \cup_{y \in Y} B_{\frac{r_y}{2}}(y).$$

We have $X \subset U$, $Y \subset V$. Moreover, U and V are disjoint. In fact, suppose the contrary that there is a point $z \in U \cap V$. Then there are points $x_0 \in X$, $y_0 \in Y$ such that

$$d(x_0, z) < \frac{r_{x_0}}{2} \quad \text{and} \quad d(z, y_0) < \frac{r_{y_0}}{2}.$$

Assume that $r_{x_0} \leq r_{y_0}$. Then we have

$$d(x_0, y_0) \leq d(x_0, z) + d(z, y_0) < \frac{r_{x_0}}{2} + \frac{r_{y_0}}{2} \leq r_{y_0},$$

i.e., $x_0 \in B_{r_{y_0}}(y_0)$. This contradicts the definition of r_{y_0}, and consequently, this shows that there is no point $z \in U \cap V$. $\square$

Definition A.4.6 *A topological space $(X, \mathcal{T})$ is said to be **metrizable** if the topology $\mathcal{T}$ can be specified by means of some metric (if $(X, \mathcal{T})$ is homeomorphic to some metric space).*

Example A.4.4 *The topological space $(X, \mathcal{T}_t)$ is not metrizable. Indeed, $(X, \mathcal{T}_t)$ is not Hausdorff, therefore cannot be a metric space, and therefore it is not metrizable.*

Finally, specifying a metric in E, we define a topology in E. Thus, any metric defines a topology, but not conversely. Therefore **all properties of topological spaces hold for metric spaces**.

A.5 Weak and Weak* topologies

A.5.1 Weak topology

Definition A.5.1 *Let X be a normed vector space. Let $X^* = \mathcal{L}(X, \mathbb{R})$ be its dual space.*

*The **weak topology** on X is the initial topology with respect to X^*. It is noted $\sigma(X, X^*)$.*

The weak topology has fewer open sets compared to the topology derived from the norm on X. We note that if X is a normed space, then the topological space $(X, \sigma(X, X^*))$ is a Hausdorff space.

These statements are equivalent:

- For all f in X^*, $f(x_n)$ converges to $f(x)$
- (x_n) converges to x in $\sigma(X, X^*)$.

Remark A.5.1 *If a topology τ is weaker than T, then τ contains less open/closed sets to compare to T, but it contains more compact sets. The compact sets are very important for theorems of existence.*

Theorem A.5.1 *For a normed space X, $(X, \sigma(X, X^*))$ is a Hausdorff not metrizable topological space.*

A.5.2 Weak* topology

Definition A.5.2 *Let X be a normed vector space. Let $X^* = \mathcal{L}(X, \mathbb{R})$ be its dual space. For all $x \in X$ we consider a linear functional $\phi_x : X^* \to \mathbb{R}$ defined by*

$$f \mapsto \phi_x(f) = f(x).$$

When x varies on X we obtain a family of linear functionals $(\phi_x)_{x \in X}$.

*The **weak* topology**, noted by $\sigma(X^*, X)$, is the coarsest topology in X^* that makes all linear functionals $(\phi_x)_{x \in X}$ continuous.*

Since $X \subset X^{**}$, the weak* topology $\sigma(X^*, X)$ is weaker than the weak topology $\sigma(X^*, X^{**})$, which is weaker than the strong topology on X^*. Clearly, the two topologies $\sigma(X^*, X^{**})$ and $\sigma(X^*, X)$ will be the same if and only if X is reflexive. In particular, if X is a Hilbert space then $X^{**} = X$, therefore the weak* topology and the weak topology coincide.

Definition A.5.3 *The weak* convergence is the convergence in the weak* topology $\sigma(X^*, X)$, denoted by $\overset{*}{\rightharpoonup}$.*

Theorem A.5.2 *Let X be a normed space. Then $(X^*, \sigma(X^*, X))$ is a Hausdorff topological space and*

$$\overline{B_1(0)} \subset (X^*, \sigma(X^*, X)) \text{ is metrizable } \iff X \text{ is separable.}$$

Appendix B

More about compactness

B.1 Compact topological spaces

Definition B.1.1 *1. A **cover** of a set A in a topological space $(X, \mathcal{T})$ is a family of sets $\{U_\alpha\}$ such that $A \subset \cup_\alpha U_\alpha$.*

*2. A cover $\{U_\alpha\}$ of A in a topological space $(X, \mathcal{T})$ is called **open** if all U_α are open in $(X, \mathcal{T})$.*

*3. A family of sets $\{V_\alpha\}$ is called a **subcover** of A in $(X, \mathcal{T})$ if*

(a) $\{V_\alpha\}$ is a subset of the cover $\{U_\alpha\}$ of A

(b) $\{V_\alpha\}$ is a cover of A.

Definition B.1.2 *A topological space $(X, \mathcal{T})$ is said to be **compact** if every open cover of $(X, \mathcal{T})$ has a finite subcover.*

Example B.1.1 *Any closed bounded subset of $\mathbb{R}^n$ is compact (see Chapter 2). On the other hand, R^n itself (e.g., the real line, a two-dimensional plane or three-dimensional space) is not compact.*

Definition B.1.3 *A system of subsets $\{U_\alpha\}$ of a set A is said to be **centered** or to have **the finite intersection property**, if every finite intersection $\cap_{k=1}^n U_k$ is nonempty.*

Theorem B.1.1 *A topological space $(X, \mathcal{T})$ is compact if and only if every centered system of closed subsets of $(X, \mathcal{T})$ has a non empty intersection.*

Proof. Suppose $(X, \mathcal{T})$ is compact, and let $\{F_\alpha\}$ be any centered system of closed subsets of $(X, \mathcal{T})$. Then the sets $U_\alpha = X \setminus F_\alpha$ are open.

Since any finite intersection $\cap_{i=1}^n F_i$ is not empty, de Morgan's law (A.2) implies

$$\cup_{i=1}^n U_i = \cup_{i=1}^n X \setminus F_i = X \setminus \cap_{i=1}^n F_i \subsetneq X$$

that there is no finite system of sets $U_i = X \setminus F_i$ which covers X. But X is compact, thus the whole system of $\{U_\alpha\}$ cannot cover X (see Definition B.1.2), and hence $\cap F_\alpha \neq \varnothing$.

Conversely, suppose every centered system of closed subsets of $(X, \mathcal{T})$ has a non-empty intersection, and let $\{U_\alpha\}$ be any open cover of X. Setting $F_\alpha = X \setminus U_\alpha$, we find using de Morgan's law (A.2) that

$$\cap_\alpha F_\alpha = \cap_\alpha (X \setminus U_\alpha) = X \setminus (\cup_\alpha U_\alpha) = X \setminus X = \varnothing.$$

By the hypothesis, this implies that the system $\{F_\alpha\}$ is not centered, *i.e.*, that there are sets $F_1, \ldots, F_n$ such that $\cap_{i=1}^n F_i = \varnothing$. But then the corresponding open sets $U_i = X \setminus F_i$ form a finite subcover of the cover $\{U_\alpha\}$. Consequently, $(X, \mathcal{T})$ is compact. $\square$

Theorem B.1.2 *Every closed subset F of a compact topological space $(X, \mathcal{T})$ is itself compact.*

Proof. The subset F is considered as a topological space with the induced topology $\mathcal{T}_F = T \cap F$. Therefore, since F is closed in $\mathcal{T}$, every set W closed in the induced topology $\mathcal{T}_F$, is also closed in the initial topology $\mathcal{T}$.

Let $\{F_\alpha\}$ be any centered system of closed subsets of the subspace $F \subset X$ by the induced topology. Then every F_α is closed in $(X, \mathcal{T})$ as well, *i.e.*, $\{F_\alpha\}$ is a centered system of closed subsets of $(X, \mathcal{T})$. Therefore $\cap_\alpha F_\alpha \neq \varnothing$, by Theorem B.1.1. But then F is compact, by Theorem B.1.1 again. $\square$

Corollary B.1.1 *Every closed subset of a compact Hausdorff space is itself a compact Hausdorff space.*

Proof. The proof follows from Theorem B.1.2 and the fact that every subset of a Hausdorff space is itself a Hausdorff space. $\square$

Theorem B.1.3 *Let $(K, \mathcal{T}_K)$ be a compact Hausdorff space and $(X, \mathcal{T}_X)$ be any Hausdorff space containing $(K, \mathcal{T}_K)$, i.e. $\mathcal{T}_K$ is the induced topology by $\mathcal{T}_X$. Then $(K, \mathcal{T}_K)$ is closed in $(X, \mathcal{T}_X)$.*

Proof.

Suppose $y \notin K$, so that $y \in X \setminus K$. Let us show that $X \setminus K$ is open in $(X, \mathcal{T}_X)$. Given any point $x \in K$, there is an open neighborhood U_x of x and an open neighborhood V_x of y (see Fig. B.1) such that

$$U_x \cap V_x = \varnothing.$$

The neighborhoods $\{U_x\}_{x \in K}$ form an open (in $\mathcal{T}_X$) cover of K. Let us recall that if a set W is closed (or open) in the initial topology $\mathcal{T}_X$, it is also closed (open) in the induced topology $\mathcal{T}_K$. Hence, by the compactness of K, the cover $\{U_x \cap K\}_{x \in K}$, open in $\mathcal{T}_K$, has a finite subcover $U_{x_1} \cap K, \ldots, U_{x_n} \cap K$. Let

$$V = V_{x_1} \cap \ldots \cap V_{x_n}.$$

Then V is an open neighborhood of the point y which does not intersect the set $U_{x_1} \cup \ldots \cup U_{x_n} \supsetneq K$, and hence $y \notin \overline{K}$. It means that

$$\forall y \in X \setminus K \ \exists \text{ open neighborhood } V \subset X \setminus K,$$

which proves that $X \setminus K$ is open in $(X, \mathcal{T}_X)$. $\square$

Theorem B.1.4 *Every compact Hausdorff space $(K, \mathcal{T}_K)$ is a normal space.*

Proof. Let X and Y be any two disjoint closed subsets of K. Let us construct open sets $V \supset X$ and $U \supset Y$ such that the normality condition holds:

$$V \cap U = \varnothing.$$

Repeating the argument given in the proof of Theorem B.1.3, we easily see that, given any point $y \in Y$, there exists an open neighborhood U_y containing y and an open set $V_y \supset X$ such that $U_y \cap V_y = \varnothing$. Since Y is compact, by Theorem B.1.2, the cover $\{U_y\}_{y \in Y}$ of the set Y has a finite subcover $U_{y_1}, \ldots, U_{y_n}$. Thus we take

$$V = V_{y_1} \cap \ldots \cap V_{y_n}, \quad U = U_{y_1} \cup \ldots \cup U_{y_n}. \qquad \square$$

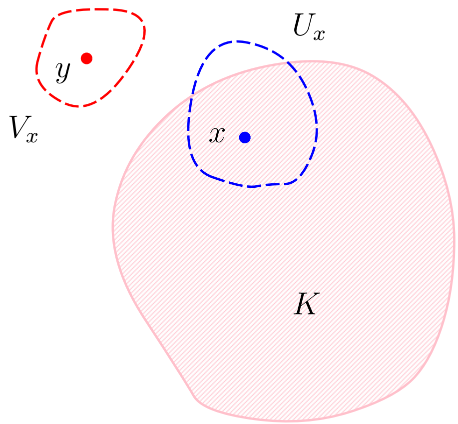

Figure B.1 – Compact K with $y \notin K$, $x \in K$ and the open neighborhoods U_x of x and V_x of y such that $U_x \cap V_x = \varnothing$.

Theorem B.1.5 *If $(X, \mathcal{T})$ is a compact space, then any infinite subset of X has at least one limit point.*

Proof. Suppose X contains an infinite set J with no limit point. Then in J there exists a countable subset

$$J_1 = \{x_1, x_2, \dots\}$$

(we remind that a set S is called countable if there exists an injective mapping from S to the set of natural numbers $\mathbb{N}$). Therefore, J_1 has no limit point. Let us show that J_1 is closed.

Let us consider $x \in X \setminus J_1$. Since J_1 has no limit point, there exists a neighborhood U of x which does not contain any point of J_1:

$$\nexists y \in J_1 : \quad y \in U.$$

Then for all $x \in X \setminus J_1$ there exists a neighborhood $U \subset X \setminus J_1$, and thus $X \setminus J_1$ is open. Consequently, we conclude that J_1 is closed.

But then the sets

$$J_n = \{x_n, x_{n+1}, \dots\} \quad n \in \mathbb{N}^*$$

form a centered system of closed sets in $(X, \mathcal{T})$ with an empty intersection

$$\cap_{n \in \mathbb{N}^*} J_n = \varnothing,$$

i.e., $(X, \mathcal{T})$ is not compact.$\square$

Remark B.1.1 *The following assertions are not equivalent in topological spaces:*

1. *any infinite subset of X has at least one limit point;*
2. *any sequence in X has a convergent subsequence.*

Definition B.1.4 *A subset $K \subset X$ of a topological space $(X, \mathcal{T})$ is called* ***countably compact*** *if every infinite subset of K has at least one limit point (in K).*

Thus Theorem B.1.5 says that every compact set is countably compact. The converse, however, is not true. For the relation between the concepts of compactness and sequentially compactness see [7]. We just formulate the following Theorem (see [7] p.95 for the proof):

Theorem B.1.6 *A topological space $(X, \mathcal{T})$ is countably compact*

1. *if and only if every countable open cover of X has a finite subcover.*

2. *if and only if every countable centered system of closed subsets of X has a nonempty intersection.*

Definition B.1.5 *A subset $K \subset X$ of a topological space $(X, \mathcal{T})$ is called **sequentially compact** if every sequence in K has a convergent subsequence.*

Definition B.1.6 *A topological space is called **locally compact** if every point has a compact neighborhood.*

Example B.1.2 *$\mathbb{R}^n$ is locally compact.*

B.2 Continuous mappings of compact spaces

Next we show that the "continuous image" of a compact space is itself a compact space:

Theorem B.2.1 *Let $(X, \mathcal{T}_X)$ be a compact space and f a continuous mapping of $(X, \mathcal{T}_X)$ in a topological space $(Y, \mathcal{T}_Y)$. Then $f(X)$ endowed with the induced topology $\mathcal{T}_y \cap f(X)$ is itself compact.*

Proof. Let $\{V_\alpha\}$ be any open (by $\mathcal{T}_y \cap f(X)$) cover of $f(X)$:

$$\cup_\alpha V_\alpha = f(X),$$

and let $U_\alpha = f^{-1}(V_\alpha)$. As f is continuous, U_α are open in $(X, \mathcal{T}_X)$. Moreover $\{U_\alpha\}$ covers the space X:

$$X = f^{-1}(\cup_\alpha V_\alpha) = \cup_\alpha f^{-1}(V_\alpha) = \cup_\alpha U_\alpha.$$

Since $(X, \mathcal{T}_X)$ is compact, $\{U_\alpha\}$ has a finite subcover $U_1, \ldots, U_n$:

$$X = \cup_{i=1}^n U_i.$$

Then the sets $V_1, \ldots, V_n$, where $V_i = f(U_i)$, cover the entire image $f(X)$. It follows that $(f(X), \mathcal{T}_Y \cap f(X))$ is compact. $\square$

Theorem B.2.2 *If f be a continuous bijection of a compact Hausdorff space $(X, \mathcal{T}_X)$ onto a compact Hausdorff space $(Y, \mathcal{T}_Y)$, then f is a homeomorphism.*

Proof. Let f be a continuous bijection between two compact Hausdorff spaces $(X, \mathcal{T}_X)$ and $(Y, \mathcal{T}_Y)$. We need to show that the inverse mapping f^{-1} is itself continuous:

$$(f^{-1})^{-1}(V) \text{ is closed for all closed } V \in X.$$

We notice that $f = (f^{-1})^{-1}$ which means that we need to prove

$$f(V) \text{ is closed for all closed } V \in X.$$

Let V be a closed set in $(X, \mathcal{T}_X)$ and $P = f(V)$ its image in $(Y, \mathcal{T}_Y)$. Then P is a compact Hausdorff space, by Theorem B.2.1. Hence, by Theorem B.1.3, P is closed in $(Y, \mathcal{T}_Y)$. Therefore, we conclude that for any closed set $P \subset Y$ the inverse image $f^{-1}(P) = V$ is closed in $(X, \mathcal{T}_X)$. Consequently, from Theorem A.2.4 it follows that f^{-1} is continuous. $\square$

B.3 Relatively compact subsets

Among the subsets of a topological space, those whose closures are compact are of special interest:

Definition B.3.1 *A subset M of a topological space $(X, \mathcal{T})$ is said to be **relatively compact** in $(X, \mathcal{T})$ if its closure $\overline{M}$ in $(X, \mathcal{T})$ is compact.*

Example B.3.1 *According to Theorem B.1.2, every subset of a compact topological space is relatively compact.*

Example B.3.2 *As we have seen in Section 2.3, every bounded subset of the real line $\mathbb{R}$ is relatively compact.*

Problem B.3.1 *A topological space $(X, \mathcal{T})$ is said to be locally compact if every point $x \in X$ has at least one relatively compact neighborhood. Show that a compact space is automatically locally compact, but not conversely. Prove that every closed subspace of a locally compact subspace is locally compact.*

Appendix C

General theory of Hilbert spaces

C.1 Series with any (countable or not) index sets

In Definitions 4.4.1 and 4.4.2, the set of index I can be an arbitrary set. It means I can be any countable or uncountable set. If I is an uncountable set, how do we define a series over I?

Definition C.1.1 *Let us consider a series*

$$\sum_{i \in I} a_i, \quad a_i \geq 0, a_i \in \mathbb{R}. \tag{C.1}$$

Let F be any finite subset of a set I. We call

$$S_F = \sum_{i \in F} a_i$$

*by a **partial sum** of the series (C.1).*

Definition C.1.2 *Series (C.1) is called convergent, if there exists*

$$S = \sup_{F \subset I} S_F \quad (F \text{ a finite subset of } I).$$

In this case the number S is called the sum of series (C.1):

$$S = \sum_{i \in I} a_i.$$

*If such S does not exist, then series (C.1) is called **divergent**.*

Lemma C.1.1 *If series (C.1) converges, then there are only a countable (or finite) number of elements a_i which are different to 0. Thus, series (C.1) is equal to an usual series with a sum over $i \in \mathbb{N}$:*

$$\sum_{i \in I} a_i = \sum_{k=1}^{\infty} a_{i_k}.$$

Proof. Let $S = \sup_{F \subset I} S_F$, where F is a finite subset of I. Then, by definition of the supremum,

$$\forall n = 1, 2, \dots \quad \exists F_n : \quad 0 \leq S - S_{F_n} \leq \frac{1}{n}.$$

Let us construct a non-decreasing sequence

$$\tilde{F}_1 \subset \tilde{F}_2 \subset \ldots \subset \tilde{F}_n \subset \ldots.$$

For instance, we can take:

$$\begin{aligned}
&\tilde{F}_1 = F_1, \\
&\tilde{F}_2 = F_1 \cup F_2, \quad (\Rightarrow \; F_1 \subset \tilde{F}_2 \text{ and } S \geq S_{\tilde{F}_2} \geq S_{F_2}) \\
&\ldots \\
&\tilde{F}_n = F_1 \cup F_2 \cup \ldots \cup F_n, \\
&\ldots
\end{aligned}$$

By construction,

$$\forall n \in \mathbb{N}^* \quad S \geq S_{\tilde{F}_n} \geq S_{F_n} \quad \text{and } S - S_{\tilde{F}_n} \leq \frac{1}{n}.$$

Moreover, since the sequence $(\tilde{F}_n)_{n\in\mathbb{N}^*}$ is non-decreasing, thus, by definition, the sequence $(S_{\tilde{F}_n})_{n\in\mathbb{N}^*}$ is a bounded (by S) non-decreasing sequence:

$$S_{\tilde{F}_1} \leq S_{\tilde{F}_2} \leq \ldots \leq S_{\tilde{F}_n} \leq \ldots \leq S.$$

Therefore, $S_{\tilde{F}_n} \to S$ for $n \to +\infty$.

We define

$$F = \cup_{n\in\mathbb{N}^*} \tilde{F}_n.$$

The set F is countable (as a countable union of finite sets).

Let us prove that

$$\sum_{i\in I} a_i = \sum_{i\in F} a_i = S,$$

or equivalently, let's prove that

$$a_i = 0 \quad \text{for } i \notin F.$$

Indeed, let's take a_{i_0} such that $i_0 \notin F$ and let's suppose that

$$a_{i_0} + \sum_{i\in F} a_i = S.$$

Then we define

$$\hat{F}_n = \tilde{F}_n \cup \{i_0\}.$$

Since $(\hat{F}_n)_{n\in\mathbb{N}^*}$ is a non-decreasing sequence of finite sets, then $(S_{\hat{F}_n})_{n\in\mathbb{N}^*}$ is also a non-decreasing sequence bounded by S. Therefore, we find that

$$S_{\hat{F}_n} = S_{\tilde{F}_n} + a_{i_0} \to S \text{ for } n \to +\infty.$$

But, at the same time, $S_{\tilde{F}_n} \to S$ for $n \to +\infty$. Thus, passing to the limit, we obtain that

$$S = S + a_{i_0},$$

from where we conclude that $a_{i_0} = 0$. $\square$

Remark C.1.1 *If we consider series (C.1) with any real (or complex) coefficients a_i, we say that the series converges if $\sum_{i\in I} |a_i|$ converges.*

Using Definition C.1.1, we can generalize the Bessel inequality for all systems (countable or not) of a Pre-Hilbert space X: Since for all finite F in I $\sum_{i\in F}|c_i|^2 \leq \|f\|^2$, thus

$$\sum_{i\in I}|c_i|^2 = \sup_{F\subset I}(\sum_{i\in F}|c_i|^2) \leq \|f\|^2.$$

Let us formulate it in the following Lemma:

Lemma C.1.2 *Let $\{v_i, i\in I\}$ be an orthonormal system in a Pre-Hilbert space X. For all $f\in X$, the Fourier coefficients of f with respect to the system $\{v_i, i\in I\}$*

$$c_i = \langle f, v_i\rangle, \quad i\in I,$$

satisfy the Bessel inequality:

$$\sum_{i\in I}|c_i|^2 \leq \|f\|^2.$$

As $|c_i|^2 \geq 0$ for all $i\in I$ and the series $\sum_{i\in I}|c_i|^2$ is bounded by $\|f\|^2$, it follows that $\sum_{i\in I}|c_i|^2$ converges. Therefore, we notice:

Corollary C.1.1 *Let $\{c_i, i\in I\}$ be the Fourier coefficients of f with respect to an orthonormal system $\{v_i, i\in I\}$ in a Pre-Hilbert space X. Then $J=\{i\in I|\ c_i\neq 0\}$ is a countable set.*

Since any convergant series is actually a countable series, in what follows we only consider the convergant series with $I=\mathbb{N}^*$.

Let us prove

Lemma C.1.3 *Let $v=\{v_i, i\in I\}$ be an orthonormal system in a Hilbert space H and let $\{c_i, i\in I\}$ be an arbitrary set of numbers satisfying the inequality:*

$$\sum_{i\in I}|c_i|^2 < \infty.$$

(Thus $J=\{i\in I|\ c_i\neq 0\}$ is a countable set.) Then in H there exists a unique vector $y=\sum_{i\in I}c_iv_i$.

Proof. Thanks to Lemma C.1.1, we can write $y=\sum_{n=1}^{\infty}c_{i_n}v_{i_n}$. Let us define the partial sum

$$S_n = \sum_{k=1}^{n} c_{i_k}v_{i_k}$$

and show that the sequence $(S_n)_{n\in\mathbb{N}^*}$ is a Cauchy sequence in H.

For all $\epsilon>0$ and $p\in\mathbb{N}$ we have

$$\|S_{n+p}-S_n\|^2 = \|\sum_{k=n+1}^{n+p} c_{i_k}v_{i_k}\|^2.$$

Since v is an orthonormal system, we find that

$$\|S_{n+p}-S_n\|^2 = \|\sum_{k=n+1}^{n+p} c_{i_k}v_{i_k}\|^2 = \sum_{k=n+1}^{n+p}|c_{i_k}|^2.$$

By the assumption, $\sum_{i\in I}|c_i|^2<\infty$, and, consequently, there exists $n_0(\epsilon)\in\mathbb{N}$ such that for all $n\geq n_0(\epsilon)$

$$\|S_{n+p}-S_n\|^2 = \sum_{k=n+1}^{n+p}|c_{i_k}|^2 \leq \epsilon,$$

from where we conclude that $(S_n)_{n\in\mathbb{N}^*}$ is a Cauchy sequence in H. As H is complete, then there exists a unique $y\in H$ such that $S_n\to y$ for $n\to+\infty$ in H (the unicity follows from the unicity of the limit). □

C.2 Hilbertian basis

Using the previous section, we can formulate the general theorem of the basis existence in a Hilbert space:

Theorem C.2.1 *Let H be a Hilbert space. Then in H there exists an orthonormal basis.*

Moreover, for an orthonormal system $e = \{e_i, i \in I\}$ in H the following assertions are equivalent:

1. *e is total,*
2. *$e^\perp = \{0\}$:*
$$\forall i \in I \quad \langle e_i, x\rangle = 0 \quad \Rightarrow \quad x = 0, \tag{C.2}$$
3. *e is an orthonormal basis.*

Proof. The proof of the existence of an orthonormal basis in a Hilbert space is based on Zorn's lemma. Instead of this general case, we will prove it for a separable Hilbert space in Theorem 4.4.2.

Let us prove the equivalence of 1), 2) and 3).

1) $\Rightarrow$ 2) Let $z \in H$ and let

$$\forall i \in I \quad z \perp e_i \quad i.e., \quad \forall i \in I \quad \langle z, e_i\rangle = 0.$$

The inner product $\langle\cdot,\cdot\rangle$ is a bilinear form, consequently

$$z \perp \mathrm{Span}(e), \quad i.e., \quad \forall \alpha_i \in \mathbb{R} \quad \langle z, \sum_{i\in I} \alpha_i e_i\rangle = 0.$$

The inner product $\langle\cdot,\cdot\rangle$ is continuous. In addition, if $x_n \to x$ in H such that $\langle z, x_n\rangle = 0$ for all n, then, by the continuity of the inner product,

$$\lim_{n\to\infty} \langle z, x_n\rangle = \langle z, x\rangle = 0,$$

i.e., the limit points of orthogonal sequences to z are orthogonal to z. We conclude that

$$z \perp \overline{\mathrm{Span}(e)} = H.$$

Then $z \perp H$, from where $z \perp z$ and hence $z = 0$.

2) $\Rightarrow$ 3) For all f in H we construct the Fourier series of f with respect to e:

$$\sum_{i\in I} \langle f, e_i\rangle e_i.$$

Thanks to the Bessel inequality (see Lemma C.1.2)

$$\sum_{i\in I} |\langle f, e_i\rangle|^2 \le \|f\|^2 < \infty,$$

we find that the numerical series $\sum_{i\in I} |\langle f, e_i\rangle|^2$ converges. We apply Lemma C.1.3: there exists a unique $y \in H$ such that

$$y = \sum_{i\in I} \langle f, e_i\rangle e_i.$$

We recall that the sum (of non-zero elements) here is at most countable (see Corollary C.1.1).

Let us show that $y = f$. Define $z = y - f$. Then we have

$$\forall j \in I \quad \langle z, e_j\rangle = \langle y - f, e_j\rangle = \langle \sum_{i\in I}\langle f, e_i\rangle e_i, e_j\rangle - \langle f, e_j\rangle = \langle f, e_j\rangle - \langle f, e_j\rangle = 0,$$

which means that

$$\forall j \in I \quad z \perp e_j.$$

Since e satisfies (C.2), this implies that $z = 0$, and therefore, $f = y$. We conclude that e is an orthonormal basis.

<u>3) ⇒ 1)</u> Let e be an orthonormal basis. Let us prove that $\overline{\mathrm{Span}(e)} = H$.

For all $f \in H$

$$f = \sum_{i\in I}\langle f, e_i\rangle e_i.$$

If $S_n = \sum_{i=1}^{n}\langle f, e_i\rangle e_i \in \mathrm{Span}(e)$ is a partial sum, then $S_n \to f$ for $n \to \infty$ in H and $f \in \overline{\mathrm{Span}(e)}$. Thus $H \subset \overline{\mathrm{Span}(e)}$.

Conversely: since $e \subset H$, then $\mathrm{Span}(e) \subset H$ and thus $\overline{\mathrm{Span}(e)} \subset \overline{H} = H$.

Finally, we conclude that $\overline{\mathrm{Span}(e)} = H$. □

Appendix D

L^p spaces: reflexivity, separability, dual spaces

D.1 Definition

Let's consider Ω an open set of $\mathbb{R}^n$ equipped with the Lebesgue measure.

Definition D.1.1 *The set of Lebesgue-integrable functions from Ω to $\mathbb{R}$ is noted $L^1(\Omega)$ or simply L^1 when no confusion is possible. For $f \in L^1(\Omega)$ we define*

$$\|f\|_{L^1(\Omega)} = \int_\Omega |f| d\mu.$$

Each element of L^1 is a classe of equivalent functions, which are equal almost everywhere.

Definition D.1.2 *Let $0 < p < \infty$, and $f : \Omega \to \mathbb{R}$ be measurable functions. The function f belongs to $L^p(\Omega)$ if $|f|^p \in L^1(\Omega)$. For $f \in L^p(\Omega)$ we define*

$$\|f\|_{L^p(\Omega)} = \left(\int_\Omega |f|^p d\mu \right)^{\frac{1}{p}}.$$

Remark D.1.1 *For $p = 1$ we obtain from Definition D.1.2 that $f \in L^1(\Omega)$ implies that*

1. *f is measurable,*
2. *$|f| \in L^1(\Omega)$,*

which exactly means that f is Lebesgue-integrable. Hence, Definition D.1.2 for $p = 1$ is equivalent to Definition D.1.1. Functions that are equal almost everywhere are "identified".

Remark D.1.2 *The space $(C(\Omega), \|\cdot\|_{L^1(\Omega)})$ is not complete. By the way,*

$$L^1(\Omega) = \overline{C(\Omega)}^{\|\cdot\|_{L^1}},$$

i.e. $L^1(\Omega)$ is the completion of $C(\Omega)$ by the norm $\|\cdot\|_{L^1}$.

Definition D.1.3 *Let f be a function from Ω to $\mathbb{R}$. We define essential supremum of f as a number*

$$\text{ess sup}_{x\in\Omega} f(x) = \inf_{\mu(A)=0} \left[\sup_{x\in\Omega\setminus A} f(x) \right] = \inf_{B\subsetneq\Omega:\ \mu(\Omega\setminus B)=0} \left[\sup_{x\in B} f(x) \right]. \tag{D.1}$$

Example D.1.1 *Let* $f(x) = \mathbb{1}_{\mathbb{Q}\cap[0,1]}(x)$*. By definition of the essential supremum,*

$$\text{ess sup}_{x\in[0,1]} f(x) = 0,$$

but $\sup_{x\in[0,1]} f(x) = 1.$

Remark D.1.3 *1. It holds*

$$\text{ess sup}_{x\in\Omega} f(x) = \inf\{M \in \mathbb{R} : \ f(x) \le M \quad a.\ e.\ in\ \Omega\}.$$

2. Let $\Omega \subsetneq \mathbb{R}^n$ *be an open set and* $f : \Omega \to \mathbb{R}$ *be a continuous function. Then it holds*

$$\text{ess sup}_{x\in\Omega} f(x) = \sup_{x\in\Omega} f(x).$$

Definition D.1.4 *We note* $L^\infty(\Omega)$ *the set of measurable functions from* Ω *to* $\mathbb{R}$ *for which there exists a real number* C *such that for almost every* x *in* Ω*,* $|f(x)| \le C$*, i.e.*

$$\text{ess sup}_{x\in\Omega} |f(x)| < \infty.$$

Functions equal almost everywhere are identified. We note

$$\|f\|_{L^\infty(\Omega)} = \text{ess sup}_{x\in\Omega} |f(x)|.$$

Problem D.1.1 *Prove using Remark D.1.3 1) that*

$$|f(x)| \le \|f\|_{L^\infty(\Omega)} \qquad a.\ e.\ in\ \Omega.$$

Proposition D.1.1 *Let* $f \in L^p(\Omega)$ *for* $0 < p \le \infty$*. Then*

$$\|f\|_{L^p(\Omega)} = 0 \quad \Longleftrightarrow \quad f = 0 \ a.e.\ on\ \Omega.$$

Proof. For $0 < p < \infty$ we have:

$$\|f\|_{L^p(\Omega)} = 0 \quad \Longleftrightarrow \quad \int_\Omega |f|^p d\mu = 0 \quad \Longleftrightarrow \quad |f|^p = 0 \text{ a.e. on } \Omega \quad \Longleftrightarrow \quad f = 0 \text{ a.e. on } \Omega.$$

Problem D.1.2 *Prove it for* $p = \infty$*:*

$$\|f\|_{L^\infty(\Omega)} = 0 \quad \Longleftrightarrow \quad f = 0 \ a.e.\ on\ \Omega.$$

There are some indications:

1. *Use Remark D.1.3 1).*
2. *Show for* $A = \|f\|_{L^\infty(\Omega)}$ *that* $\Omega_{A+\frac{1}{n}} = \{x \in \Omega : f(x) \le A + \frac{1}{n}\}$ *has Lebesgue measure* $\mu(\Omega_{A+\frac{1}{n}}) = \mu(\Omega)$ *for* $n = 1, 2, \ldots.$
3. *Consider* $\Omega_A = \cap_{n=1}^\infty \Omega_{A+\frac{1}{n}}$ *which is a full Lebesgue measure:* $\mu(\Omega_A) = \mu(\Omega)$*.*

Problem D.1.3 *Show that $L^p(\Omega)$ for $0 < p \leq \infty$ is a linear vector space. Indication: use the numerical inequality*

$$\forall a, b \geq 0 \quad and\ p > 0 \quad (a+b)^p \leq 2^p(a^p + b^p).$$

For example, if $a \geq b$, then

$$(a+b)^p \geq (2a)^p = 2^p a^p \leq 2^p(a^p + b^p).$$

Definition D.1.5 *Let $p \in [1, \infty]$.*

A function f belongs to $L^{p,loc}(\Omega)$ when $f\mathbb{1}_K$ belongs to $L^p(\Omega)$ for every compact $K \subset \Omega$, where $\mathbb{1}_K$ is the characteristic function of K: $\mathbb{1}_K(x) = 1$ if $x \in K$ and 0 otherwise.

Definition D.1.6 *Let $p \in [1, \infty]$.*

*We call **Hölder conjugate (or dual index)** of p, the number $p' = 1 + \frac{1}{p-1}$ so that $\frac{1}{p} + \frac{1}{p'} = 1$ (see Section 2.1.1) (if $p = 1$ then $p' = \infty$ and $p = \infty$ then $p' = 1$).*

Note that the Hölder conjugate of 2 is 2.

Proposition D.1.2 ***(Hölder's Inequality)*** *Let $p \in [1, \infty]$ and p' be its Hölder conjugate. Let $f \in L^p(\Omega)$ and $g \in L^{p'}(\Omega)$. Then $fg \in L^1$ and*

$$\|fg\|_{L^1(\Omega)} \leq \|f\|_{L^p(\Omega)} \|g\|_{L^{p'}(\Omega)}.$$

Problem D.1.4 *Using results of Section 2.1.1, prove Hölder's inequality.*

If we summarize all the results, we have

Corollary D.1.1 *Let $p \in [1, \infty]$. $\|\cdot\|$ is a norm on $L^p(\Omega)$.*

D.1.1 Applications of Hölder's inequality

Proposition D.1.3 *Let $0 < p \leq q \leq \infty$, $\Omega \subsetneq \mathbb{R}^n$, $\mu(\Omega) < \infty$. Then*

1. *$L^q(\Omega) \subset L^p(\Omega)$,*
2. *$\|f\|_{L^p(\Omega)} \leq \mu(\Omega)^{\frac{1}{p} - \frac{1}{q}} \|f\|_{L^q(\Omega)}$.*

Proof. We denote by $Q = \frac{q}{p} \geq 1$ and calculate the dual index of Q:

$$Q' = \frac{Q}{Q-1} = \frac{q}{q-p}.$$

We now apply Hölder's inequality with Q and Q':

$$\begin{aligned}
\|f\|^p_{L^p(\Omega)} &= \int_\Omega |f|^p d\mu = \int_\Omega |f|^p \cdot 1 d\mu \\
&\leq \||f|^p\|_{L^Q} \cdot \|1\|_{L^{Q'}} = \left(\int_\Omega (|f|^p)^Q d\mu\right)^{\frac{1}{Q}} \left(\int_\Omega 1 d\mu\right)^{\frac{1}{Q'}} = \left(\int_\Omega |f|^q d\mu\right)^{\frac{p}{q}} \mu(\Omega)^{\frac{q-p}{q}},
\end{aligned}$$

from where

$$\|f\|_{L^p(\Omega)} = \left(\int_\Omega |f|^p d\mu\right)^{\frac{1}{p}} \leq \left(\int_\Omega |f|^q d\mu\right)^{\frac{1}{q}} \mu(\Omega)^{\frac{q-p}{pq}},$$

which gives 2). If $f \in L^q$, thanks to 2), $\|f\|_{L^p(\Omega)} < \infty$, what implies that $f \in L^p$, *i.e.* 1) holds. $\square$

Let us notice that the condition $\mu(\Omega) < \infty$ is very important. For instance,

Problem D.1.5 *For $\Omega = \mathbb{R}^+$ give an example of a function f such that for $0 < p < q \le \infty$*

$$f \in L^q(\mathbb{R}^+), \quad \textit{but } f \notin L^p(\mathbb{R}^+).$$

Proposition D.1.4 *Let $0 < p_1 < p < p_2 \le \infty$. Let's define α by the equality*

$$\frac{1}{p} = \frac{\alpha}{p_1} + \frac{1-\alpha}{p_2} \tag{D.2}$$

Note that for $p = p_1$ $\alpha = 1$ and for $p = p_2$ $\alpha = 0$. Such $\alpha \in]0,1[$ exists and it is unique. Then from $f \in L^{p_1}(\Omega) \cap L^{p_2}(\Omega)$ follows that $f \in L^p(\Omega)$ and

$$\|f\|_{L^p(\Omega)} \le \|f\|^{\alpha}_{L^{p_1}(\Omega)} \|f\|^{1-\alpha}_{L^{p_2}(\Omega)}. \tag{D.3}$$

Proof. Let us consider $p_2 < \infty$. We denote by $q = \frac{p_1}{\alpha p}$ and calculate the dual index of q using (D.12):

$$\begin{aligned}
&1 = \frac{\alpha p}{p_1} + \frac{(1-\alpha)p}{p_2} \quad \Rightarrow \quad \frac{\alpha p}{p_1} < 1 \Rightarrow \ q > 1 \\
&q' = \frac{q}{q-1} \Rightarrow \frac{1}{q'} = 1 - \frac{1}{q} = 1 - \frac{\alpha p}{p_1} = \frac{(1-\alpha)p}{p_2}, \\
&\text{i.e. } q' = \frac{p_2}{(1-\alpha)p}.
\end{aligned}$$

We now apply Hölder's inequality with q and q':

$$\begin{aligned}
&\|f\|^p_{L^p(\Omega)} = \int_\Omega |f|^p d\mu = \int_\Omega |f|^{\alpha p} \cdot |f|^{(1-\alpha)p} d\mu \\
&\le \left(\int_\Omega (|f|^{\alpha p})^{\frac{p_1}{\alpha p}} d\mu \right)^{\frac{\alpha p}{p_1}} \left(\int_\Omega (|f|^{(1-\alpha)p})^{\frac{p_2}{(1-\alpha)p}} d\mu \right)^{\frac{(1-\alpha)p}{p_2}} = \left(\int_\Omega |f|^{p_1} d\mu \right)^{\frac{\alpha p}{p_1}} \left(\int_\Omega |f|^{p_2} d\mu \right)^{\frac{(1-\alpha)p}{p_2}},
\end{aligned}$$

from where

$$\begin{aligned}
&\|f\|_{L^p(\Omega)} = \left(\int_\Omega |f|^p d\mu \right)^{\frac{1}{p}} \le \left(\int_\Omega |f|^{p_1} d\mu \right)^{\frac{\alpha}{p_1}} \left(\int_\Omega |f|^{p_2} d\mu \right)^{\frac{1-\alpha}{p_2}} \\
&= \|f\|^{\alpha}_{L^{p_1}(\Omega)} \|f\|^{1-\alpha}_{L^{p_2}(\Omega)},
\end{aligned}$$

which gives (D.3).

Problem D.1.6 *Prove (D.3) for $p_2 = \infty$.*

We finish with the interpolation inequality (see [2]):

Proposition D.1.5 *Let $\{f_i, i \in I\}$ be a family of functions with $f_i \in L^{p_i}(\Omega)$ and $\frac{1}{p} = \sum \frac{1}{p_i} \le 1$. Then $\prod f_i \in L^p(\Omega)$ and*

$$\|\prod f_i\|_{L^p(\Omega)} \le \prod \|f_i\|_{L^{p_i}(\Omega)}$$

Corollary D.1.2 *If $f \in L^p(\Omega) \cap L^q(\Omega)$ then $f \in L^r(\Omega)$ for any r such that $p \le r \le q$.*

D.1.2 Completeness of L^p spaces

Are L^p Banach spaces? The answer is given in the Fischer-Riesz Theorem:

Theorem D.1.1 ***(Fischer-Riesz)*** *Let $p \in [1, \infty]$, $\Omega \subset \mathbb{R}^n$, $\mu(\Omega) > 0$. Then $L^p(\Omega)$ is a Banach space.*

Note that: $\mu(\Omega)$ is the measure of Ω.

Remark D.1.4 *The theorem can be reformulated in terms of Cauchy sequences as:*

For $p \in [1, \infty]$, let $(f_k)_{k\in\mathbb{N}}$ be a Cauchy sequence in $L^p(\Omega)$. Then there exists $f \in L^p(\Omega)$ such that $\lim_{k\to\infty} \|f_k - f\|_{L^p(\Omega)} = 0$.

Proof. Proof for $p = \infty$. Let (f_n) be a Cauchy sequence in $L^\infty(\Omega)$, thus we can write

$$\|f_m - f_n\|_{L^\infty(\Omega)} \to 0 \quad m, n \to \infty,$$

what means that

there exists $\Omega_{nm} \subset \Omega$ of full measure ($\mu(\Omega \setminus \Omega_{nm}) = 0$) such that

$$\sup_{x\in\Omega_{nm}} |f_m(x) - f_n(x)| \to 0 \quad m, n \to \infty.$$

We note $\Omega' = \cap_{n,m}^{\infty} \Omega_{nm} \subset \Omega$, then $\mu(\Omega \setminus \Omega') = 0$ and $\Omega' \subset \Omega_{nm}$ for all n and m in $\mathbb{N}$. Therefore,

$$\sup_{x\in\Omega'} |f_m(x) - f_n(x)| \to 0 \quad m, n \to \infty.$$

Thus, $(f_n(x))$ is a Cauchy sequence (in $\mathbb{R}$) for all $x \in \Omega'$. Thanks to Cauchy criteria for numeric sequences, we conclude that there exists

$$\lim_{n\to\infty} f_n(x) = f(x) \quad \forall x \in \Omega',$$

i.e. almost everywhere on Ω.

We need to verify that $f \in L^\infty(\Omega)$ and that $f_n \to f$ for $n \to \infty$ in $L^\infty(\Omega)$.

For any k in $\mathbb{N}^*$, there exists an integer N such that $n > m > N$ implies

$$|f_m(x) - f_n(x)| < 1/k \quad \forall x \in \Omega'.$$

Taking m to the limit yields

$$|f(x) - f_n(x)| < 1/k \quad \forall n > N(k) \quad \forall x \in \Omega'.$$

Therefore,

$$\sup_{x\in\Omega'} |f(x) - f_n(x)| < 1/k \quad \forall n > N(k),$$

i.e.

$$\|f - f_n\|_{L^\infty} = \inf_{\Omega'\subset\Omega:\mu(\Omega\setminus\Omega')=0} \sup_{x\in\Omega'} |f(x) - f_n(x)| \le 1/k \quad \forall n > N(k).$$

We summarize:

- $f - f_n \in L^\infty \quad \forall n > N(k)$,
- $\|f - f_n\|_{L^\infty} \le 1/k \quad \forall n > N(k) \quad \Rightarrow f_n \to f$ for $n \to \infty$ in $L^\infty(\Omega)$.

Since L^∞ is a linear vector space and in addition $f_n \in L^\infty(\Omega)$, it follows that

$$f = f_n + (f - f_n) \in L^\infty(\Omega).$$

Proof for $p \in [1, \infty[$

Let (f_n) be a Cauchy sequence in $L^p(\Omega)$, *i.e.*

$$\|f_m - f_n\|_{L^p(\Omega)} \to 0 \quad m, n \to \infty.$$

Step 1– There exists a subsequence (f_{n_k}) such that

$$\|f_{n_{k+1}} - f_{n_k}\|_{L^p(\Omega)} < \frac{1}{2^k}.$$

Since (f_n) is a Cauchy sequence, there exists $N_1 \in \mathbb{N}$ such that $m > n > N_1$ implies $\|f_n - f_m\|p < \frac{1}{2}$. Let $n_1 = N_1$. There exists $N_2 \in \mathbb{N}$ such that $m > n > N_2$ implies $\|f_n - f_m\|p < \frac{1}{2^2}$. Let $n_2 = \max\{n_1, N_2\}$.

(f_{n_k}) is constructed by induction.

Step 2– (f_{n_k}) converges in L^p: Let

$$g_m = \sum_{k=1}^{m} |f_{n_{k+1}} - f_{n_k}|.$$

We note that $g_m \geq 0$. We have

$$\|g_m\|_{L^p} = \| \sum_{k=1}^{m} |f_{n_{k+1}} - f_{n_k}| \|_{L^p} \leq \sum_{k=1}^{m} \|f_{n_{k+1}} - f_{n_k}\|_{L^p} \leq 1 - \frac{1}{2^m} < 1.$$

Then the theorem of Beppo-Levi (see [4]) provides $g \in L^p$ such that $g_m(x) \to g(x)$ a.e. and $\|g_m - g\|_{L^p} \to 0$.

For l and k with $l > k$, we have

$$|f_{n_l}(x) - f_{n_k}(x)| \leq |f_{n_l}(x) - f_{n_{l-1}}(x)| + \ldots + |f_{n_k+1}(x) - f_{n_k}(x)| \leq g(x) - g_{k-1}(x) \tag{D.4}$$

Consequently $(f_{n_k}(x))$ is a Cauchy sequence in $\mathbb{R}$. Note $f(x)$ its limit.

Taking l to the limit gives

$$|f(x) - f_{n_k}(x)| \leq g(x) - g_{k-1}(x) \leq g(x)$$

Hence:

- $|f(x) - f_{n_k}(x)|^p \to 0$ for a.e.x,
- $|f(x) - f_{n_k}(x)|^p \leq g(x)^p$.

From the Dominated Convergence Theorem (see [4]) we derive

$$f = f_{n_k} + (f - f_{n_k}) \in L^p$$

and $\|f - f_{n_k}\|_p \to 0$.

Let $\epsilon > 0$ be arbitrary and $n \in \mathbb{N}$. Let $m > n$ be such that $m = n_k$ for a given k. Since (f_n) is a Cauchy sequence, $\|f_n - f_m\|_{L^p} < \epsilon/2$. Since (f_{n_k}) converges toward f, $\|f_m - f\|_{L^p} < \epsilon/2$. Hence

$$\|f_n - f\|_{L^p} \leq \|f_n - f_m\|_{L^p} + \|f_m - f\|_{L^p} < \epsilon.$$

Thus (f_n) converges toward f in L^p. $\square$

Remark D.1.5 *We note that we can use Theorem 2.6.1 to prove the case $p = \infty$ in Fischer-Riesz Theorem: Let f_n be an absolutely convergent series. We define*

$$\Omega_n = \{x \in \Omega | \ f_n(x) < \|f_n\|\}$$

of full measure and take

$$\Omega' = \bigcap_n \Omega_n.$$

Since,

$$\mu(\Omega \setminus \Omega') = \bigcup_n \mu(\Omega \setminus \Omega_n) = 0,$$

and $\sum_n f_n(x)$ converges on Ω' since it is absolutely convergent in $\mathbb{R}$.

D.2 Study of L^p $(1 < p < \infty)$

D.2.1 Reflexivity

First we prove the Clarkson inequality for $2 \le p < \infty$:

Proposition D.2.1 ***(Clarkson inequality)*** *For $2 \le p < \infty$, we have*

$$\forall f, g \in L^p \quad \left\| \frac{f+g}{2} \right\|_{L^p}^p + \left\| \frac{f-g}{2} \right\|_{L^p}^p \le \frac{1}{2} \left(\|f\|_{L^p}^p + \|g\|_{L^p}^p \right). \tag{D.5}$$

Proof. It is suficient to show (since a norm is a function with values in $\mathbb{R}^+$) that for $2 \le p < \infty$

$$\forall a, b \in \mathbb{R} \quad \left| \frac{a+b}{2} \right|^p + \left| \frac{a-b}{2} \right|^p \le \frac{1}{2}(|a|^p + |b|^p).$$

We have

$$\forall \alpha, \beta \ge 0 \quad \alpha^p + \beta^p \le (\alpha^2 + \beta^2)^{\frac{p}{2}}.$$

Indeed, for all positive fixed β, the function $y = (x^2 + \beta^2)^{\frac{p}{2}} - x^p - \beta^p$ is increasing on $]0, +\infty[$ for $2 \le p < \infty$:

$$\begin{aligned} y' \ge 0 \quad &\Leftrightarrow \quad px(x^2+\beta^2)^{\frac{p-2}{2}} - px^{p-1} \ge 0 \quad \Leftrightarrow \quad (x^2+\beta^2)^{\frac{p-2}{2}} - x^{p-2} \ge 0 \\ \Leftrightarrow \quad &(x^2+\beta^2)^{p-2} \ge x^{2(p-2)} \quad \Leftrightarrow \quad x^2 + \beta^2 \ge x^2 \quad \Leftrightarrow \quad \beta^2 \ge 0. \end{aligned}$$

We notice that for $x = 0$ $y' = 0$ and that if we fix α instead of β, by the symmetry of the formula, we also obtain an increasing function $(x^2 + \alpha^2)^{\frac{p}{2}} - x^p - \alpha^p$.

We chose $\alpha = \frac{a+b}{2}$ and $\beta = \frac{a-b}{2}$ and find that

$$\left| \frac{a+b}{2} \right|^p + \left| \frac{a-b}{2} \right|^p \le \left(\left| \frac{a+b}{2} \right|^2 + \left| \frac{a-b}{2} \right|^2 \right)^{\frac{p}{2}} = \left(\frac{a^2}{2} + \frac{b^2}{2} \right)^{\frac{p}{2}} \le \frac{1}{2}(|a|^p + |b|^p).$$

Here we have used that the function $x \mapsto |x|^{\frac{p}{2}}$ is convex for $p \ge 2$. $\square$

Remark D.2.1 *For $1 < p \le 2$ it holds the following Clarkson inequality:*

$$\forall f, g \in L^p \quad \left\| \frac{f+g}{2} \right\|_{L^p}^{p'} + \left\| \frac{f-g}{2} \right\|_{L^p}^{p'} \le \left(\frac{1}{2}\|f\|_{L^p}^p + \frac{1}{2}\|g\|_{L^p}^p \right)^{\frac{1}{p-1}}, \tag{D.6}$$

where p' is the Hölder conjugate of p (see [4]), i.e., $\frac{1}{p} + \frac{1}{p'} = 1$. The proof of (D.6) can be found in [6].

Theorem D.2.1 *Let* $p \in]1, \infty[$. L^p *is uniformly convex.*

Proof. We know, thanks to the Fischer-Riesz Theorem, that L^p is a Banach space.

Let us prove that L^p is uniformly convex (see Definition 3.5.3) for $2 \le p < \infty$ using (D.5) (the proof of the case $1 < p \le 2$ is analogous and is based on (D.6)).

Given $\epsilon > 0$, we suppose that

$$\|f\|_{L^p} \le 1, \quad \|g\|_{L^p} \le 1 \text{ and } \|f - g\|_{L^p} > \epsilon.$$

From (D.5) we find that

$$\left\|\frac{f+g}{2}\right\|_{L^p}^p < 1 - \left(\frac{\epsilon}{2}\right)^p,$$

and thus

$$\left\|\frac{f+g}{2}\right\|_{L^p} < 1 - \delta$$

with

$$\delta = 1 - \left(1 - \left(\frac{\epsilon}{2}\right)^p\right)^{\frac{1}{p}} > 0. \quad \square$$

Corollary D.2.1 *Let* $p \in]1, \infty[$. L^p *is reflexive.*

Proof. Since L^p is uniformly convex for $p \in]1, \infty[$, then by the Milman-Pettis Theorem 3.5.2 L^p is reflexive.

Let us give the proof that L^p is reflexive for $p \in]1, 2]$ using the fact that L^p is reflexive for $p \in [2, \infty[$.

<u>Define $T : L^p \to (L^{p'})^*$ as follows:</u>

For a fixed $u \in L^p$, the mapping $Tu : \ f \in L^{p'} \mapsto \int uf \ \in \mathbb{R}$ is a linear functional on $L^{p'}$, which is continuous thanks the Hölder inequality:

$$|\int uf| \le \|u\|_{L^p}\|f\|_{L^{p'}} \quad \Rightarrow \quad \exists C = \|u\|_{L^p} \ge 0 : |\int uf| \le C\|f\|_{L^{p'}} \quad (\text{as } u \in L^p,\ \|u\|_{L^p} < \infty).$$

Thus, in our notations,

$$\forall f \in L^{p'} \quad \langle Tu, f\rangle = \int uf,$$

where $T : \ u \in L^p \mapsto Tu \in (L^{p'})^*$.

<u>Let us prove that $\|Tu\|_{(L^{p'})^*} = \|u\|_{L^p}$.</u>

By the Hölder inequality, we have

$$|\langle Tu, f\rangle| \le \|u\|_{L^p}\|f\|_{L^{p'}} \quad \Rightarrow \quad \|Tu\|_{(L^{p'})^*} \le \|u\|_{L^p}.$$

On the other hand, let us take

$$f_0(x) = |u(x)|^{p-2}u(x) \quad (f_0(x) = 0 \quad \text{if } u(x) = 0).$$

For f_0 we have

$$f_0 \in L^{p'}, \quad \|f_0\|_{L^{p'}} = \|u\|_{L^p}^{p-1}, \quad \langle Tu, f_0\rangle = \|u\|_{L^p}^p.$$

Thus, we find that

$$\|Tu\|_{(L^{p'})^*} \ge \frac{\langle Tu, f_0\rangle}{\|f_0\|} = \|u\|_{L^p}.$$

Since

$$\|u\|_{L^p} \le \|Tu\|_{(L^{p'})^*} \le \|u\|_{L^p},$$

we conclude that $\|Tu\|_{(L^{p'})^*} = \|u\|_{L^p}$.

L^p is reflexive: From $\|Tu\|_{(L^{p'})^*} = \|u\|_{L^p}$ it follows that T is an isometry of L^p to a closed (as L^p is complete) sub-space of $(L^{p'})^*$. Indeed, let us show that $T(L^p)$ is a closed sub-space of $(L^{p'})^*$. Since L^p is complete, for all Cauchy sequence (f_k) in L^p there exists $f \in L^p$ such that

$$f_k \to f \text{ in } L^p.$$

Thus, $Tf \in T(L^p) \subset (L^{p'})^*$ and, as T is linear and continuous operator from L^p to $(L^{p'})^*$, it follows that

$$f_k \to f \text{ in } L^p \quad \Rightarrow \quad Tf_k \to Tf \text{ in } (L^{p'})^*.$$

Therefore, $T(L^p)$ contains all its limit points, and hence is closed in $(L^{p'})^*$.

Since $p \in]1, 2]$, then $p' \in [2, \infty[$ and we know, thanks to the Clarkson inequality and the Milman-Pettis Theorem 3.5.2, that $L^{p'}$ is reflexive. Therefore, by Theorem 3.5.1, $(L^{p'})^*$ is reflexive.

Consequently, by Theorem 5.1.3, each bounded sequence in $(L^{p'})^*$ contains a subsequence which converges weakly in $(L^{p'})^*$. Since $T(L^p)$ is a closed sub-space of $(L^{p'})^*$ endowed with the norm of $(L^{p'})^*$, thus $T(L^p)$ is a Banach space and, in addition, each bounded sequence in $T(L^p)$ contains a subsequence which converges weakly in $T(L^p)$. Thus, by Theorem 5.1.3, $T(L^p)$ is reflexive, from where L^p is reflexive too. $\square$

D.2.2 Dual space

Theorem D.2.2 ***(Riesz)*** *Let $p \in]1, \infty[$ and $\phi \in (L^p)^*$. There exists a unique $u \in L^{p'}$ such that*

$$\forall f \in L^p, \quad \langle \phi, f \rangle = \int uf.$$

In addition

$$\|u\|_{L^{p'}} = \|\phi\|_{L^{p*}}$$

Proof. For $u \in L^{p'}$, we define $T : L^{p'} \to (L^p)^*$ by

$$\forall f \in L^p \quad \langle Tu, f \rangle = \int uf.$$

As in Corollary D.2.1 (replace p by p' and p' by p), we have that T is a linear continuous operator which is also isometric:

$$\forall u \in L^{p'} \quad \|Tu\|_{(L^p)^*} = \|u\|_{L^{p'}}.$$

We now prove the surjectivity of T.

Define $E = T(L^{p'})$, it is closed since $L^{p'}$ is complete and

$$\|Tu\|_{(L^p)^*} = \|u\|_{L^{p'}}.$$

Let us prove E is dense in $(L^p)^*$.

We use Corollary 3.4.4 of the Hahn-Banach Theorem.

Suppose the converse, *i.e.*, E is not dense in $(L^p)^*$:

$$\exists h \in (L^p)^{**} \quad h \neq 0 \text{ such that } \langle h, Tu \rangle = 0 \quad \forall (Tu) \in E.$$

Since L^p is reflexive, then $h \in L^p$. Let $\langle Tu, h\rangle = 0$ for all u in $L^{p'}$ (it is equivalent to $\langle h, Tu\rangle = 0$ for all $(Tu) \in E$). Therefore

$$\forall u \in L^{p'} \quad \int uh = 0.$$

For $u = |h|^{p-2}h$, we obtain

$$\int uh = \|h\|_{L^p} = 0.$$

Hence $h = 0$. This is a contradiction with the assumption that $h \neq 0$. Therefore, E is dense in $(L^p)^*$ and at the same time E is closed, *i.e.*

$$E = \overline{E} = (L^p)^*.$$

Therefore, T is a linear operator which is an isometry surjective of $L^{p'}$ to $(L^p)^*$. Thus, it is a bijection and $(L^p)^*$ can be identified with $L^{p'}$. □

Corollary D.2.2 $(L^p)^*$ *can be identified with* $L^{p'}$.

D.2.3 Separability

Theorem D.2.3 *Let* $p \in [1, \infty[$ *and* Ω *be a domain in* $\mathbb{R}^n$. $L^p(\Omega)$ *is separable.*

Proof. Let $R = (R_i)_{i\in I}$ be the (countable) family of multidimensional intervals with rational ends (an element of R is $\prod_{k=1}^{n}]a_k, b_k[$ with a_k and b_k in $\mathbb{Q}$) included in Ω.

Let E be the vector space over $\mathbb{Q}$ generated as a span of characteristic functions associated with elements of R. It means that E contains all finite linear combinations with rational coefficients of functions $\mathbb{1}_{R_i}$:

$$\text{if } \phi \in E \text{ then } \phi(x) = \sum_{j=1}^{N} \alpha_j \mathbb{1}_{R_{i_j}}$$

for a finite N, rational α_j and $R_{i_j} \in R$.

By its definition, E is countable.

Let us prove that E is dense in $L^p(\Omega)$.

Let fix $f \in L^p$ and a real number $\epsilon > 0$. Since $C_0(\Omega)$ is dense in $L^p(\Omega)$, there exists a continuous function g such that

$$\|f - g\|_{L^p} \leq \frac{\epsilon}{2}.$$

Let Ω' be an open bounded set such that

$$\operatorname{supp} g \subset \Omega' \subset \Omega.$$

As $g \in C_0(\Omega')$ and Ω' is a bounded open subset of Ω, we can build $\phi \in E$ such that

$$|g(x) - \phi(x)| \leq \frac{\epsilon}{|\Omega'|^{\frac{1}{p}}} \quad a.e. \text{ in } \Omega',$$

where by $|\Omega'|$ we denote the volume (measure) of Ω', which is finite.

Indeed, by definition of functions in E, $\phi(x)$ has the form of $\sum_{j=1}^{N} \alpha_j \mathbb{1}_{R_j}$. Let us take a cover of $\operatorname{supp} g$ by a disjoint finite union of R_j, $\coprod_{j=1}^{N} R_j$ such that the oscillation of g is smaller than $\frac{\epsilon}{|\Omega'|^{\frac{1}{p}}}$:

$$\forall (x, y) \in \left(\coprod_{j=1}^{N} R_j\right)^2 \quad |g(x) - g(y)| < \frac{\epsilon}{|\Omega'|^{\frac{1}{p}}}.$$

For instance, we chose $\alpha_j = g(x_j)$ for a $x_j \in R_j$. Let us notice, that the construction of ϕ is quite similar to the construction of a Riemann sum for a continuous bounded function on a compact domain.

Then

$$\|g - \phi\|_{L^p} \leq \frac{\epsilon}{2}.$$

Thus we find

$$\|f - \phi\|_{L^p} \leq \|f - g\|_{L^p} + \|g - \phi\|_{L^p} \leq \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon. \quad \square$$

Remark D.2.2 *The proof fails when $p = \infty$ due to the lack of density of C_0 in L^∞. But it does not prove L^∞ is not separable!*

Remark D.2.3 *When $p = 2$. For f and g in L^2 note $\langle f, g\rangle = \int fg$*

Equipped with this inner product, L^2 is a Hilbert space.

As expected $(L^2)^* = L^2$.

D.3 Study of L^1

D.3.1 Separability

From Theorem D.2.3 we directly have

Theorem D.3.1 *L^1 is separable.*

D.3.2 Dual space

We recall that Ω is a domain in $\mathbb{R}^n$.

Theorem D.3.2 *Let $\phi \in (L^1)^*(\Omega)$. Then there exists a unique $u \in L^\infty(\Omega)$ such that*

$$\forall f \in L^1(\Omega), \quad \langle \phi, f\rangle = \int_\Omega fu.$$

We further have

$$\|u\|_{L^\infty} = \|\phi\|_{(L^1)^*}.$$

Subsequently, we can identify $(L^1)^*$ and L^∞.

Proof. Let $\varphi \in (L^1(\Omega))^*$.

1. <u>Existence of u:</u>
   Let's fix $w \in L^2(\Omega)$ such that for all compact $K \subset \Omega$, there exists $\alpha_K > 0$ such that

   $$w > \alpha_K \quad \text{a.e. on } K.$$

   Such a function exists, take for instance:

   $$w(x) = \frac{1}{1 + \|x\|^{\frac{(n+1)}{2}}},$$

where $\|x\|$ is the Euclidean norm in $\mathbb{R}^n$.

The mapping

$$f \in L^2(\Omega) \mapsto \langle \phi, wf \rangle \in \mathbb{R}$$

is a linear continuous functional on L^2. (Since, w and f are in L^2, thus by the Cauchy-Schwartz inequality (or the Hölder inequality) $wf \in L^1$. Therefore, as $\varphi \in (L^1(\Omega))^*$, $\phi(wf) = \langle \phi, wf \rangle$ is well defined).

Let us apply Theorem D.2.2 of Riesz for $p = 2$: there exists unique $v \in L^2(\Omega)$ such that

$$\forall f \in L^2(\Omega) \quad \int_\Omega vf = \langle \varphi, fw \rangle. \tag{D.7}$$

We note

$$\forall x \in \Omega \quad u(x) = \frac{v(x)}{w(x)}.$$

The definition of u is correct since $w(x) > 0$ for all $x \in \Omega$ and, as in addition v and w are in L^2, thus u is measurable (as a product of two measurable functions v and $\frac{1}{w}$).

2. <u>$u \in L^\infty$:</u>

From (D.7) and the fact that $fw \in L^1$, comes

$$\forall f \in L^2(\Omega) \quad \left| \int_\Omega vf \right| \leq \|\varphi\|_{(L^1)^*} \|fw\|_{L^1}. \tag{D.8}$$

Let

$$C > \|\varphi\|_{(L^1)^*}. \tag{D.9}$$

Note

$$A = \{x \in \Omega | \; |u(x)| > C\}.$$

Let us show that <u>A is a null set</u> (of Lebesgue measure $\mu(A) = 0$). This will prove that $u \in L^\infty$: $|u(x)| \leq C$ for a.e. $x \in \Omega$ implies that $u \in L^\infty$.

Suppose the converse: let A not be a null set.

There exists B, a subset of A, which is not a null set and has a finite measure $\mu(B) > 0$. Consider f defined by:

$$f(x) = \begin{cases} \frac{u(x)}{|u(x)|}, & \text{if } x \in B \\ 0, & \text{if } x \in \Omega \setminus B \end{cases}.$$

Thus, by definition of u, $v = uw$ and

$$vf = uwf = \begin{cases} \frac{u^2 w}{|u|} = |u|w, & \text{if } x \in B \\ 0, & \text{if } x \in \Omega \setminus B \end{cases}. \tag{D.10}$$

Therefore, since $B \subset A$ and for all $x \in A$ $|u(x)| > C$, we have

$$\int_\Omega vf d\mu = \int_B |u| w d\mu > C \int_B w d\mu. \tag{D.11}$$

Since B is of finite measure, then B is bounded and $\overline{B}$ is compact. By our construction, $w > 0$ a.e. in B, hence $|w| = w$ a.e. in B and

$$\int_B w d\mu = \|w\|_{L^1(B)} > 0.$$

We further find, using the definition of f, that

$$\|fw\|_{L^1(\Omega)} = \int_B \left|\frac{u}{|u|}w\right| d\mu = \int_B |w| d\mu = \int_B w d\mu = \|w\|_{L^1(B)}. \tag{D.12}$$

Since $vf \geq 0$ in Ω by (D.10), then $\int_\Omega vf d\mu \geq 0$, and finally,

$$|\int_\Omega vf d\mu| = \int_\Omega vf d\mu = \int_\Omega |vf| d\mu = \|vf\|_{L^1(\Omega)}.$$

Putting (D.11) and (D.12 in (D.7) we obtain

$$C\|w\|_{L^1(B)} \leq \|vf\|_{L^1(\Omega)} = |\int_\Omega vf d\mu| \leq \|\varphi\|_{(L^1)^*(\Omega)}\|fw\|_{L^1(\Omega)} = \|\varphi\|_{(L^1)^*}\|w\|_{L^1(B)}.$$

Therefore $C \leq \|\varphi\|_{(L^1)^*}$. It is in contradiction with (D.9). Consequently, A is null set and thus $u \in L^\infty$.

3. <u>$\|u\|_{L^\infty(\Omega)} \leq \|\varphi\|_{(L^1)^*(\Omega)}$:</u>
   For all $C > \|\varphi\|_{(L^1)^*(\Omega)}$ we have

   $$|u(x)| \leq C \text{ for a.e. } x \in \Omega.$$

   Thus

   $$\forall \epsilon > 0, \quad \|u\|_{L^\infty(\Omega)} \leq \|\varphi\|_{(L^1)^*(\Omega)} + \epsilon$$

   Therefore

   $$\|u\||_{L^\infty(\Omega)} \leq \|\varphi\|_{(L^1)^*(\Omega)}.$$

4. <u>$\|u\|_{L^\infty(\Omega)} \geq \|\varphi\|_{(L^1)^*(\Omega)}$:</u>
   From (D.7) and $v = uw$, we have

   $$\forall f \in L^2(\Omega), \quad <\varphi, fw>= \int_\Omega uwf d\mu. \tag{D.13}$$

   For any $g \in C_0(\Omega)$, $f = \frac{g}{w}$ is in L^2, since $w \geq \epsilon > 0$ on $\operatorname{supp} g$ and g is bounded in Ω. Therefore, for $f = \frac{g}{w}$, from (D.13) we obtain

   $$\forall g \in C_0(\Omega), \quad <\varphi, g>= \int_\Omega ug d\mu. \tag{D.14}$$

   But $C_0(\Omega)$ is dense in $L^1(\Omega)$, thus

   $$\forall g \in L^1(\Omega) \quad \langle\varphi, g\rangle = \int_\Omega ug d\mu.$$

   Hence

   $$\forall g \in L^1(\Omega) \quad |\langle\varphi, g\rangle| \leq \int_\Omega |ug| d\mu \leq \|u\|_{L^\infty(\Omega)}\|g\|_{L^1(\Omega)},$$

   and therefore

   $$\|\varphi\|_{(L^1)^*(\Omega)} \leq \|u\|_{L^\infty(\Omega)}.$$

5. <u>$\|u\|_{L^\infty(\Omega)} = \|\varphi\|_{(L^1)^*(\Omega)}$:</u> it follows from the points 3) and 4).

6. Uniqueness of u:

   Suppose the converse. Let u_1 and u_2 in $L^\infty(\Omega)$ be such that:

$$\forall f \in L^1(\Omega), \quad \langle \varphi, f \rangle = \int_\Omega f u_1 d\mu \quad \text{and} \quad \langle \varphi, f \rangle = \int_\Omega f u_2 d\mu,$$

   from where we have that

$$\forall f \in L^1(\Omega) \quad \int_\Omega f(u_1 - u_2) d\mu = 0.$$

   As $z = u_1 - u_2 \in L^\infty(\Omega)$ (u_1 and u_2 are in $L^\infty(\Omega)$ and $L^\infty(\Omega)$ is a vector space), therefore $z \in L^1_{loc}(\Omega)$: for all compact subset K in Ω

$$\|z\|_{L^1(K)} = \int_K |z| d\mu \leq \mu(K) \|z\|_{L^\infty(\Omega)} < +\infty.$$

   As $C_0(\Omega)$ is a subset (even dense) in $L^1(\Omega)$

$$\forall f \in L^1(\Omega) \quad \int_\Omega z f d\mu = 0 \quad \Rightarrow \quad \forall g \in C_0(\Omega) \quad \int z g d\mu = 0.$$

   Then, by Lemma 7.1.3, $z = 0$ a.e. on Ω. $\square$

D.3.3 Reflexivity

Proposition D.3.1 *Let Ω be a domain in $\mathbb{R}^n$. $L^1(\Omega)$ is not reflexive.*

Proof.

Let x be in the interior of Ω, *i.e.*, there exists U an open neigborhood of x such that $U \subset \Omega$. Moreover, since $\Omega \subset \mathbb{R}^n$ which is a metric and a normed space, it is equivalent to the existance of $N \in \mathbb{N}^*$ such that $B_{\frac{1}{N}}(x) \subset \Omega$, where $B_{\frac{1}{N}}(x)$ is a ball centered in x with the radius $\frac{1}{N}$.

For $m \geq N$, let

$$f_m = \frac{\mathbb{1}_{B_{\frac{1}{m}}(x)}}{|B_{\frac{1}{m}}(x)|},$$

where by $|B|$ is denoted its volume. We have

$$\|f_m\|_{L^1(\Omega)} = 1.$$

Suppose L^1 is reflexive. Then the unit ball is compact for the weak topology $\sigma(L^1, L^\infty)$ (see Kakutani Theorem 5.1.4). Thus there exists a subsequence of (f_m), denoted by (f_{m_k}), and a function $f \in L^1(\Omega)$ such that (f_{m_k}) converges to f weakly in L^1 (see also Theorem 5.1.3). It means that

$$\forall \varphi \in (L^1)^* = L^\infty, \quad \int_\Omega f_{m_k} \varphi d\mu \to \int_\Omega f \varphi d\mu. \tag{D.15}$$

For $\varphi \in C_0(\Omega \setminus \{x\})$ there exists $k_0 \in \mathbb{N}^*$ such that

$$\forall k \geq k_0, \quad \int_\Omega f_{m_k} \varphi d\mu = 0.$$

Hence

$$\forall \varphi \in C_0(\Omega \setminus \{x\}) \quad \int_\Omega f \varphi d\mu = 0.$$

A point $\{x\}$ has the Lebesgue measure equal to zero, thus

$$\int_\Omega f\varphi d\mu = \int_{\Omega\setminus\{x\}} f\varphi d\mu = 0.$$

Applying Lemma 7.1.3 for an open set $\Omega\setminus\{x\}$ ($f \in L^1(\Omega)$ implies that $f \in L^1_{loc}(\Omega)$), we obtain that $f = 0$ a.e. on $\Omega\setminus\{x\}$ and thus $f = 0$ a.e. on Ω. If we take $\varphi \equiv 1$ in (D.15), then

$$\int_\Omega f_{m_k} d\mu = 1 \to \int_\Omega f d\mu = 0,$$

which is a contradiction.

Therefore "L^1 is reflexive" is false. Hence L^1 is not reflexive. □

D.4 Study of L^∞

D.4.1 Reflexivity

Proposition D.4.1 *Let Ω be a domain in $\mathbb{R}^n$. $L^\infty(\Omega)$ is not reflexive.*

Proof. If it was, so would be $L^1(\Omega)$. □

D.4.2 Separability

We know from Example 2.2.9 that ℓ^∞ is not separable. Let now prove

Proposition D.4.2 *Let Ω be a domain in $\mathbb{R}^n$. $L^\infty(\Omega)$ is not separable.*

Proof.

As Ω is a domain in $\mathbb{R}^n$, it is an open connected set. For any x in the (open) set Ω we define the distance between x and the boundary of Ω

$$r(x) = d(x, R^n \setminus \Omega).$$

Consequently, (see Fig. D.1) for all $x \in \Omega$ all open balls

$$B_{r(x)}(x) = \{y \in \mathbb{R}^n |\ \|x-y\|_{\mathbb{R}^n} < r(x)\} \subset \Omega.$$

Therefore, let us define

$$\forall x \in \Omega \quad f_x = \mathbb{1}_{B_{r(x)}(x)}.$$

The set $\{f_x,\ x \in \Omega\}$ is not countable subset of $L^\infty(\Omega)$ (since Ω is not countable) such that

$$\forall x \in \Omega \quad \|f_x\|_{L^\infty(\Omega)} = 1,$$

and if $x \neq y$ (x and y in Ω) then the distance between f_x and f_y is equal to 1:

$$\|f_x - f_y\|_{L^\infty(\Omega)} = 1.$$

For all $x \in \Omega$ let's take in $L^\infty(\Omega)$, an open ball centered in f_x of radius $\frac{1}{2}$:

$$B_{\frac{1}{2}}(f_x) = \{h \in L^\infty |\ \|h - f_x\|_{L^\infty(\Omega)} < \frac{1}{2}\}.$$

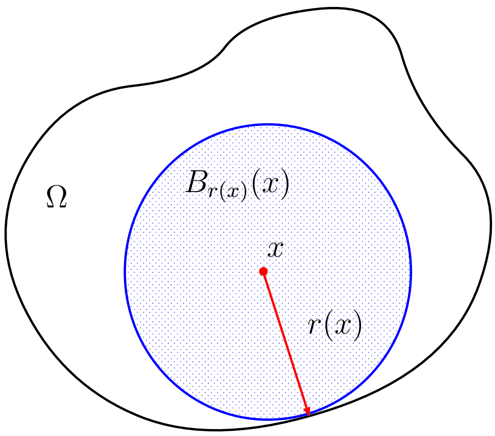

Figure D.1 – Example of $B_{r(x)}(x) \subset \Omega$.

Then for all $x \neq y$ in Ω we have $B(f_x, \frac{1}{2}) \cap B_{\frac{1}{2}}(f_y) = \varnothing$.

Suppose a set E be dense in $L^\infty(\Omega)$. Thus (see Apendix A) for all $x \in \Omega$

$$B_{\frac{1}{2}}(f_x) \cap E \neq \varnothing.$$

Hence, E can not be countable and therefore, $L^\infty(\Omega)$ is not separable. $\square$

D.5 Recap

SPACE	D DENSE	REFLEXIVE	SEPARABLE	DUAL SPACE
p = 1	YES	NO	YES	L^∞
1 < p < ∞	YES	YES	YES	$L^{p'}$
p = ∞	NO	NO	NO	strictly containing L^1

A Hilbert space for p=2

Figure D.2 – Properties of L^p spaces.

Let's summarize the properties of L^p space in Fig. D.2.

Appendix E

General Sobolev embedding theorems

E.1 Sobolev embeddings

Sobolev's embedding theorems give the fundamental properties of the Sobolev spaces.

These properties are connected with the degree of smoothness that can be expected in a Sobolev space.

The larger the product mp ... the smoother the function!

The critical value is the space dimension n. If $mp > n$ then all functions in $W^{m,p}$ are continuous (usual disclaimer: we are dealing with classes).

Proposition E.1.1 *For any domain Ω in $\mathbb{R}^n$ (bounded or not), we have*

$$W^{m,p}(\Omega) \subset L^q(\Omega) \quad \text{for } 1/q = 1/p - m/n \text{ if } mp < n$$

and

$$W^{m,p}(\Omega) \subset L^q(\Omega) \quad \text{for any } q \in [1, \infty[\text{ if } mp = n.$$

We recall that if X and Y are two normed vector spaces, then a linear application $A : X \to Y$ is **compact** if from any bounded sequence (x_k) of elements of X one can extract a subsequence (x_{k_l}) such that its image (Ax_{k_l}) is strongly convergent in Y.

Definition E.1.1 *If there exists a compact linear application from X to Y, we note $X \subset\subset Y$*

Theorem E.1.1 *Let Ω be a bounded domain in $\mathbb{R}^n$ with locally Lipschitz boundary $\partial\Omega$. Then*

1. *$W^{m,p}(\Omega) \subset\subset L^r(\Omega)$ for any $r \in [1, q[$, where $1/q = 1/p - m/n$ provided $mp < n$.*
2. *$W^{m,p}(\Omega) \subset\subset L^q(\Omega)$ for any $q \in [1, \infty[$ provided $mp = n$.*
3. *$W^{m,p}(\Omega) \subset\subset C(\Omega)$ if $mp > n$.*

Example E.1.1 *Let Ω be a a bounded domain of $\mathbb{R}^2$ $(n = 2)$ with a regular boundary and let take $H^1(\Omega) = W^{1,2}(\Omega)$.*

As $1 \times 2 = 2$, therefore $W^{1,2}(\Omega) \subset\subset L^q(\Omega)$ for any $q \in [1, \infty[$.

Thus, if $\|u_k\|_{H^1(\Omega)}$ *is bounded then, one can extract a subsequence in* $L^2(\Omega)$ *that converges strongly in* $L^2(\Omega)$*. The unit ball of* $H^1(\Omega)$ *is compact in* $L^2(\Omega)$ *but not in* $H^1(\Omega)$*.*

We conclude this section by

Theorem E.1.2 *Let* Ω *be a bounded domain in* $\mathbb{R}^n$ *with a locally Lipschitz boundary* $\partial\Omega$*. If* $k > l$*,* $(k-l)p < n$ *and* $1/q = 1/p - (k-l)/n$*, then* $W^{k,p}(\Omega) \subset W^{l,q}(\Omega)$ *and the embedding is continuous.*

Appendix F

Examples for elliptic PDEs

Let us study the $2D$ deformations of an elastic membrane. We will consider two cases of boundary conditions:

- Dirichlet boundary condition: $u|_\Gamma = 0$ (the membrane is attached on Γ, this is what we do in the video lecture),
- Neumann boundary condition: $\frac{\partial u}{\partial n}|_\Gamma = 0$ (then membrane is free on Γ),

where $\frac{\partial u}{\partial n}$ is the normal derivative over the external normal n defined by the scalar product $\nabla u \cdot n$.

F.1 Dirichlet boundary-valued problem for the Poisson equation

Let $\overline{\Omega}$ be a compact connex subset of $\mathbb{R}^n$ with a boundary $\partial\Omega$: $\overline{\Omega} = \Omega \bigsqcup \partial\Omega$. We consider

$$-\Delta u = f(x), \quad x = (x_1, \dots, x_n) \in \Omega, \tag{F.1}$$

$$u|_{\partial\Omega} = 0. \tag{F.2}$$

Let $f \in L^2(\Omega)$. As the trace of u on $\partial\Omega$ is zero, we define the following variational problem in $H_0^1(\Omega)$: find $u \in H_0^1(\Omega)$ such that

$$\forall \phi \in H_0^1(\Omega) \qquad \int_\Omega \nabla u \nabla \phi \mathrm{d}x = \int_\Omega f(x)\phi(x)\mathrm{d}x. \tag{F.3}$$

To obtain the variational problem, if u is regular, for instance $u \in C^2(\Omega) \cap C^1(\overline{\Omega})$, we multiply (F.1) by ϕ, integrate the equality over Ω (by Lebesgue measure) and integrate by parts:

$$-\int_\Omega \Delta u \phi \mathrm{d}x = -\int_{\partial\Omega} \frac{\partial u}{\partial n}\phi \mathrm{d}x + \int_\Omega \nabla u \nabla \phi \mathrm{d}x = \int_\Omega f(x)\phi(x)\mathrm{d}x.$$

Here $\int_{\partial\Omega} \frac{\partial u}{\partial n}\phi \mathrm{d}x = 0$ since $\phi \in H_0^1(\Omega)$ ($\phi \in H_0^1(\Omega)$ implies that $\phi|_{\partial\Omega} = 0$).

Definition F.1.1 *$u \in H_0^1(\Omega)$ is called **weak solution** of problem (F.1)–(F.2), if u is a solution of the variational problem (F.3).*

Remark F.1.1 *We have shown that from (F.1)–(F.2) we obtain the variational problem (F.3). Let's now show that the converse: if $u \in H_0^1(\Omega)$ is a solution of the variational problem (F.3), then u is a solution of (F.1)–(F.2).*

Indeed, since (F.3) is true for all $\phi \in H_0^1(\Omega)$ and $\mathcal{D}(\Omega) \subsetneq H_0^1(\Omega)$, then (F.3) also holds for all $\phi \in \mathcal{D}(\Omega)$. Thus,

$$\forall \phi \in \mathcal{D}(\Omega) \qquad \int_\Omega \nabla u \nabla \phi \mathrm{d}x = \int_\Omega f(x)\phi(x)\mathrm{d}x.$$

Since ∇u and f are in $L^2(\Omega) \subsetneq L^1_{loc}(\Omega)$, we obtain that in sense of distributions

$$\forall \phi \in \mathcal{D}(\Omega) \qquad \int_\Omega (-\Delta u - f)\phi \mathrm{d}x = 0.$$

Thus, by du Bois-Reymond Lemma 7.1.3, we obtain that

$$\Delta u + f = 0 \quad a.e.\ in\ \Omega.$$

On the other hand, since $u \in H_0^1(\Omega)$, it means that its trace on $\partial\Omega$ is equal to zero (see point 6 of Theorem 8.3.1).

We conclude that problem (F.1)–(F.2) and the variational problem (F.3) are equivalent.

Definition F.1.2 *Function $u \in C^2(\Omega) \cap C(\overline{\Omega})$, satisfying (F.1)–(F.2), is called a* ***classical*** *solution of problem (F.1)–(F.2).*

Proposition F.1.1 *Let $u \in C^2(\Omega) \cap C(\overline{\Omega})$ be a weak solution of (F.1)–(F.2). Then u is a classical solution of (F.1)–(F.2).*

Proof. Since $u \in C^2(\Omega) \cap C(\overline{\Omega})$ is a weak solution of (F.1)–(F.2) then $u \in H_0^1(\Omega)$ and it holds (F.3). In particular,(F.3) holds for all $\phi \in \mathcal{D}(\Omega)$. We integrate by parts:

$$\forall \phi \in \mathcal{D}(\Omega) \quad \int_\Omega \Delta u \phi \mathrm{d}x + \int_\Omega f(x)\phi(x)\mathrm{d}x = 0,$$

or equivalently,

$$\forall \phi \in \mathcal{D}(\Omega) \quad \int_\Omega (\Delta u + f)\phi \mathrm{d}x = 0.$$

As, by assumption, $f \in L^2(\Omega)$ and $\Delta u \in C(\Omega)$, then $\Delta u + f \in L^1_{loc}(\Omega)$. Moreover, by du Bois-Reymond Lemma 7.1.3, we obtain that

$$\Delta u + f = 0 \quad \text{a.e. in } \Omega.$$

In other words, $f = -\Delta u \in C(\Omega)$ a.e. in Ω, *i.e.*, f is equivalent to the continuous function. Therefore, for the continuous representative of the class of f this equation (see (F.1)) holds for all $x \in \Omega$:

$$-\Delta u = f \in C(\Omega) \quad \forall x \in \Omega.$$

Since $u \in C(\overline{\Omega}) \cap H_0^1(\Omega)$ (see also point 6 of Theorem 8.3.1), thus we have that the trace of u is defined in the classical sense and it is $u|_{\partial\Omega} = 0$.

We conclude that u is a classical solution of (F.1)–(F.2). $\square$

Let's show the following proposition:

Proposition F.1.2 *Problem (F.1)–(F.2) is well-posed: there exists an unique weak solution $u \in H_0^1(\Omega)$ and for all $f \in L^2(\Omega)$ there exists a constant $C > 0$ such that*

$$\|u\|_{H_0^1(\Omega)} \leq C\|f\|_{L^2(\Omega)}. \tag{F.4}$$

Estimation (F.4) shows the stability of u as a function of f.

Proof In Chapter 8 in the proof of Theorem 8.3.3 it was shown that the following norms on $H_0^1(\Omega)$ are equivalent:

$$\|u\|_{H_0^1(\Omega)} = \sqrt{\int_\Omega (u^2 + |\nabla u|^2)\mathrm{d}x},$$
$$\|u\|_{H_0^1(\Omega)}^{new} = \sqrt{\int_\Omega |\nabla u|^2 \mathrm{d}x} = \|\nabla u\|_{L^2(\Omega)}.$$

Since these norms are associated with the inner products:

$$\|u\|_{H_0^1(\Omega)} = \sqrt{\langle u, u\rangle_{H_0^1(\Omega)}}, \quad (u,\phi)_{H_0^1(\Omega)} = \int_\Omega (u\phi + \nabla u \nabla \phi)\mathrm{d}x, \tag{F.5}$$
$$\|u\|_{H_0^1(\Omega)}^{new} = \sqrt{\{u, u\}_{H_0^1(\Omega)}}, \quad \{u,\phi\}_{H_0^1(\Omega)} = \int_\Omega \nabla u \nabla \phi \mathrm{d}x, \tag{F.6}$$

it implies that $\langle \cdot, \cdot \rangle$ and $\{\cdot, \cdot\}$ are equivalent inner products in $H_0^1(\Omega)$.

Therefore, we will apply the Theorem of Riesz to show the existence of a unique weak solution of (F.3). We write (F.3) in the form:

$$\forall \phi \in H_0^1(\Omega) \quad \{u, \phi\}_{H_0^1(\Omega)} = \langle f, \phi\rangle_{L^2(\Omega)}. \tag{F.7}$$

The space $(H_0^1(\Omega), \|\cdot\|^{new})$ with the norm $\|\cdot\|^{new}$, and with the associated inner product $\{\cdot,\cdot\}$, is a Hilbert space.

For fixed $f \in L^2(\Omega)$, we define on L^2 a linear functional

$$l_f(\phi) = \langle f, \phi\rangle_{L^2(\Omega)} = \int_\Omega f\phi. \tag{F.8}$$

As $f \in L^2(\Omega)$, the linear functional $l_f(\phi)$ is well defined for all $\phi \in L^2(\Omega)$, and consequently, $l_f(\phi)$ is also well defined for $\phi \in H_0^1(\Omega) \subset L^2(\Omega)$. Let us show that l_f is continuous on $(H_0^1(\Omega), \|\cdot\|^{new})$.

Let $\phi_n \in H_0^1(\Omega)$ and $\phi_n \to \phi$ in $(H_0^1(\Omega), \|\cdot\|^{new})$. As the norm $\|\cdot\|^{new}$ is equivalent to the canonical norm $\|\cdot\|$ of $H_0^1(\Omega)$, the sequence ϕ_n also converges by the norm $\|\cdot\|$ of $H_0^1(\Omega)$ and, hence, the sequence ϕ_n also converges by the norm of $L^2(\Omega)$. Thanks to the continuity of the inner product in L^2, it means that

$$\lim_{n\to\infty} l_f(\phi_n) = \lim_{n\to\infty} \langle f, \phi_n\rangle_{L^2(\Omega)} = \langle f, \phi\rangle_{L^2(\Omega)} = l_f(\phi).$$

Applying the Theorem of Riesz in the Hilbert space $(H_0^1(\Omega), \|\cdot\|^{new})$, we conclude that for all $\phi \in H_0^1(\Omega)$ there exists a unique solution $u \in H_0^1(\Omega)$ of equation (F.7).

To finish the proof, we need to show the stability estimation (F.4). As (F.7) holds for all $\phi \in H_0^1(\Omega)$, we are able to take $\phi = u$. Using Cauchy-Schwartz and Poincaré inequalities, we find

$$(\|u\|_{H_0^1(\Omega)}^{new})^2 = \int_\Omega |\nabla u|^2 \mathrm{d}x \le \|f\|_{L^2(\Omega)} \|u\|_{L^2(\Omega)} \le C(\Omega)\|f\|_{L^2(\Omega)} \|u\|_{H_0^1(\Omega)}^{new},$$

from which directly follows (F.4). $\square$

F.2 Robin boundary-valued problem for the Poisson equation

Let $\Omega \subset \mathbb{R}^n$ be a bounded domain with a regular boundary (otherwise Ω admits the extension of elements in $H^1(\Omega)$ to a wider domain keeping the the same regularity properties). We consider the Poisson equation in Ω with the homogeneous Robin boundary condition:

$$-\Delta u = f(x), \quad x = (x_1, \ldots, x_n) \in \Omega, \tag{F.9}$$
$$\frac{\partial u}{\partial \vec{n}} + \sigma(x) u|_{\partial\Omega} = 0, \tag{F.10}$$

where $\vec{n}$ is the external normal defined at all points of $\partial\Omega$. We suppose that $f(x) \in L^2(\Omega)$, and that $\sigma(x) \in C(\partial\Omega)$ is a measurable bounded function, separated from zero

$$\sigma(x) \geq \sigma_0 > 0.$$

Let us show that the corresponding variational problem is to find $u(x) \in H^1(\Omega)$ such that

$$\forall \phi(x) \in H^1(\Omega) \quad \int_\Omega \nabla u \nabla \phi \mathrm{d}x + \int_{\partial\Omega} \sigma u \phi ds = \int_\Omega f(x)\phi(x)\mathrm{d}x. \tag{F.11}$$

In fact, it is a corollary of the formula of the integration by parts using the boundary condition:

$$-\int_\Omega \Delta u \phi \mathrm{d}x = -\int_{\partial\Omega} \frac{\partial u}{\partial n}\phi ds + \int_\Omega \nabla u \nabla \phi \mathrm{d}x = \int_\Omega \nabla u \nabla \phi \mathrm{d}x + \int_{\partial\Omega} \sigma u \phi ds.$$

Definition F.2.1 *$u \in H^1(\Omega)$ is called **weak solution** of problem (F.9)–(F.10), if u is a solution of the variational problem (F.11).*

Problem F.2.1 *Prove that problem (F.1)–(F.2) and the variational problem (F.3) are equivalent.*

Definition F.2.2 *Function $u \in C^2(\Omega) \cap C^1(\overline{\Omega})$, satisfying (F.9)–(F.10), is called a **classical** solution of problem (F.9)–(F.10).*

Proposition F.2.1 *Let $u \in C^2(\Omega) \cap C^1(\overline{\Omega})$ be a weak solution of (F.9)–(F.10). Then u is a classical solution of (F.9)–(F.10).*

Proof. Let's just notice that since Ω is bounded and $u \in C^1(\overline{\Omega})$, it follows that $u \in H^1(\Omega)$. In addition, we have the following inclusions

$$\mathcal{D}(\Omega) \subsetneq H_0^1(\Omega) \subsetneq H^1(\Omega).$$

As u is a weak solution of (F.9)–(F.10), then for all $\phi \in \mathcal{D}(\Omega)$ it holds (F.11). In the same way as for the Dirichlet boundary condition, we integrate by parts:

$$\forall \phi \in \mathcal{D}(\Omega) \quad \int_\Omega (\Delta u + f)\phi \mathrm{d}x = 0.$$

By assumption, $f \in L^2(\Omega)$ and $\Delta u \in C(\Omega)$, then $\Delta u + f \in L^1_{loc}(\Omega)$. Moreover, by du Bois-Reymond Lemma 7.1.3, we obtain that

$$\Delta u + f = 0 \quad \text{a.e. in } \Omega.$$

In other words, $f = -\Delta u \in C(\Omega)$ a.e. in Ω, *i.e.*, f is equivalent to the continuous function. Therefore, for the continuous representative of the class of f this equation (see (F.9)) holds for all $x \in \Omega$:

$$-\Delta u = f \in C(\Omega) \quad \forall x \in \Omega.$$

Let us prove that u satisfies (F.10). Since $\partial\Omega$ is regular, then (see Proposition 8.3.2)

$$C^\infty(\overline{\Omega}) \subsetneq H^1(\Omega).$$

We take $\phi \in C^\infty(\overline{\Omega})$ in (F.11) and integrate by parts:

$$\begin{aligned}
&\forall\phi(x) \in C^\infty(\overline{\Omega}) \quad \int_\Omega \nabla u \nabla\phi \mathrm{d}x + \int_{\partial\Omega} \sigma u\phi ds - \int_\Omega f(x)\phi(x)\mathrm{d}x = 0, \quad \Rightarrow \\
&\forall\phi(x) \in C^\infty(\overline{\Omega}) \quad \int_\Omega (-\Delta u - f)\phi \mathrm{d}x + \int_{\partial\Omega} \left(\frac{\partial u}{\partial \nu} + \sigma u\right)\phi ds = 0.
\end{aligned}$$

By the previously proven (F.9), we find

$$\forall\phi(x) \in C^\infty(\overline{\Omega}) \quad \int_{\partial\Omega} \left(\frac{\partial u}{\partial \nu} + \sigma u\right)\phi ds = 0.$$

Since $\left(\frac{\partial u}{\partial \nu} + \sigma u\right)|_{\partial\Omega} \in C(\partial\Omega)$, and $\mathcal{D}(\partial\Omega) \subsetneq C^\infty(\overline{\Omega})$, thus we can apply du Bois-Reymond Lemma 7.1.3, and obtain that

$$\frac{\partial u}{\partial \nu} + \sigma u\bigg|_{\partial\Omega} = 0.$$

We conclude that u is a classical solution of (F.9)–(F.10). □

To proceed to the existence of a solution of (F.9)-(F.10), we firstly show the following Lemma:

Lemma F.2.1 *The norm*

$$\|u\|_\sigma^2 = \int_\Omega |\nabla u|^2 \mathrm{d}x + \int_{\partial\Omega} \sigma u^2 ds \tag{F.12}$$

is equivalent to the canonical norm $\|\cdot\|$ *of* $H^1(\Omega)$.

Proof. We prove it in two steps:

1. We show that $\|\cdot\|_\sigma$ is equivalent to the norm

$$\|u\|_1^2 = \int_\Omega |\nabla u|^2 \mathrm{d}x + \int_{\partial\Omega} u^2 ds. \tag{F.13}$$

2. Using an analogue of the Poincaré inequality in $H^1(\Omega)$

$$\int_\Omega u^2 \mathrm{d}x \leq C \left(\left|\int_{\partial\Omega} u ds\right|^2 + \int_\Omega |\nabla u|^2 \mathrm{d}x\right) \tag{F.14}$$

   we show the equivalence of $\|\cdot\|_1$ and the canonical norm of $H^1(\Omega)$.

It is easy to see that, since $\sigma(x) \geq \sigma_0 > 0$, if

$$C_1 = \min(1, \sigma_0), \quad C_2 = \max(1, \sup_{x\in\partial\Omega} \sigma(x)),$$

then $C_1\|u\|_1 \leq \|u\|_\sigma \leq C_2\|u\|_1$. Let us show the existence of C_1 and C_2 such that

$$C_1\|u\|_1 \leq \|u\|_{H^1(\Omega)} \leq C_2\|u\|_1.$$

As the boundary $\partial\Omega$ is regular, the embedding of $H^1(\Omega)$ in $L^2(\partial\Omega)$ is bounded:

$$\|u\|_{L^2(\partial\Omega)} \leq C\|u\|_{H^1(\Omega)}.$$

Consequently,

$$\|u\|_1^2 = \|\nabla u\|_{L^2(\partial\Omega)}^2 + \|u\|_{L^2(\partial\Omega)}^2 \leq \|\nabla u\|_{L^2(\partial\Omega)}^2 + C\|u\|_{H^1(\Omega)}^2 \leq \tilde{C}\|u\|_{H^1(\Omega)}^2,$$

and $C_1 = 1/\tilde{C}$.

Using (F.14), we find

$$\begin{aligned}
\|u\|_{H^1(\Omega)}^2 &\leq \|\nabla u\|_{L^2(\Omega)}^2 + C\left(\left|\int_{\partial\Omega} uds\right|^2 + \|\nabla u\|_{L^2(\Omega)}^2\right) \\
&= (1+C)\|\nabla u\|_{L^2(\Omega)}^2 + C\left|\int_{\partial\Omega} 1\cdot uds\right|^2 \\
&\leq (1+C)\|\nabla u\|_{L^2(\Omega)}^2 + M\,\mathrm{mes}(\partial\Omega)\int_{\partial\Omega} u^2 ds \leq C_2\|u\|_1^2,
\end{aligned}$$

where $C_2 = \max(M\,\mathrm{mes}(\partial\Omega), 1+M)$. □

Proposition F.2.2 *Problem (F.9)–(F.10) is well-posed: there exists an unique weak solution $u \in H^1(\Omega)$ and for all $f \in L^2(\Omega)$ there exists a constant $C > 0$ such that*

$$\int_\Omega |\nabla u|^2 \mathrm{d}x + \int_{\partial\Omega} \sigma u^2 ds \leq C\int_\Omega |f|^2 dx. \tag{F.15}$$

Estimation (F.4) shows the stability of u as a function of f.

Proof. Using the result of the previous Lemma F.2.1, we apply (see the Poisson problem with Dirichlet boundary conditions) the Riesz theorem in $(H^1(\Omega), \|\cdot\|_\sigma)$ and obtain the existence of the unique weak solution of (F.11) which is also stable (F.4). □

Bibliography

[1] R. A. Adams and J. J. F. Fournier. *Sobolev spaces*. Academic Press, 2003.

[2] H. Brezis. *Analyse fonctionnelle : Théorie et applications*. Sciences SUP, 2005.

[3] L. C. Evans. *Partial Differential Equations*. Graduate Studies in Mathematics, 1994.

[4] L. Gabet. *Analyse*. Ecole Centrale Paris (Polycopié du cours de la prémière année), 2014.

[5] F. Golse, Y. Laszlo, F. Pacard, and C. Viterbo. *Analyse réelle et complexe*. École Polytechnique, 2014.

[6] E. Hewitt and K. Stromberg. *Real and abstract analysis*. Springer, 1965.

[7] A.N. Kolmogorov and S.V. Fomin. *Introductory Real Analysis*. Dover publications, INC., 1975.

[8] P. Lafitte. *Analyse théorique et numérique des équations aux dérivées partielles*. École Centrale Paris, Polycopié 1ère année, 2014.

[9] G. Teschl. *Functional Analysis*. Universität Wien, Austria, 2006.

[10] V.S. Vladimirov. *Equations of Mathematical Physics*. Pure and Applied Mathematics, 1971.

Index

K

L

M

N

O

P

R

S

T

U

W