People's Democratic Republic of Algeria

Ministry of Higher Education and Scientific Research

University: **Abdelhafid Boussouf. Mila**

Institute of **Mathematics and Computer Scences**

Departement of **MathematicsS**

Academic division: Mathematics

S5

NORMED VECTOR SPACES

A course given by: Rakia Ahmed-Yahia.

CONTENTS

Introduction **3**

1 BANACH SPACES **5**

1.1 Normed vector spaces 5

1.1.1 Norm 5

1.1.2 Norm proprieties 6

1.1.3 Some common examples 9

1.1.4 Equivalent norms 16

1.2 Banach spaces 17

1.2.1 Cauchy sequences 17

1.3 Normed finite-dimensional vector spaces 20

1.3.1 Equivalence of norms 20

1.3.2 Compactness of the balls 23

1.4 Linear mappings 25

1.5 Dual of vector normed space 28

2 HILBERT SPACES **30**

2.1 Inner product 30

2.1.1 Definitions . . . 30
2.2 Elementary properties . . . 33
2.2.1 Cauchy-Schwarz inequality . . . 33
2.2.2 Parallelogram equality . . . 35
2.2.3 Orthogonality . . . 36
2.2.4 Hilbert spaces . . . 37
2.3 The Projection Theorem and its consequences . . . 38
2.3.1 Projection Theorem . . . 38
2.3.2 Consequences . . . 40
2.3.3 Representation of the dual . . . 46
2.3.4 Adjoint of an operator . . . 47
2.4 Orthonormal bases . . . 48
2.4.1 Separable spaces . . . 48
2.4.2 Orthonormal systems . . . 50
2.4.3 Orthonormal bases . . . 52
2.4.4 Existence of orthonormal bases . . . 54
2.5 Separability of $L^2(0,1)$ . . . 55
2.5.1 Stone-Weierstrass theorem . . . 55
2.5.2 Complex case . . . 64
2.5.3 The trigonometric system . . . 65

INTRODUCTION

Functional Analysis studies the functions, as elements of certain spaces, called functional. It has been developed based mainly on the study of normed vector spaces; and, more particularly, complete normed vector spaces (Banach spaces). The benefit of having complete normed spaces is that we have properties of existence within them: for any sequence of Cauchy, there is an element which is its limit. This course is aimed to 3rd year mathematics students, during semester 5. It requires a good knowledge of general Topology. Having some knowledge of complex variable functions, integrals, series, linear algebra,..etc. will also be useful.

CHAPTER 1

BANACH SPACES

In this chapter, we will give some generalities about abstract normed spaces, with examples, and we will treat the case of finite-dimensional spaces. Some topological notions are added in order to simplify comprehension and resolution of exercises in fifth and sixth semesters. This will also be an opportunity to set certain notations.

1.1 Normed vector spaces

1.1.1 Norm

Definition 1.1.1. *Let E be a real or complex vector space. A norm on E is an application, most often denoted $\|.\|$:*

$$\|\cdot\| : E \longrightarrow \mathbb{R}_+ = [0, +\infty[$$

having the following three properties:

1. *a)* $\|x\| \geq 0$ *for all* $x \in E$ *and b)* $\boxed{\|x\| = 0 \Longleftrightarrow x = 0}$;
2. $\boxed{\|\lambda x\| = |\lambda| \|x\|}$, $\forall x \in E, \forall \lambda \in \mathbb{K}$ *(homogeneity);*
3. $\|x + y\| \leqslant \|x\| + \|y\|, \forall x, y \in E$ *(triangular inequality).*

If we delete 1) b), we say that $\|.\|$ is a semi-norm. Note that then 2) nevertheless results in $\|0\| = 0$.

1.1.2 Norm proprieties

Proposition 1.1.1. *The function $x \in E \mapsto \|x\| \in \mathbb{R}_+$ is continuous*

Proof. Just use inequality $|\,\|x\| - \|y\|\,| \leq \|x - y\|$. □

From a norm, we obtain a distance on E by setting $\boxed{d(x, y) = \|x - y\|}$.

We then define the :

- Open balls : $\overset{\circ}{B}(x, r) = \{y \in E; \|x - y\| < r\}$;
- Closed balls : $B(x, r) = \{y \in E; \|x - y\| \leqslant r\}$,

which makes it possible to define a topology on E; a part A of E is open (and we also say that A is an open set of E) if for all $x \in A$ there exists a ball centered at x, of radius $r = r_x > 0$, contained in A. There is no need to specify whether it is an open ball or a closed ball. Indeed, if A contains the closed ball $B(x, r)$, it contains a fortiori the open ball $\overset{\circ}{B}(x, r)$; and, conversely, if A contains the open ball $\overset{\circ}{B}(x, r)$, it contains the closed ball $B(x, r')$, for all $r' < r$. Note that the empty set $\varnothing$ is an open set (since there is no x in A, the property defining open set is trivially verified). The entire space E is clearly an open space. It follows from the definition that any union of open parts is an open set. Any intersection of a finite number of open sets is an open one. If $x \in A = A_1 \cap \cdots \cap A_n$, and $\overset{0}{B}(x, r_k) \subseteq A_k$, then $\overset{\circ}{B}(x, r) \subseteq A$, with $r = \min(r_1, \ldots, r_n)$

A part V containing the point $x_0 \in E$ is a neighborhood of x_0 if it contains a ball (open or closed) with center x_0, and with radius r > 0.

A part is closed (we also say that it is closed set) if its complement is open. By complementarity, we obtain that $\varnothing$ and E are closed, that the intersection of any family of closed sets is still closed one, as well as any union of a finite number of closed sets. If $A \subseteq E$ is a part of E, we call interior of A, and we write $\overset{\circ}{A}$, or $int(A)$, the largest open set contained in

A (it is the union of all the open sets contained in A), and we call closure, of A the smallest closed set containing A (it is the intersection of all closed sets containing A). We denote by $\bar{A}$ the closure of A. We recall (it's easy to see) that $x \in \bar{A}$ if, and only if, there exists a sequence of elements of A converging to x. We say that A is dense in E if $\bar{A} = E$.

Proposition 1.1.2. *Any open ball is an open set and any closed ball is a closed set.*

Proof. 1) Let $x \in \overset{\circ}{B}(x_0, r_0)$ and let $0 < r < r_0 - \|x - x_0\| > 0$. For $\|x - y\| \leqslant r$, we have $\|y - x_0\| \leq \|y - x\| + \|x - x_0\| \leq r + \|x - x_0\| < r_0$; therefore $B(x, r) \subseteq \overset{\circ}{B}(x_0, r_0)$.
2) Let $x \notin B(x_0, r_0)$ and let $0 < r < \|x - x_0\| - r_0$. Since, if $\|y - x\| \leq r$, we have $\|y - x_0\| \geq \|x_0 - x\| - \|x - y\| \geq \|x_0 - x\| - r > r_0$, then $B(x, r) \subseteq [B(x_0, r_0)]^c$. □

Remarks.

(i) Closed subsets are important while studying the solution of equation, where one looks for approximate solutions by constructing sequences of approximations, of all which belong to a set Y of functions with certain properties. If Y is a closed set and if the sequence is convergent, the limit also belongs to Y, giving a convergent sequence of approximations in the solution set Y.

(ii) It is clear that $Y \subset \bar{Y}$, and $Y = \bar{Y}$ if and only if Y is closed.

All the preceding topological notions do not involve the fact that E is a vector space, nor that the distance is defined from a norm; they are therefore valid in any metric space. On the other hand, we have a specific property in normed spaces, which justifies the notation of open balls: the interior of $B(r, r)$ is the open ball $\overset{\circ}{B}(x, r)$ and the closure of the open ball $\overset{\circ}{B}(x, r)$ is the closed ball $\overline{B(x, r)}$ (see below).

Definition 1.1.2. *When a vector space E is endowed with a norm and the topology associated with this norm, ws say that it is a normed vector space, or, more simply, a normed space.*

Notation. We will denote by B_E the closed ball $B(0, 1)$ with center 0 and radius 1.
We will say that it is the **unit ball** of E.

Proposition 1.1.3. *If E is a normed space, then the mapping:*

$$\begin{aligned} +: \quad & E \times E \to E & \quad \textit{and} \quad & \mathbb{K} \times E \to E \\ & (x, y) \mapsto x + y & & (\lambda, x) \mapsto \lambda x \end{aligned}$$

are continuous.

Definition 1.1.3. *Let E be a real or complex vector space, provided with a topology. We say that E is a topological vector space (t.v.s) if the maps:*

$$\begin{aligned} +: \quad & E \times E \to E & \quad \textit{and} \quad & \mathbb{K} \times E \to E \\ & (x, y) \mapsto x + y & & (\lambda, x) \mapsto \lambda x \end{aligned}$$

are continuous.

We say that a topological vector space is locally convex space (l.c.s), if every point has a base of convex neighborhoods.

Balls are convex, and any normed space is a (t.v.s) locally convex.

Corollary 1.1.1. *The translations:*

$$\begin{aligned} \tau_a : \quad & E \longrightarrow E & \quad (a \in E) \\ & x \longmapsto x + a \end{aligned}$$

and the dilations:

$$\begin{aligned} h_\lambda : \quad & E \longrightarrow E & \quad (\lambda \in \mathbb{K}) \\ & x \longmapsto \lambda x \end{aligned}$$

are continuous. These are homeomorphisms (if $\lambda \neq 0$ for dilations).

Corollary 1.1.2. *All closed balls of radius $r > 0$ are homeomorphic to each other, therefore to B_E. All open balls of radius $r > 0$ are homeomorphic to each other.*

Corollary 1.1.3. *The closure of the open ball* $\overset{\circ}{B}(x,r)$ *is the closed ball* $B(x,r)$ *and the interior of the closed ball* $B(x,r)$ *is the open ball* $\overset{\circ}{B}(x,r)$ *.*

Proof. 1) The closure of the open ball is obviously contained in the closed ball, since the latter is closed in E. Conversely, if $y \in B(x,r)$, we have
$y_n = \frac{1}{n}x + \left(1-\frac{1}{n}\right)y \in \overset{\circ}{B}(x,r)$ car $\left\|x - \left[\frac{1}{n}x + \left(1-\frac{1}{n}\right)y\right]\right\| = \left(1-\frac{1}{n}\right)\|x-y\| < r$; as $y = \lim_{n\to\infty} y_n$, we obtain $y \in \overset{\circ}{B}(x,r)$.

2) Being open in E, the open ball is contained in the interior of the closed ball. To show the reverse inclusion, show that if y is not in the open ball, then no ball $B(y,\rho)$ of center y and radius $\rho > 0$ is contained in $B(x,r)$. But if y is not in $\overset{\circ}{B}(x,r)$, then we have $\|y-x\| \geq r$. For all $\rho > 0$, the vector $z = y + \frac{\rho}{\|y-x\|}(y-x)$ is in $B(y,\rho)$, since $\|z-y\| = \frac{\rho}{\|y-x\|}\|y-x\| = \rho$, but is not in $B(x,r)$, because $\|z-x\| = \left\|y + \frac{\rho}{\|y-x\|}(y-x) - x\right| = \left(1+\frac{\rho}{\|y-x\|}\right)\|y-x\| \geq \left(1+\frac{\rho}{|y-x|}\right)r > r$. So y is not in the interior of $B(x,r)$. □

Corollary 1.1.4. *If F is a vector subspace of E, then its closure $\bar{F}$ is a vector subspace too.*

Proof. Let be $x, y \in \bar{F}$ and $a, b \in \mathbb{K}$. There exist $x_n, y_n \in F$ such that $x_n \underset{n\to\infty}{\longrightarrow} x$ and $y_n \underset{n\to\infty}{\longrightarrow} y$. By Proposition 1.4.1, we have $ax+by = \lim_{n\to\infty}(ax_n + by_n)$; and as $ax_n + by_n \in F$, we obtain $ax+by \in \bar{F}$. □

1.1.3 Some common examples

Spaces of sequences

1) a) It is immediate to see that if we put, for $x = (x_1, \ldots, n) \in \mathbb{K}^n$:

$$\begin{cases} \|x\|_1 = |x_1| + \cdots + |x_n| \\ \|x\|_\infty = \max\{|x_1|, \ldots, |x_n|\} \end{cases}$$

Then $\|.\|_1$ and $\|.\|_\infty$ are two norms on $\mathbb{K}^n$.

We note $\ell_1^n = (K^n, |.|_1)$ and $\ell_\infty^n = (K^n, |.|_\infty)$.

b) If p is an real number that satifies $1 < p < \infty$, a norm on $\mathbb{R}^n$ is obtained when we put:

$$\boxed{\|x\|_p = \left(\sum_{k=1}^{n} |x_k|^p\right)^{1/p}}$$

We note $\ell_p^n = (K^n, |.|_p)$. Only triangular inequality:

$$\boxed{\left(\sum_{k=1}^{n} |x_k + y_k|^p\right)^{1/p} \leqslant \left(\sum_{k=1}^{n} |x_k|^p\right)^{1/p} + \left(\sum_{k=1}^{n} |y_k|^p\right)^{1/p},}$$

called **Minkowski** inequality, is not obvious; it can be demonstrated as follows: By convexity of the function $t \in \mathbf{R}_+ \mapsto t^p$, we have $[\alpha u + (1-\alpha)v]^p \leqslant \alpha u^p + (1-\alpha)v^p$ if $0 \leqslant \alpha \leqslant 1$ and $u, v \geqslant 0$. Take $\alpha = \frac{\|x\|_p}{\|x\|_p + \|y\|_p}$ (such that $1 - \alpha = \frac{\|y\|_p}{\|x\|_p + \|y\|_p}$),$u = \frac{|x_k|}{\|x\|_p}$ and $v = \frac{|y_k|}{\|y\|_p}$ (if $\|x\|_p = 0$ or $\|y\|_p = 0$, the result is abvious). By summing, we get

$$\frac{1}{(\|x\|_p + \|y\|_p)^p} \sum_{k=1}^{n} (|x_k| + |y_k|)^p \leqslant 1,$$

which gives the result, since $|x_k + y_k| \leqslant |x_k| + |y_k|$ for all $k = 1, \ldots, n$.

A very useful inequality is the **Hölder's inequality** . Recall that if $1 < p < \infty$, **the conjugate exponent** of p is the number q satisfying $\boxed{\frac{1}{p} + \frac{1}{q} = 1}$. Explicitly, $\boxed{q = \frac{p}{p-1}}$. We have $1 < q < \infty$, and p is the conjugate exponent of q. They are also linked by the equality $\boxed{(p-1)(q-1) = 1}$. Hölder's inequality is then stated as follows if $1 < p < \infty$ and q is the conjugate exponent of p, then, for all $x_1, ..., x_n, y_1, ..., y_n \in \mathbb{K}$, we have:

$$\boxed{\sum_{k=1}^{n} |x_k y_k| \leqslant \left(\sum_{k=1}^{n} |x_k|^p\right)^{1/p} \left(\sum_{k=1}^{n} |y_k|^q\right)^{1/q}}$$

If $p = 2$, then $q = 2$: this is **the Cauchy-Schwarz inequality** (due, in this form, to Cauchy in 1821).

To show Hölder's inequality, we start from the inequality $ab \leq \frac{a^p}{p} + \frac{b^q}{q}$, for $a, b \geq 0$ (this is a consequence of the convexity of the function $t \in \mathbb{R}_+ \mapsto \frac{t^p}{p}$ and of the fact that its derivative $t \mapsto t^{P-1}$ is the reciprocal (or inverse) of the derivative $t \mapsto t^{q-1}$ of $t \mapsto \frac{t^q}{q}$, as we can see it just as simply, for example by studying the variations of the function $t \mapsto \frac{t^p}{p} + \frac{b^q}{q} - bt$); we apply it with $a = \frac{|x_k|}{\|x\|_p}$ and $b = \frac{|y_k|}{\|y\|_q}$ (we can assume $\|x\|_p > 0$ and $\|y\|_q > 0$), and we add up. We obtain $\frac{1}{\|x\|_p \|y\|_q} \sum_{k=1}^n |x_k y_k| \leq \frac{1}{p} + \frac{1}{q} = 1$, hence Hölder's inequality.

2) These examples generalize to infinite dimension.

a)Let:

$$\boxed{c_0 = \left\{ x = (x_n)_{n \geqslant 1} \in \mathbb{K}^{N^*} ; \lim_{n \to \infty} x_n = 0 \right\}},$$

and:

$$\boxed{\ell_\infty = \left\{ x = (x_n)_{n \geqslant 1} \in \mathbb{K}^{\mathbf{N}^*} ; (x_n)_n \text{ is bounded } \right\}};$$

we provide them with the norm defined by:

$$\boxed{\|x\|_\infty = \sup_{n \geq 1} |x_n|}.$$

b) for $1 \leq p < \infty$, we set:

$$\boxed{\ell_p = \left\{ x = (x_n)_{n \geqslant 1} \in \mathbb{K}^{N^*} ; \sum_{n=1}^{\infty} |x_n|^p < +\infty \right\}};$$

it is provided with the norm defined by:

$$\boxed{\|x\|_p = (\sum_{n=1}^{\infty} |x_n|^p)^{\frac{1}{p}}}.$$

The fact that ℓ_p is a vector subspace of the space of sequences, and that $\|.\|_p$, i.e. a norm on ℓ_p is deduced from the Minkowski inequality (obvious when $p = 1$) generalized as follows:

$$\left(\sum_{n=1}^{\infty} |x_n + y_n|^p \right)^{1/p} \leqslant \left(\sum_{n=1}^{\infty} |x_n|^p \right)^{1/p} + \left(\sum_{n=1}^{\infty} |y_n|^p \right)^{1/p},$$

For all $x_1, x_2, \ldots, y_1, y_2, \ldots \in \mathbf{K}$. We obtain it from the previous one by making the number of terms tend towards infinity: for all $N \geqslant 1$, we have: $\left(\sum_{n=1}^{N} |x_n + y_n|^p \right)^{1/p} \leqslant$

$\left(\sum_{n=1}^{N}|x_n|^p\right)^{1/p}+\left(\sum_{n=1}^{N}|y_n|^p\right)^{1/p}\leqslant(\sum_{n=1}^{\infty}|x_n|^p)^{1/p}+(\sum_{n=1}^{\infty}|y_n|^p)^{1/p}$

Hölder's inequality is generalized in the same way. If $1<p<\infty$ and if q is the conjugate exponent of p, we have:

$$\sum_{n=1}^{\infty}|x_ny_n|\leqslant\left(\sum_{n=1}^{\infty}|x_n|^p\right)^{1/p}\left(\sum_{n=1}^{\infty}|y_n|^q\right)^{1/q}.$$

In particular, when $x=(x_n)_n\in\ell_p$ and $y=(y_n)_n\in\ell_q$, we have $xy\in\ell_1$ and $\|xy\|_1\leq\|x\|_p\|y\|_q$.

The spaces ℓ_p are in fact special cases of the Lebesgue spaces $L^p(m)$, whose definition we will recall below, corresponding to the counting measure on N^* .

Function spaces

1) a) Let A be a set and let the space $\mathscr{F}_b(A)$ be the space (which we also note $\ell_\infty(A)$ If we want to focus on the 'family of elements' aspect) of functions bounded on A, with values in $\mathbb{K}=\mathbb{R}$ or $\mathbb{C}$. If we set:

$$\boxed{\|f\|_\infty=\sup_{x\in A}|f(x)|},$$

Then we have a norm, called the uniform norm. The topology associated with this norm is the topology of uniform convergence; indeed, it is clear that $\|f_n-f\|_\infty\underset{n\to\infty}{\longrightarrow}0$ if and only if $(f_n)_n$ converges uniformly on A to f .

b) Let K be a compact space and $\boxed{\mathscr{C}(K)}$ the space of continuous functions on K (with scalar values). Any continuous function on a compact being bounded, $\mathscr{C}(K)$ is a vector subspace of $\mathscr{F}_b(K)$. It is usually provided with the induced norm $\|f\|_\infty=\sup_{x\in k}|f(x)|$.

Note that, when $K=[0,1]$, for example, we can also provide $\mathscr{C}([0,1])$ with the norm defined by:

$$\|f\|_1=\int_0^1|f(t)|dt,$$

that verifies $|f|_1\leq|f|_\infty$.

c) On the space $\mathscr{C}([0,1])$ of functions k times continuously derivable on $[0,1]$, the norm can be set as fellown:

$$\|f\|^{(k)} = \max\left\{\|f\|_\infty, \|f'\|_\infty, \ldots, \left\|f^{(k)}\right\|_\infty\right\}$$

2) Lebesgue spaces.

Let $(S, \mathscr{F}, m)$ be a measured space; for $1 < p < \infty$, we denote $\boxed{\mathscr{L}^p(m)}$ the space of all measurable functions $f : S \mapsto \mathbb{K} = \mathbb{R}$ or $\mathbb{C}$ such that :

$$\int_S |f(t)|^p dm(t) < +\infty,$$

and we put:

$$\boxed{\|f\|_p = \Big(\int_S |f(t)|^p dm(t)\Big)^{\frac{1}{p}}}.$$

Note that $\|f\|_p = 0$ if and only $f = 0$ m-almost everywhere.

Theorem 1.1.1 (Minkowki Inequality). *Set $1 \leq p < \infty$. for $f, g \in \mathscr{L}^p(m)$, we have the* ***Minkowki Inequality*** *:*

$$\boxed{\Big(\int_S |f+g|^p dm\Big)^{1/p} \leqslant \Big(\int_S |f|^p dm\Big)^{1/p} + \Big(\int_S |g|^p dm\Big)^{1/p}}$$

It follows that $\mathscr{L}^p(m)$ is a vector subspace of the space of measurable functions and that $\|.\|_p$, is a semi-norm on $\mathscr{L}^p(m)$. For $p = 1$, the inequality is obvious.

Proof. The proof is the same as for the sequences. We place ourselves in the case $p > 1$. We can assume $\|f\|_p > 0$ and $\|g\|_p > 0$(because otherwise $f = 0$ m-a.e. and then $f + g = g$ m-a.e, or $g = 0$ m-a.e and then $f + g =$ m-a.e). We apply the convexity inequality $[\alpha u + (1-\alpha)v]^p \leqslant \alpha u^p + (1-\alpha)v^p$ with $\alpha = \|f\|_p / (\|f\|_p + \|g\|_p) \in [0, 1]$, $u = |f(t)|/\|f\|_p$ and $v = |g(t)|/\|g\|_p$. Since $\alpha/\|f\|_p = (1-\alpha)/\|g\|_p = 1/(\|f\|_p + \|g\|_p)$, we have $\left(\frac{|f(t)|+|g(t)|}{\|f\|_p+\|g\|_p}\right)^p \leqslant$

$\frac{\alpha}{\|f\|_p^p}|f(t)|^p + \frac{1-\alpha}{\|g\|_p^p}|g(t)|^p$, hence, by integrating:

$$\int_S \frac{(|f(t)|+|g(t)|)^p}{(\|f\|_p+\|g\|_p)^p} dm(t) \leqslant \frac{\alpha}{\|f\|_p^p}\int_S |f(t)|^p dm(t) + \frac{1-\alpha}{\|g\|_p^p}\int_S |g(t)|^p dm(t)$$
$$= \alpha + (1-\alpha) = 1$$

this gives the result since $|f(t)+g(t)| \leq |f(t)|+|g(t)|$. □

We saw that $\|.\|_p$, is not a norm in general, since $\|f\|_p = 0$ if and only $f = 0$ m-almost everywhere. If$\mathscr{N}$ denotes the space of measurable functions $f : S \mapsto \mathbb{K}$ null m-almost everywhere, the quotient space $\boxed{L^p(m) = \mathscr{L}^p(m)/\mathscr{N}}$ is then normed if we set $\|\tilde{f}\|_p = \|f\|_p$.
In practice, we will not distinguish between the function and its m-equivalence class almost everywhere $\tilde{f}$, and we will therefore write $f \in L^p(m)$ instead of $f \in \mathscr{L}^p(m)$. However, sometimes one have to be careful, especially when handling non-countable quantities of functions. This distinction may already occur for questions of measurability. We can also see this in the following example:
Let $\mathscr{F}$ be the set of all finite parts of $[0, 1]$; for all $A \in \mathscr{F}$, we have, in terms of the Lebesgue measure, $\mathbf{1}_A = 0$ a.e.; so $\tilde{\mathbf{1}_A} = \tilde{0}$. But, on the other hand, $sup_{A\in\mathscr{F}}\mathbf{1}_A(x) = 1$ for all $x \in [0, 1]$; so $(sup_{A\in\mathscr{F}}\tilde{\mathbf{1}}_A) = \tilde{\mathbf{1}}$.
As mentionned for sequences, Hölder's inequality is very useful.

Theorem 1.1.2 (Hölder Inequality)**.** *If $1 < p < \infty$ and if q is the conjugate exponent of p, then we have, for $f \in \mathscr{L}^P(m)$and $g \in$ mathscrL$^q(m)$, Hölder's inequality:*

$$\boxed{\int_S |fg|^p dm \leqslant \left(\int_S |f|^p dm\right)^{1/p} \left(\int_S |g|^p dm\right)^{1/p}}$$

For $p = q = 2$, we call it Cauchy-Schwarz inequality:

$$\boxed{\int_S |fg|^2 dm \leqslant \left(\int_S |f|^2 dm\right)^{1/2} \left(\int_S |g|^2 dm\right)^{1/2}},$$

if $f, g \in \mathscr{L}^2(m)$. (It was demonstrated by Bouniakowski in 1859 and re-proven by Schwarz in 1885; it generalizes the inequality for sums demonstrated by Cauchy). It is demonstrated in the same way as for sums, by integrating instead of adding.

Proof. We can assume $\|f\|_p > 0$ et $\|g\|_q > 0$ because otherwise $f = 0$ $m - a.e.$ or $g = 0$ $m - a.e.$, and then $fg = 0m - a.e.$ We use the inequality $ab \leqslant \frac{a^p}{p} + \frac{b^q}{q}$ with $a = |f(t)|/\|f\|_p$ and $b = |g(t)|/\|g\|_q$. By integrating, we obtain:

$$\int_S \frac{|f(t)g(t)|}{\|f\|_p\|g\|_q}dm(t) \leqslant \frac{1}{p}\int_S \frac{|f(t)|^p}{\|f\|_p^p}dm(t) + \frac{1}{q}\int_S \frac{|g(t)|^q}{\|g\|_q^q}dm(t) = \frac{1}{p} + \frac{1}{q} = 1$$

hence the result. □

As an application we have the following result.

Proposition 1.1.4. *Let $(S, \mathscr{F}, m)$be a measured space of finite measure. Then, for $1 < p_1 < p_2 < \infty$, we have $\mathscr{L}^{p_2}(m) \subseteq \mathscr{L}^{p_1}(m) \subseteq \mathscr{L}^(1(m)$. Moreover if $m(S) = 1$ (i.e. m is a probability measure), then $\|f\|_1 < \|f\|_{p_1} < \|f\|_{p_2}$ for all $f \in \mathscr{L}^{p_2}(m)$.*

Proof. We can assume $p_1 < p_2$ Let $p = \frac{p_2}{p_1}$. As $p > 1$, we can use Hölder's inequality:

$$\int_S |f|^{p_1} dm \leqslant \left(\int_S 1^q dm\right)^{1/q} \left(\int_S (|f|^{p_1})^p \, dm\right)^{1/p} = [m(S)]^{1/q} \left(\int_S |f|^{p_2} dm\right)^{1/p};$$

hence $\|f\|_{p_1} \leq [m(S)]^{\frac{1}{p_1} - \frac{1}{p_2}} \|f\|_{p_2}$.

The second inclusion is obtained by replacing p_2 by p_1 and taking $p_1 = 1$. □

Remark 1. *On the contrary, for spaces ℓ_p, we have the opposite inclusions; for $1 < p_1 < p_2 < \infty$:*

$$\ell_1 \subseteq \ell_{p_1} \subseteq \ell_{p_2} \subseteq c_0 \subseteq \ell_\infty.$$

In addition, $\|x\|_\infty \leq \|x\|_{p_2} \leq \|x\|_{p_1} \leq \|x\|_1$ for all $x \in \ell_1$.

Indeed, if $x \in \ell_{p_2}$,$\sum_{n=1}^\infty |x_n|^{p_2} < +\infty$; so $x_n \underset{n\to\infty}{\longrightarrow} 0$. Moreover, for all $n \geq 1$, $|x_n| \leqslant (\sum_{n=1}^\infty |x_n|^{p_2})^{1/p_2} = \|x\|_{p_2}$; therefore $\|x\|_\infty = \sup_{n\geqslant 1} |x_n| \leqslant \|x\|_{p_2}$. Now, if $x \in \ell_{p_1}$ is not zero, let's set $x' = x/\|x\|_{p_1}$. We have $\|x'\|_{p_1} = 1$, that is to say $\sum_{n=1}^\infty |x'_n|^{p_1} = 1$. It follows that $|x'_n|^{p_1} \leqslant 1$, and therefore

$|x'_n| \leqslant 1$, *for all* $n \geqslant 1$. *Then, for all* $n \geqslant 1, |x'_n|^{p_2} \leqslant |x'_n|^{p_1}$, *since* $p_2 \geqslant p_1$. *It follows that* $\sum_{n=1} |x'_n|^{p_2} \leqslant \sum_{n=1} |x'_n|^{p_1} = 1$, *that is,* $\sum_{n=1} |x_n|^{p_2} \leqslant \|x\|_{p_1}^{p_2}$. *So* $x \in \ell_{p_2}$ *and* $\|x\|_{p_2} \leqslant \|x\|_{p_1}$.

Remark 2. *On the other hand, it is important to note that* $\boxed{\mathscr{L}^{p_1}(\mathbb{R}) \nsubseteq \mathscr{L}^{p_2}(\mathbb{R})}$ *for all* $p_1 \notin p_2$.
Indeed, if $p_1 < p_2$, *the function defined by* $f(t) = 1/t^{1/p_2}$ *for* $0 < t \leqslant 1$, *and by* $f(t) = 0$ *elsewhere, is in* $\mathscr{L}^{p_1}(\mathbb{R})$ *because* $p_1/p_2 < 1$, *but not in* $\mathscr{L}^{p_2}(\mathbb{R})$. *If* $p_1 > p_2$, *then the function defined by* $f(t) = 1/t^{1/p_2}$ *for* $t \geqslant 1$, *and* $f(t) = 0$ *for* $t < 1$, *is in* $\mathscr{L}^{p_1}(\mathbb{R})$ *because this time* $p_1/p_2 > 1$, *but is not in* $\mathscr{L}^{p_2}(\mathbb{R})$.

1.1.4 Equivalent norms

Definition 1.1.4. *Let* E *be a vector space with two norms* $\|.\|$ *and* $|||.|||$. *We say that* $|||.|||$*is finer than* $\|.\|$*(and that* $\|.\|$ *is less fine than* $|||.|||$*) if it exists a constant* $K > 0$ *such that:*

$$\|x\| \leq K|||x|||, \quad \forall x \in E.$$

This is equivalent to saying that the identity application:

$$\mathrm{id}_E : (E, |||.|||) \to (E, \|.\|)$$

is continuous.

This is still equivalent to saying that:

$$B_{|||.|||}(0, r/K) \subseteq B_{\|.\|}(0, r);$$

the balls for $|||.|||$ *are therefore "smaller" than the balls for* $\|.\|$*: they separate the points better; more precisely, the topology defined by* $|||.|||$ *is finer than that defined by* $\|.\|$ *(there are more open sets).*

Exemple. In $\mathscr{C}([0, 1])$, the norm $\|.\|_\infty$ is finer than the norm $\|.\|_1$.

Definition 1.1.5. *We say that two norms $\|.\|$ and $|||.|||$ on the vector space E are equivalent if there exist two constants $K_1, K_2 > 0$ such that:*

$$K_1\|x\| \leq |||x||| \leq K_2\|x\|, \quad \forall x \in E.$$

In other words, each is thinner than the other.

This amounts to saying that ***the identity application*** *I_d carries out an isomorphism of E on itself (or rather of E equipped with $\|.\|$ on E equipped with $|||.|||$). This also means that $\|.\|$and $|||.|||$. define the same topology on E.*

Exemples.

1) In $\mathbb{K}^n$ the norms $\|.\|_p$ for $1 \leq p \leq \infty$ are equivalent:

$$\|x\|_\infty \leq \|x\|_p \leq \|x\|_1 \leq n\|x\|_\infty.$$

We will see that in fact all the norms on $\mathbb{K}^n$ are equivalent to each other.

2) In $\mathscr{C}([0,1])$, the norms $\|.\|_\infty$ and $\|x\|_1$ are not equivalent.

1.2 Banach spaces

1.2.1 Cauchy sequences

Definition 1.2.1. *A sequence $(x_k)_k$ of elements of a normed space E is called a Cauchy sequence if:*

$$(\forall \varepsilon > 0) \quad (\exists N \geqslant 1) \quad k, l \geqslant N \quad \Longrightarrow \quad \|x_k - x_l\| \leqslant \varepsilon.$$

Any convergent sequence is Cauchy sequence.

Definition 1.2.2. *We say that a normed space is complete if every Cauchy sequence is convergent. We call Banach space any complete normed space.*

Exemples.

a) It is immediate to see that $\ell_p^n = ((K)^n, \|.\|_p)$ is complete for $1 \leq p \leq \ infty$.

b) The spaces c_0 and ℓ_p, for $1 \leq p \leq \infty$ are complete .

c) $(\mathscr{C}(K), \|.\|_\infty)$ is complete: any uniformly Cauchy sequence is uniformly convergent, and if they are continuous, the limit is continuous too.

On the other hand, $(\mathscr{C}([0,1]), \|.\|_1)$ is not complete.

d) $(\mathscr{C}^k([0,1]), \|.\|_\infty)$ is not complete for $k \geq 1$, but $(\mathscr{C}^k([0,1]), \|.\|^{(k)})$ is complete.

e) Lebesgue spaces are complete. This is the subject of the following theorem.

Theorem 1.2.1 (Riesz-Fisher theorem)**.** *For any measured space $(S, \mathscr{T}, m)$, and for $1 \leq p < \infty$, the space $L^p(m)$ is a Banach space.*

E. Fisher et. F. Riesz actually demonstrated, independently, in 1907 that $L^2([0,1])$ is isomorphic to ℓ_2; this essentially relies on the fact that $L^2([0,1])$ is complete (see Chapter 2 on Hilbert spaces); this is why we give the name Riesz-Fisher to this theorem, proven in fact, for $L^p([0,1])$ and $1 < p < \infty$, by F. Riesz in 1910 (and to distinguish it from the many other theorems due to F. Riesz).

proof of lemma. Let $(F_n)_n$ be a Cauchy sequence in $L^P(m)$. Let's choose a representative $f_n \in \mathscr{L}^p(m)$ de F_n.

a) As the sequence is Cauchy, we can construct a subsequence $(f_{n_k})_j$ with $(n_1 < n_2 < \ldots)$ such that:

$$\|f_{n_{k+1}} - f_{n_k}\|_p \leqslant \frac{1}{2^k} \quad \forall k \geqslant 1.$$

Let's put:

$$\begin{cases} g_k = \sum_{j=1}^{k} \left| f_{n_{j+1}} - f_{n_j} \right| \\ g = \sum_{j=1}^{\infty} \left| f_{n_{j+1}} - f_{n_j} \right| \end{cases}$$

These functions are measurable and we have:

$$\|g_k\|_p \leqslant \sum_{j=1}^{k} \left\| \left| f_{n_{j+1}} - f_{n_j} \right| \right\|_p = \sum_{j=1}^{k} \left\| f_{n_{j+1}} - f_{n_j} \right\|_p \leqslant \sum_{j=1}^{k} \frac{1}{2^j} \leqslant 1.$$

Fatou's Lemma, applied to the sequence $(g_k^p)_{k\geqslant 1}$, gives:

$$\int_S g^p dm \leqslant \liminf_{k\to\infty} \int_S g_k^p dm = \liminf_{k\to\infty} \|g_k\|_p^p \leqslant 1.$$

The function g^p is therefore integrable. In particular it is finished almost everywhere; so g too. This means that the series $\sum_{k\geq 1}(f_{n_{k+1}}(t) - f_{n_k}(t))$ converges absolutely, for almost all $t \in S$.

Hence, let's put:

$$f(t) = \begin{cases} f_{n_1}(t) + \sum_{k=1}^{\infty} \left(f_{n_{k+1}}(t) - f_{n_k}(t)\right) & \text{si } g(t) < +\infty \\ 0 & \text{otherwise} \end{cases}$$

Then f is measurable and:

$$f(t) = \lim_{k\to\infty} f_{n_k}(t) \quad \forall t \in S \textit{ almost everywhere} .$$

b) It remains to be seen that the sequence is from cauchy, there exists an integer $N \geq 1$ such that:

$$n, k \geqslant N \quad \Longrightarrow \quad \|f_n - f_k\|_p \leqslant \varepsilon.$$

For $k \geqslant N$, Fatou's Lemma gives:

$$\int_S |f - f_k|^p \, dm \leqslant \liminf_{j\to\infty} \int_S \left|f_{n_j} - f_k\right|^p dm \leqslant \varepsilon^p.$$

We first deduce that $(f - f_k) \in \mathscr{L}^p(m)$, therefore that $f = (f - f_k) + f_k \in \mathscr{L}^p(m)$; and then, since $\varepsilon > 0$ is arbitrary, that $\lim_{k\to\infty} \|f - f_k\|_p = 0$.

c) Finally, if we write $F \in L^p(m)$ for the m-almost everywhere equivalence class of f, we have $\lim_{k\to\infty} \|F - F_k\|_p = \lim_{k\to\infty} \|f - f_k\|_p = 0$.

■

Remark. It is worth mentioning that the following underlined and very important result has been demonstrated (we will no longer make a distinction between a function and its equivalence class almost everywhere):

Theorem 1.2.2. *If* $f_n \underset{n\to\infty}{\mapsto} f$ *in* $L^p(m)$*, with* $1 \leqslant p < \infty$*, then we can extract a sub- sequence* $(f_{n_k})_k$ *which converges almost everywhere to* f*.*

Remarks. a) It's possible that the sequence itself will not converge anywhere. For example, on $S =]0,1]$, i.e. $f_n = \mathbf{I}_{]\frac{1}{2^k},\frac{l+1}{2^k}]}$ when $n = 2^k + l, 0 \leqslant l \leqslant 2^k - 1$:

Hence, $\|f_n\|_p = \frac{1}{2^{k/p}}$ for $2^k \leqslant n \leqslant 2^{k+1} - 1$; it fellows that, $f_n \underset{n\to\infty}{\longrightarrow} 0$ in $L^p([0,1])$, but for no $t \in]0.1[$, the sequence $(f_n(t))_n$ is convergent. However, the sub-sequence $(f_{2k})_{k\geqslant 0}$, for example, pointwise convergences a.e.

b) It may be noted that $\mathscr{C}[-1,1]$ is not a closed subspace in $L_2[-1,1]$.

c) The space $\mathscr{C}(\Omega)$ is a dense subspace of $L_2(\Omega)$.

d) The set of all polynomials is dense in $L_2(\Omega)$.

1.3 Normed finite-dimensional vector spaces

1.3.1 Equivalence of norms

Theorem 1.3.1. *On a finite-dimensional vector space, all norms are equivalent to each other.*

Proof. 1) Let $\|.\|$ be an arbitrary norm on E. We will show that it is equivalent to a particular norm on E, so that, by transitivity, two arbitrary norms will be equivalent.

2) Let $\{e_1, \ldots, e_n\}$ be a base of E. If $x = \sum_{k=1}^n \xi_k e_k$, we set:

$$|\|x\mid\| = \max\{|\xi_1|, \ldots, |\xi_n|\}.$$

Thus, $(E, |||\cdot|||)$ is isometric to $(\mathbb{K}^n, \|\cdot\|_\infty) = \ell_\infty^n(\mathbb{K})$, by the application

$$V: \quad \mathbb{K}^n \longmapsto E$$
$$a = (a_1, \dots, a_n) \quad \longmapsto V(\xi) = \sum_{k=1}^{n} a_k e_k.$$

We also have:

$$\|x\| \leqslant \sum_{k=1}^{n} |\xi_k| \, \|e_k\| \leqslant \left(\sum_{k=1}^{n} \|e_k\| \right) \cdot \max_{1 \leqslant k \leqslant n} |\xi_k| = K |||x|||.$$

3) This means that the identity mapping $\mathrm{i}d_E : (E, |||.|||) \to (E, \|\cdot\|)$ is continuous . So, the application:

$$N: \quad (E, |||\cdot|||) \longrightarrow \mathbb{R}_+$$
$$x \longmapsto \|x\|$$

is also continuous, by the Proposition 1.1.1

4) Or:

$$S_n = \{a = (a_1, \dots, a_n) \in \mathbf{K}^n ; \|a\|_\infty = 1\}.$$

It is a closed and bounded, therefore compact, part of $\mathbf{K}^n$ (note that the norm $\|\cdot\|_\infty$ defines the usual topology on $\mathbf{K}^n$). so :

$$S = \{x \in E ; |||x||| = 1\}$$

is a compact part of $(E, |||.|||)$ (by isometry: $S = V(S_\infty)$).

5) It follows that there exists $x_0 \in S$ such that $\|x_0\| = N(x_0) = \inf_{x \in S} N(x) = \inf_{x \in S} \|x\|$. Since $x_0 \neq 0$ (since $\|\|x_0\| = 1$), we have $c = \|x_0\| > 0$. It means that :

$$(\forall x \in S) \quad \|x\| \geqslant c.$$

By homogeneity (for all $x \neq 0, x' = x/|||x\| \in S$), we obtain:

$$(\forall x \in E) \quad \|x\| \geqslant c|||x|||,$$

□

This was to be demonstrated.

Remark. In passing, we showed:

Corollary 1.3.1. *Any finite-dimensional normed space n is isomorphic to $\mathbb{K}^n$, equiped with one of its usual norms.*

It follows:

Corollary 1.3.2. *If E is a finite-dimensional normed space, its bounded closed parts are compact.*

Corollary 1.3.3. *1) Any finite-dimensional normed space is complete.*

2) Any vector subspace of finite dimension in a normed space is closed in this space.

3) If E is a finite-dimensional normed space, then any linear mapping $T : E \to F$ in an arbitrary normed space F is continuous.

Proof. 1) follows immediately from the Corollary 1.3.1, and 2) from the fact that everything under complete space is closed. For 3), it suffices to notice that if $e_1, \ldots, e_n$ is a base of E, and $|||.|||$ the associated norm as in the proof of the Theorem 1.3.1, then, for $x = \sum_{k=1}^{n} a_k e_k \in E$, we have:

$$\|T(x)\|_F \leqslant \left(\sum_{k=1}^{n} \|T(e_k)\|_F \right) \max_{k \leqslant n} |a_k| = C \mid \|x\| \leqslant CK\|x\|_E$$

since $|||.|||$ and $\|.\|_E$ are equivalent. □

We will be careful, on the other hand, that if it is the arrival space which is of finite dimension, continuity is not automatic (since there exist non-continuous linear forms, if E is of infinite dimension.

1.3.2 Compactness of the balls

We have seen in the proof of Theorem 1.3.1 that the essential point (via the Corollary 1.3.2) is that the closed bounded parts of a finite-dimensional normed space are compact. Note that it is equivalent to saying that all closed balls are compact. We will see that this actually only happens in finite dimension.

Theorem 1.3.2 (Riesz Theorem, 1918)**.** *If a normed space E has a compact ball $B(x_o, r)$, of radius $r > 0$, then it is of finite dimension.*

We deduce that in an infinite dimensional space, the compacts are "very thin":

Corollary 1.3.4. *If E is a normed space of infinite dimension, then every compact of E has an empty interior.*

Indeed, if K is a non-empty interior compact, it contains a closed ball of radius $r > 0$, which is therefore compact, and therefore E is of finite dimension.

Note that if a ball is compact, it is necessarily a closed ball. On the other hand, if a ball, of radius $r > 0$, is compact, then all closed balls are, since they are homeomorphic to each other (those of zero radius being compact anyway). It is therefore sufficient to show that if E is of infinite dimension, then its unit ball B_E is not compact. To do this, we will use a lemma.

Lemma 1.3.1 (Riesz's lemma)**.** *Let F be a closed vector subspace of a normed space E, which is not an entire E. Then, for any number δ such that $0 < \delta < 1$, there exists $x \in E$ such that:*

$$\begin{cases} \|x\| = 1 \\ \operatorname{dist}(x, F) \geqslant 1 - \delta \end{cases}$$

Let's remember that :

$$\operatorname{dist}(x, F) = \inf_{y \in F} \|x - y\|.$$

If F is of finite dimension, an argument of compactness makes it possible to show that in fact we can choose such a $x \in E$, of norm 1, with dist $(x, F) = 1$, but we won't need it. In the case of the Euclidean space $(\mathbb{R}^n, \|\cdot\|_2)$, it suffices to take x of norm 1 and orthogonal a F (because then, for all $y \in F$, we have $\|x-y\|_2^2 = \|x\|_2^2 + \|y\|_2^2$, by the Pythagorean Theorem; therefore dist $(x, F) \geqslant \|x\|_2 = 1$, hence the equality because $|x\| \geqslant$ dist (x, F), since $0 \in F$). This is why this lemma is sometimes called the Quasi-Perpendicular Lemma.

proof of Riesz Theorem. Let E be a normed space of infinite dimension. Let's set a number $\delta \in]0, 1[$,; for example $\delta = \frac{1}{2}$.
Let us start from a $x_1 \in E$, of norm 1, and take for F the vector subspace F_1 generated by x_1. As it is of dimension 1, it is closed, and is not equal to E, since E is of infinite dimension. The lemma gives a $x_2 \in E$, of norm 1 such that:

$$\|x_2 - x_1\| \geq dist(x_2, F_1) \geq \frac{1}{2}.$$

Let us then take for F the vector subspace F_2 generated by x_1 and x_2 . It is of dimension 2 (because $x_2 \notin F_1$), and is therefore closed, and different from E; there therefore exists $x_3 \in E$, of norm 1 such that:

$$\|x_3 - x_1\| et \|x_3 - x_2\| \geq dist(x_3, F_2) \geq \frac{1}{2}.$$

As E is of infinite dimension, we can iterate the process indefinitely. We obtain a sequence $(x_k)_{k\geq 1}$ of vectors of norm 1 such that:

$$\|x_k - x_l\| \geq \frac{1}{2}, \quad \forall k \neq l.$$

This sequence cannot have any convergent subsequence. As it is contained in the unit ball of E, this ball is not compact. ■

Remark. We have in fact demonstrated a little more than what was stated, namely that if E is of infinite dimension, its unit sphere S_E is not compact (note that S_E is closed in B_E; so if B_E is compact, so is S_E).

proof of lemma. Like $F \neq E$, we can find $x_0 \in E$ such that $x_0 \notin F$. As F is closed, we have:

$$d = dist(x_0, F) > 0.$$

As $0 < \delta < 1$, we have $\frac{d}{1-\delta} > d$ and we can therefore find a $y_0 \in F$ such that $\|x_0 - y_0\| \leq \frac{d}{(1-\delta)}$. All that remains is to "correct" x_o by y_0 and to normalize this vector: let $x = \frac{x_0 - y_0}{\|x_0 - y_0\|}$; it is indeed a vector of norm 1 and, like $y_0 + \|x_0 - y\| y \in F$,

we have :

$$\begin{aligned} \|x - y\| &= \frac{1}{\|x_0 - y_0\|} \|(x_0 - y_0) - \| x_0 - y_0 \| y\| \\ &\geqslant \frac{1}{\|x_0 - y_0\|} \operatorname{dist}(x_0, F) = \frac{d}{\|x_0 - y_0\|} \geqslant 1 - \delta, \end{aligned}$$

for all $y \in F$. ■

1.4 Linear mappings

For linear mapings, we have a very simple and very useful criterion of continuity.

Proposition 1.4.1. *Let $(E, \|.\|_E)$ and $(F, \|.\|_F)$ be two normed spaces and let $T : E \mapsto F$ be a linear map. Then T is continuous if and only if there exists a constant $K \geq 0$ such that:*

$$\boxed{\|T(x)\|_F \leq K\|x\|_E, \quad \forall x \in E.}$$

Proof. It is clear that this property causes the continuity of T because we have, thanks to linearity:

$$\|T(x) - T(y)\|_F \leq K\|x - y\|_E, \quad \forall x, y \in E;$$

T is therefore even Lipschitzian.

Conversely, if T is continuous at 0, we have, by definition:

$$(\exists K > 0) \quad \|y - 0\|_E = \|y\|_E \leqslant 1/K \quad \Longrightarrow \|T(y)\|_F = \|T(y) - T(0)\|_F \leqslant 1$$

For all $x \in E$, non-zero, let $y = \frac{1}{K\|x\|_E} x$; we have $\|y\|_E = 1/K$ and the implication above gives, thanks to the homogeneity of T and the norm:

$$\frac{1}{K\|x\|_E}\|T(x)\|_F \leqslant 1,$$

hence $\|T(x)\|_F \leqslant K\|x\|_E$. As this inequality is obviously true for $x = 0$, this shows Proposition 1.4.1 .

We therefore have $\sup_{x\neq 0} \frac{\|T(x)\|_F}{\|x\|_E} < +\infty$. The following proposition is then obvious: □

Proposition 1.4.2. *Let $T : E \to F$ be a continuous linear mapping. If we put* $\boxed{\|T\| = \sup_{x\neq 0} \frac{\|T(x)\|_F}{\|x\|_E}}$*, then:*

$$\boxed{\|T(x)\|_F \leqslant \|T\|\|x\|_E, \quad \forall x \in E}$$

$\|T\|$ is therefore the smallest constant $K \geqslant 0$ appearing in the Proposition1.4.1.

Proposition 1.4.3. *We also have*

$$\boxed{\|T\| = \sup_{\|x\|_E \leq 1} \|T(x)\|_F = \sup_{\|x\|_E = 1} \|T(x)\|_F.}$$

Proof. Let's call S the first expression and S_1 the next one. We of course have $S_1 \leq S$, and also $S \leq \|T\|$, since $\|T(x)\|_F \leq \|T\|$ if $\|x\|_E \leq 1$, by definition of $\|T\|$. It remains to be seen that $\|T\| \leq S_1$; but :

$$\|T\| = \sup_{x\notin 1} \frac{\|T(x)\|_F}{\|x\|_E} = \sup_{x\notin 1} \|T(\frac{x}{\|x\|_E})\|_F \leq S_1,$$

Because $\frac{x}{\|x\|_E}$ is of norm 1. □

Proposition 1.4.4. *Let $\mathscr{L}(E, F)$ be the space of all continuous linear mappings of E in F. The mapping $T \mapsto \|T\|$ is a norm on $\mathscr{L}(E, F)$, called the operator norm.*

If F is complete, so is $\mathscr{L}(E, F)$.

Proof. The fact that this is a norm is easy to verify.

Let F be complete, and let $(T_n)_n$ be a Cauchy sequence in $\mathscr{L}(E, F)$. Then, for all $x \in E$, the sequence $(T_n(x))_n$, is Cauchy in F, by virtue of the inequality:

$$\|T_n(x) - T_k(x)\|_F \leqslant \|T_n - T_k\| \|x\|_E, \tag{1.1}$$

It therefore converges to an element $T(x) \in F$. It is easy to see that then $T : E \to F$ is linear. It is continuous because:

$$\|Tx\|_F = \lim_{n\to\infty} \|T_n x\|_F \leqslant \limsup_{n\to\infty} \|T_n\| \, \|x\|_E \leqslant \left(\sup_{n\geqslant 1} \|T_n\|\right) \|x\|_E$$

and because $(\sup_{n\geqslant 1} \|T_n\|) < +\infty$ since any Cauchy sequence is bounded. Finally, by making k tend towards infinity in 1.1, we obtain:

$$\|T_n(x) - T(x)\|_F \leqslant \left(\limsup_{k\to\infty} \|T_n - T_k\|\right) \|x\|_E$$

when n tends to infinity, since $(T_n)_n$ is Cauchy. So $(T_n)_n$ converges to T for the operator norm.

In particular, if $F = \mathbb{K}, \mathscr{L}(E, \mathbb{K})$ is always complete. □

Definition 1.4.1. *If the linear mapping $T : E \mapsto F$ is bijective continuous and if $T^{-1} : F \mapsto E$ is continuous, we say that T is an isomorphism (of normed spaces) between E and F.*

We say that E and F are isomorphic if there exists an isomorphism between E and F; we says that they are isometric if there exists an isometric isomorphism $T : E \to F$.

Note that saying that a mapping $T : E \to F$ is isometric means that we have $\|T(x_1) - T(x_2)\|_F = \|x_1 - x_2\|_E$ for all $x_1, x_2 \in E$. When T is linear, this is expressed by $\|T(x)\|_F = \|x_e\|_E$ for all $x \in E$;T is therefore in particular of norm $\|T\|$ =1. All isometry is injective; to say that it is bijective is therefore equivalent to saying that it is surjective. In this case, T^{-1} is also an isometry; it is therefore automatically continuous.

Saying that T is an isomorphim means that T is bijective linear and that there exist two *constants* $0 < \alpha < \beta < \infty$ *such that:*

$$\boxed{\alpha\|x\|_E \leq \|Tx\|_F \leq \beta\|x\|_E} \quad for\ all\ x \in E$$

Indeed, if T *is an isomorphism, the continuity of* T^{-1} *allows us to write:* $\|T^{-1}y\|_E \leq \|T^{-1}\|\|y\|_F$ *for all* $y \in F$ *, i.e.* $\|x\|_E \leq \|T^{-1}\|\|Tx\|_F$ *for all* $x \in E$*. We therefore have the double inequality, with* $\alpha = \frac{1}{\|T^{-1}\|}$ *and* $\beta = \|T\|$ *. Conversely, if we have this double inequality, then* T *is continuous and* $\|T\| \leq \beta$ *and* T^{-1} *is continuous and* $\|T^{-1}\| \leq \frac{1}{\alpha}$*, since* $\alpha\|T^{-1}y\|_E \leq \|y\|_E$ *for all* $y \in F$*.*

We can also notice that the left inequality results in the injectivity of T*.*

1.5 Dual of vector normed space

Definition 1.5.1. $\mathscr{L}(E, \mathbb{K})$ *is denoted by* E^* *and is called the dual of* E*. It is still a Banach space.*

Note that E^* is the topological dual of E, and is strictly smaller than the algebraic dual - the space of all linear functionals -, of E, at least if E is infinite dimension. The norm of $\varphi \in E^*$ is therefore defined by:

$$\boxed{\|\varphi\| = \|\varphi\|_{E^*} = \sup_{x \neq 0} \frac{|\varphi(x)|}{\|x\|} = \sup_{\|x\| \leqslant 1} |\varphi(x)| = \sup_{\|x\|=1} |\varphi(x)|.}$$

Notation. We often use the notation $\boxed{\langle \varphi, x \rangle = \varphi(x)}$.

Theorem 1.5.1. *The dual space* X' *of a normed space* X *is a Banach space (whether or not* X *is).*

Remark.

- Other terms are dual, adjoint space and conjugate space.

- Algebraic dual space $X^\star$ of X is the vector space of all linear functionals on X.

Example 1.1.
- *Space* $\mathbb{R}^n$*: The dual space of* $\mathbb{R}^n$ *is* $\mathbb{R}^n$*.*

- *Space* ℓ^1. *The dual space of* ℓ^1 *is* ℓ^∞.

- *Space* ℓ^p. *The dual space of* ℓ^p *is* ℓ^q; *here,* $1 \leq p \leq +\infty$ *and* q *is the conjugate of* p, *that is,* $\frac{1}{p} + \frac{1}{q} = 1$.

CHAPTER 2

HILBERT SPACES

The concept of Hilbert space extends the methods of linear algebra by generalizing notions of Euclidean space. These spaces owe their name to the German mathematician David Hilbert. Hilbert spaces are a special case of Banach spaces. In this chapter, we generally take $\mathbb{K} = \mathbb{C}$.

2.1 Inner product

2.1.1 Definitions

Definition 2.1.1. *Let H be a real vector space, resp. complex. We call inner product on H any symmetric bilinear form, resp. hermitian, which is positive-definite. We will denote by $< x|y >$ the inner product of the vectors $x, y \in H$.*

This means that the application:

$$< .|. >: \quad H \times H \longrightarrow \mathbb{K} = \mathbb{R} \text{ or } \mathbb{C}$$

$$(x, y) \longmapsto < x \mid y >$$

fulfilled the conditions:

1) for all $y \in H$, the map $x \in H \mapsto < x \mid y > \in \mathbb{K}$ is a linear form;

2) for all $x, y \in H$, we have:

$$\begin{cases} < y \mid x > = < x \mid y > & \text{if the space is real} \\ < y \mid x > = \overline{< x \mid y >} & \text{(complex conjugation), if the space is complex;} \end{cases}$$

3) for all $x \in H$, we have $< x \mid x > \geqslant 0$ and $< x \mid x > = 0$ if and only if $x = 0$.

Remark 3. *1) This means that, in the complex case, we therefore have, for $x, y \in H$ and $\lambda \in \mathbb{C}$:*

$$(a) \quad \langle \alpha x + \beta y, z \rangle = \alpha \langle x, z \rangle + \beta \langle y, z \rangle$$

$$(b) \quad \langle x, \alpha y + \beta z \rangle = \bar{\alpha} \langle x, y \rangle + \bar{\beta} \langle x, z \rangle$$

2) (a) shows that the inner product is linear in the first factor. Since in (c) we have complex conjugates $\bar{\alpha}$ and $\bar{\beta}$ on the right, we say that the inner product is conjugate linear in the second factor. Expressing both properties together, we say that the inner product is sesquilinear. This means "$1\frac{1}{2}$ times linear".

> **Definition 2.1.2.** *If the vector space H is endowed with an inner product, we say that it is an inner product space **pre-Hilbert space**.*

Exemples. 1) a) The usual inner product of $\mathbb{R}^n$ is defined by:

$$\boxed{< x|y > = x_1 y_1 + \ldots + x_n y_n}$$

for $x = (x_1, \ldots, x_n), y = (y_1, \ldots, y_n) \in \mathbb{R}^n$.

The usual inner product of $\mathbb{C}^n$ is defined by:

$$\boxed{< x|y > = x_1 \bar{y_1} + \ldots + x_n \bar{y_n}}$$

for $x = (x_1, \ldots, x_n), y = (y_1, \ldots, y_n) \in \mathbb{C}^n$.

b) We can define other inner products on $\mathbb{K}^n$ by giving weights, that is to say numbers

$w_1, \dots, w_n > 0$, and by setting:

$$\begin{cases} < x|y >= \sum_{k=1}^{n} w_k x_k y_k, & \text{if } \mathbb{K} = \mathbb{R} \\ < x \mid y >= \sum_{k=1}^{n} w_k x_k \bar{y}_k, & \text{if } \mathbb{K} = \mathbb{VS} \end{cases}$$

2) Si$(S, \mathscr{T}, m)$ is a measured space, we provide $H = L^2(m)$ with a inner product (which we will call of natural) by setting, for $f, g \in L^2(m)$:

$$\boxed{< (f \mid g >= \int_S fgdm} \text{ in the real case,}$$

And :

$$\boxed{< f \mid g >= \int_S f\bar{g}dm} \text{ in the complex case.}$$

In particular, on ℓ_2, we have the natural inner product defined by:

$$\boxed{< x \mid y >= \sum_{n=1}^{\infty} x_n y_n} \text{ in the real case,}$$

And :

$$\boxed{< x \mid y >= \sum_{n=1}^{\infty} x_n \bar{y}_n} \text{ in the complex case.}$$

for $x = (x_n)_{n \geqslant 1}, y = (y_n)_{n \geqslant 1} \in \ell_2$. Consider the sequences

$$x = (1, 1, 1, \dots) \quad \text{and} \quad y = \left(1, \frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \dots\right)$$

Since $1^2 + 1^2 + \cdots$ does not converge, x does not belong to in ℓ_2. On the other hand, the series $1^2 + \left(\frac{1}{2}\right)^2 + \left(\frac{1}{4}\right)^2 + \cdots$ does converge; hence y belongs to ℓ_2.

2.2 Elementary properties

Notation. since $< x|x > \geq 0$, we can put:

$$\boxed{\|x\| = \sqrt{(x|x)}.}$$

Proposition 2.2.1. *for all $x, y \in H$:*

a) $\boxed{\|x+y\|^2 = \|x\|^2 + \|y\|^2 + 2 < x|y >}$ *(real case);*

b) $\boxed{\|x+y\|^2 = \|x\|^2 + \|y\|^2 + 2Re < x|y >}$ *(complex case).*

Proof. Just expand:

$$\|x+y\|^2 =< x+y \mid x+y >=< x \mid x > + < y \mid y > + < x \mid y > + < y \mid x >,$$

and use the fact that $< x \mid y > + < y \mid x >=< x \mid y > + \overline{< x \mid y >} = 2 < x \mid y >$ in the real case , and $= 2\,\mathrm{Re} < x \mid y >$ in the complex case. □

2.2.1 Cauchy-Schwarz inequality

Theorem 2.2.1 (Cauchy-Schwarz inequality)**.** *For all $x, y \in H$:*

$$\boxed{| < x \mid y > | \leqslant \|x\|\|y\|.}$$

Exemple. In the case where $H = L^2(m)$, it is equivalent to the Cauchy-Schwarz inequality for integrals:

$$\left|\int_S fgdm\right| \leqslant \int_S |fg|dm \leqslant \left(\int_S |f|^2 dm\right)^{1/2} \left(\int_S |g|^2 dm\right)^{1/2}$$

Proof. We will only do it in the complex case; it's a little easier in the real case (we consider the sign of the inner product instead of its argument). In fact the proof is valid even for inner semi-products, that is to say if the symmetric bilinear form (resp. Hermitian) is only positive (that is to say that we do not ask that $(x|x) = 0$ implies $x = 0$).

Let $\theta \in \mathbb{R}$ such that:

$$\left(\mathrm{e}^{-i\theta} x \mid y\right) = \mathrm{e}^{-i\theta} < x \mid y > \in \mathbb{R}_+$$

(if $< x \mid y > \neq 0, \theta$ is the argument of the complex number $< x \mid y >$). Let $x' = \mathrm{e}^{-i\theta} x$. For all $t \in \mathbb{R}$, we have, by the Proposition 2.2.1:

$$\|x'\|^2 + 2\,\mathrm{Re}\,(x' \mid y)\, t + \|y\|^2 t^2 = \|x' + ty\|^2 \geqslant 0.$$

If $\|y\| = 0$, we have $\|x'\|^2 + 2\,\mathrm{Re}\,(x' \mid y)\, t \geqslant 0$ for all $t \in \mathbb{R}$; this is only possible if $\mathrm{Re}\,(x' \mid y) = 0$. If $\|y\| \neq 0$, we have a second degree trinomial in t, which is always positive or zero; its discriminant must be negative or zero:

$$\mathrm{Re}\,(x' \mid y) - \|x'\|^2 \, \|y\|^2 \leqslant 0.$$

As :

$$(x' \mid y) = \mathrm{e}^{-i\theta} < x \mid y > = | < x \mid y > | \in \mathbb{R}_+$$

we have :

$$\mathrm{Re}\,(x' \mid y) = (x' \mid y) = | < x \mid y > |.$$

Since, in addition, $\|x'\| = \|x\|$, we obtain the announced inequality. □

Corollary 2.2.1. *The expression $\|x\| = \sqrt{(x|x)}$ defines a norm on H, called the* ***Hilbert*** *norm.*

Proof. Just check the triangle inequality:

$$\|x + y\|^2 = \|x\|^2 + \|y\|^2 + 2Re(x|y) \leq \|x\|^2 + \|y\|^2 + 2\|x\|\|y\| = (\|x\| + \|y\|)^2,$$

thanks to the Cauchy-Schwarz inequality. □

Corollary 2.2.2. *for each $y \in H$, the linear form:*

$$\boxed{\begin{array}{cccc} \Phi_y : & H & \longrightarrow & \mathbb{K} = \mathbb{R} \text{ or } \mathbb{C} \\ & x & \longmapsto & < x \mid y > \end{array}}$$

is <u>continue</u>. Its norm in H^* is $\boxed{\|\Phi_y\| = \|y\|}$.

Proof. We can assume $y \neq 0$. The Cauchy-Schwarz inequality says that:

$$|\Phi_y(x)| = | < x \mid y > | \leqslant \|y\|\|x\|;$$

this proves that Φ_y is continuous and that $\|\Phi_y\| \leqslant \|y\|$.

Since $\Phi_y(y) = \|y\|^2$, we have $\|\Phi_y\| \geqslant \frac{|\Phi_y(y)|}{\|y\|} = \|y\|$. ☐

important Remark . Case of equality in the Cauchy-Schwarz inequality. When we look at the proof of the inequality (in the case of a dot product), we see that we have $|(x|y)| = \|x\|\|y\|$ if and only if $y = 0$ or if $y \neq 0$ and the discriminant of the second degree trinomial at t is zero; this means that this trinomial has a (double) root: there exists $t_0 \in \mathbb{R}$ such that $\|x' + t_0 y\| = 0$; in other words $e^{-i\theta}x + t_0 y = 0$: the vectors x and y are **linearly related**.

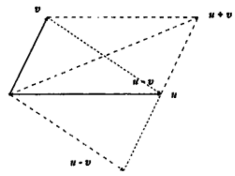
Conversely, if x and y are linearly dependent, it is clear that we have equality.

2.2.2 Parallelogram equality

Lemma 2.2.1 (parallelogram identity)**.** *For all $u, v \in H$:*

$$\boxed{\|u + v\|^2 + \|u - v\|^2 = 2(\|u\|^2 + \|v\|^2)}. \tag{2.1}$$

The proof is immediate, with the Proposition 2.2.1. This means that the sum of the squares of the diagonals of a parallelogram is equal to the sum of the squares of the four sides.



We conclude that if a norm does not satisfy 2.1, it cannot be obtained from an inner product. Such norms do exist; We conclude that, not all normed spaces are inner product spaces.

2.2.3 Orthogonality

Definition 2.2.1. *We say that two vectors x and y of a pre-Hilbert space H are orthogonal if $(x|y) = 0$. We note* $\boxed{x \perp y}$.

Exemple. In $H = \mathbb{R}^2$, for the usual inner product, we have $(-1, 1) \perp (1, 1)$.

Note that the orthogonality relation is symmetric: if $x \perp y$, then $y \perp x$ (because $(y|x) = (\bar{x}|y)$).

According to Proposition 2.2.1, we have, **in the real case**:

$$\boxed{x \perp y \iff \|x + y\|^2 = \|x\|^2 + \|y\|^2},$$

what we can call the Pythagorean Theorem.

In the complex case:

$$x \perp y \iff [\|x + y\|^2 = \|x\|^2 + \|y\|^2 \ \ \textit{and} \ \ \|x + iy\|^2 = \|x\|^2 + \|y\|^2].$$

Indeed, for any complex number a, we have $Im(a) = Re(-ia)$ and consequently $Im(x|y) = Re(x|iy)$.

Parts $A, B \subseteq H$ are called **orthogonal** if all $x \in A$ is orthogonal to all $y \in B$:

$$x \perp y, \quad \forall x \in A, \forall y \in B.$$

We also say that one is orthogonal to the other.

Definition 2.2.2. *The orthogonal of a part $A \subseteq H$ is the set:*

$$\boxed{A^{\perp} = \{y \in H; y \perp x, \forall x \in A\}}$$

We have $B^{\perp} \subseteq A^{\perp}$ if $A \subseteq B$; therefore in particular $(\bar{A})^{\perp} \subseteq A^{\perp}$; but the continuity of applications $\boxed{\Phi_y : x \mapsto < x \mid y >}$ leads to that $(\bar{A})^{\perp} = A^{\perp}$.

Proposition 2.2.2. *For any part A of $H, A^{\perp}$ is orthogonal to A; it is the largest orthogonal part to A. Moreover $A^{\perp}$ is a closed vector subspace of H.*

Proof. The beginning is clear. For the rest, note that:

$$\boxed{A^{\perp} = \bigcap_{x \in A} \ker \Phi_x}$$

and that each vector subspace ker $\Phi_x = \Phi_x^{-1}(\{0\})$ is closed since Φ_x is continuous. □

2.2.4 Hilbert spaces

Definition 2.2.3. *If a pre-Hilbert space is complete, for the norm induced by its inner product, we say that it is* ***a Hilbert space****.*

It is therefore a special case of Banach space.

Exemples.

1 Any pre-Hilbert space of finite dimension is a Hilbert space. When it is a $\mathbb{R}-$ vector space, we say that it is an Euclidean space.

2 For any positive measure m, $\boxed{L^2(m) \text{ is a Hilbert space}}$, by virtue of the Riesz-Fisher

theorem, since the norm $\|.\|_2$:

$$\|f\|_2 = \left(\int_S |f(t)|^2 dm(t)\right)^{1/2}$$

is associated with the usual inner product:

$$(f \mid g) = \int_S f(t)\overline{g(t)}dm(t).$$

In particular, $\boxed{\ell_2 \text{ is a Hilbert space}}$.

2.3 The Projection Theorem and its consequences

2.3.1 Projection Theorem

It is thanks to this theorem that we obtain all the good properties of Hilbert spaces.

Let us first recall that a part C of a vector space is said to be convex if the segment $[x, y]$ is contained in C since $x, y \in C$:

$$x, y \in C \Longrightarrow [x, y] \subseteq C,$$

where $[x, y] = tx + (1-t)y; t \in [0, 1]$.

vector subspace is convex; every ball is convex.

Theorem 2.3.1 (Projection theorem)**.** *Let H be a* ***Hilbert*** *space and let C be a non-empty* ***convex*** *and* ***closed*** *part of H. Then, for all $x \in H$, there exists a unique $y \in C$ such that:*

$$\|x - y\| = dist(x, C).$$

We say that $y = P_C(x)$ is the projection of x onto C. It is characterized by the property:

$$y \in C \quad and \quad Re(x - y|z - y) \leq 0, \forall z \in C. \qquad (*)$$

Note that the completeness of H is not absolutely essential: we can remove it, but assuming that it is C which is complete.

Proof. 1) Existence.

Let $d = dist(x, C) = inf_{x \in C} \|x - z\|$.

Note that if $d = 0$, then $x \in C$ (because C is closed), and $y = x$ is the unique point of C such that $\|x - y\| = d$.

For all $n > 1$, there exists $z_n \in C$ such that:

$$\|x - z_n\|^2 \leqslant d^2 + \frac{1}{n}$$

Let us then apply, for $n, p \geqslant 1$, the identity of the parallelogram to $u = x - z_n$ and $v = x - z_p$; we obtain :

$$4 \left\| x - \frac{z_n + z_p}{2} \right\|^2 + \|z_n - z_p\|^2 = 2 \left(\|x - z_n\|^2 + \|x - z_p\|^2 \right).$$

But, C being convex, we have $\frac{z_n + z_p}{2} \in C$; therefore:

$$\left\| x - \frac{z_n + z_p}{2} \right\| \geqslant d$$

so that we obtain:

$$\|z_n - z_p\|^2 \leqslant 2 \left(d^2 + \frac{1}{n} + d^2 + \frac{1}{p} \right) - 4d^2 = 2 \left(\frac{1}{n} + \frac{1}{p} \right).$$

The sequence $(z_n)_n$ is therefore a Cauchy sequence. As H is complete, it therefore converges to an element $y \in H$. But since C is closed, we have in fact, since the z_n are in $C, y \in C$.

Moreover, the fact that $\|x - z_n\|^2 \leqslant d^2 + 1/n$ leads, passing to the limit, that $\|x - y\| \leqslant d$. We therefore have $\|x - y\| = d$, since $y \in C$.

2) Uniqueness. If $\|x - y_1\| = \|x - y_2\| = d$, with $y_1, y_2 \in C$, then, as above, the identity

of the parallelogram gives:

$$\begin{aligned} 4d^2 + \|y_1 - y_2\|^2 &\leqslant 4\left\|x - \frac{y_1+y_2}{2}\right\|^2 + \|y_1 - y_2\|^2 \\ &= 2\left(\|x-y_1\|^2 + \|x-y_2\|^2\right) = 2\left(d^2+d^2\right) \end{aligned}$$

hence $\|y_1 - y_2\|^2 \leqslant 0$, which is only possible if $y_1 = y_2$.

3) Proof of $(*)$.

a) If $z \in C$, we have $(1-t)y + tz \in C$ for $0 \leqslant t \leqslant 1$, by the convexity of C; SO :

$$\|x - (1-t)y - tz\|^2 \geqslant \|x-y\|^2,$$

or by expanding $\|x-(1-t)y-tz\|^2 = \|(x-y)+t(y-z)\|^2$ with Proposition 2.2.1:

$$t^2\|y-z\|^2 + 2t\,\mathrm{Re}(x-y \mid y-z) \geqslant 0.$$

For $t \neq 0$, divide by t, then let t tend to 0; it comes $\mathrm{Re}(x-y \mid y-z) \geqslant 0$, or:

$$\mathrm{Re}(x-y \mid z-y) \leqslant 0.$$

b) Conversely, if y satisfies $(*)$, we have, for all $z \in C$:

$$\begin{aligned} \|x-z\|^2 &= \left\|\left(x-y+(y-z)\right)\right\|^2 = \left\|x-y\right\|^2 + \left\|y-z\right\|^2 + 2\,\mathrm{Re}(x-y \mid y-z) \\ &= \|x-y\|^2 + \|y-z\|^2 - 2\,\mathrm{Re}(x-y \mid z-y) \geqslant \|x-y\|^2; \end{aligned}$$

therefore $y = P_C(x)$, by uniqueness. □

2.3.2 Consequences

Proposition 2.3.1. *The map $P_C : H \to C$ is continuous; more precisely, we have, for all $x_1, x_2 \in H$:*

$$\|P_C(x_1) - P_C(x_2)\| \leqslant \|x_1 - x_2\|.$$

Proof. Let $y_1 = P_C(x_1)$ and $y_2 = P_C(x_2)$; the condition $(*)$ gives:

$$\begin{cases} \operatorname{Re}(x_1 - y_1 \mid z - y_1) \leqslant 0 & \forall z \in C; \\ \operatorname{Re}(x_2 - y_2 \mid z' - y_2) \leqslant 0 & \forall z' \in C. \end{cases}$$

Taking $z = y_2$ and $z' = y_1$, and adding, it comes:

$$\operatorname{Re}([x_1 - y_1] - [x_2 - y_2] \mid y_2 - y_1) \leqslant 0.$$

We therefore obtain:

$$\begin{aligned} \|y_1 - y_2\|^2 = \operatorname{Re}\|y_1 - y_2\|^2 &= \operatorname{Re}([y_2 - x_2] + [x_2 - x_1] + [x_1 - y_1] \mid y_2 - y_1) \\ &= \operatorname{Re}([x_1 - y_1] - [x_2 - y_2] \mid y_2 - y_1) + \operatorname{Re}(x_2 - x_1 \mid y_2 - y_1) \\ &\leqslant \operatorname{Re}(x_2 - x_1 \mid y_2 - y_1) \\ &\leqslant |(x_2 - x_1 \mid y_2 - y_1)| \leqslant \|x_2 - x_1\| \, \|y_2 - y_1\| \end{aligned}$$

by the Cauchy-Schwarz inequality. It follows, by dividing by $\|y_2 - y_1\|$ (which we can assume is not zero, because otherwise the result is obvious), that we have indeed
$\|y_1 - y_2\| \leqslant \|x_2 - x_1\|$ □

In the case where the convex C is a vector subspace, we have better properties.

Theorem 2.3.2. *If F is a closed vector subspace of the Hilbert space H, then the mapping $P_F : H \to F$ is a continuous <u>linear</u> mapping, and $P_F(x)$ is the unique point $y \in F$ such that:*

$$\boxed{y \in F \quad et \quad x - y \in F^\perp}.$$

Proof. First, if $y \in F$ et $x - y \in F^\perp$, then we have:

$$\operatorname{dist}(x, F)^2 = \inf_{z \in F} \|x - z\|^2 = \inf_{z \in F} \left[\|x - y\|^2 + \|y - z\|^2\right] = \|x - y\|^2;$$

so $\|x - y\| = \operatorname{dist}(x, F)$ et $y = P_F(x)$. The converse results from the condition (*):

$$\operatorname{Re} < x - y \mid z - y > \leqslant 0, \quad \forall z \in F;$$

in fact, as F is a vector subspace, we have:

$$z = y + \lambda w \in F, \quad \forall w \in F \quad \text{et} \quad \forall \lambda \in \mathbb{K}.$$

When H is real, we therefore have, for all $w \in F$:

$$\lambda < x - y \mid w >=< x - y \mid \lambda w > \leqslant 0, \quad \forall \lambda \in \mathbb{R}$$

which is only possible if $< x - y \mid w >> 0$. When the space H is complex, we have, in the same way, for all $w \in F$:

$$\lambda \operatorname{Re} < x - y \mid w >= \operatorname{Re} < x - y \mid \lambda w > \leqslant 0, \quad \forall \lambda \in \mathbb{R},$$

and, with $z = y + i\lambda w$:

$$\lambda \operatorname{Im} < x - y \mid w >= \operatorname{Re} < x - y \mid i\lambda w) > \leqslant 0, \quad \forall \lambda \in \mathbb{R}$$

which, again, is only possible if $< x - y \mid w >= 0$. The linearity of P_F is then easy to see, thanks to the uniqueness; indeed, if

$y_1 = P_F\,(x_1)$,$y_2 = P_F\,(x_2)$, then $(x_1 - y_1)\,,(x_2 - y_2) \in F^{\perp}$; so, for $a_1, a_2 \in \mathbb{K}$, $(a_1x_1 + a_2x_2) - (a_1y_1 + a_2y_2) \in F^{\perp}$;

hence $P_F\,(a_1x_1 + a_2x_2) = a_1y_1 + a_2y_2$. □

Note that continuity was seen in Proposition 2.3.1, and that by taking $x_2 = 0$ in this proposition, we have: $\|P_F(x)\| \leqslant \|x\|$ for all $x \in H$; the norm of P_F is therefore $\leqslant 1$. But since $P_F(x) = x$ for all $x \in F$, we obtain, if $F \neq \{0\}$, that $\boxed{\|P_F\| = 1}$.

As an exercise, we can show that, for a closed convex C, P_C is linear if and only if C is a vector subspace.

Theorem 2.3.3. *If H is a **Hilbert** space, then, for any closed vector space, we have:*

$$\boxed{H = F \oplus F^{\perp}},$$

and the projection onto F parallel to the associated $F^{\perp}$ is P_F. It is therefore continuous, so that the direct sum is a direct topological sum.

We say that P_F is the orthogonal projection on F.

The fact that H is the direct sum of F and $F^{\perp}$ means that all $x \in H$ is uniquely written as $\boxed{x = y + z}$, with $\boxed{y \in F, z \in F^{\perp}}$. Note that, since F and $F^{\perp}$ are orthogonal, we have: $\|x\|^2 = \|y\|^2 + \|z\|^2$; in other words:

$$\boxed{\|x\|^2 = \|P_F(x)\|^2 + \|x - P_F(x)\|^2 .}$$

We find the fact that P_F is continuous and of norm 1, if $F \neq \{0\}$. We also see that $\|Id_H - P_F\| = 1$, if $F^{\perp} \neq \{0\}$; but we will see just after that in fact $Id_H - P_F$ is the orthogonal projection on $F^{\perp}$.

Proof. We have $x = P_F(x) + (x - P_F(x))$, with $x - P_F(x) \in F^{\perp}$, by Theorem 2.3.2.
On the other hand, if $x \in F \cap F^{\perp}$, we have, in particular, $< x \mid x >= 0$; so $x = 0$. □

Remark. The Theorem 2.3.3 is really specific to Hilbert spaces.

The following result can be shown directly, but it is easily obtained from Theorem 2.3.3

Corollary 2.3.1. *We have $F^{\perp\perp} = \bar{F}$ for every vector subspace F of the Hilbert space H.*

Proof. As $F^{\perp}$ is a closed vector subspace, by Proposition 2.2.2, we can apply Theorem 2.3.3: $H = F^{\perp} \oplus F^{\perp\perp}$, which can also be written: $H = F^{\perp\perp} \oplus F^{\perp}$

On the other hand, we can also apply this theorem to the closed vector subspace $\bar{F} : H = \bar{F} \oplus (\bar{F})^{\perp} = \bar{F} \oplus F^{\perp}$. It follows, since we know that $\bar{F} \subseteq F^{\perp\perp}$, that $F^{\perp\perp} = \bar{F}$. □

Note that in general a vector subspace has an infinity of supplementaries; but it only has one orthogonal supplement.

We deduce, since $H^{\perp} = \{0\}$ and $0^{\perp} = H$, the following **very practical** density criterion.

Corollary 2.3.2. *Let H be a Hilbert space, and F be a vector subspace of H. Then F is dense in H if and only if $F^{\perp} = 0$.*

Thus, to show that an vector subspace F is **dense** in H, it suffices to verify that:

$$\boxed{[(x|y) = 0, \quad \forall x \in F] \implies y = 0}.$$

Let's see an example application. Recall that the support of $f : \mathbb{R} \to \mathbb{C}$, denoted *supp$f$*, is the adhesion of $\{x \in \mathbb{R}; f(x) \neq 0\}$.

Theorem 2.3.4. *The space $\mathscr{K}(\mathbb{R})$ of continuous functions on $\mathbb{R}$ with compact support is dense in $L^2(\mathbb{R})$.*

This theorem is demonstrated, in a more general form, in any Integration course (see also Theorem III.1.2); but it is a question here, even if the result is important in itself, of seeing how to apply the Corollary 2.3.2

Note that $\mathscr{K}(\mathbb{R})$ is not really contained in $L^2(R)$, since the latter is a space of equivalence classes of functions, but, as two continuous maps which are equal almost everywhere, for the Lebesgue measure, are in fact everywhere, the canonical map $j : \mathscr{K}(\mathbb{R}) \to L^2(\mathbb{R})$, which associates each function with its equivalence class, is injective; we can therefore identify each $f \in \mathscr{K}(\mathbb{R})$ with its equivalence class $j(f)$, that is to say $\mathscr{K}(\mathbb{R})$ with $j[\mathscr{K}(\mathbb{R})]$.

Proof. Let $g \in L^2(\mathbb{R})$ such that:

$$< f \mid g >= \int_{\mathbb{R}} f\bar{g}d\lambda = 0, \quad \forall f \in \mathscr{K}(\mathbb{R}).$$

We want to show that $g = 0$.

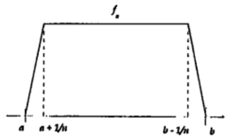
Taking the real and imaginary parts, we can assume that g is real-valued, and we write $g = g^+ - g^-$. We have, for all $f \in \mathscr{K}(\mathbb{R})$:

$$\int_{\mathbb{R}} f(t)g^+(t)dt = \int_{\mathbb{R}} f(t)g^-(t)dt.$$

Let $a < b$. There exist $f_n \in \mathscr{K}(\mathbb{R})$ such that:

$$\begin{cases} 0 \leqslant f_n \leqslant \mathbf{1}_{]a,b[} \\ f_n(t) \underset{n\to\infty}{\longrightarrow} \mathbf{1}_{]a,b[}(t) \text{ for } t \in \mathbb{R}, \end{cases}$$

and such that the sequence $(f_n)_n$ is increasing.



The Monotone Convergence Theorem gives:

$$\int_a^b g^+(t)dt = \lim_{n\to\infty} \uparrow \int_{\mathbb{R}} f_n(t)g^+(t)dt = \lim_{n\to\infty} \uparrow \int_{\mathbb{R}} f_n(t)g^-(t)dt = \int_a^b g^-(t)dt.$$

This means that positive measures $\mu = g^+.\lambda$ and $\nu = g^-.\lambda$ are equal on all intervals $]a, b[$ and take finite values there:

$$\int_a^b g^+(t)dt \leqslant \int_a^b |g(t)|dt = \int_{\mathbb{R}} |g(t)|\mathbf{1}_{]a,b[}(t)dt \leqslant \sqrt{b-a}\|g\|_2 < +\infty,$$

by the Cauchy-Schwarz inequality. The Uniqueness of Measures Theorem then says that $\mu = \nu$. This means that $g^+ = g^-$ almost everywhere, i.e. $g = 0$ in $L^2(\mathbb{R})$. □

Corollary 2.3.3. $\mathscr{C}([0,1])$ *is dense in* $L^2(0,1)$.

Proof. Let $f \in L^2(0,1)$. Let's extend it to $\tilde{f}$ on $\mathbb{R}$ by 0 outside $[0,1]$. We have $\tilde{f} \in L^2(\mathbb{R})$. For all $\varepsilon > 0$, there exists $g \in \mathscr{K}(\mathbb{R})$ such that $\|f - g\|_{L^2(\mathbb{R})} \leqslant \varepsilon$. Let $h = g_{|[0,1]}$ be the restriction of g to $[0,1]$. We have, on the one hand, $h \in \mathscr{C}([0,1])$ and, on the other hand, $\|f - h\|_{L^2(0,1)} \leqslant \|\tilde{f} - g\|_{L^2(\mathbb{R})} \leqslant \varepsilon$. □

2.3.3 Representation of the dual

Recall that the dual is:

$$H^* = \{\Phi : H \to \mathbb{K}; \quad \Phi \text{ continuous linear}\},$$

where $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$ is the coordinate space .

Knowing how to give a concrete representation of the dual of a functional space often allows us to solve problems about the space itself. In the case of Hilbert spaces, it is particularly simple.

Let us first recall that we have seen that, for all $y \in H$, the linear form $\Phi_y : x \in H \to (x|y)$ is continuous, i.e. say is an element of the dual H^*, and that $\|\Phi_y\| = \|y\|$. It turns out that all elements of the dual are of this form.

Theorem 2.3.5 (Fréchet-Riesz representation theorem)**.** *Let H be a Hilbert space. For all $\Phi \in H^*$, there exists a (unique) $y \in H$ such that* $\boxed{\Phi(x) = (x|y)}$ *for all x in$H$.*

This theorem was independently proven by M. Fréchet and F. Riesz in 1907, for $H = L^2(0,1)$; both articles were published, coincidentally, in the same issue of Notes aux Comptes de l'Académie des Sciences. Another way to see this theorem is to say that the application:

$$\begin{array}{cccc} J: & H & \longrightarrow & H^* \\ & y & \longmapsto & \Phi_y = J(y) \end{array}$$

is surjective. It is therefore bijective because it is an isometry (in the sense of metric spaces): $\|J(y) - J(y')\| = \|\Phi_y - \Phi_{y'}\| = \|\Phi_{y-y'}\| = \|y - y'\|$.

Note that in the real case, J is linear, but that in the complex case, it is only semi-linear.

Proof. We already know that J is a metric isometry; this proves uniqueness. What we need to see is surjectivity.

Let $\Phi \in H^*$ be non-zero. As Φ is continuous, the vector subspace $F = \ker \Phi$ is closed. So:

$$H = (\ker \Phi) \oplus (\ker \Phi)^{\perp}.$$

But since Φ is a non-zero linear form, $\ker \Phi$ is of codimension 1; therefore $(\ker \Phi)^{\perp}$ is of dimension 1.

Let $u \in (\ker \Phi)^{\perp}$, of norm 1, and let $y = \overline{\Phi(u)}u$. Then, like $y \in (\ker \Phi)^{\perp}$, Φ_y is zero on $\ker \Phi$; but, on the other hand:

$$\Phi_y(u) =< u \mid y >= \Phi(u) < u \mid u >= \Phi(u)\|u\|^2 = \Phi(u)$$

Thus we have $\Phi = \Phi_y$. □

Remark. The value $y = \Phi(\bar{u})u$ may seem to "fall from the sky". In fact, if we want to have $\Phi(x) = (x|y)$ for all $x \in H$, we must have it for $x \in ker\Phi$; so y must be in $(ker\Phi)^{\perp}$. Thus $y = cu$, and the equality $\Phi(u) = (u|y)$ results in $\Phi(u) = \bar{c}(u|u) = \bar{c}\|u\|^2 = \bar{c}$. We therefore necessarily have $y = \Phi(\bar{u})u$.

2.3.4 Adjoint of an operator

We call an operator on H any continuous linear map $T : H \to H$.

Proposition 2.3.2. *Let H be a Hilbert space. For all $T \in \mathscr{L}(H)$, there exists another operator, denoted T^*, and called the adjoint of T, such that:*

$$\boxed{(Tx|y) = (x|T^*y)}, \quad \forall x, y \in H.$$

Moreover, $\|T^*\| = \|T\|$.

Proof. Let $y \in H$. The mapping:

$$\begin{aligned} \Phi_y \circ T : \quad & H \longrightarrow \mathbb{K} \\ & x \longmapsto < Tx \mid y > \end{aligned}$$

is a continuous linear form on H; there therefore exists, by the FréchetRiesz Theorem, a unique element of H, which we will denote T^*y, such that:

$$\langle x \mid T^* y\rangle =< Tx \mid y >, \quad \forall x \in H.$$

Because of uniqueness, the map $T^* : y \in H \mapsto T^* y \in H$ is clearly linear: if $y_1, y_2 \in H$ and $a_1, a_2 \in \mathbb{K}$, we have, for all $x \in H$:

$$\begin{aligned}\langle x \mid T^* (a_1 y_1 + a_2 y_2)\rangle &= \langle Tx \mid a_1 y_1 + a_2 y_2\rangle = \bar{a}_1 \langle Tx \mid y_1\rangle + \bar{a}_2 \langle Tx \mid y_2\rangle \\ &= \bar{a}_1 \langle x \mid T^* y_1\rangle + \bar{a}_2 \langle x \mid T^* y_2\rangle = \langle x \mid a_1 T^* y_1 + a_2 T^* y_2\rangle\end{aligned}$$

therefore $T^* (a_1 y_1 + a_2 y_2) = a_1 T^* y_1 + a_2 T^* y_2$.

On the other hand, the Cauchy-Schwarz inequality gives:

$$|(\Phi_y \circ T)(x)| = | < Tx \mid y > | \leqslant \|Tx\| \|y\| \leqslant \|T\| \|x\| \|y\|;$$

therefore $\|T^* y\| = \|\Phi_y \circ T\| \leqslant \|T\| \|y\|$. This proves that the linear map T^* is continuous and that $\|T^*\| \leqslant \|T\|$.

To see that $\|T\| \leqslant \|T^*\|$, notice that T^* itself has an adjoint T^{**}, and that we have $T^{**} = T$:

$$\langle y \mid T^{**} x\rangle = \langle T^* y \mid x\rangle =< y \mid Tx >$$

for all $x, y \in H$; this implies that $T^{**} x = Tx$ for all $x \in H$. Then $\|T\| = \|T^{**}\| \leqslant \|T^*\|$ □

2.4 Orthonormal bases

2.4.1 Separable spaces

Definition 2.4.1. *A topological space E is said to be separable if there exists a part $D \subseteq E$ which is countable and dense in E : $\bar{D} = E$.*

In the case of normed spaces, we have an equivalent notion.

Proposition 2.4.1. *Let E be a normed vector space. For E to be separable, it is necessary and sufficient that there exists in E a part Δ which is countable and* ***total*** *in E.*

We say that a part Δ of a normed vector space E is total when the vector subspace vect (Δ) generated by this part is dense.

Proof. The $\mathbb{Q}$-vector subspace (respectively the $(\mathbb{Q}+i\mathbb{Q})$-vector subspace) generated by Δ is countable and its adherence is the same as that of vect (Δ). □

Exemples. 1) Any vector space of finite dimension is separable.

2) The spaces c_0 and ℓ_p, for $1 \leq p < \infty$, are separable, because if

$$e_n = (0, \ldots, 0, 1, 0, \ldots)$$

$$\uparrow$$

$$n^{th} \text{ position}$$

then $\Delta = \{e_n; n \geqslant 1\}$ is totale, since, for all $x = (\xi_1, \xi_2, \ldots) \in \ell_p$, we have:

$$\|x - (\xi_1 e_1 + \cdots + \xi_n e_n)\|^p = \sum_{k=n+1}^{\infty} |\xi_k|^p \underset{n\to\infty}{\longrightarrow} 0$$

and when $x \in c_0$:

$$\|x - (\xi_1 e_1 + \cdots + \xi_n e_n)\|_\infty = \sup_{k\geqslant n+1} |\xi_k| \underset{n\to\infty}{\longrightarrow} 0.$$

It can be shown that ℓ_∞ is not separable.

Proposition 2.4.2. *Any subspace of a separable metric space is separable.*

Proof. Let E be a separable metric space, $D = \{x_n; n \geqslant 1\}$ a part of dense countable E, and $F \subseteq E$. For any pair of integers $n, k \geqslant$ such that $F \cap B(x_n, 1/k)$ is not empty, let us choose an element $y_{n,k}\ in F \cap B(x_n, 1/k)$; otherwise (for notational purposes), let $y_{n,k} = y_0$, where y_0 is a given fixed element of F (we can assume F not empty) . Then $D_F = \{y_{n,k}; n, k \geqslant 1\}$ is a countable part of F, and it is dense in F: let $y \in F$; there exists, for all $k \geqslant 1$, an integer $n \geqslant 1$ such that $d(y, x_n) \leqslant 1/k$; we therefore have $y \in B(x_n, 1/k)$; therefore $F \cap B(x_n, 1/k) \neq \emptyset$, and $y_{n,k} \in F \cap B(x_n, 1/k)$; then $d(y, y_{n,k}) \leqslant d(y, x_n) + d(x_n, y_{n,k}) \leqslant 2/k$. $\square$

2.4.2 Orthonormal systems

We will assume in the following that H is a pre-Hilbert space, of **infinite dimension**.

Definition 2.4.2. *Let $(u_i)_{i\in I}$ be a family of elements of H, indexed by an arbitrary set I, non-empty. We say that it is an **orthonormal family**, or an **orthonored** system, if:*

1)$\|u_i\| = 1, \forall i \in I$;

2)$u_i \perp u_j, \forall i \neq j$.

Note that every subsystem $(u_i)_{i\in J}$ $(J \subseteq I)$ of an orthonormal system $(u_i)_{i\in I}$ is still orthonormal.

Exemples. 1) In ℓ_2, the sequence $(e_n)_{n\geq 1}$ is orthonormal.

2) In $L^2(0,1)$, we put:

$$\boxed{e_n(t) = e^{2\pi int}}, \quad n \in \mathbb{Z};$$

the system $(e_n)_{n\in\mathbb{Z}}$ is orthonormal; we say that it is the **trigonometric system**.

Proposition 2.4.3. *If the finite system $(u_1, ..., u_n)$ is orthonormal, then, for all $a_1, ..., a_n \in \mathbb{K}$:*

$$\left\| \sum_{k=1}^{n} a_k u_k \right\|^2 = \sum_{k=1}^{n} |a_k|^2$$

Proof. Just develop using Proposition II.1.3:

$$\left\| \sum_{k=1}^{n} a_k u_k \right\|^2 = \sum_{k=1}^{n} \|a_k u_k\|^2 + \sum_{k\neq j} \langle a_k u_k \mid a_j u_j \rangle ,$$

and use that $\|a_k u_k\| = |a_k| \|u_k\| = |a_k|$ and that, for $k \neq j$, $\langle a_k u_k \mid a_j u_j \rangle = a_k \bar{a}_j \langle u_k \mid u_j \rangle = 0$ $\square$

Corollary 2.4.1. *Any orthonormal family is <u>free</u> (that is to say that the vectors composing it are linearly independent).*

Proposition 2.4.4 (Bessel inequality)**.** *Let H be a pre-Hilbert space. For any orthonormal family $(u_i)_{i\in I}$ in H, we have, for all $x \in H$:*

$$\boxed{\sum_{i\in I} |\langle x \mid u_i\rangle|^2 \leqslant \|x\|^2.}$$

In the inequality above, the sum on the first member is defined as follows: if $(a_i)_{i\in I}$ is a family of positive real numbers, then:

$$\sum_{i\in I} a_i \overset{\text{def}}{=} \sup_{J\subseteq I, J \text{ finite}} \sum_{i\in J} a_i$$

If $\ell_2(I) = \left\{(a_i)_{i\in I} \in \mathbb{K}^I; \sum_{i\in I} |a_i|^2 < +\infty\right\}$, Bessel's inequality leads to an application:

$$S: \quad H \longrightarrow \ell_2(I)$$
$$x \longmapsto (< x|u_i >)_{i\in I}$$

it is linear, and Bessel's inequality further says that it is continuous, and of norm $\leqslant 1$.

Proof. If $\xi_i = \langle x \mid u_i\rangle$, we have, since the family is orthonormal, for any finite part J of I:

$$0 \leqslant \left\| x - \sum_{i\in J} \xi_i u_i \right\|^2 = \|x\|^2 - 2\sum_{i\in J} \operatorname{Re}\langle x \mid \xi_i u_i\rangle + \sum_{i\in J} |\xi_i|^2$$

which gives the result car $\langle x \mid \xi_i u_i\rangle = \bar{\xi}_i \langle x \mid u_i\rangle = \bar{\xi}_i \xi_i = |\xi_i|^2$. □

Proposition 2.4.5. *Let H be a pre-Hilbert space and let $(u_n)_{n\geq 1}$ be an orthonormal sequence in H. If a vector $x \in H$ can be written $x = \sum_{n=1}^{\infty} \xi_n u_n$, then we necessarily have* $\boxed{\xi_n =< x|u_n >}$ *for all $n \geq 1$.*

Here sequence means countable family.

Proof. For each $k \geqslant 1$, the linear form Φ_{u_k} is continuous; So :

$$\langle x \mid u_k\rangle = \Phi_{u_k}(x) = \sum_{n=1}^{\infty} \Phi_{u_k}(\xi_n u_n) = \sum_{n=1}^{\infty} \xi_n \langle u_n \mid u_k\rangle = \xi_k.$$

□

Proposition 2.4.6. *Let $(u_n)_{n\geqslant 1}$ be an orthonormal sequence and $x = \sum_{n=1}^{\infty} \xi_n u_n$. Let F_n be the vector subspace generated by $u_1, \ldots, u_n$. So :*

$$P_{F_n}(x) = \sum_{k=1}^{n} \xi_k u_k$$

Proof. As we have $\xi_k = \langle x \mid u_k \rangle$, by the previous proposition, we obtain that $\langle x - \sum_{k=1}^{n} \xi_k u_k \mid u_j \rangle = 0$ for all $j \leqslant n$; so if $y_n = \sum_{k=1}^{n} \xi_k u_k$, we have $x - y_n \in F_n^{\perp}$. Like $y_n \in F_n$, the characterization of Theorem II.2.4 says that $y_n = P_{F_n}(x)$. □

Proposition 2.4.7. *If H is un Hilbert space, and $(u_n)_{n\geqslant 1}$ is an orthonormal sequence in H, then, for every sequence $(\xi_n)_{n\geqslant 1} \in \ell_2$, the series $\sum_{n=1}^{\infty} \xi_n u_n$ converges in H.*

In other words (using Proposition 2.4.5), the linear map continues:

$$S: \quad H \longrightarrow \ell_2$$

$$x \longmapsto (< x|u_n >)_{n\geq 1}$$

is surjective.

Proof. Just note note that the series satisfies the Cauchy criterion, because the Proposition 2.4.3 gives:

$$\| \sum_{k=n}^{n+p} \xi_k u_k \|^2 = \sum_{k=n}^{n+p} |\xi_n|^2 \underset{n\to\infty}{\longrightarrow} 0,$$

uniformly in p □

2.4.3 Orthonormal bases

Definition 2.4.3. *We say that an orthonormal sequence $(u_n)_{n\geq 1}$ in a pre-Hilbert space H is an* ***orthonormal basis*** *of H if the set $u_n; n \geq 1$ is total in H. We also say that $(u_n)_{n\geq 1}$ is a* ***Hilbert basis****.*

Note that, as we have restricted ourselves to taking countable families, the space H will necessarily be separable.

On the other hand, it should be noted that this notion of orthonormal base is, in infinite dimension, different from the notion of base, in the algebraic sense of the term: a family of vectors of a vector space is a base if any vector can s 'write, uniquely, as a linear combination of an finite number of terms of the family; but the following theorem says that, for an orthonormal base, any element is written as the sum of an series, which involves all the terms of the orthonormal base.

Theorem 2.4.1. *Let H be a pre-Hilbert space and let $(u_n)_{n\geq 1}$ be an orthonormal basis of H. Then, any element $x \in H$ is written:*

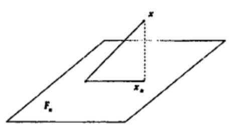
$$\boxed{x = \sum_{n=1}^{\infty} \xi_n u_n}, \quad \textit{with} \quad \boxed{\xi_n = \langle x \mid u_n \rangle}$$

*Moreover, for all $x, y \in H$, we have the **Parseval formulas**:*

1) $\boxed{\|x\|^2 = \sum_{n=1}^{\infty} |\langle x \mid u_n \rangle|^2}$;

2) $\boxed{< x \mid y >= \sum_{n=1}^{\infty} \langle x \mid u_n \rangle \overline{\langle y \mid u_n \rangle}}$ *, the series absolutely converges.*

Proof. Let F_n denote the vector subspace generated by $u_1, \ldots, u_n$, and let $x_n = P_{F_n}(x)$.



The set $u_n; n \geq 1$ being total, the subspace $\cup_{n\geq 1} F_n$ is dense in H; then, the sequence $(F_n)_{n\geq 1}$ being increasing, we have:

$$\|x - x_n\| = \text{dist}\,(x, F_n) \underset{n\to\infty}{\longrightarrow} 0.$$

On the other hand, according to the Corollary 2.4.1, $\{u_1, \ldots, u_n\}$ is a base, in the usual sense, of F_n; and, by Proposition 2.4.5, we therefore have:

$$x_n = \sum_{k=1}^{n} \langle x_n \mid u_k \rangle\, u_k.$$

But $(x - x_n) \in F_n^{\perp}$; therefore, for $k \leqslant n, \langle x_n \mid u_k \rangle = \langle x \mid u_k \rangle = \xi_k$ does not depend on n. We therefore have:

$$x = \lim_{n \to \infty} \sum_{k=1}^{n} \xi_k u_k = \sum_{k=1}^{\infty} \xi_k u_k.$$

Likewise $y = \sum_{k=1}^{\infty} \zeta_k u_k$, with $\zeta_k = \langle y \mid u_k \rangle$. Then, by continuity (Corollary 2.2.2):

$$< x \mid y >= \left\langle \sum_{k=1}^{\infty} \xi_k u_k \mid y \right\rangle = \sum_{k=1}^{\infty} \xi_k \langle u_k \mid y \rangle = \sum_{k=1}^{\infty} \xi_k \overline{\zeta_k},$$

which gives the other identity when $y = x$. □

It follows from Theorem 2.4.1 and Proposition 2.4.7 that we have:

Corollary 2.4.2. *Let H be a separable Hilbert space, and let $(u_n)_{n\geq 1}$ be an orthonormal basis of H. Then the linear map:*

$$S: \quad H \longrightarrow \ell_2$$
$$x \longmapsto (< x|u_n >)_{n\geq 1}$$

is an isomorphism of Hilbert spaces, that is to say an isomorphism preserving the inner product:$(S(\xi)|S(\zeta)) = (\xi|\zeta)$ for all $\xi, \zeta \in \ell_2$.

It is in particular an isometry $\|S(x)\| = \|x\|$ for all $x \in H$. When H is not complete, we always have an isometry preserving the inner product, but it is not surjective.

The reciprocal isomorphism is:

$$S^{-1}: \quad \ell_2 \longrightarrow H$$
$$(\xi_n)_{n\geq 1} \longmapsto \sum_{n=1}^{\infty} \xi_n u_n$$

We will see that in fact every separable Hilbert space has orthonormal bases, and therefore the previous corollary applies to all separable Hilbert spaces.

2.4.4 Existence of orthonormal bases

Theorem 2.4.2. *Every separable Hilbert space has orthonormal bases.*

In fact, completeness is not useful here (because at each step, we only work in vector subspaces of finite dimension, therefore complete).

We obtain, as a consequence of Theorem 2.4.2 and of the Corollary 2.4.2, the following essential result, in which, this time the hypothesis of completeness cannot be omitted.

Theorem 2.4.3. *All separable Hilbert spaces, of infinite dimension, are isomorphic to each other, and in particular to* ℓ_2 .

Proof of the Theorem 2.4.2 . We simply use the Gram-Schmidt orthonormalization process.

Consider a countable part $\{v_n; n \geq 1\}$ total. We can assume that the $v_n, n \geq 1$, are linearly independent (by removing those which are a linear combination of the previous ones).

Let F_n be the vector subspace generated by $v_1, \ldots, v_n$. We set $u_1 = \frac{v_1}{\|v_1\|}$, and

$$u'_{n+1} = P_{F_n^{\perp}}(v_{n+1}), \quad u_{n+1} = \frac{u'_{n+1}}{\|u'_{n+1}\|}$$

Then the sequence $(u_n)_{n\geqslant 1}$ is orthonormal, and the set $\{u_n; n \geqslant 1\}$ is total because the vector subspace generated by $u_1, \ldots, u_n$ is F_n. Indeed, by Theorem 2.3.2, for $2 \leqslant k \leqslant n$, we have $u'_k - v_k \in F_{k-1}^{\perp\perp} = F_{k-1}$, and therefore $u'_k \in F_k$ since $v_k \in F_k$ and $F_{k-1} \subseteq F_k$.

2.5 Separability of $L^2(0,1)$

2.5.1 Stone-Weierstrass theorem

It is a density theorem in the space $\mathscr{C}_{\mathbb{R}}(K)$ or $\mathscr{C}_{\mathbb{C}}(K)$ of functions continues $f : K \to \mathbb{R}$ or $\mathbb{C}$, where K is a compact space. Depending on whether the space is real or complex, it is not stated in the same way: a hypothesis must be added in the complex case.

Real case

Theorem 2.5.1 (Stone-Weierstrass theorem, real case)**.** *Let K be an compact space and A be a* ***subalgebra*** *of the real Banach algebra $\mathscr{C}_{\mathbb{R}}(K)$.*

We further assume that:

a) A separates the points of K;

b) A contains constants.

Then A is dense in $\mathscr{C}_{\mathbb{R}}(K)$.

Remarks.

1) A subalgebra of $\mathscr{C}(K)$ is a multiplication-stable vector subspace.

2) Saying that A separates the points of K means that if $x, y \in K$ are distinct, then there exists $f \in A$ such that $f(x) \neq f(y)$.

3) The assumption that A contains constant functions is only made to eliminate the case of subalgebras $A = \{f \in \mathscr{C}(K); f(a) = 0\}$ for a given $a \in K$.

Note that, A being a vector subspace, A contains the constants if and only if $\mathbf{1} \in A$. We obtain the following immediate consequence.

Theorem 2.5.2. *Let K be a compact part of R^d; then the set $\mathscr{P}_{\mathbb{R}}(K)$ of all* ***real polynomials*** *with d variables, restricted to K, is dense in $\mathscr{C}_{\mathbb{R}}(K)$.*

Theorem 2.5.3. *The real space $L^2_{\mathbb{R}}(0,1)$ is* ***separable****.*

Proof. We know that $\mathscr{C}_R([0,1])$ is dense in $L^2_R(0,1)$. On the other hand, Theorem II.4.2 tells us that $\mathscr{P}_{\mathbf{R}}([0,1])$ is dense in $\mathscr{C}_{\mathbf{R}}([0,1])$. So $\mathscr{P}_{\mathbf{R}}([0,1])$ is dense in $L^2_{\mathbb{R}}(0,1)$, because the norm uniform on $\mathscr{C}_{\mathbf{R}}([0,1])$ is finer than the norm of $L^2_{\mathbb{R}}(0,1)$: for all $f \in L^2_{\mathbf{R}}(0,1)$ and all $\varepsilon > 0$, there exists $g \in \mathscr{C}_{\mathbf{R}}([0,1])$ such that $\|f-g\|_2 \leqslant \varepsilon/2$; then there exists $p \in \mathscr{P}_{\mathbb{R}}([0,1])$ such that $\|g-p\|_\infty \leqslant \varepsilon/2$; but then $\|g-p\|_2 \leqslant \|g-p\|_\infty \leqslant \varepsilon/2$, and therefore $\|f-p\|_2 \leqslant \varepsilon$.

It only remains to notice that $\mathscr{P}_{\mathbf{R}}([0,1])$ is generated by the sequence defined by:

$$p_0(t) = 1, \quad p_1(t) = t, \quad p_2(t) = t^2, \quad , \quad p_n(t) = t^n,$$

to obtain the separability of $L^2_{\mathbb{R}}(0,1)$.

Note by the way that, we have proven the separability of $\mathscr{C}_{mathbbR}([0,1])$. □

Corollary 2.5.1. *$L^2_{\mathbb{R}}(0,1)$ is isomorphic to the real space ℓ_2 .*

This is the theorem demonstrated by Fisher and Riesz in 1907. The essential point being the fact that $L^2_{\mathbb{R}}(0,1)$ is complete.

Proof of the Stone-Weierstrass Theorem.

It is done in several stages.

Step 1. There exists a sequence of real polynomials $(r_n)_{n\geqslant 0}$ which converges uniformly on $[0,1]$ to the square root function $r : t \mapsto \sqrt{t}$.

Proof. We define $(r_n)_{n\geqslant 0}$ by induction, starting from $r_0 = 0$ and setting, for all $n \geqslant$0:

$$r_{n+1}(t) = r_n(t) + \frac{1}{2}\left(t - [r_n(t)]^2\right).$$

It is clear, by induction, that the r_n are polynomials. Moreover, for all $n \geqslant 0$, we have $0 \leqslant r_n(t) \leqslant \sqrt{t}$; indeed, by induction: we have, on the one hand, $t - [r_n(t)]^2 \geqslant 0$ and therefore $r_{n+1}(t) \geqslant r_n(t) \geqslant 0$, and on the other hand:

$$\sqrt{t} - r_{n+1}(t) = \left[\sqrt{t} - r_n(t)\right]\left[1 - \frac{1}{2}\left(\sqrt{t} + r_n(t)\right)\right] \geqslant 0$$

because $\sqrt{t} + r_n(t) \leqslant \sqrt{t} + \sqrt{t} = 2\sqrt{t} \leqslant 2$. Note that in passing, we saw that the sequence $(r_n)_{n\geqslant 0}$ is increasing.

Being increasing and increasing, it converges towards a limit $r(t)$. The recurrence relation shows that $r(t) = \sqrt{t}$. It remains to be seen that there is uniform convergence. First method: by hand. Let us set $\varepsilon_n(t) = \sqrt{t} - r_n(t)$. We saw above, since $r_n(t) \geqslant 0$, that:

$$0 \leqslant \varepsilon_{n+1}(t) = \varepsilon_n(t)\left[1 - \frac{1}{2}\left(\sqrt{t} + r_n(t)\right)\right] \leqslant \varepsilon_n(t)\left(1 - \frac{\sqrt{t}}{2}\right)$$

SO :

$$
\begin{aligned}
0 \leqslant \varepsilon_n(t) \leqslant \varepsilon_0(t)\left(1-\frac{\sqrt{t}}{2}\right)^n &= \sqrt{t}\left(1-\frac{\sqrt{t}}{2}\right)^n \\
&\leqslant \sup_{0 \leqslant x \leqslant 1/2} 2(1-x)x^n \quad (\text{ pose } x = 1-\sqrt{t}/2) \\
&= 2x_n\left(1-x_n\right)x_n^n \quad \text{ with } x_n = n/(n+1) \\
&= \frac{2}{n+1}x_n^n \leqslant \frac{2}{n+1}.
\end{aligned}
$$

□

Second method. Just use the following theorem.

Theorem 2.5.4 (Dini's theorem). *Let K be a compact space.*

If $(u_n)_{n\geq 1}$ is an increasing sequence of continuous functions $u_n : K \to \mathbb{R}$ which simply converges to a continuous function $u : K \to \mathbb{R}$, the convergence is uniform.

This is of course obviously false if we do not assume the continuous limit.

Proof. Let $\varepsilon > 0$.

For each $x \in K$, there exists an integer $N(x)$ such that:

$$n \geqslant N(x) \quad \Longrightarrow \quad 0 \leqslant u(x) - u_n(x) \leqslant \varepsilon/3.$$

As u and $u_{N(x)}$ are continuous, there exists a neighborhood of x, which can be taken to be open, such that:

$$x' \in V(x) \Longrightarrow \begin{cases} |u(x) - u(x')| \leqslant \varepsilon/3 \\ \left|u_{N(x)}(x') - u_{N(x)}(x)\right| \leqslant \varepsilon/3. \end{cases}$$

As K is compact, there exists $x_1, \ldots, x_m \in K$ such that:

$$K = \bigcup_{i=1}^{m} V(x_i).$$

If $N = \max\{N(x_1), \ldots, N(x_m)\}$, we have, for $n \geqslant N$:

$$0 \leqslant u(x) - u_n(x) \leqslant \varepsilon, \quad \forall x \in K,$$

because x belongs to one of $V(x_i)$ and $n \geqslant N(x_i)$; So :

$$\begin{aligned}0 \leqslant u(x)-u_n(x) &\leqslant u(x)-u_{N(x_i)}(x)\\ &\leqslant (u(x)-u(x_i))+\Big(u(x_i)-u_{N(x_i)}(x_i)\Big)+\Big(u_{N(x_i)}(x_i)-u_{N(x_i)}(x)\Big)\\ &\leqslant \frac{\varepsilon}{3}+\frac{\varepsilon}{3}+\frac{\varepsilon}{3}=\varepsilon.\end{aligned}$$

□

<u>2nd step</u>. If $f \in A$, then $|f| \in \bar{A}$.

Proof. Indeed, we can assume $f \neq 0$. Let $a = \|f\|_\infty$. We have $[f(x)]^2/a^2 \in [0,1]$ for all $x \in K$. But, since r_n is a polynomial, and A is an algebra, we have $r_n(f^2/a^2) \in A$ if $f \in A$. Passing to the limit, we obtain:

$$|f| = a \lim_{n\to\infty} r_n\Big(f^2/a^2\Big) \in \bar{A}$$

the limit being uniform, that is to say taken for the norm of $\mathscr{C}_{\mathbb{R}}(K)$. □

<u>Step 3</u>. If $f, g \in A$, then $\max\{f,g\}, \min\{f,g\} \in \bar{A}$.

Proof. It is enough to note that:

$$\begin{cases}\max\{f,g\} = \frac{1}{2}(f+g+|f-g|)\\ \min\{f,g\} = \frac{1}{2}(f+g-|f-g|)\end{cases}$$

and use Step 2 (as well as the fact that $\bar{A}$ is a vector subspace). □

<u>Step 3a</u>. If $f, g \in \bar{A}$, then $\max\{f,g\}, \min\{f,g\} \in \bar{A}$.

Proof. This results from the fact that $\bar{A}$ satisfies the conditions required for A: it remains a subalgebra (recall that the convergence in $\mathscr{C}(K)$ is the uniform convergence), and, since A contains the constants and separates the points of K, it is *a* fortiori the same for $\bar{A}$. □

Of course, by recurrence:

$$f_1, \dots, f_n \in \bar{A} \quad \Longrightarrow \quad \max\{f_1, \dots, f_n\} \in \ barA \quad \text{and} \quad \min\{f_1, \dots, f_n\} \in \bar{A}.$$

Step 4 . If $x, y \in K$ and $x \neq y$, then:

$$(\forall \alpha, \beta \in \mathbb{R}) \quad (\exists h \in A) \quad h(x) = \alpha \quad \text{and} \quad h(y) = \ beta$$

This is the first step in the approximation: we can obtain with a function of A values given at two given points distinct from K.

Proof. As A separates the points, there exists $g \in A$ such that $g(x) \neq g(y)$. Assume:

$$h = \alpha \mathbb{I} + \frac{\beta - \alpha}{g(y) - g(x)}(g - g(x)\mathbb{I}).$$

We have $h(x) = \alpha, h(y) = \beta$, and $h \in A$, because $g \in A, \mathbb{I} \in A$, and A is a vector subspace. □

Step 5. For all $f \in \mathscr{C}(K)$, for all $x \in K$, and all $\varepsilon > 0$, there exists $g \in \bar{A}$ such that:

$$g(x) = f(x) \quad \text{and} \quad g(y) \leqslant f(y) + \varepsilon, \forall y \in K.$$

Proof. For all $z \in K$ such that $z \neq x$, it exists, by Step 4, taking $\alpha = f(x)$ and $\beta = f(z)$, a $h_z \in A$ such that $h_z(x) = f(x)$ and $h_z(z) = f(z)$.

Let h_x denote the constant function equal to $f(x)$ I. Then:

$$(\forall z \in K) \quad h_z(x) = f(x) \quad \text{and} \quad h_z(z) = f(z).$$

The continuity of f and that of h_z give a neighborhood, which can be taken to be open, V_z of z such that:

$$y \in V(z) \quad \Longrightarrow \quad h_z(y) \leqslant f(y) + \varepsilon.$$

As K is compact, there exists a finite number of elements $z_1, \ldots, z_m \in K$ such that:

$$K = V(z_1) \cup \cdots \cup V(z_m).$$

Then $g = \inf\{h_{z_1}, \ldots, h_{z_m}\} \in \bar{A}$, by Step 3a, and l 'we have, for all $y \in K$: $g(y) \leqslant f(y) + \varepsilon$, since y belongs to one of $V(z_i)$. □

Step 6. We have $\bar{A} = \mathscr{C}_{\mathbf{R}}(K)$.

Proof. Let $f \in \mathscr{C}_{\mathrm{R}}(K)$, and let $\varepsilon \geq 0$.

For all $x \in K$, there exists $g_x \in \bar{A}$ satisfying the conditions given in Step 5.

The continuity of f and that of g_x give a neighborhood, which we can choose to be open, $U(x)$ of x such that:

$$y \in U(x) \quad \Longrightarrow \quad g_x(y) \geqslant f(y) - \varepsilon.$$

The compactness of K makes it possible to find a finite number of elements $x_1, \ldots, x_p \in K$ such that:

$$K = U(x_1) \cup \cdots \cup U(x_p).$$

Then $\varphi = \max\{g_{x_1}, \ldots, g_{x_p}\} \in \bar{A}$, thanks to Step 3a; and she checks:

$$f(y) - \varepsilon \leqslant \varphi(y) \leqslant f(y) + \varepsilon, \quad \forall y \in K,$$

because each $y \in K$ is in one of $U(x_j)$.

This means that $\|f - \varphi\|_\infty \leqslant \varepsilon$.

As $\varepsilon > 0$ was arbitrary, we have $f \in \overline{(\bar{A})} = \bar{A}$.

This completes the proof of Theorem 2.5.1. □

Bernstein's proof for a compact interval of $\mathbb{R}$

The general form of the Stone-Weierstrass Theorem was given by Stone in 1948. Originally, Weierstrass had shown, in 1885, that any continuous function on a closed bounded interval of $\mathbb{R}$ could be approximated there. uniformly by polynomials. To do this, he used a convolution product.

In 1913, Bernstein gave a beautiful probabilistic proof, which we will present below. Let us first note that, by changing the variable, we can assume that the interval in question is [0, 1]. The initial idea is as follows: we fix $t \in [0,1]$ (also, if we want, we can take only $0 < t < 1$), and we consider independent random variables $X_1, \ldots, X_n$ all following Bernoulli's law with parameter t. Then $S_n = X_1 + \cdots + X_n$ follows the binomial law $\mathscr{B}(n,t)$ with parameters n and t. The weak law of large numbers says that $\frac{S_n}{n} \underset{n\to\infty}{\longrightarrow} t = \mathbb{E}\left(X_1\right)$ in probability. Then, for any function f continuous on $[0,1]$, we have $\mathbb{E}\left[f\left(\frac{S_n}{n}\right)\right] \underset{n\to\infty}{\longrightarrow} f(t)$. Indeed, if $\varepsilon > 0$ is given, the uniform continuity of f on $[0,1]$ makes it possible to find $\delta > 0$ such that $|f(x) - f\left(x'\right)| \leqslant \varepsilon$ for $|x - x'| \leqslant \delta$; convergence in probability then gives a $N \geqslant 1$ such that $\mathbb{P}\left(\left|\frac{S_n}{n} - t\right| > \delta\right)$ *leqslant*ε if $n \geqslant N$. Then, for $n \geqslant N$, we have:

$$\begin{aligned}\left|\mathbb{E}\left[f\left(\frac{S_n}{n}\right)\right] - f(t)\right| &= \int_{\left\{\left|\frac{s_n}{n}-t\right|>\delta\right\}} \left|f\left[\frac{S_n(\omega)}{n}\right] - f(t)\right| d\mathbb{P}(\omega) \\ &+ \int_{\left\{\left|\frac{s_n}{n}-t\right|\leqslant\delta\right\}} \left|f\left[\frac{S_n(\omega)}{n}\right] - f(t)\right| d\mathbb{P}(\omega) \\ &\leqslant 2\|f\|_\infty \varepsilon + \varepsilon\end{aligned}$$

Or $\mathbb{E}\left[f\left(\frac{S_n}{n}\right)\right] = \sum_{k=0}^{n} C_n^k t^k (1-t)^{n-k} f\left(\frac{k}{n}\right)$. We set:

$$\left[B_n(f)\right](t) = \sum_{k=0}^{n} C_n^k t^k (1-t)^{n-k} f\left(\frac{k}{n}\right);$$

It is a polynomial of degree n. It is called the n^{th} Bernstein polynomial of f.

We have just seen that we have a simple convergence of $B_n(f)$ to f.
We will see that, thanks to a uniform estimation of the variance of Bernoulli variables, the proof of the weak law of large numbers for these variables allows to obtain the uniform convergence of $B_n(f)$ towards f.

Let us first recall that if X is a random variable following Bernoulli's law

of parameter t, then its variance is $Var(X) = t(1-t)$. We have, by the Bienaymé-

Tchebychev inequality, for all $\delta > 0$:

$$\begin{aligned}
\mathbb{P}\left(\left|\frac{S_n}{n} - t\right| > \delta\right) &= \mathbb{P}\left(\left|\frac{S_n}{n} - \mathbb{E}(X)\right| > \delta\right) = \mathbb{P}\left(\left|\frac{S_n}{n} - \ mathbbE\left(\frac{S_n}{n}\right)\right| > \delta\right) \\
&\leqslant \frac{1}{\delta^2}\operatorname{Var}\left(\frac{S_n}{n}\right) = \frac{1}{n^2\delta^2}\operatorname{Var}(S_n) \\
&= \frac{1}{n^2\delta^2}\sum_{j=1}^{n}\operatorname{Var}(X_j) \qquad \text{(by independence)} \\
&= \frac{\text{Var(X)}}{n\delta^2} = \frac{t(1-t)}{n\delta^2} \leqslant \ frac1/4n\delta^2.
\end{aligned}$$

Consider the continuity modulus of f, defined by:

$$\omega_f(h) = \sup\left\{\left|f(t) - f(t')\right|; |t - t'| \leqslant h\right\}.$$

Saying that f is uniformly continuous means that $\omega_f(h) \underset{h\to 0}{\longrightarrow} 0$. Let's set a $\delta > 0$, which we will specify later. We have, for all $t \in [0,1]$ (we will be careful to differentiate the occurrence $\omega \in \Omega$ from the continuity module ω_f; we could have modified these notations , but these are the ones usually used!):

$$\begin{aligned}
|f(t) - [B_n(f)](t)| &= \left|\mathbb{E}\left[f(t) - f\left(\frac{S_n}{n}\right)\right]\right| \\
&\leqslant \mathbb{E}\left(\left|f(t) - f\left(\frac{S_n}{n}\right)\right|\right) = \int_\Omega \left|f(t) - f\left(\frac{S_n(\omega)}{n}\right)\right| d\mathbb{P}(\omega) \\
&= \int_{\left\{\left|t - \frac{s_n(\omega)}{n}\right| \leqslant \delta\right\}} + \int_{\left\{\left|t - \frac{s_n(\omega)}{n}\right| > \delta\right\}} \\
&\leqslant \omega_f(\delta) + 2\|f\|_\infty \mathbb{P}\left(\left|t - \frac{S_n}{n}\right| > \delta\right) \\
&\leqslant \omega_f(\delta) + 2\|f\|_\infty \frac{1}{4n\delta^2}
\end{aligned}$$

For any given $\varepsilon > 0$, let's now choose δ so that $\omega_f(\delta) \leqslant \varepsilon/2$, then $N \geqslant 1$ such that $\|f\|_\infty \frac{1}{2N\delta^2} \leqslant \varepsilon/2$. We will have, for $n \geqslant N$, $\mid f(t) - [B_n(f)]\, t) \mid \leqslant \varepsilon$ for all $t \in [0,1]$, which proves that $B_n(f)$ tends uniformly towards f.

2.5.2 Complex case

As it stands, the statement of Theorem 2.5.1 is false for complex-valued function spaces. For example, if K is the closed unit disk $\bar{\mathbb{D}}$ of the complex plane, any uniform limit on K of polynomials p_n, is holomorphic in the open disk $\mathbb{D}$, thanks to the Weierstrass Theorem on the uniform convergence of sequences of holomorphic functions. The adherence of the algebra of polynomials is therefore not $\mathscr{C}_{\mathbb{C}}(K)$ for example, the function $z \to \bar{z}$ n It's not in there. In fact, this example is essentially the only case that needs to be considered; in fact, we have:

Theorem 2.5.5 (Stone-Weierstrass theorem, complex case)**.** *Let K be a compact space and let A be a complex subalgebra of the complex Banach space $\mathscr{C}_{\mathbb{C}}(K)$. If:*

a) A separates the points of K;

b) A contains constant functions;

c) A is stable by conjugation: $f \in A \implies \bar{f} \in A$,

then A is dense in $\mathscr{C}_{\mathbb{C}}(K)$.

Note that here $\bar{f}$ denotes the function $t \in K \mapsto f(\bar{t}) \in \mathbb{C}$, where $f(\bar{t})$ is the conjugate complex number of $f(t)$.

Proof. Condition c) allows us to say that:

$$f \in A \implies Ref = \frac{f+\bar{f}}{2} \in A \quad and \quad Imf = \frac{f-\bar{f}}{2i} \in A.$$

Either :

$$A_{\mathbb{R}} = \{f \in A; f(t) \in \mathbb{R}, \forall t \in K\}.$$

The above remark allows us to say that:

$$A = A_{\mathbb{R}} + iA_{\mathbf{R}}.$$

Furthermore, $A_{\mathbb{R}}$ is a subalgebra of $\mathscr{C}_{\mathbb{R}}(K)$, which contains the constant (real) functions, and separates the points of K : if $u \neq v$, there exists $f \in A$ such that $f(u) \neq f(v)$; but then $\operatorname{Re} f(u) \neq \operatorname{Re} f(v)$ or $\operatorname{Im} f(u) \neq \operatorname{Im} f(v)$, and $\operatorname{Re} f, \operatorname{Im} f \in A_{\mathbb{R}}$. It follows from the real case that $A_{\mathbb{R}}$ is dense in $\mathscr{C}_{\mathbb{R}}(K)$. But then, $A = A_{\mathbb{R}} + iA_{\mathbb{R}}$ is dense in $\mathscr{C}_{\mathbf{C}}(K) = \mathscr{C}_{\mathbb{R}}(K) + i\mathscr{C}_{\mathbb{R}}(K)$. □

Exemple. Let K be a compact part of $\mathbb{C}$. The set of polynomials, with complex coefficients, in the two variables z and $\bar{z}$ is dense in $\mathscr{C}_{\mathbf{C}}(K)$.

Note that it is also, by identifying $\mathbb{C}$ with $\mathbb{R}^2$, the set of polynomials, with complex coefficients, in the two real variables x and y, by identifying $z = x + iy \in \mathbb{C}$ with$(x, y) \in \mathbb{R}^2$.

2.5.3 The trigonometric system

We will consider here functions $f : \mathbb{R} \to \mathbb{C}$ **periodic**, with period 1 on $\mathbb{R}$.

The surjective application:

$$\begin{aligned} e_1 : \quad \mathbb{R} &\longrightarrow \mathbb{U} = \{z \in \mathbb{C}; |z| = 1\} \\ t &\longmapsto e^{2\pi i t} = u \end{aligned}$$

allows them to be identified with functions defined on $\mathbb{U}$. We can also identify them with the functions defined on the torus $\mathbb{T} = \mathbb{R}/\mathbb{Z}$.

Furthermore, we know that for any continuous function f on $\mathbb{R}$ of period 1, there exists a unique continuous function $\tilde{f} : \mathbb{U} \to \mathbb{C}$ such that $f = \tilde{f} \circ e_1$ (resp. $\ddot{f} : \mathbb{T} \to \mathbb{C}$ such that $f(x) = \ ddot f(x + \mathbb{Z}))$. The space $\mathscr{C}_1(\mathbb{R})$ of continuous functions on $\mathbb{R}$ of period 1, provided with the norm $\|f\|_\infty = \sup_{x \in \mathbb{R}} |f(x)|$, is therefore identified with the space $\mathscr{C}(\mathbb{U})$ of continuous functions on the compact $\mathbb{U}$. It is also identified with the subspace $\tilde{\mathscr{C}} = \{f \in \mathscr{C}([0, 1]); f(0) = f(1)\}$. These identifications are isometric since:

$$\sup_{x \in \mathbb{R}} |f(x)| = \sup_{x \in [0,1]} |f(x)| = \sup_{u \in \mathbb{U}} |\tilde{f}(u)| = \sup_{\xi \in \mathbb{T}} |\ddot{f}(\xi)|.$$

Definition 2.5.1. *We call any finite sum a trigonometric polynomial.*

$$\sum_{n=N_1}^{N_2} a_n \mathrm{e}^{2\pi int}$$

with $a_n \in \mathbb{C}$ *and* $N_1, N_2 \in \mathbb{Z}, N_1 \leqslant N_2$.

Note that by adding zero coefficients if necessary, we can always write a trigonometric polynomial in the symmetric form:

$$\sum_{n=-N}^{N} a_n e^{2\pi int},$$

where N is a positive integer.

We will note, for all $n \in Z$:

$$\boxed{e_n(t) = e^{2\pi int}}, \quad t \in \mathbb{R}$$

The set $e_n; n \in Z$ is called the **trigonometric system**.

Trigonometric polynomials are identified with the usual polynomials in u and $\bar{u}$ on $\mathbb{U}$, since all $u \in \mathbb{U}$ is written in the form $u = e_1(t) = e^{2\pi it}$, and then $u^n e^{2\pi int} = e_n(t)$, and that $\bar{u} = e^{-2\pi int} = e_{-n}(t)$. The complex Stone-Weierstrass theorem applied to $\mathscr{C}_{\mathbb{C}}(\mathbb{U})$ therefore gives:

Theorem 2.5.6. *The set of trigonometric polynomials is dense in the space of continuous functions of period 1 on* $\mathbb{R}$.

Now consider the space of measurable functions $f : \mathbb{R} \to \mathbb{C}$ of period 1 as :

$$\int_0^1 |f(t)|^2 dt < +\infty.$$

When quotiented by the subspace of negligible functions, this quotient is identified as $L^2(0,1) =$

$L^2_{\mathbb{C}}(0,1)$; indeed, for any measurable function $g:[0,1]\to\mathbb{C}$, the measurable function:

$$\tilde{g}: \quad [0,1] \longrightarrow \mathbb{C}$$
$$t \mapsto \begin{cases} g(t) \text{ if } 0 \leqslant t < 1; \\ g(0) \text{ if } t = 1 \end{cases}$$

extends by periodicity into a measurable function $f:\mathbb{R}\to\mathbb{C}$ of period 1, and $\int_0^1 |g(t)|^2 dt = \int_0^1 |f(t)|^2 dt$.

These identifications having been made, we can state:

Theorem 2.5.7. *The trigonometric system is an orthonormal base of* $L^2_{\mathbb{C}}(0,1)$

Corollary 2.5.2. *The real space* $L^2_{\mathbb{R}}(0,1)$ *has an orthonormal basis formed by the functions:*

$$1, \sqrt{2}\cos(2\pi t), \sqrt{2}\cos(4\pi t), \ldots, \sqrt{2}\cos(2\pi n t), \ldots$$
$$\sqrt{2}\sin(2\pi t), \sqrt{2}\sin(4\pi t), \ldots, \sqrt{2}\sin(2\pi n t), \ldots$$

Remarks 1) $\mathbb{Z}$ being countable, we could re-index the trigonometric system with positive integers.

2) The theorem means that, for all $f \in L^2_{\mathbf{C}}(0,1)$, we have:

$$\lim_{N\to\infty} \left\| f - \sum_{n=-N}^{N} \widehat{f}(n) e_n \right\|_2 = 0$$

where the inner products:

$$\boxed{\widehat{f}(n) = \langle f \mid e_n \rangle = \int_0^1 f(t)\overline{e_n(t)}dt = \ int_0^1 f(t)\mathrm{e}^{-2\pi i n t} dt},$$

for $n \in \mathbb{Z}$, are called the Fourier coefficients of f. Parseval's formula is then written $\boxed{\int_0^1 |f(x)|^2 dx = \sum_{n\in\mathbf{Z}} |\widehat{f}(n)|^2}$.

We know that <u>exists</u> then a strictly increasing sequence of integers $(l_n)_{n\geq 1}$ such as :

$$\lim_{n\to\infty} \sum_{k=-l_n}^{l_n} \widehat{f}(k)e^{2\pi ikt} = f(t)$$

for <u>almost all</u> $t \in [0,1]$.

Proof of theorem 2.5.7 It is first easy to see that $\{e_n; n \in \mathbb{Z}\}$ is orthonormal:

$$\langle e_n \mid e_p \rangle = \int_0^1 e^{2\pi int}\mathrm{e}^{-2\pi ipt}dt = \ int_0^1 \mathrm{e}^{2\pi i(n-p)t}dt = \begin{cases} 1 \text{ if } n = p \\ 0 \text{ if } n \neq p \end{cases}$$

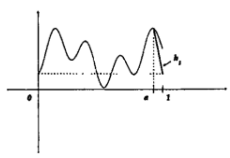
It is total because the trigonometric polynomials are dense in $\mathscr{C}(\mathbb{U})$ and $\|\cdot\|_\infty \geqslant \|\cdot\|_2$, using the following lemma:

Lemma 2.5.1. *The set $\mathscr{C}_1(\mathbb{R})$ of continuous functions on $\mathbb{R}$ of period 1, identified à$\tilde{\mathscr{C}} = \{f \in \mathscr{C}([0,1]); f(0) = f(1)\}$, is dense in $L^2(0,1)$.*

Indeed, if $f \in L^2(0,1)$, then there exists, for all $\varepsilon > 0, g \in \tilde{\mathscr{C}} \cong \mathscr{C}(\mathbb{U})$ such that $\|f - g\|_2 \leqslant \varepsilon/2$; there then exists a trigonometric polynomial p such that $\|g - p\|_\infty \leqslant \varepsilon/2$; but $\|g - p\|_2 \leqslant \|g - p\|_\infty \leqslant \varepsilon/2$; therefore $\|f - p\|_2 \leqslant \varepsilon$.

Proof of the lemma. Let $f \in L^2(0,1)$ and let $\varepsilon > 0$. We know (Corollary II.2.9) that there exists $h \in \mathscr{C}([0,1])$ such that $\|f - h\|_2 \leqslant \varepsilon/2$. Let $M > 0$ such that $|h(t)| \leqslant M$ for all $t \in [0,1]$, and let us note $a = 1 - (\frac{\varepsilon}{4M})^2$.

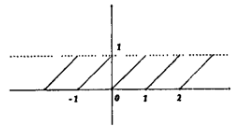
We will modify h on $[a,1]$ by setting $h_1(1) = h(0)$ and taking h_1 affine between a and 1 . Then $h_1 \in \tilde{\mathscr{C}}, \|h_1\|_\infty \leqslant M$, and:



$$\|h-h_1\|_2 = \left(\int_a^1 |h(t)-h_1(t)|^2\,dt\right)^{1/2}$$
$$\leqslant (1-a)^{1/2} \sup_{a\leqslant t\leqslant 1} (|h(t)|+|h_1(t)|)$$
$$\leqslant \frac{\varepsilon}{4M}\times(M+M) = \frac{\varepsilon}{2}.$$

We therefore have $\|f-h_1\|_2 \leq \varepsilon$.

Application example. Let $f:\mathbb{R}\to\mathbb{R}$ be the function defined by $f(t)=t$ for $0\leq t<1$, and extended by periodicity on $\mathbb{R}$.



Then $f\in L^2(0,1)$ and:

$$\|f\|_2^2 = \int_0^1 t^2 dt = \frac{1}{3}.$$

The Fourier coefficients of f are:

$$\widehat{f}(n) = \int_0^1 t\mathrm{e}^{-2\pi int}dt, \quad n\in\mathbb{Z}$$

For $n=0$: $\int_0^1 tdt = 1/2$; for $n\neq 0$:

$$\widehat{f}(n) = \left[\frac{t\mathrm{e}^{-2\pi int}}{-2\pi in}\right]_0^1 - \ int_0^1\frac{\mathrm{e}^{-2\pi int}}{-2\pi in}dt = \frac{1}{-2\pi in} = \frac{i}{2\pi n}.$$

Parseval's formula $\|f\|_2^2 = \sum_{n\in\mathbf{Z}} |\widehat{f}(n)|^2$ therefore gives:

$$\frac{1}{3} = \frac{1}{4} + \sum_{n\neq 0}\frac{1}{4\pi^2 n^2} = \frac{1}{4} + \frac{1}{2\pi^2}\sum_{n=1}^{\infty}\frac{1}{n^2};$$

from where :

$$\sum_{n=1}^{\infty}\frac{1}{n^2} = 2\pi^2\left(\frac{1}{3}-\frac{1}{4}\right) = \frac{\pi^2}{6}$$

Fourier coefficient of functions of $L^1(0,1)$

For any measurable $f : [0,1] \to \mathbb{C}$, the Cauchy-Schwarz inequality:

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

says that $\mathscr{L}^2([0,1]) \subseteq \mathscr{L}^1([0,1])$. We therefore have a natural injection of $L^2(0,1)$ into $L^1(0,1)$. By identifying $L^2(0,1)$ with its image in $L^1(0,1)$, we will write: $\boxed{L^2(0,1)\ subseteq L^1(0,1)}$

For any $f \in L^1(0,1)$, we can define the Fourier coefficients:

$$\boxed{\widehat{f}(n) = \int_0^1 f(t)\mathrm{e}^{-2\pi int} dt, \quad n \in \mathbb{Z}}$$

since $|\mathrm{e}^{-2\pi int}| = 1$. We have $|\widehat{f}(n)| \leqslant \|f\|_1$ for all $n \in \mathbb{Z}$. Moreover :

Theorem 2.5.8 (Riemann-Lebesgue Lemma). *For any function $f \in L^1(0,1)$, its Fourier coefficients tend to 0 when $|n|$tends to infinity:*

$$\boxed{\hat{f}(n) \underset{|n|\to\infty}{\longrightarrow} 0}$$

Proof. If $g \in L^2(0,1)$, Parseval's formula:

$$\|g\|_2^2 = \sum_{n\in\mathbf{Z}} |\widehat{g}(n)|^2$$

shows that we have, in particular, $\widehat{g}(n) \underset{|n|\to\infty}{\longrightarrow} 0$.

Now, if $f \in L^1(0,1)$, there exists, for all $\varepsilon > 0$, a function $g \in L^2(0,1)$ (by example g stepped, or g continuous) such that $\|f-g\|_1 \leqslant \varepsilon$. As we have $|\widehat{f}(n) - \widehat{g}(n)| \leqslant \|f-g\|_1 \leqslant \varepsilon$, we obtain $|\widehat{f}(n)| \leqslant |\widehat{g}(n)| + \varepsilon$; SO :

$$\limsup_{|n|\to\infty} |\widehat{f}(n)| \leqslant \limsup_{|n|\to\infty} |\widehat{g}(n)| + \varepsilon = \varepsilon$$

□

Hilbert space exhibits two phenomena (not occurring in Euclidean m-space) described in the examples below:

Example 2.1. *Consider the sequence* $\langle p_n \rangle$ *of points in Hilbert space where* $p_k = \langle a_{1k}, a_{2k}, \ldots \rangle$ *is defined by* $a_{ik} = \delta_{ik}$*; i.e.* $a_{ik} = 1$ *if* $i = k$*, and* $a_{ik} = 0$ *if* $i \neq k$*. Observe, as illustrated below, that the projection* $\langle \pi_i (p_n) \rangle$ *of* $\langle p_n \rangle$ *into each coordinate space converges to zero:*

$$
\begin{aligned}
p_1 =& \langle 1, 0, 0, 0, \ldots \rangle \\
p_2 =& \langle 0, 1, 0, 0, \ldots \rangle \\
p_3 =& \langle 0, 0, 1, 0, \ldots \rangle \\
p_4 =& \langle 0, 0, 0, 1, \ldots \rangle \\
& \ldots\ldots \\
& \ldots\ldots \\
& \cdot \downarrow \downarrow \\
0 =& \langle 0, 0, 0, 0, \ldots \rangle
\end{aligned}
$$

But the sequence $\langle p_n \rangle$ *does not converge to* $\mathbf{0}$*, since* $d(p_k, 0) = 1$ *for every* $k \in \mathbf{N}$*; in fact,* $\langle p_n \rangle$ *has no convergent subsequence.*

Example 2.2. *Let* $\mathscr{H}$ *denote the proper subspace of* $\mathbf{H}$ *which consists of all points in* $\mathbf{H}$ *whose first coordinate is zero. Observe that the function* $f : \mathbf{H} \to \mathscr{H}$ *defined by* $f(\langle a_1, a_2, \ldots \rangle) = \langle 0, a_1, a_2, \ldots \rangle$ *is one-one, onto and preserves distances. Hence Hilbert space is isometric to a proper subspace of itself.*

BIBLIOGRAPHY

[1] Daniel Li, Cours d'analyse fonctionnelle avec 200 exercices corrigés, Ellipse, 2013.

[2] Erwin Kreiszig, Introductory functional analysis with applications, John Wiley & Sons, 1978.

[3] Haiim Brezis, Analyse fonctionnelle, Theorie et applications, Masson, 1987.