

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 $\langle x | y \rangle$ the inner product of the vectors $x, y \in H$.

This means that the application:

$$\begin{aligned}\langle . | . \rangle: H \times H &\longrightarrow \mathbb{K} = \mathbb{R} \text{ or } \mathbb{C} \\ (x, y) &\longmapsto \langle x | y \rangle\end{aligned}$$

fulfilled the conditions:

- 1) for all $y \in H$, the map $x \in H \mapsto \langle x | y \rangle \in \mathbb{K}$ is a linear form;

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

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

3) for all $x \in H$, we have $\langle x | x \rangle \geq 0$ and $\langle x | x \rangle = 0$ 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 "1/2 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:

$$\langle x | y \rangle = x_1 y_1 + \dots + x_n y_n$$

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

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

$$\langle x | y \rangle = x_1 \bar{y}_1 + \dots + x_n \bar{y}_n$$

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

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:

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

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

$$\langle f | g \rangle = \int_S f g dm \text{ in the real case,}$$

And :

$$\langle f | g \rangle = \int_S f \bar{g} dm \text{ in the complex case.}$$

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

$$\langle x | y \rangle = \sum_{n=1}^{\infty} x_n y_n \text{ in the real case,}$$

And :

$$\langle x | y \rangle = \sum_{n=1}^{\infty} x_n \bar{y}_n \text{ in the complex case.}$$

for $x = (x_n)_{n \geq 1}, y = (y_n)_{n \geq 1} \in \ell_2$. 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 $1^2 + 1^2 + \dots$ does not converge, x does not belong to ℓ_2 . On the other hand, the series $1^2 + \left(\frac{1}{2}\right)^2 + \left(\frac{1}{4}\right)^2 + \dots$ does converge; hence y belongs to ℓ_2 .

2.2 Elementary properties

Notation. since $\langle x|x \rangle \geq 0$, we can put:

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

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

- a) $\|x+y\|^2 = \|x\|^2 + \|y\|^2 + 2\langle x|y \rangle$ (real case);
- b) $\|x+y\|^2 = \|x\|^2 + \|y\|^2 + 2\operatorname{Re} \langle x|y \rangle$ (complex case).

Proof. Just expand:

$$\|x+y\|^2 = \langle x+y | x+y \rangle = \langle x|x \rangle + \langle y|y \rangle + \langle x|y \rangle + \langle y|x \rangle,$$

and use the fact that $\langle x|y \rangle + \langle y|x \rangle = \langle x|y \rangle + \overline{\langle x|y \rangle} = 2\operatorname{Re} \langle x|y \rangle$ in the real case, and $= 2\operatorname{Re} \langle x|y \rangle$ in the complex case. $\square$

2.2.1 Cauchy-Schwarz inequality

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

$$|\langle x|y \rangle| \leq \|x\| \|y\|.$$

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

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

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 $\langle x|x \rangle = 0$ implies $x = 0$).

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

$$\langle e^{-i\theta} x | y \rangle = e^{-i\theta} \langle x | y \rangle \in \mathbb{R}_+$$

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

$$\|x'\|^2 + 2 \operatorname{Re} \langle x' | y \rangle t + \|y\|^2 t^2 = \|x' + ty\|^2 \geq 0.$$

If $\|y\| = 0$, we have $\|x'\|^2 + 2 \operatorname{Re} \langle x' | y \rangle t \geq 0$ for all $t \in \mathbb{R}$; this is only possible if $\operatorname{Re} \langle x' | y \rangle = 0$.
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:

$$\operatorname{Re} \langle x' | y \rangle - \|x'\|^2 \|y\|^2 \leq 0.$$

As :

$$\langle x' | y \rangle = e^{-i\theta} \langle x | y \rangle = |\langle x | y \rangle| \in \mathbb{R}_+$$

we have :

$$\operatorname{Re} \langle x' | y \rangle = \langle x' | y \rangle = |\langle x | y \rangle|.$$

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

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

Proof. Just check the triangle inequality:

$$\|x + y\|^2 = \|x\|^2 + \|y\|^2 + 2 \operatorname{Re} \langle x | y \rangle \leq \|x\|^2 + \|y\|^2 + 2 \|x\| \|y\| = (\|x\| + \|y\|)^2,$$

thanks to the Cauchy-Schwarz inequality. $\square$

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

$$\begin{array}{ccc} \Phi_y: & H & \longrightarrow \mathbb{K} = \mathbb{R} \text{ or } \mathbb{C} \\ & x & \longmapsto \langle x | y \rangle \end{array}$$

is continue. Its norm in H^ is $\|\Phi_y\| = \|y\|$.*

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

$$|\Phi_y(x)| = |\langle x | y \rangle| \leq \|y\| \|x\|;$$

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

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

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 $|\langle x | y \rangle| = \|x\| \|y\|$ 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.

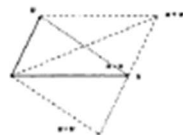
2.2.2 Parallelogram equality

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

$$\|u + v\|^2 + \|u - v\|^2 = 2(\|u\|^2 + \|v\|^2). \quad (2.1)$$

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 $x \perp y$.

Example. 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 $(y|x) = \overline{(x|y)}$).

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

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

what we can call the Pythagorean Theorem.

In the complex case:

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

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:

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

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 $\Phi_y : x \mapsto \langle x | y \rangle$ 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:

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

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 , $L^2(m)$ is a Hilbert space, by virtue of the Riesz-Fisher

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

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

is associated with the usual inner product:

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

In particular, ℓ_2 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 $x, y \in C$:

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

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:*

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

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

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

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

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:

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

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

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

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

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

so that we obtain:

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

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 $\|x - z_n\|^2 \leq d^2 + 1/n$ leads, passing to the limit, that $\|x - y\| \leq d$. 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:

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

hence $\|y_1 - y_2\|^2 \leq 0$, 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 $0 \leq t \leq 1$, by the convexity of C ; SO :

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

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

$$t^2\|y-z\|^2 + 2t\operatorname{Re}(x-y | y-z) \geq 0.$$

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

$$\operatorname{Re}(x-y | z-y) \leq 0.$$

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

$$\begin{aligned} \|x-z\|^2 &= \|(x-y) + (y-z)\|^2 = \|x-y\|^2 + \|y-z\|^2 + 2\operatorname{Re}(x-y | y-z) \\ &= \|x-y\|^2 + \|y-z\|^2 - 2\operatorname{Re}(x-y | z-y) \geq \|x-y\|^2; \end{aligned}$$

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

2.3.2 Consequences

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

$x_1, x_2 \in H$:

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

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

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

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

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

We therefore obtain:

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

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

$$\|y_1 - y_2\| \leq \|x_2 - x_1\|$$

□

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 $P_F : H \rightarrow F$ is a continuous linear mapping, and $P_F(x)$ is the unique point $y \in F$ such that:*

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

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

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

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

$$\operatorname{Re}(x - y | z - y) \leq 0, \quad \forall z \in F;$$

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 $w \in F$:

$$\lambda \langle x - y | w \rangle = \langle x - y | \lambda w \rangle \leq 0, \quad \forall \lambda \in \mathbb{R}$$

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

$$\lambda \operatorname{Re} \langle x - y | w \rangle = \operatorname{Re} \langle x - y | \lambda w \rangle \leq 0, \quad \forall \lambda \in \mathbb{R},$$

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

$$\lambda \operatorname{Im} \langle x - y | w \rangle = \operatorname{Re} \langle x - y | i\lambda w \rangle \leq 0, \quad \forall \lambda \in \mathbb{R}$$

which, again, is only possible if $\langle x - y | w \rangle = 0$. 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: $\|P_F(x)\| \leq \|x\|$ for all $x \in H$; the norm of P_F is therefore ≤ 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:*

$$H = F \oplus F^\perp,$$

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 $x = y + z$, with $y \in F, z \in F^\perp$. Note that, since F and $F^\perp$ are orthogonal, we have: $\|x\|^2 = \|y\|^2 + \|z\|^2$; in other words:

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

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, $\langle x | x \rangle = 0$; so $x = 0$. $\square$

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 $F : H = \bar{F} \oplus (\bar{F})^\perp = \bar{F} \oplus F^\perp$. It follows, since we know that $\bar{F} \subseteq F^{\perp\perp}$, that $F^{\perp\perp} = \bar{F}$. $\square$

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:

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

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

Theorem 2.3.4. *The space $\mathcal{X}(\mathbb{R})$ 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 $\mathcal{X}(\mathbb{R})$ is not really contained in $L^2(\mathbb{R})$, 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 $j : \mathcal{X}(\mathbb{R}) \rightarrow L^2(\mathbb{R})$, which associates each function with its equivalence class, is injective; we can therefore identify each $f \in \mathcal{X}(\mathbb{R})$ with its equivalence class $j(f)$, that is to say $\mathcal{X}(\mathbb{R})$ with $j[\mathcal{X}(\mathbb{R})]$.

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

$$(f|g) = \int_{\mathbb{R}} f\bar{g}d\lambda = 0, \quad \forall f \in \mathcal{X}(\mathbb{R}).$$

We want to show that $g = 0$.

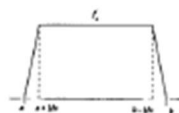
Taking the real and imaginary parts, we can assume that g is real-valued, and we write $g = g^+ - g^-$. We have, for all $f \in \mathcal{X}(\mathbb{R})$:

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

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

$$\begin{cases} 0 \leq f_n \leq \mathbf{1}_{]a,b[} \\ f_n(t) \xrightarrow{n \rightarrow \infty} \mathbf{1}_{]a,b[}(t) \text{ for } t \in \mathbb{R}, \end{cases}$$

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



The Monotone Convergence Theorem gives:

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

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

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

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 $L^2(\mathbb{R})$. $\square$

Corollary 2.3.3. $\mathcal{C}([0, 1])$ 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 $g \in \mathcal{X}(\mathbb{R})$ such that $\|f - g\|_{L^2(\mathbb{R})} \leq \varepsilon$. Let $h = g|_{[0,1]}$ be the restriction of g to $[0, 1]$. We have, on the one hand, $h \in \mathcal{C}([0, 1])$ and, on the other hand, $\|f - h\|_{L^2(0,1)} \leq \|\tilde{f} - g\|_{L^2(\mathbb{R})} \leq \varepsilon$. $\square$

2.3.3 Representation of the dual

Recall that the dual is:

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

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 $\Phi_y : x \in H \rightarrow (x|y)$ 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 $\Phi(x) = (x|y)$ 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:

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

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:

$$\Phi_y(u) = \langle u | y \rangle = \Phi(u) \langle u | u \rangle = \Phi(u) \|u\|^2 = \Phi(u)$$

Thus we have $\Phi = \Phi_y$. □

Remark. The value $y = \overline{\Phi(u)}u$ 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 $y = \overline{\Phi(u)}u$.

2.3.4 Adjoint of an operator

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

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

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

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

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

$$\begin{aligned} \Phi_y \circ T : H &\longrightarrow \mathbb{K} \\ x &\longmapsto \langle Tx | y \rangle \end{aligned}$$

is a continuous linear form on H ; there therefore exists, by the Fréchet-Riesz Theorem, a

unique element of H , which we will denote T^*y , such that:

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

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 $x \in H$:

$$\begin{aligned} \langle x | T^*(a_1y_1 + a_2y_2) \rangle &= \langle Tx | a_1y_1 + a_2y_2 \rangle = \bar{a}_1 \langle Tx | y_1 \rangle + \bar{a}_2 \langle Tx | y_2 \rangle \\ &= \bar{a}_1 \langle x | T^*y_1 \rangle + \bar{a}_2 \langle x | T^*y_2 \rangle = \langle x | a_1T^*y_1 + a_2T^*y_2 \rangle \end{aligned}$$

therefore $T^*(a_1y_1 + a_2y_2) = a_1T^*y_1 + a_2T^*y_2$.

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

$$|(\Phi_y \circ T)(x)| = |\langle Tx | y \rangle| \leq \|Tx\| \|y\| \leq \|T\| \|x\| \|y\|;$$

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

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

$$\langle y | T^{**}x \rangle = \langle T^*y | x \rangle = \langle y | Tx \rangle$$

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

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 $\text{vect}(\Delta)$ 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 $\text{vect}(\Delta)$. $\square$

Examples. 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

$$\begin{array}{c} e_n = (0, \dots, 0, 1, 0, \dots) \\ \uparrow \\ n^{\text{th}} \text{ position} \end{array}$$

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

$$\|x - (\xi_1 e_1 + \dots + \xi_n e_n)\|^p = \sum_{k=n+1}^{\infty} |\xi_k|^p \xrightarrow{n \rightarrow \infty} 0$$

and when $x \in c_0$:

$$\|x - (\xi_1 e_1 + \dots + \xi_n e_n)\|_{\infty} = \sup_{k \geq n+1} |\xi_k| \xrightarrow{n \rightarrow \infty} 0.$$

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, $D = \{x_n; n \geq 1\}$ a part of dense countable E , and $F \subseteq E$. For any pair of integers $n, k \geq 1$ 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 $D_F = \{y_{n,k}; n, k \geq 1\}$ is a countable part of F , and it is dense in F : let $y \in F$; there exists, for all $k \geq 1$, an integer $n \geq 1$ such that $d(y, x_n) \leq 1/k$; 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 $d(y, y_{n,k}) \leq d(y, x_n) + d(x_n, y_{n,k}) \leq 2/k$. $\square$

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.

Examples. 1) In ℓ_2 , the sequence $(e_n)_{n \geq 1}$ is orthonormal.
 2) In $L^2(0, 1)$, we put:

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

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 $(u_1, \dots, u_n)$ is orthonormal, then, for all $a_1, \dots, a_n \in \mathbb{K}$:

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

Proof. Just develop using Proposition II.1.3:

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

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

Corollary 2.4.1. Any orthonormal family is free (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 $x \in H$:*

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

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:

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

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

$$\begin{aligned} S: H &\longrightarrow \ell_2(I) \\ x &\longmapsto (\langle x | u_i \rangle)_{i \in I} \end{aligned}$$

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

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

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

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

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*

$$\boxed{\xi_n = \langle x | u_n \rangle} \text{ for all } n \geq 1.$$

Here sequence means countable family.

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

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

□

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

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

Proof. As we have $\xi_k = \langle x | u_k \rangle$, by the previous proposition, we obtain that $\langle x - \sum_{k=1}^n \xi_k u_k | u_j \rangle = 0$ for all $j \leq n$; 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 $(u_n)_{n \geq 1}$ is an orthonormal sequence in H , then, for every sequence $(\xi_n)_{n \geq 1} \in \ell_2$, 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:

$$\begin{aligned} S: H &\longrightarrow \ell_2 \\ x &\longmapsto (\langle x | u_n \rangle)_{n \geq 1} \end{aligned}$$

is surjective.

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

$$\left\| \sum_{k=n}^{n+p} \xi_k u_k \right\|^2 = \sum_{k=n}^{n+p} |\xi_k|^2 \xrightarrow{n \rightarrow \infty} 0,$$

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 be written, 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:*

$$x = \sum_{n=1}^{\infty} \xi_n u_n, \quad \text{with} \quad \xi_n = \langle x | u_n \rangle$$

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

$$\begin{aligned} 1) \quad & \|x\|^2 = \sum_{n=1}^{\infty} |\langle x | u_n \rangle|^2; \\ 2) \quad & \langle x | y \rangle = \sum_{n=1}^{\infty} \langle x | u_n \rangle \overline{\langle y | u_n \rangle}, \text{ the series absolutely converges.} \end{aligned}$$

Proof. Let F_n denote the vector subspace generated by $u_1, \dots, u_n$, 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:

$$\|x - x_n\| = \text{dist}(x, F_n) \xrightarrow{n \rightarrow \infty} 0.$$

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

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

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

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

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

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

which gives the other identity when $y = x$. $\square$

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:*

$$\begin{aligned} S: H &\longrightarrow \ell_2 \\ x &\longmapsto (\langle x | u_n \rangle)_{n \geq 1} \end{aligned}$$

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:

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

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 $v_1, \dots, v_n$. 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 $(u_n)_{n \geq 1}$ is orthonormal, and the set $\{u_n; n \geq 1\}$ is total because the vector subspace generated by $u_1, \dots, u_n$ is F_n . Indeed, by Theorem 2.3.2, for $2 \leq k \leq n$, we have $u'_k - v_k \in F_{k-1}^\perp = F_{k-1}$, 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 $\mathcal{C}_\mathbb{R}(K)$ or $\mathcal{C}_\mathbb{C}(K)$ of functions continues $f: K \rightarrow \mathbb{R}$ 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 $\mathcal{C}_{\mathbb{R}}(K)$.*

We further assume that:

- a) A separates the points of K ;
- b) A contains constants.

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

Remarks.

- 1) A subalgebra of $\mathcal{C}(K)$ 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 $A = \{f \in \mathcal{C}(K); f(a) = 0\}$ for a given $a \in K$.

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

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

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

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

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

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

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

Note by the way that, we have proven the separability of $\mathcal{C}_{\text{mathbb{R}}}([0, 1])$. $\square$

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 $(r_n)_{n \geq 0}$ which converges uniformly on $[0, 1]$ to the square root function $r : t \mapsto \sqrt{t}$.

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

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

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

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

because $\sqrt{t} + r_n(t) \leq \sqrt{t} + \sqrt{t} = 2\sqrt{t} \leq 2$. Note that in passing, we saw that the sequence $(r_n)_{n \geq 0}$ 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 $r_n(t) \geq 0$, that:

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

so :

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

□

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 $u_n : K \rightarrow \mathbb{R}$ which simply converges to a continuous function $u : K \rightarrow \mathbb{R}$, 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:

$$n \geq N(x) \implies 0 \leq u(x) - u_n(x) \leq \varepsilon/3.$$

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

$$x' \in V(x) \implies \begin{cases} |u(x) - u(x')| \leq \varepsilon/3 \\ |u_{N(x)}(x') - u_{N(x)}(x)| \leq \varepsilon/3. \end{cases}$$

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

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

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

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

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

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

□

2nd step. 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:

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

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

□

Step 3. 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).

□

Step 3a. 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 $\mathcal{C}(K)$ 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:

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

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. $\square$

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

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

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 $f(x)$. 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:

$$y \in V(z) \implies h_z(y) \leq f(y) + \varepsilon.$$

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

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

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

Step 6. We have $\bar{A} = \mathcal{C}_R(K)$.

Proof. Let $f \in \mathcal{C}_R(K)$, 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:

$$y \in U(x) \implies g_x(y) \geq f(y) - \varepsilon.$$

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

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

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

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

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

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

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

This completes the proof of Theorem 2.5.1. $\square$

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 $X_1, \dots, X_n$ all following Bernoulli's law with parameter t . Then $S_n = X_1 + \dots + X_n$ follows the binomial law $\mathcal{B}(n, t)$ with parameters n and t . The weak law of large numbers says that $\frac{S_n}{n} \xrightarrow[n \rightarrow \infty]{} t = \mathbb{E}(X_1)$ in probability. Then, for any function f continuous on $[0, 1]$, we have $\mathbb{E}\left[f\left(\frac{S_n}{n}\right)\right] \xrightarrow[n \rightarrow \infty]{} f(t)$. Indeed, if $\varepsilon > 0$ is given, the uniform continuity of f on $[0, 1]$ makes it possible to find $\delta > 0$ such that $|f(x) - f(x')| \leq \varepsilon$ for $|x - x'| \leq \delta$; convergence in probability then gives a $N \geq 1$ such that $\mathbb{P}\left(\left|\frac{S_n}{n} - t\right| > \delta\right) \leq \varepsilon$ if $n \geq N$. Then, for $n \geq N$, we have:

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

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:

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

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 $\text{Var}(X) = t(1-t)$. We have, by the Bienaymé-

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

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

Consider the continuity modulus of f , defined by:

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

Saying that f is uniformly continuous means that $\omega_f(h) \xrightarrow{h \rightarrow 0} 0$. 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!):

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

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

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 $\mathcal{C}_{\mathbb{R}}(K)$. But then, $A = A_{\mathbb{R}} + iA_{\mathbb{R}}$ is dense in $\mathcal{C}_{\mathbb{C}}(K) = \mathcal{C}_{\mathbb{R}}(K) + i\mathcal{C}_{\mathbb{R}}(K)$. $\square$

Example. 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 $\mathcal{C}_{\mathbb{C}}(K)$.

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 $f : \mathbb{R} \rightarrow \mathbb{C}$ **periodic**, with period 1 on $\mathbb{R}$.

The surjective application:

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

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 $\tilde{f} : \mathbb{U} \rightarrow \mathbb{C}$ such that $f = \tilde{f} \circ e_1$ (resp. $\tilde{f} : \mathbb{T} \rightarrow \mathbb{C}$ such that $f(x) = \tilde{f}(x + \mathbb{Z})$). The space $\mathcal{C}_1(\mathbb{R})$ 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 $\mathcal{C}(\mathbb{U})$ of continuous functions on the compact $\mathbb{U}$. It is also identified with the subspace $\mathcal{C}^{\circ} = \{f \in \mathcal{C}([0, 1]); f(0) = f(1)\}$. These identifications are isometric since:

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

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

$$\sum_{n=N_1}^{N_2} a_n e^{2\pi i n t}$$

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

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

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

where N is a positive integer.

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

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

The set $e_n; n \in \mathbb{Z}$ 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 $u = e_1(t) = e^{2\pi i t}$, and then $u^n e^{2\pi i n t} = e_n(t)$, and that $\bar{u} = e^{-2\pi i t} = e_{-1}(t)$. The complex Stone-Weierstrass theorem applied to $\mathcal{C}(\mathbb{U})$ 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 $f: \mathbb{R} \rightarrow \mathbb{C}$ 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 $g : [0, 1] \rightarrow \mathbb{C}$, the measurable function:

$$\tilde{g} : [0, 1] \rightarrow \mathbb{C}$$

$$t \mapsto \begin{cases} g(t) & \text{if } 0 \leq t < 1; \\ g(0) & \text{if } t = 1 \end{cases}$$

extends by periodicity into a measurable function $f : \mathbb{R} \rightarrow \mathbb{C}$ 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:*

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

$$\sqrt{2} \sin(2\pi t), \sqrt{2} \sin(4\pi t), \dots, \sqrt{2} \sin(2\pi n t), \dots$$

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

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

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

where the inner products:

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

for $n \in \mathbb{Z}$, are called the Fourier coefficients of f . Parseval's formula is then written

$$\int_0^1 |f(x)|^2 dx = \sum_{n \in \mathbb{Z}} |\hat{f}(n)|^2.$$

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

$$\lim_{n \rightarrow \infty} \sum_{k=-l_n}^{l_n} \hat{f}(k) e^{2\pi i k t} = f(t)$$

for almost all $t \in [0, 1]$.

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

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

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

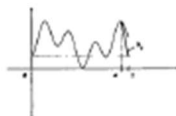
Lemma 2.5.1. *The set $\mathcal{C}_1(\mathbb{R})$ of continuous functions on $\mathbb{R}$ of period 1, identified à $\tilde{\mathcal{C}} = \{f \in \mathcal{C}([0, 1]); f(0) = f(1)\}$, is dense in $L^2(0, 1)$.*

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

Proof of the lemma. Let $f \in L^2(0, 1)$ and let $\varepsilon > 0$. We know (Corollary II.2.9) that there exists $h \in \mathcal{C}([0, 1])$ such that $\|f - h\|_2 \leq \varepsilon/2$. Let $M > 0$ such that $|h(t)| \leq M$ 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.

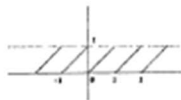
Then $h_1 \in \tilde{\mathcal{C}}$, $\|h_1\|_\infty \leq M$, and:



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

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

Application example. Let $f : \mathbb{R} \rightarrow \mathbb{R}$ 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:

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

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

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

Parseval's formula $\|f\|_2^2 = \sum_{n \in \mathbb{Z}} |\hat{f}(n)|^2$ 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 $f : [0, 1] \rightarrow \mathbb{C}$, the Cauchy-Schwarz inequality:

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

says that $\mathcal{L}^2([0, 1]) \subseteq \mathcal{L}^1([0, 1])$. 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: $L^2(0, 1) \subseteq L^1(0, 1)$

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

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

since $|e^{-2\pi i n t}| = 1$. We have $|\hat{f}(n)| \leq \|f\|_1$ 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:*

$$\hat{f}(n) \xrightarrow{|n| \rightarrow \infty} 0$$

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

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

shows that we have, in particular, $\hat{g}(n) \xrightarrow{|n| \rightarrow \infty} 0$.

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 $\|f - g\|_1 \leq \varepsilon$. As we have $|\hat{f}(n) - \hat{g}(n)| \leq \|f - g\|_1 \leq \varepsilon$, we obtain $|\hat{f}(n)| \leq |\hat{g}(n)| + \varepsilon$; SO :

$$\limsup_{|n| \rightarrow \infty} |\hat{f}(n)| \leq \limsup_{|n| \rightarrow \infty} |\hat{g}(n)| + \varepsilon = \varepsilon$$

□

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 $p_k = \langle a_{1k}, a_{2k}, \dots \rangle$ 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:

$$p_1 = \langle 1, 0, 0, 0, \dots \rangle$$

$$p_2 = \langle 0, 1, 0, 0, \dots \rangle$$

$$p_3 = \langle 0, 0, 1, 0, \dots \rangle$$

$$p_4 = \langle 0, 0, 0, 1, \dots \rangle$$

.....

.....

$\cdot \downarrow \downarrow$

$$0 = \langle 0, 0, 0, 0, \dots \rangle$$

But the sequence $\langle p_n \rangle$ does not converge to 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 $\mathcal{H}$ denote the proper subspace of $\mathbf{H}$ which consists of all points in $\mathbf{H}$ whose first coordinate is zero. Observe that the function $f : \mathbf{H} \rightarrow \mathcal{H}$ defined by $f(\langle a_1, a_2, \dots \rangle) = \langle 0, a_1, a_2, \dots \rangle$ 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 Kreyszig, Introductory functional analysis with applications, John Wiley & Sons, 1978.
- [3] Haïm Brezis, Analyse fonctionnelle, Theorie et applications, Masson, 1987.