Espaces de Hilbert

François DE MARÇAY
Département de Mathématiques d'Orsay
Université Paris-Saclay, France

Dans ce chapitre, on commence par étudier l'espace de Hilbert concret et tangible $\ell^2_{\mathbb{C}} \equiv \mathbb{C}^\infty$, avant d'introduire le concept général au moyen d'une définition mathématique abstraite.

1. Espace vectoriel complexe hermitien concret $\mathbb{C}^n$ en dimension finie

Les espaces vectoriels sont omniprésents en mathématiques : algèbre linéaire ; matrices ; déterminants ; schémas numériques pour les équations différentielles ; et en physique : principe de superposition en mécanique quantique ; équations différentielles ordinaires ; électromagnétisme. Ce chapitre vise principalement l'étude des espaces vectoriels qui sont de dimension *infinie*. Ils interviennent fréquemment en Analyse.

Le corps des scalaires de tous les espaces vectoriels considérés sera toujours supposé égal à $\mathbb{R}$ ou, le plus souvent, à $\mathbb{C}$. Soit donc E un espace vectoriel sur $\mathbb{C}$ (ou sur $\mathbb{R}$). Les éléments de E sont des *vecteurs* que l'on peut :

- additionner : $v + w \in E$ si $v, w \in E$;
- multiplier par des scalaires : $\lambda\, v \in E$ si $\lambda \in \mathbb{C}$ (ou $\lambda \in \mathbb{R}$) et $v \in E$.

Définition 1.1. Une *norme* $\|\cdot\|$ sur E est une application à valeurs positives :

$$E \ni v \longmapsto \|v\| \in [0, \infty[$$

satisfaisant :

- positivité stricte : $\|v\| = 0$ si et seulement si $v = 0$;
- invariance par dilatation : $\|\lambda v\| = |\lambda|\, \|v\|$ pour tout $\lambda \in \mathbb{C}$ (ou $\lambda \in \mathbb{R}$) et tout $v \in E$;
- inégalité triangulaire : $\|v + w\| \leqslant \|v\| + \|w\|$ pour tous $v, w \in E$.

La positivité stricte exclut que certains vecteurs non nuls puissent avoir une longueur nulle. Ensuite, l'invariance par dilatation est exigible pour des raisons à la fois structurales (compatibilité avec la multiplication par des scalaires λ) et « physiques » (changement d'unité de mesure). Enfin, l'inégalité triangulaire n'est pas seulement calquée sur la géométrie élémentaire dans le plan, elle exprime surtout une compatibilité du comportement avec l'addition $v + w$, et aussi, elle jouera un rôle technique crucial en tant que moyen de comparaison dans tous les calculs de majorations qui vont suivre.

Par exemple, pour tout entier $n \geqslant 1$, l'espace :

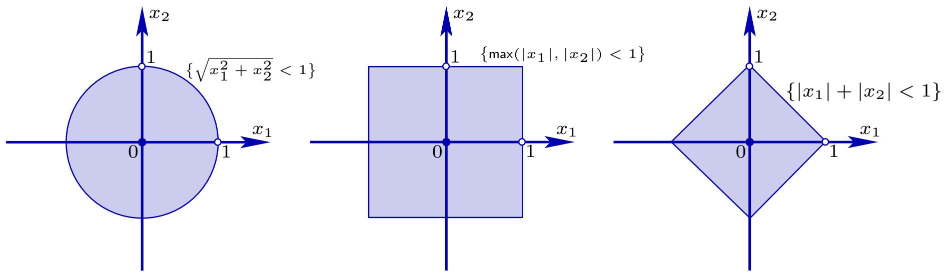
$$\mathbb{R}^n = \{(x_1, x_2, \ldots, x_n) \colon x_i \in \mathbb{R}\}$$

est canoniquement muni de la norme dite *euclidienne* :

$$\|(x_1, x_2, \dots, x_n)\| := \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2}.$$

On peut aussi munir $\mathbb{R}^n$ d'autres normes (vérifier que les trois axiomes sont satisfaits) :

$$\max\big(|x_1|, \dots, |x_n|\big) \qquad \text{ou encore :} \qquad |x_1| + \cdots + |x_n|.$$



En fait, sur un espace vectoriel réel ou complexe de dimension finie, toutes les normes sont *équivalentes* entre elles au sens suivant.

Proposition 1.2. [Équivalence des normes en dimension finie] *Soient $\|\cdot\|_1$ et $\|\cdot\|_2$ deux normes quelconques sur un espace vectoriel réel ou complexe E. Si la dimension de E est finie :*

$$\dim E < \infty,$$

alors il existe deux constantes $0 < c \leqslant C < \infty$ telles que :

$$c\,\|v\|_2 \leqslant \|v\|_1 \leqslant C\,\|v\|_2,$$

pour tout vecteur $v \in E$.

Géométriquement parlant, les boules unités $\big\{\|v\|_1 \leqslant 1\big\}$ et $\big\{\|v\|_2 \leqslant 1\big\}$ pour les deux normes deviennent contenues l'une dans l'autre après une dilatation (contraction) appropriée : disque dans un carré ; carré dans un disque.

Cet énoncé exprime que, de toute majoration :

$$\|v\|_1 \leqslant \text{quelque chose},$$

obtenue avec la première norme, on peut déduire une majoration analogue :

$$\|v\|_2 \leqslant \text{même chose à la constante } C \text{ près},$$

avec la deuxième norme, et *vice-versa*.

Démonstration. On traite le cas d'un $\mathbb{R}$-espace vectoriel E. Si $n := \dim E \geqslant 0$, on sait que $E \cong \mathbb{R}^n$. Puisque le cas $n = 0$ est trivial, supposons $n \geqslant 1$.

Fixons alors une norme de référence sur $\mathbb{R}^n$:

$$|x| := \max\big(|x_1|, \dots, |x_n|\big),$$

laquelle le munit de sa topologie standard. Il suffit de faire voir que toute norme quelconque $\|x\|$ sur $\mathbb{R}^n$ est équivalente à $|x|$, puisque « être des normes équivalentes » est une relation d'équivalence (exercice).

Si $(e_1, \dots, e_n)$ est la base canonique de $\mathbb{R}^n$, en termes de laquelle tout vecteur $x = (x_1, \dots, x_n)$ s'écrit $x = x_1\,e_1 + \cdots + x_n\,e_n$, il vient par inégalité triangulaire :

$$\begin{aligned} \|x\| &\leqslant |x_1|\,\|e_1\| + \cdots + |x_n|\,\|e_n\| \\ &\leqslant C\,|x|, \end{aligned}$$

où la constante $C := \sum_i \|e_i\|$ est finie. Il reste à trouver $0 < c < \infty$ telle que $c\,|x| \leqslant \|x\|$ pour tout x.

Observons que l'inégalité valable pour tous $x, y \in \mathbb{R}^n$:

$$\big\|x - y\big\| \;\leqslant\; C\,\big|x - y\big|,$$

montre la lipschitzianité, donc la continuité de l'application :

$$\mathbb{R}^n \;\ni\; z \;\longmapsto\; \|z\| \;\in\; \mathbb{R}_+.$$

Soit maintenant la 'sphère' unité pour la norme de référence :

$$S_{|\cdot|} \;:=\; \big\{x \in \mathbb{R}^n\colon\; |x| = 1\big\},$$

qui est frontière de l'hypercube compact $[-1, 1]^n$. Étant image inverse du fermé $\{1\}$, cette 'sphère' $S_{|\cdot|}$ est fermée, donc compacte.

Par conséquent, la restriction à $S_{|\cdot|}$ de l'application *continue* $z \longmapsto \|z\|$ y atteint son minimum :

$$c \;:=\; \min_{|x|=1} \|x\| \;=\; \|x_*\|,$$

en au moins un certain point $x_* \in S_{|\cdot|}$.

Or $x_* \neq 0$ n'est *pas* l'origine de $\mathbb{R}^n$ — puisque $0 \notin S_{|\cdot|}$ —, donc $\|x_*\| > 0$, d'où nous déduisons que l'inégalité visée est au moins vraie sur $S_{|\cdot|}$:

$$c \cdot 1 \;=\; c\,|x| \;\leqslant\; \|x\| \qquad\qquad (\forall\,|x| = 1)$$

Pour un $x \in \mathbb{R}^n \backslash \{0\}$ non nul quelconque, en appliquant cette inégalité au vecteur $\frac{x}{|x|}$ qui appartient à $S_{|\cdot|}$, il vient :

$$c\,\Big|\tfrac{x}{|x|}\Big| \;\leqslant\; \Big\|\tfrac{x}{|x|}\Big\|,$$

d'où $c|x| \leqslant \|x\|$ après pulvérisation des dénominateurs.

Lorsque E est un $\mathbb{C}$-espace vectoriel, avec $n := \dim_{\mathbb{C}} E$, on se ramène à $\mathbb{R}^{2n}$. □

Quand on travaille en dimension finie $n \geqslant 1$ sur le corps des nombres complexes $\mathbb{C} = \mathbb{R} + \sqrt{-1}\,\mathbb{R}$ au lieu de $\mathbb{R}$, il faut remplacer la norme euclidienne ci-dessus par la norme dite *hermitienne* :

$$\|(z_1, z_2, \dots, z_n)\| := \sqrt{\big(|z_1|^2 + |z_2|^2 + \cdots + |z_n|^2\big)},$$

où l'on a bien entendu :

$$|z_k|^2 := x_k^2 + y_k^2 \qquad\qquad (k = 1 \cdots n),$$

si chaque nombre complexe z_k est décomposé en partie réelle et partie imaginaire :

$$z_k = x_k + \sqrt{-1}\,y_k \qquad\qquad (k = 1 \cdots n).$$

Contrairement à ce qui va se passer en dimension infinie pour $\mathbb{C}^\infty$, dans $\mathbb{C}^n$, la somme $\sum_{i=1}^n |z_i|^2$ converge toujours, et de plus, on peut définir un produit scalaire dit *hermitien* par la formule :

$$\big\langle z, w\big\rangle := z_1\overline{w}_1 + z_2\overline{w}_2 + \cdots + z_n\overline{w}_n,$$

où pour tout $w_k = u_k + \sqrt{-1}\,v_k$, le nombre complexe conjugué est :

$$\overline{w}_k := u_k - \sqrt{-1}\,v_k \qquad\qquad (k = 1 \cdots n),$$

et ce produit scalaire hermitien redonne visiblement la norme hermitienne introduite à l'instant :

$$\|z\| = \sqrt{\langle z,\, z\rangle}.$$

On constate aisément que ce produit hermitien satisfait les propriétés suivantes.

Lemme 1.3. [Propriétés du produit hermitien sur $\mathbb{C}^n$] *Pour tous $z, z', z'', w, w', w'' \in \mathbb{C}^n$ et tous $\lambda, \mu \in \mathbb{C}$, on a :*

(i) $\langle w, z \rangle = \overline{\langle z, w \rangle}$;

(ii) $\langle z' + z'', w \rangle = \langle z', w \rangle + \langle z'', w \rangle$,
$\langle z,\, w' + w'' \rangle = \langle z, w' \rangle + \langle z, w'' \rangle$;

(iii) $\langle \lambda z, w \rangle = \lambda \langle z, w \rangle$,
$\langle z, \mu w \rangle = \overline{\mu} \langle z, w \rangle$;

(iv) $\langle z, z \rangle \in \mathbb{R}_+$;

(v) $\langle z, z \rangle = 0 \Longrightarrow z = 0$. □

En renversant les rôles, ces cinq propriétés élémentaires peuvent aussi être envisagées comme cinq axiomes définissant la notion de produit hermitien abstrait sur un espace vectoriel complexe de dimension finie.

Dans le cas d'un $\mathbb{R}$-espace vectoriel, les barres de conjugaison disparaissent.

Maintenant, dehors la dimension finie !

2. Espace hermitien concret $\ell^2_{\mathbb{C}} \equiv \mathbb{C}^\infty$ en dimension infinie

Car oui ! Notre objectif est d'étudier les espaces vectoriels normés qui sont de dimension *infinie*, i.e. qui ne sont pas de dimension finie, car ils seront très utiles en Analyse où l'on regarde des espaces de fonctions.

Contrairement à la dimension finie, sur les espaces vectoriels de dimension *infinie*, toutes les normes ne sont *pas* équivalentes.

Exemple 2.1. Soit $\mathscr{C}^0 := \mathscr{C}^0([0,1], \mathbb{R})$ l'espace des fonctions continues sur l'intervalle $[0,1] \subset \mathbb{R}$, muni des deux normes :

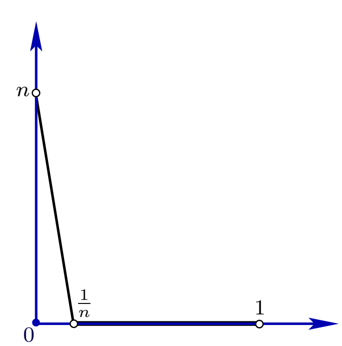
$$\|f\|_{L^1} = \int_0^1 |f(x)|\, dx \qquad \text{et} \qquad \|f\|_{\mathscr{C}^0} := \max_{0 \leqslant x \leqslant 1} |f(x)|.$$

Bien qu'une première inégalité soit vraie et souvent utile :

$$\|f\|_{L^1} = \int_0^1 |f(x)|\, dx \leqslant \int_0^1 \max |f|\, dx = 1 \cdot \|f\|_{\mathscr{C}^0},$$

nous affirmons qu'il n'existe pas de constante $0 < C < \infty$ réalisant l'autre inégalité nécessaire à une équivalence :

$$\|f\|_{\mathscr{C}^0} \leqslant C\, \|f\|_{L^1} \qquad (\forall\, f \in \mathscr{C}^0).$$



En effet, pour $n \geqslant 1$ entier, soit la suite de fonctions continues positives :

$$f_n(x) \,:=\, \begin{cases} -\,n^2x+n & \text{lorsque } \, 0 \leqslant x \leqslant \frac{1}{n}, \\ 0 & \text{lorsque } \, \frac{1}{n} \leqslant x \leqslant 1. \end{cases}$$

Alors :

$$\big\|f_n\big\|_{\mathscr{C}^0} \,=\, n \,\,\longrightarrow\,\, \infty,$$

tandis que par la formule donnant l'aire d'un triangle :

$$\big\|f_n\big\|_{L^1} \,=\, \tfrac{1}{2}\,\tfrac{1}{n}\,n \,=\, \tfrac{1}{2},$$

donc en particulier, pour ces $f_n \in \mathscr{C}^0$, une telle constante $0 < C < \infty$ n'existe pas. $\square$

Les espaces vectoriels normés de dimension infinie les plus 'simples' sont ceux qu'on appelle *espaces de Hilbert*, car leur norme dérive d'un produit scalaire entre vecteurs et le produit scalaire permet de faire beaucoup de géométrie comme en dimension finie, avec une notion naturelle d'orthogonalité, des projections orthogonales, des bases orthonormées, un bon comportement de la dualité, *etc.* En un certain sens, les espaces de Hilbert sont les espaces de dimension infinie les plus simples possibles, car ils sont dotés de la structure géométrique la plus riche.

L'espace $\mathbb{C}^n$ de dimension quelconque $n \geqslant 1$ sur les nombres complexes est de dimension finie, mais c'est en partant de lui que nous allons nous diriger vers la dimension infinie. Le modèle le plus simple d'un espace hermitien de type $\mathbb{C}^n$ et de dimension infinie serait bien entendu $\mathbb{C}^\infty$, i.e. $\mathbb{C}^n$ avec $n = \infty$. Il a un autre nom : on l'appelle classiquement ℓ^2, « petit-ℓ^2 » à l'oral pour le distinguer des espaces « grand-L^2 » de fonctions de carré intégrable :

$$L^2(\mathbb{R}) := \Big\{ f \,\text{ fonctions mesurables sur } \mathbb{R} \text{ telles que } \textstyle\int_{\mathbb{R}} |f(x)|^2\,dx < \infty \Big\},$$

dont nous reparlerons dans peu de temps.

Définition 2.2. **[Espace $\ell^2_{\mathbb{C}}$]** L'espace ℓ^2 est constitué des suites *infinies* dénombrables :

$$z := (z_1, z_2, \dots, z_i, \dots) \in \mathbb{C}^\infty$$

de nombres complexes $z_i \in \mathbb{C}$ telles que :

$$\sum_{i=1}^{\infty} |z_i|^2 < \infty.$$

Bien entendu, cette quantité $\sum_{i=1}^\infty |z_i|^2$, lorsqu'elle est finie, satisfait pour tout nombre complexe $\lambda \in \mathbb{C}$:

$$\sum_{i=1}^{\infty} \big|\lambda\, z_i\big|^2 = |\lambda|^2 \sum_{i=1}^{\infty} |z_i|^2.$$

Une définition similaire pour $\ell^2_{\mathbb{R}}$ peut aisément être formulée, mais nous travaillerons principalement avec $\ell^2_{\mathbb{C}}$, que nous noterons parfois ℓ^2.

Immédiatement, une première question surgit :

- *ℓ^2 ainsi défini est-il un espace vectoriel ?*

Ce n'est pas évident, car il se pourrait très bien que $\sum_{i=1}^\infty |z_i + w_i|^2$ ne soit pas $< \infty$ lorsque les deux sommes $\sum_{i=1}^\infty |z_i|^2$ et $\sum_{i=1}^\infty |w_i|^2$ sont toutes deux $< \infty$. Autre question :

- *Existe-t-il aussi un produit scalaire hermitien sur $\ell^2 \equiv \mathbb{C}^\infty$ comme c'était le cas sur $\mathbb{C}^n$?*

Encore une fois, cela n'est pas clair, car il se pourrait que $\sum_{i=1}^{\infty} z_i\overline{w}_i$ ne soit pas $< \infty$ lorsque $\sum_{i=1}^{\infty} |z_i|^2 < \infty$ et $\sum_{i=1}^{\infty} |w_i|^2 < \infty$.

Étudions donc cet espace ℓ^2 de suites infinies dont la somme des modules au carré converge.

Lemme 2.3. *Pour tous $z, w \in \ell^2_{\mathbb{C}}$, on a la convergence de :*

$$\langle z,\, w\rangle \;:=\; \langle z,\, w\rangle_{\ell^2} \;:=\; \sum_{i=1}^{\infty} z_i\,\overline{w}_i \;<\; \infty.$$

Démonstration. En effet, pour un indice quelconque fixé $i \geqslant 1$, on a la positivité :

$$0 \leqslant \big(|z_i| - |w_i|\big)^2 = |z_i|^2 + |w_i|^2 - 2\,|z_i|\,|w_i|, \tag{2.4}$$

d'où en faisant passer $|z_i|\,|w_i| = |z_i\,\overline{w}_i|$ à gauche :

$$\big|z_i\,\overline{w}_i\big| \leqslant \tfrac{1}{2}\left(|z_i|^2 + |w_i|^2\right).$$

Sommons alors ces inégalités pour $i = 1$ jusqu'à l'infini, ce qui nous donne :

$$\sum_{i=1}^{\infty} \big|z_i\,\overline{w}_i\big| \leqslant \tfrac{1}{2}\sum_{i=1}^{\infty}\left(|z_i|^2 + |w_i|^2\right) < \infty, \tag{2.5}$$

et donc par inégalité triangulaire infinie, on obtient bien la finitude de :

$$\big|\langle z,\, w\rangle_{\ell^2}\big| := \left|\sum_{i=1}^{\infty} z_i\,\overline{w}_i\right| \;\leqslant\; \sum_{i=1}^{\infty}\big|z_i\,\overline{w}_i\big| \;<\; \infty. \qquad \square$$

Ainsi, $\sum_{i=1}^{\infty} z_i\,\overline{w}_i$ constitue un candidat naturel pour définir un produit scalaire hermitien sur $\ell^2_{\mathbb{C}}$ qui sera formellement analogue à celui $\sum_{i=1}^{n} z_i\,\overline{w}_i$ qui existait sur $\mathbb{C}^n$, mais pour l'instant, on ignore encore que $\ell^2_{\mathbb{C}}$ possède une structure d'espace vectoriel !

En tout cas, on peut d'ores et déjà introduire la quantité :

$$||z|| \;:=\; ||z||_{\ell^2} \;:=\; \left(\sum_{i=1}^{\infty}|z_i|^2\right)^{\frac{1}{2}} \;<\; \infty,$$

laquelle va, dans un instant, s'avérer être une vraie norme sur le vrai espace vectoriel ℓ^2.

Un examen plus approfondi du calcul précédent va nous montrer que la valeur absolue de la quantité $\langle z, w\rangle = \sum_{i=1}^{\infty} z_i\,\overline{w}_i$ est toujours majorée par le produit $||z||\,||w||$. Cette inégalité élémentaire, en tout point analogue à l'inégalité :

$$|\vec{v}\cdot\vec{w}| \leqslant ||\vec{v}||\,||\vec{w}||,$$

satisfaite par le produit scalaire entre deux vecteurs $\vec{v}, \vec{w} \in \mathbb{R}^2$ du plan euclidien, est tellement importante qu'elle a reçu un nom distinctif.

Théorème 2.6. [Inégalité de Cauchy-Schwarz] *Pour tous $z, w \in \ell^2_{\mathbb{C}}$, on a :*

$$\left|\sum_{i=1}^{\infty} z_i\overline{w}_i\right| \leqslant \left(\sum_{i=1}^{\infty}|z_i|^2\right)^{\frac{1}{2}}\left(\sum_{i=1}^{\infty}|w_i|^2\right)^{\frac{1}{2}},$$

c'est-à-dire en notation abrégée :

$$\big|\langle z, w\rangle\big| \leqslant ||z||\,||w||.$$

La démonstration qui suit est valable aussi bien en dimension finie sur $\mathbb{C}^n$ que sur $\ell^2 \equiv \mathbb{C}^\infty$ en dimension infinie. On donnera même plus loin une autre démonstration de cette inégalité de Cauchy-Schwarz, en ajoutant l'information supplémentaire que l'inégalité $\leqslant$ devient une égalité $=$ si et seulement si z et w sont colinéaires.

Démonstration. Puisque l'inégalité est trivialement satisfaite lorsque $z = 0$ ou $w = 0$, on peut supposer que $z \neq 0 \neq w$, ce qui permet d'introduire les deux éléments :

$$z' := \frac{z}{\|z\|} = \left(\frac{z_1}{\|z\|}, \frac{z_2}{\|z\|}, \dots, \frac{z_i}{\|z\|}, \dots\right)$$

et :

$$w' := \frac{w}{\|w\|} = \left(\frac{w_1}{\|w\|}, \frac{w_2}{\|w\|}, \dots, \frac{w_i}{\|w\|}, \dots\right)$$

de ℓ^2 construits en renormalisant convenablement z et w afin de se ramener à des éléments qui sont de 'norme' égale à 1 :

$$\|z'\|^2 = \sum_{i=1}^{\infty} \frac{|z_i|^2}{\|z\|^2} = \tfrac{\|z\|^2}{\|z\|^2} = 1 \qquad \text{et} \qquad \|w'\|^2 = \sum_{i=1}^{\infty} \frac{|w_i|^2}{\|w\|^2} = \tfrac{\|w\|^2}{\|w\|^2} = 1.$$

Lorsque nous aurons établi que $\|\cdot\|$ est bien une norme sur l'espace vectoriel ℓ^2, nous serons autorisés à dire que cette renormalisation nous ramène à considérer deux vecteurs z' et w' de ℓ^2 qui sont de norme 1.

Appliquons alors à z' et w' l'inégalité (2.5), ce qui nous donne formellement :

$$\sum_{i=1}^{\infty} \left|z_i'\,\overline{w}_i'\right| \leqslant \tfrac{1}{2} \sum_{i=1}^{\infty} \left(|z_i'|^2 + |w_i'|^2\right).$$

Mais on peut immédiatement remplacer à gauche les valeurs des z_i' et des $\overline{w}_i'$:

$$\begin{aligned} \sum_{i=1}^{\infty} \left|\frac{z_i}{\|z\|}\,\frac{\overline{w}_i}{\|w\|}\right| &\leqslant \tfrac{1}{2}\left(\|z'\|^2 + \|w'\|^2\right) \\ &= \tfrac{1}{2}\left(1^2 + 1^2\right) = 1, \end{aligned}$$

et ensuite, si l'on fait basculer à droite le produit des normes $\|z\|\,\|w\|$ qui apparaît en facteur commun dans les dénominateurs de la somme à gauche, on obtient bien :

$$\left|\langle z,\, w\rangle\right| \;=\; \left|\sum_{i=1}^{\infty} z_i\,\overline{w}_i\right| \;\leqslant\; \sum_{i=1}^{\infty} \left|z_i\,\overline{w}_i\right| \;\leqslant\; \|z\|\,\|w\|. \qquad \square$$

À présent, Cauchy-Schwarz nous va nous permettre d'établir que ℓ^2 est stable sous l'opération somme, donc que ℓ^2 est un espace vectoriel ; de plus, l'inégalité triangulaire va être vraie.

Lemme 2.7. *Comme en dimension finie, la quantité :*

$$\|z\| := \left(\sum_{i=1}^{\infty} |z_i|^2\right)^{\frac{1}{2}}$$

définie pour tous les éléments $z = (z_1, z_2, \dots, z_i, \dots)$ de ℓ^2 satisfait :

$$\|z + w\| \;\leqslant\; \|z\| + \|w\| \qquad (\forall\, z, w \in \ell^2).$$

On en déduit alors que $\|\cdot\|$ constitue une vraie norme sur le vrai espace vectoriel $\ell^2_{\mathbb{C}}$, puisque les deux autres conditions pour être une norme sont trivialement satisfaites, en particulier l'homogénéité $\|\lambda\, z\| = |\lambda|\, \|z\|$ sous dilatation.

Démonstration. En élevant au carré l'inégalité à établir :

$$\|z+w\|^2 \overset{?}{\leqslant} \big(\|z\|+\|w\|\big)^2 = \|z\|^2 + \|w\|^2 + 2\,\|z\|\,\|w\|,$$

on se ramène à considérer une inégalité qui lui est équivalente. Maintenant, si on explicite toutes les sommes infinies impliquées :

$$\sum_{i=1}^{\infty} (z_i + w_i)(\overline{z}_i + \overline{w}_i) \overset{?}{\leqslant} \sum_{i=1}^{\infty} z_i\overline{z}_i + \sum_{i=1}^{\infty} w_i\overline{w}_i + 2\left(\sum_{i=1}^{\infty} |z_i|^2\right)^{\frac{1}{2}} \left(\sum_{i=1}^{\infty} |w_i|^2\right)^{\frac{1}{2}},$$

si on développe le membre de gauche :

$$\sum_{i=1}^{\infty} z_i\overline{z}_i + \sum_{i=1}^{\infty} w_i\overline{w}_i + \sum_{i=1}^{\infty} \big(z_i\overline{w}_i + w_i\overline{z}_i\big) \overset{?}{\leqslant} \sum_{i=1}^{\infty} z_i\overline{z}_i + \sum_{i=1}^{\infty} w_i\overline{w}_i + 2\left(\sum_{i=1}^{\infty} |z_i|^2\right)^{\frac{1}{2}} \left(\sum_{i=1}^{\infty} |w_i|^2\right)^{\frac{1}{2}},$$

et si on soustrait de part et d'autre $\sum_i z_i\overline{z}_i + \sum_i w_i\overline{w}_i$, il reste à établir :

$$\sum_{i=1}^{\infty} z_i\overline{w}_i + \sum_{i=1}^{\infty} w_i\overline{z}_i \overset{?}{\leqslant} 2\left(\sum_{i=1}^{\infty} |z_i|^2\right)^{\frac{1}{2}} \left(\sum_{i=1}^{\infty} |w_i|^2\right)^{\frac{1}{2}}.$$

Mais cette dernière inégalité, vraie, est conséquence évidente de Cauchy-Schwarz. □

Corollaire 2.8. *L'espace ℓ^2 est un espace vectoriel.* □

Nous pouvons maintenant revenir à la définition possible $\langle z, w\rangle = \sum_{i=1}^{\infty} z_i\overline{w}_i$ d'un produit scalaire hermitien en dimension infinie. Grâce aux résultats précédents, on peut vérifier facilement que les relations suivantes sont satisfaites, car toutes les sommes infinies qui y sont impliquées convergent.

Lemme 2.9. [Propriétés du produit hermitien sur $\ell^2_{\mathbb{C}}$] *Pour tous $z, z', z'', w, w', w'' \in \ell^2_{\mathbb{C}}$ et tous $\lambda, \mu \in \mathbb{C}$, on a :*

(i) $\langle w, z\rangle = \overline{\langle z, w\rangle}$;

(ii) $\langle z' + z'', w\rangle = \langle z', w\rangle + \langle z'', w\rangle$,
$\langle z,\, w' + w''\rangle = \langle z, w'\rangle + \langle z, w''\rangle$;

(iii) $\langle \lambda z, w\rangle = \lambda\,\langle z, w\rangle$,
$\langle z, \mu w\rangle = \overline{\mu}\,\langle z, w\rangle$;

(iv) $\langle z, z\rangle \in \mathbb{R}_+$;

(v) $\langle z, z\rangle = 0 \Longrightarrow z = 0$. □

On déduit naturellement de (ii) et de (iii) que :

$$\langle \lambda' z' + \lambda'' z'',\, w\rangle = \lambda'\langle z', w\rangle + \lambda''\langle z'', w\rangle,$$
$$\langle z, \mu' w' + \mu'' w''\rangle = \overline{\mu}'\langle z, w'\rangle + \overline{\mu}''\langle z, w''\rangle.$$

Nous verrons dans un instant que ces propriétés peuvent être envisagées comme cinq axiomes abstraits pour définir les espaces hermitiens de dimension infinie.

Résumons ainsi toutes les propriétés que nous avons établies au sujet de $\ell^2_{\mathbb{C}} \equiv \mathbb{C}^\infty$ qui sont en tout point analogues à celles des espaces $\mathbb{C}^n$ de dimension finie.

Théorème 2.10. **[Structure de $\ell^2_{\mathbb{C}}$]** *L'espace $\ell^2_{\mathbb{C}}$ des suites* infinies *dénombrables :*

$$z := (z_1, z_2, \dots, z_i, \dots) \in \mathbb{C}^\infty$$

de nombres complexes $z_i \in \mathbb{C}$ dont la somme des modules au carré converge :

$$\|z\|^2 := \sum_{i=1}^{\infty} |z_i|^2 < \infty$$

est un espace vectoriel complexe normé dont la norme :

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

dérive d'un produit scalaire hermitien naturel :

$$\langle z, w\rangle := \sum_{i=1}^{\infty} z_i\, \overline{w}_i \qquad\qquad (z, w \in \ell^2).$$

La valeur absolue de ce produit scalaire hermitien est toujours contrôlée par le produit des normes grâce à l'inégalité dite de Cauchy-Schwarz :

$$\big|\langle z, w\rangle\big| \leqslant \|z\|\,\|w\| \qquad\qquad (z, w \in \ell^2),$$

et de plus, cette inégalité est une égalité lorsque, et seulement lorsque z et w sont colinéaires, *à savoir : lorsqu'il existe deux constantes $\lambda,\ \mu \in \mathbb{C}$ telles que $0 = \lambda\, z + \mu\, w$.*

Démonstration. Seul le cas d'égalité dans Cauchy-Schwarz n'a pas encore été établi. Nous allons en fait redémontrer l'inégalité d'une manière différente et traiter ensuite le cas d'égalité. Puisque tout est trivial lorsque $z = 0$ ou $w = 0$, on supposera que $z \neq 0$ et que $w \neq 0$.

Soit θ l'argument du nombre complexe $\langle z, w\rangle$. En remplaçant w par $w' := e^{i\theta}\, w$, l'inégalité de Cauchy-Schwarz reste inchangée (puisque $|e^{i\theta}| = 1$) et le produit hermitien $\langle z, w'\rangle$ devient un nombre *réel* (puisque $\langle z, w'\rangle = e^{-i\theta}\langle z, w\rangle$). On peut donc supposer que $\langle z, w\rangle$ est réel.

Ensuite, pour $t \in \mathbb{R}$, le polynôme du second degré en t à coefficients réels :

$$\begin{aligned} 0 \leqslant \|z + tw\|^2 &= \big\langle z + tw,\ z + tw\big\rangle \\ &= \langle z, z\rangle + \overline{t}\,\langle z, w\rangle + t\,\langle w, z\rangle + t\overline{t}\,\langle w, w\rangle \\ &= \langle z, z\rangle + t\,\langle z, w\rangle + t\,\overline{\langle z, w\rangle} + t^2\,\langle w, w\rangle = \underbrace{\|z\|^2 + 2t\,\langle z,\ w\rangle + t^2\,\|w\|^2}_{=:\,P(t)}, \end{aligned}$$

a toujours des valeurs positives, puisque c'est un carré. Il est donc nécessaire que son discriminant soit négatif :

$$4\,\langle z,\ w\rangle^2 - 4\,\|z\|^2\,\|w\|^2 \leqslant 0 \iff \text{Cauchy-Schwarz},$$

ce qui redémontre l'inégalité.

Enfin, on a l'égalité $\langle z,\ w\rangle = \|z\|\,\|w\|$ qui signifie l'annulation du discriminant si et seulement si ce polynôme $P(t)$ a une racine double t_0, et alors dans ce cas, pour cette racine t_0, on a :

$$0 = P(t_0) = \|z + t_0 w\|^2,$$

c'est-à-dire justement, z et w sont colinéaires ! □

Définition 2.11. [Complétude] Un espace vectoriel normé $(E, \|\cdot\|)$ sur $\mathbb{C}$ (ou sur $\mathbb{R}$) est dit *complet* si les suites de Cauchy sont en fait convergentes dans E, c'est-à-dire :

Pour toute suite $(v_k)_{k=1}^{\infty}$ d'éléments $v_k \in E$ satisfaisant :

$$\forall\, \varepsilon > 0 \quad \exists\, \mathrm{K} \geqslant 1 \quad \text{tel que} \quad \Big(k_1, k_2 \geqslant \mathrm{K} \Longrightarrow \|\, v_{k_1} - v_{k_2} \| \leqslant \varepsilon\Big),$$

Il existe une (unique) limite $v_\infty \in E$, à savoir qui satisfait :

$$\forall\, \varepsilon > 0 \quad \exists\, \mathrm{K} \geqslant 1 \quad \text{tel que} \quad \Big(k \geqslant K \Longrightarrow \|v_k - v_\infty\| \leqslant \varepsilon\Big).$$

L'hypothèse que $(v_k)_{k=1}^{\infty}$ est une suite de Cauchy exprime que tous les v_{k_1} et v_{k_2} deviennent arbitrairement proches entre eux pourvu que k_1 et k_2 soient assez grands. Mais cela ne veut pas dire que les v_k convergent vers une valeur, car l'endroit où les v_k devraient « atterrir » pourrait fort bien ne pas appartenir à l'espace E dans lequel ils se situent.

Ainsi, demander que $(E, \|\cdot\|)$ soit complet, cela revient à demander que toute suite d'éléments de E qui est faite pour converger converge effectivement vers un élément (unique) *appartenant à* E. Intuitivement, E n'a aucun « trou » en lui-même, il n'a aucune trace d'« incomplétude ».

Rappelons aussi que toute suite convergente est de Cauchy (exercice de révision). Il est connu, aussi, que $\mathbb{R}^n$ et $\mathbb{C}^n \simeq \mathbb{R}^{2n}$ sont complets pour tout $n \geqslant 1$ (exercice de révision).

Qu'en est-il de $\ell^2_{\mathbb{C}}$? L'énoncé suivant montre que les propriétés de $\ell^2_{\mathbb{C}}$ sont très analogues à celles de $\mathbb{C}^n$.

Théorème 2.12. [Complétude de$\ell^2_{\mathbb{C}}$] *L'espace $\ell^2_{\mathbb{C}}$ est complet.*

Démonstration. Soit donc :

$$z_k \,=\, \big(z_{k,1}, z_{k,2}, \dots, z_{k,i}, \dots\big) \,\in\, \ell^2 \qquad\qquad (k \geqslant 1),$$

une suite de Cauchy dans ℓ^2, c'est-à-dire satisfaisant :

$$\forall\, \varepsilon > 0 \quad \exists\, \mathrm{K} \geqslant 1 \quad \text{tel que} \quad \Big(k_1,\ k_2 \geqslant \mathrm{K} \Longrightarrow \|z_{k_1} - z_{k_2}\|_{\ell^2} \leqslant \varepsilon\Big).$$

Alors pour tout i, on a la majoration triviale du i-ème terme par une série positive infinie, ce qui nous permet de déduire :

$$\begin{aligned}\big|z_{k_1,i} - z_{k_2,i}\big| \,=\, \big(|z_{k_1,i} - z_{k_2,i}|^2\big)^{\frac{1}{2}} \,&\leqslant\, \bigg(\sum_{i=1}^{\infty} |z_{k_1,i} - z_{k_2,i}|^2\bigg)^{\frac{1}{2}} \\ &\leqslant\, \varepsilon.\end{aligned}$$

Grâce à cette inégalité, pour tout i fixé, la suite $(z_{k,i})_{k=1}^{\infty}$ d'éléments de $\mathbb{C}$ est de Cauchy. Puisque $\mathbb{C}$ est complet, il existe une unique limite :

$$z_{\infty,i} = \lim_{k \to \infty} z_{k,i}.$$

Posons alors :

$$z_\infty := \big(z_{\infty,1}, z_{\infty,2}, \dots, z_{\infty,i}, \dots\big).$$

On vient ainsi de démontrer que *chaque composante* $z_{k,i}$ de la suite $z_k = (z_{k,1}, z_{k,2}, \dots, z_{k,i}, \dots)$ de ℓ^2 converge vers un certain nombre complexe $z_{\infty,i}$. Maintenant, le problème qui se pose à nous, c'est de savoir si ce z_∞ appartient vraiment à ℓ^2, et ce n'est pas encore clair.

En effet, il se pourrait très bien que la somme $\sum_{i=1}^{\infty} |z_{\infty,i}|^2$ diverge vers ∞ et c'est bien là que le problème de la complétude se pose : l'élément-limite peut « sortir » de l'espace dans lequel la suite se situe. Il faut donc des arguments supplémentaires.

À cette fin, fixons maintenant un entier arbitraire $\text{I} \geqslant 1$. Pour tous $k_1, k_2 \geqslant \text{K}$, à nouveau pour la raison triviale que $\sum_{i=1}^{I} \leqslant \sum_{i=1}^{\infty}$, on a une majoration simple :

$$\sum_{i=1}^{\text{I}} \big|z_{k_1,i} - z_{k_2,i}\big|^2 \ \leqslant\ \|z_{k_1} - z_{k_2}\|^2 \ \leqslant\ \varepsilon^2.$$

Or à présent, nous pouvons faire tendre k_2 vers ∞, ce qui nous donne :

$$\sum_{i=1}^{\text{I}} \big|z_{k_1,i} - z_{\infty,i}\big|^2 \ \leqslant\ \varepsilon^2.$$

Mais l'entier I étant arbitraire, on peut le prendre arbitrairement grand, la somme $\sum_{i=1}^{I}$ reste toujours majorée par ε^2, et alors pour $\text{I} = \infty$, en renotant k au lieu de k_1, on obtient tout simplement que pour tout $k \geqslant \text{K}$:

$$\sum_{i=1}^{\infty} \big|z_{k,i} - z_{\infty,i}\big|^2 \ \leqslant\ \varepsilon^2.$$

Ceci montre en particulier que $z_k - z_\infty$ appartient à ℓ^2 !

Mais comme nous avons déjà établi que ℓ^2 est un espace vectoriel, nous en déduisons que :

$$z_\infty = z_k - (z_k - z_\infty) \in \ell^2.$$

Enfin, $\varepsilon > 0$ étant arbitraire, on conclut immédiatement de $\|z_k - z_\infty\|^2 \leqslant \varepsilon^2$ pour $k \geqslant \text{K}$ que $\lim_{k\to\infty} \|z_k - z_\infty\|^2 = 0$. □

Dans le cas de $\mathbb{R}^\infty \equiv \ell^2_{\mathbb{R}}$, des raisonnements très analogues, et d'ailleurs légèrement plus simples, conduisent à un résultat qu'il vaut la peine d'énoncer explicitement afin de bien en apercevoir les petites nuances.

Théorème 2.13. **[Structure de $\ell^2_{\mathbb{R}}$]** *L'espace $\ell^2_{\mathbb{R}}$ des suites* infinies *dénombrables :*

$$x := (x_1, x_2, \dots, x_i, \dots) \in \mathbb{R}^\infty$$

de nombres réels $x_i \in \mathbb{R}$ dont la somme des carrés converge :

$$\|x\|^2 := \sum_{i=1}^{\infty} x_i^2 \ <\ \infty,$$

est un $\mathbb{R}$-espace vectoriel normé complet *dont la norme :*

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

dérive du produit scalaire euclidien infini :

$$\langle x, y\rangle := \sum_{i=1}^{\infty} x_i\, y_i \qquad\qquad (x, y \in \ell^2).$$

En sus de l'inégalité triangulaire, il jouit de l'inégalité de Cauchy-Schwarz :

$$\big|\langle x, y\rangle\big| \leqslant \|x\|\,\|y\| \qquad\qquad (x, y \in \ell^2),$$

qui n'est égalité que lorsque x et y sont colinéaires. □

3. Espaces préhilbertiens

Toujours pour fixer les idées, nous travaillerons avec des $\mathbb{C}$-espaces vectoriels en précisant parfois les simplifications agréables dont bénéficient les $\mathbb{R}$-espaces vectoriels.

La plupart du temps, ces espaces seront entendus dans l'intuition comme étant de dimension infinie, puisque la dimension finie est (trop) simple !

Définition 3.1. Une application $f\colon E \longrightarrow \mathbb{C}$ définie sur un $\mathbb{C}$-espace vectoriel E est dite *anti-linéaire* si elle satisfait :

$$\begin{aligned} f(z+w) &= f(z)+f(w) && {\scriptstyle (z,\,w\,\in\,E)}, \\ f(\lambda\, z) &= \overline{\lambda}\, z && {\scriptstyle (\lambda\,\in\,\mathbb{C},\, z\,\in\,E)}. \end{aligned}$$

Autrement dit, les multiplications par des scalaires sont *perturbées* par une conjugaison complexe.

Définition 3.2. Une application $\varphi\colon E\times E \longrightarrow \mathbb{C}$ est dite *sesquilinéaire* si elle est linéaire par rapport à la première variable et anti-linéaire par rapport à la seconde variable :

$$\begin{aligned} \varphi(\lambda z+\mu w, t) &= \lambda\, \varphi(z,t)+\mu\,\varphi(w,t), \\ \varphi(z, \mu w+\nu t) &= \overline{\mu}\,\varphi(z,w)+\overline{\nu}\,\varphi(z,t), \end{aligned}$$

pour tous $z, w, t \in E$ et tous $\lambda, \mu, \nu \in \mathbb{C}$.

Lorsque le corps de l'espace vectoriel est $\mathbb{R}$, une forme sesquilinéaire :

$$\varphi\colon E\times E \longrightarrow \mathbb{R}$$

est tout simplement une forme $\mathbb{R}$-bilinéaire

Définition 3.3. Une telle forme sesquilinéaire est dite *hermitienne* si l'on a de plus :

$$\varphi(z,w) = \overline{\varphi(w,z)},$$

pour tous $z, w \in E$.

Sur $\mathbb{R}$, une forme 'hermitienne' est tout simplement une forme symétrique :

$$\varphi(x,y) = \varphi(y,x) \qquad {\scriptstyle (\forall\, x,\,y\,\in\,E)}.$$

Proposition 3.4. *Toute forme sesquilinéaire hermitienne $\varphi\colon E\times E \longrightarrow \mathbb{C}$ est déterminée de manière unique par sa restriction à la diagonale :*

$$\big\{(z,z)\in E\times E\colon\ z\in E\big\},$$

autrement dit par la connaissance de la fonction :

$$\mathbb{C} \ni z \longmapsto \varphi(z,z) \in \mathbb{C},$$

grâce à la formule dite de polarisation *:*

$$4\,\varphi(z,w) = \varphi\big(z+w, z+w\big)-\varphi\big(z-w, z-w\big)+\sqrt{-1}\,\varphi\big(z+\sqrt{-1}\,w,\ z+\sqrt{-1}\,w\big)-\sqrt{-1}\,\varphi\big(z-\sqrt{-1}\,w,\ z-\sqrt{-1}\,w\big).$$

Démonstration. Vérifier cette formule qui est non seulement belle mais s'avérera très souvent utile ultérieurement est un jeu d'enfant, puisqu'il suffit de développer le membre de :

$$\begin{aligned}
\text{droite} \;&=\; \varphi(z,z)+\varphi(z,w)+\varphi(w,z)+\varphi(w,w)- \\
&\quad -\;\varphi(z,z)+\varphi(z,w)+\varphi(w,z)-\varphi(w,w)+ \\
&\quad +\sqrt{-1}\,\varphi(z,z)+\varphi(z,w)-\varphi(w,z)-\sqrt{-1}\,\varphi(w,w)- \\
&\quad -\sqrt{-1}\,\varphi(z,z)+\varphi(z,w)-\varphi(w,z)+\sqrt{-1}\,\varphi(w,w) \\
&= \qquad 0 \qquad +4\,\varphi(z,w)+ \quad 0 \quad + \quad 0,
\end{aligned}$$

et d'additionner chacune des 4 colonnes. □

Observons au passage que cette démonstration n'a pas utilisé l'hermitianité de φ.

Lemme 3.5. *Sur $\mathbb{R}$, toute forme bilinéaire symétrique est déterminée par ses valeurs sur la diagonale grâce à la formule* de polarisation *:*

$$4\,\varphi(x,y) \;=\; \varphi\big(x+y,x+y\big)-\varphi\big(x-y,x-y\big).$$

Démonstration. Mais ici, on utilise vraiment le caractère symétrique de φ :

$$\begin{aligned}
\text{droite} \;&=\; \varphi(x,x)+\varphi(x,y)+\varphi(y,x)+\varphi(y,y)- \\
&\quad -\;\varphi(x,x)+\varphi(x,y)+\varphi(y,x)-\varphi(y,y) \\
&= \qquad 0 \quad +2\,\varphi(x,y)+2\,\varphi(y,x)- \quad 0,
\end{aligned}$$

sinon, le calcul n'aboutit pas. □

En effet par (contre-)exemple, la forme *anti*-symétrique bilinéaire non nulle sur $\mathbb{R}^2$:

$$\big((x_1,x_2),\,(y_1,y_2)\big) \;\longmapsto\; x_1\,y_2-x_2\,y_1$$

s'annule identiquement sur la diagonale. Mais de toute manière, nous ne nous intéresserons qu'aux formes symétriques ou hermitiennes.

Définition 3.6. Un *produit scalaire hermitien* sur un $\mathbb{C}$-espace vectoriel E — de dimension finie ou de dimension infinie — est une forme sesquilinéaire hermitienne :

$$\varphi\colon \quad E\times E\longrightarrow\mathbb{C}$$

qui est *définie positive* au sens où :

$$\varphi(z,z)>0,$$

pour tout $z\in E\backslash\{0\}$ non nul.

La notion analogue existe également pour les $\mathbb{R}$-espaces vectoriels.

On définit aussi une forme hermitienne (faiblement) positive par :

$$\varphi(z,z)\;\geqslant\;0 \qquad\qquad (\forall\,z\in E).$$

Classiquement, on note de manière abrégée $\langle\cdot,\cdot\rangle$ au lieu de $\varphi(\cdot,\cdot)$ un produit scalaire, complexe hermitien ou réel symétrique.

Définition 3.7. Un *espace pré-hilbertien* sur $\mathbb{R}$ ou sur $\mathbb{C}$ est un $\mathbb{R}$- ou $\mathbb{C}$-espace vectoriel E muni d'un produit scalaire $\langle\cdot,\cdot\rangle$.

Si E est défini sur $\mathbb{R}$, on peut lui associer son *complexifié* $E^{\mathbb{C}} = E + \sqrt{-1}\, E$, qui est le $\mathbb{C}$-espace vectoriel constitué des élements :

$$x + \sqrt{-1}\, y \qquad\qquad (x, y \in E)$$

que l'on additionne de manière évidente, et que l'on multiplie par des constantes complexes :

$$(\alpha + \sqrt{-1}\, \beta) \cdot \big(x + \sqrt{-1}\, y\big) \ := \ \alpha\, x - \beta\, y + \sqrt{-1}\, \big(\alpha\, y + \beta\, x\big).$$

Si $\langle \cdot, \cdot \rangle_E$ est un produit scalaire (réel) sur E, alors la formule :

$$\big\langle x + \sqrt{-1}\, y,\ u + \sqrt{-1}\, v \big\rangle_{E^{\mathbb{C}}} \ := \ \langle x, u \rangle_E + \langle y, v \rangle_E - \sqrt{-1}\, \langle x, v \rangle_E + \sqrt{-1}\, \langle y, u \rangle_E$$

offre un produit scalaire (hermitien : exercice !) sur $E^{\mathbb{C}}$.

Proposition 3.8. *Soit E un espace vectoriel complexe de dimension finie ou infinie muni d'un produit scalaire hermitien $\langle \cdot, \cdot \rangle$. Alors l'application :*

$$z \longmapsto \sqrt{\langle z, z \rangle} \ =: \ ||z||$$

définit une norme *sur E qui satisfait l'inégalité de Cauchy-Schwarz :*

$$\big|\langle z, w \rangle\big| \leqslant ||z|| \cdot ||w|| \qquad\qquad (\forall\, z, w \in E),$$

avec égalité seulement lorsque z et w sont colinéaires.

Démonstration. Les arguments sont les mêmes que dans la démonstration du Théorème 2.10 — sans changer les symboles ! —, à ceci près que le produit scalaire est maintenant abstrait, général. □

Il importe de noter que l'inégalité de Cauchy-Schwarz reste vraie sans hypothèse de stricte positivité, car cela est parfois utilisé.

Proposition 3.9. *Pour toute forme hermitienne faiblement positive $\varphi(z, z) \geqslant 0$ sur un $\mathbb{C}$-espace vectoriel E :*

$$\big|\varphi(z, w)\big| \ \leqslant \ \sqrt{\varphi(z, z)} \cdot \sqrt{\varphi(w, w)} \qquad\qquad (\forall\, z, w \in E).$$

Démonstration. Le même argument utilisant un polynôme du second degré fonctionne (exercice), mais les cas d'égalités peuvent être (beaucoup) plus nombreux que les circonstances de colinéarité, à cause de la présence de vecteurs dits *isotropes* sur lesquels $\varphi(z, z) = 0$ dégénère. □

Proposition 3.10. *Sur E muni d'un produit scalaire $\langle \cdot, \cdot \rangle$, on a l'inégalité triangulaire :*

$$||z + w|| \leqslant ||z|| + ||w|| \qquad\qquad (\forall\, z, w \in E),$$

avec égalité lorsque, et seulement lorsque $z = 0$, ou $w = 0$, ou encore lorsque $w = c\, z$ avec $c \in \mathbb{R}_+^*$.

Démonstration. On calcule, on majore :

$$\begin{aligned}
\big\|z + w\big\|^2 \ &= \ \big\langle z + w, z + w \big\rangle \\
&= \ \langle z, z \rangle + \langle z, w \rangle + \langle w, z \rangle + \langle w, w \rangle \\
&= \ ||z||^2 + 2\, \mathrm{Re}\, \langle z, w \rangle + ||w||^2 \\
\text{[Cauchy-Schwarz !]} \qquad &\leqslant \ ||z||^2 + 2\, ||z||\, ||w|| + ||w||^2 \\
&= \ \big(||z|| + ||w||\big)^2,
\end{aligned}$$

et on retrouve la triangulaire au carré !

Puisqu'on a trivialement égalité lorsque $z = 0$ ou $w = 0$, on peut supposer que $z \neq 0 \neq w$.

Si $w = c\,z$ avec $c > 0$, alors clairement :

$$\|z + w\| \;=\; \big\|(1 + c)\,z\big\| \;=\; (1 + c)\,\|z\| \;=\; \|z\| + c\,\|z\| \;=\; \|z\| + \|w\|.$$

Réciproquement, une inspection du calcul qui précède — lequel n'incorporait qu'une seule inégalité — montre qu'on aura égalité $\|z + w\| = \|z\| + \|w\|$ seulement si on a eu égalité au moment d'appliquer Cauchy-Schwarz, d'où la colinéarité $w = c\,z$ avec $c \in \mathbb{R}^*$, et si donc on demande que :

$$\big\|(1 + c)\,z\big\| \overset{?}{=} \|z\| + |c|\,\|z\|,$$

à savoir que $|1 + c| = 1 + |c|$, cela force $c > 0$. □

Théorème 3.11. *Dans un $\mathbb{C}$- ou $\mathbb{R}$-espace vectoriel E de dimension finie ou infinie muni d'un produit scalaire $\langle \cdot, \cdot \rangle$, pour tous éléments $z, w \in E$, les identités remarquables suivantes sont satisfaites.*

- **Développement d'un carré :**

$$\|z + w\|^2 \;=\; \|z\|^2 + 2\,\mathrm{Re}\,\langle z, w\rangle + \|w\|^2.$$

- **Identité de polarisation sur $\mathbb{C}$:**

$$4\,\langle z, w\rangle \;=\; \big\|z + w\big\|^2 - \big\|z - w\big\|^2 + \sqrt{-1}\,\big\|z + \sqrt{-1}\,w\big\|^2 - \sqrt{-1}\,\big\|z - \sqrt{-1}\,w\big\|^2.$$

- **Identité de polarisation sur $\mathbb{R}$:**

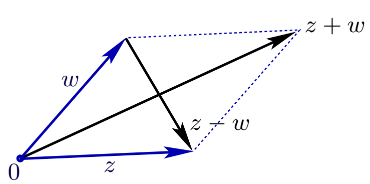
$$4\,\langle x, y\rangle \;=\; \big\|x + y\big\|^2 - \big\|x - y\big\|^2.$$

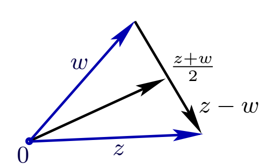
- **Identité du parallélogramme :**

$$\big\|z + w\big\|^2 + \big\|z - w\big\|^2 \;=\; 2\,\|z\|^2 + 2\,\|w\|^2.$$

- **Formule de la médiane :**

$$2\,\Big\|\tfrac{z+w}{2}\Big\|^2 + \tfrac{1}{2}\,\big\|z - w\big\|^2 \;=\; \|z\|^2 + \|w\|^2.$$





Démonstration. Les trois premières identités ont déjà été vues et expliquées.

Pour le parallélogramme, il suffit de développer sans réfléchir :

$$\begin{aligned}
\big\|z + w\big\|^2 + \big\|z - w\big\|^2 &= \big\langle z + w, z + w\big\rangle + \big\langle z - w, z - w\big\rangle \\
&= \langle z, z\rangle + \underline{\langle z, w\rangle}_{\circ} + \underline{\langle w, z\rangle}_{\circ\circ} + \langle w, w\rangle + \langle z, z\rangle - \underline{\langle z, w\rangle}_{\circ} - \underline{\langle w, z\rangle}_{\circ\circ} + \langle w, w\rangle \\
&= 2\,\|z\|^2 + 2\,\|w\|^2.
\end{aligned}$$

Chose amusante, la médiane et le parallélogramme sont une seule et même chose à dilatation près :

$$\begin{aligned} 2\left\|\tfrac{z+w}{2}\right\|^2+\tfrac{1}{2}\left\|z-w\right\|^2 &= \tfrac{2}{2^2}\left\|z+w\right\|^2+\tfrac{1}{2}\left\|z-w\right\|^2 \\ &= \|z\|^2+\|w\|^2, \end{aligned}$$

et donc, il est tout à fait permis de les confondre ! □

Définition 3.12. Deux éléments $z, w \in E$ d'un espace vectoriel muni d'un produit scalaire $\langle\cdot,\cdot\rangle$ sont dits *orthogonaux*, ce qu'on abrège par $z \perp w$, lorsque :

$$0 = \langle z, w\rangle.$$

Même en dimension infinie, le plus antique des héros mathématiques parle d'or.

Théorème 3.13. [de Pythagore] *Si $z \perp w$, alors :*

$$\left\|z+w\right\|^2 = \|z\|^2+\|w\|^2.$$ □

4. L'espace de Hilbert $L^2(\mathbb{R}^d)$

Il y a au moins deux raisons profondes qui expliquent l'importance et l'omniprésence des espaces de Hilbert dans les mathématiques.

Premièrement, ils apparaissent comme généralisation naturelle des espaces euclidiens de dimension finie : de même que le plan $\mathbb{R}^2$, l'espace $\mathbb{R}^3$ ou l'espace $\mathbb{R}^n$ (voire dans $\mathbb{C}^n$), ils bénéficient des propriétés familières telles que le produit scalaire, l'orthogonalité, les projections (orthogonale), le théorème de Pythagore, *etc.* Crucialement, on demande que les espaces de Hilbert abstraits soient *complets* pour la topologie qui dérive de leur norme. Cette condition de complétude est cruciale, et on la requiert afin que les constructions géométriques et les passages à la limite se passent aussi bien en dimension infinie qu'en dimension finie, et aussi, afin qu'il existe de bonnes bases orthonormées — de cardinal infini — le long desquelles on puisse décomposer chaque vecteur.

En fait, la complétude est une hypothèse naturelle qui est satisfaite par les objets mathématiques connus, puisque les espaces $\ell^2_{\mathbb{R}}$ et $\ell^2_{\mathbb{C}}$ sont complets, et puisque l'espace des fonctions $f\colon \mathbb{R}^d \longrightarrow \mathbb{C}$ mesurables dites *de carré intégrable :*

$$\int_{\mathbb{R}^d}\left|f(x)\right|^2 dx < \infty,$$

espace paradigmatique auquel cette section est consacrée, est tout aussi complet. Ici bien entendu, le point :

$$x = \big(x_1, \dots, x_d\big),$$

possède d coordonnées $\in \mathbb{R}$, et :

$$dx = dx_1 \cdots dx_d$$

désigne la mesure de Lebesgue.

Deuxièmement, la théorie des espaces de Hilbert s'est avérée au fil du temps être un cadre conceptuel très souple et très puissant, presque un langage à part entière voire une sorte d'*esperanto* mathématique, grâce auquel on peut formuler et résoudre élégamment des problèmes d'analyse mathématique, en les plaçant dans un espace fonctionnel adéquat riche de structures géométriques.

Soit donc un entier :

$$d \geqslant 1.$$

Après $\ell^2_{\mathbb{R}}$, le deuxième exemple canonique et ultra-concret d'espace de Hilbert est donc :

$$L^2(\mathbb{R}^d) := \left\{ f \colon \mathbb{R}^d \to \mathbb{C} \text{ mesurables} \colon \int_{\mathbb{R}^d} |f(x)|^2 \, dx < \infty \right\}.$$

De manière analogue, l'exemple des fonctions L^2 sur le cercle unité, ou, de manière équivalente, des fonctions 2π-périodiques sur $\mathbb{R}$:

$$L^2\big([-\pi, \pi]\big),$$

effectue un lien historique profond et originaire entre les séries de Fourier et les espaces de Hilbert (voir prochain chapitre).

En tout cas, sur l'espace $L^2(\mathbb{R}^d)$ qui se comporte au mieux, la norme est naturellement définie par :

$$\big|\big|f\big|\big|_{L^2(\mathbb{R}^d)} := \left(\int_{\mathbb{R}^d} \big|f(x)\big|^2 \, dx \right)^{\frac{1}{2}}.$$

Une différence *cruciale* entre l'espace $L^1(\mathbb{R}^d)$ des fonctions Lebesgue-intégrables sur $\mathbb{R}^d$ et cet espace $L^2(\mathbb{R}^d)$, c'est l'existence d'un *produit scalaire :*

$$\langle f, g \rangle_{L^2(\mathbb{R}^d)} := \int_{\mathbb{R}^d} f(x) \, \overline{g(x)} \, dx,$$

entre deux fonctions quelconques :

$$\begin{aligned} f \colon \quad \mathbb{R}^d &\longrightarrow \mathbb{C}, \\ g \colon \quad \mathbb{R}^d &\longrightarrow \mathbb{C}, \end{aligned}$$

à partir duquel on retrouve la norme :

$$\big|\big|f\big|\big|_{\mathbb{L}^2(\mathbb{R}^d)} = \big(\langle f, f \rangle_{L^2(\mathbb{R}^d)}\big)^{1/2}.$$

On démontre qu'il n'est en aucune manière possible de définir un produit scalaire sur l'espace de Lebesgue $L^1(\mathbb{R}^d)$ qui redonnerait sa norme naturelle $\int |f|$.

Toutefois, comme pour l'espace $L^1(\mathbb{R}^d)$, la condition :

$$\big|\big|f\big|\big|_{L^2(\mathbb{R}^d)} = 0$$

implique que $f(x) = 0$ pour presque tout $x \in \mathbb{R}^d$. Si donc on identifie comme il se doit les fonctions qui sont égales à un ensemble de mesure nulle près, l'espace $L^2(\mathbb{R}^d)$ est alors muni d'une vraie norme $\|\cdot\|_{L^2(\mathbb{R}^d)}$.

Or pour que la définition du produit scalaire :

$$\int_{\mathbb{R}^d} f(x) \, \overline{g(x)} \, dx$$

ait un sens, il faut au moins s'assurer que la fonction $f \, \overline{g}$ soit intégrable sur $\mathbb{R}^d$, à savoir que $f \, \overline{g} \in L^1(\mathbb{R}^d)$. Tel est bien le cas, grâce à des arguments très élémentaires.

Proposition 4.1. *L'espace* $L^2(\mathbb{R}^d)$ *jouit des propriétés suivantes.*

(i) $L^2(\mathbb{R}^d)$ *est un espace vectoriel.*

(ii) *$x \longmapsto f(x)\,\overline{g(x)}$ est intégrable chaque fois que $f, g \in L^2(\mathbb{R}^d)$, et l'inégalité de Cauchy-Schwarz :*

$$\big|\langle f,\, g\rangle_{L^2(\mathbb{R}^d)}\big| \;\leqslant\; \big|\big|f\big|\big|_{L^2(\mathbb{R}^d)}\,\big|\big|g\big|\big|_{L^2(\mathbb{R}^d)}$$

est généralement satisfaite.

(iii) *Si une fonction $g \in L^2(\mathbb{R}^d)$ est fixée, alors l'application :*

$$L_g\colon\quad \begin{array}{rcl} L^2(\mathbb{R}^d) & \longrightarrow & \mathbb{C} \\ f & \longmapsto & \langle f, g\rangle_{L^2(\mathbb{R}^d)} \end{array}$$

est $\mathbb{C}$-linéaire continue :

$$\big|L_g(f)\big| \;\leqslant\; \underbrace{\big|\big|g\big|\big|_{L^2(\mathbb{R}^d)}}_{\text{constante}}\,\big|\big|f\big|\big|_{L^2(\mathbb{R}^d)},$$

avec de plus :

$$\langle f,\, g\rangle_{L^2(\mathbb{R}^d)} = \overline{\langle g,\, f\rangle}_{L^2(\mathbb{R}^d)}.$$

(iv) *L'inégalité du triangle est satisfaite :*

$$\big|\big|f+g\big|\big|_{L^2(\mathbb{R}^d)} \;\leqslant\; \big|\big|f\big|\big|_{L^2(\mathbb{R}^d)} + \big|\big|g\big|\big|_{L^2(\mathbb{R}^d)}.$$

Dans la suite, on s'autorisera à admettre l'équivalence notationnelle flottante :

$$||\cdot|| \;\equiv\; ||\cdot||_{L^2} \;\equiv\; \big|\big|\cdot\big|\big|_{L^2(\mathbb{R}^d)},$$

ainsi que :

$$\langle\cdot,\cdot\rangle \;\equiv\; \langle\cdot,\cdot\rangle_{L^2} \;\equiv\; \big\langle\cdot,\cdot\big\rangle_{L^2(\mathbb{R}^d)},$$

suivant que l'on souhaite alléger l'écriture, ou que l'on préfère notifier explicitement de quelle norme précise il s'agit.

Démonstration. Pour $f, g \in L^2(\mathbb{R}^d)$, et $x \in \mathbb{R}^d$, de l'inégalité :

$$\begin{aligned} \big|f(x)+g(x)\big| &\;\leqslant\; |f(x)| + |g(x)| \\ &\;\leqslant\; 2\max\big(|f(x)|,\, |g(x)|\big), \end{aligned}$$

on déduit :

$$\begin{aligned} \big|f(x)+g(x)\big|^2 &\;\leqslant\; 4\max\big(|f(x)|^2 + |g(x)|^2\big) \\ &\;\leqslant\; 4\,|f(x)|^2 + 4\,|g(x)|^2, \end{aligned}$$

d'où :

$$\begin{aligned} \int \big|f+g\big|^2 &\leqslant 4\int |f|^2 + 4\int |g|^2 \\ &< \infty, \end{aligned}$$

ce qui montre que $f+g \in L^2(\mathbb{R}^d)$.

Aussi, puisque pour tout $\lambda \in \mathbb{C}$:

$$\big|\big|\lambda\, f\big|\big|_{L^2(\mathbb{R}^d)} = |\lambda|\,\big|\big|f\big|\big|_{L^2(\mathbb{R}^d)},$$

on conclut que $L^2(\mathbb{R}^d)$ est bien un espace vectoriel.

Ensuite, concernant $(\mathbf{ii})$, pour voir que $f\,\overline{g}$ appartient à $L^1(\mathbb{R}^d)$, il suffit de se rappeler — réminiscence des mathématiques babyloniennes — que pour tous $A, B \geqslant 0$ réels positifs, on a :

$$2\,AB \;\leqslant\; A^2 + B^2,$$

de telle sorte que :

$$\begin{aligned}\int_{\mathbb{R}^d} \big|f(x)\,\overline{g}(x)\big|\,dx \;&\leqslant\; \frac{1}{2}\Big[\big\|f\big\|^2_{L^2(\mathbb{R}^d)} + \big\|g\big\|^2_{L^2(\mathbb{R}^d)}\Big]\\ &<\;\infty.\end{aligned}$$

Pour démontrer l'inégalité de Cauchy-Schwarz :

$$\bigg|\int_{\mathbb{R}^d} f(x)\,\overline{g(x)}\,dx\bigg| \;\overset{?}{\leqslant}\; \big\|f\big\|_{L^2(\mathbb{R}^d)}\,\big\|g\big\|_{L^2(\mathbb{R}^d)},$$

observons d'abord que si $\|f\| = 0$ ou si $\|g\| = 0$, alors $f\,\overline{g} = 0$ presque partout, d'où $\langle f, g\rangle = 0$ et l'inégalité est triviale : $0 \leqslant 0$. Nous pouvons donc supposer que $\|f\| > 0$ et que $\|g\| > 0$, donc en divisant f et g par leurs normes respectives :

$$\widetilde{f} := \frac{f}{\|f\|} \qquad\text{et}\qquad \widetilde{g} := \frac{g}{\|g\|},$$

ce qui ne change pas l'inégalité à démontrer puisqu'elle est doublement homogène, l'inégalité à démontrer devient :

$$\int_{\mathbb{R}^d} \big|\widetilde{f}(x)\,\overline{\widetilde{g}(x)}\big|\,dx \;\overset{?}{\leqslant}\; 1,$$

mais l'inégalité que nous venons d'obtenir, appliquée à $\widetilde{f}$ et à $\widetilde{g}$, donne justement le résultat voulu :

$$\int_{\mathbb{R}^d} \big|\widetilde{f}(x)\,\overline{\widetilde{g}(x)}\big|\,dx \;\leqslant\; \frac{1}{2}\Big[\underbrace{\big\|\widetilde{f}\big\|^2_{L^2(\mathbb{R}^d)}}_{=1^2} + \underbrace{\big\|\widetilde{g}\big\|^2_{L^2(\mathbb{R}^d)}}_{=1^2}\Big],$$

puisque $\frac{1}{2}[1^2 + 1^2] = 1$!

L'assertion **(iii)** découle de la linéarité de l'intégrale, et la continuité de L_g découle de l'inégalité de Cauchy-Schwarz (exercice de révision mentale sur les opérateurs linéaires continus d'un espace vectoriel normé, résolu en détail page 42).

Enfin, pour démontrer l'inégalité triangulaire **(iv)**, on applique l'inégalité de Cauchy-Schwarz simplement comme suit :

$$\begin{aligned}\big\|f+g\big\|^2 &= \langle f+g,\, f+g\rangle\\ &= \|f\|^2 + \langle f, g\rangle + \langle g, f\rangle + \|g\|^2\\ &\leqslant \|f\|^2 + 2\,\big|\langle f, g\rangle\big| + \|g\|^2\\ &\leqslant \|f\|^2 + 2\,\|f\|\,\|g\| + \|g\|^2\\ &= \big(\|f\| + \|g\|\big)^2\end{aligned}$$

et il suffit de prendre la racine carrée de l'inégalité obtenue. □

Tournons maintenant notre attention vers la notion de limite dans l'espace $L^2(\mathbb{R}^d)$. La norme $\|\cdot\|_{L^2(\mathbb{R}^d)}$ induit une métrique naturelle en déclarant que la *distance* entre deux fonctions :

$$f \,\in\, L^2(\mathbb{R}^d) \qquad\text{et}\qquad g \,\in\, L^2(\mathbb{R}^d)$$

est la norme de leur différence :

$$\mathsf{dist}\,\big(f, g\big) := \big\|f - g\big\|_{L^2(\mathbb{R}^d)}.$$

Une suite de fonctions :

$$\big(f_n\big)_{n=1}^{\infty} \,\in\, L^2(\mathbb{R}^d)$$

est dite *de Cauchy* si :

$$\mathsf{dist}\big(f_{n_1},\, f_{n_2}\big) \,\longrightarrow\, 0 \qquad \text{lorsque } \; n_1,\, n_2 \,\longrightarrow\, \infty.$$

On dit aussi qu'une telle suite converge vers une fonction-limite :

$$f \,\in\, L^2(\mathbb{R}^d),$$

si l'on a :

$$0 = \lim_{n\,\to\,\infty} \mathsf{dist}\,\big(f_n, f\big).$$

Théorème 4.2. *L'espace $L^2(\mathbb{R}^d)$ est complet pour sa métrique :*

$$\mathsf{dist}\big(f, g\big) = \big\| f - g \big\|_{L^2(\mathbb{R}^d)}.$$

En d'autres termes, toute suite de Cauchy de fonctions de $L^2(\mathbb{R}^d)$ admet une fonction-limite qui appartient encore à $L^2(\mathbb{R}^d)$. Ce théorème, qui contraste fortement avec la situation de l'intégrale de Riemann, montre à merveille toute la puissance et toute l'utilité de l'intégrale de Lebesgue.

Démonstration. L'argument est très analogue à la démonstration de la complétude de $L^1(\mathbb{R}^d)$, et nous en détaillons les ingrédients afin de mieux (ré-)intégrer mentalement la théorie.

Étant donné une suite arbitraire de fonctions mesurables de carré intégrable :

$$\big(f_n\big)_{n=1}^{\infty} \,\in\, L^2(\mathbb{R}^d),$$

qui est *de Cauchy* pour la norme L^2 :

$$\begin{aligned}
&\forall\, \varepsilon > 0 \quad \exists\, \mathrm{N} = \mathrm{N}(\varepsilon) \gg 1 \\
&\Big(n,\; m \,\geqslant \mathrm{N}\Big) \;\Longrightarrow\; \Big(\big\| f_n - f_m \big\|_{L^2(\mathbb{R}^d)} \,\leqslant\, \varepsilon\Big),
\end{aligned} \tag{4.3}$$

l'objectif, pour établir la complétude de L^2, est de trouver ou de construire une certaine fonction $f \in L^2(\mathbb{R}^d)$ telle que :

$$0 \,=\, \lim_{n \to \infty} \big\| f_n - f \big\|_{L^2(\mathbb{R}^d)}.$$

À cette fin, extrayons tout d'abord une sous-suite auxiliaire :

$$\big(f_{n_k}\big)_{k=1}^{\infty},$$

ayant la propriété de converger assez rapidement en norme L^2 :

$$\big\| f_{n_{k+1}} - f_{n_k} \big\|_{L^2(\mathbb{R}^d)} \,\leqslant\, \frac{1}{2^k} \qquad\qquad (\forall\, k \geqslant 1),$$

pour que la fonction (mesurable) $F \colon \mathbb{R}^d \longrightarrow \mathbb{R}_+ \cup \{\infty\}$ définie par :

$$F(x) \,:=\, f_{n_1}(x) + \sum_{k=1}^{\infty} \big| f_{n_{k+1}}(x) - f_{n_k}(x) \big|$$

ait une norme L^2 finie :

$$\begin{aligned}\big\|F\big\|_{L^2(\mathbb{R}^d)} &\leqslant \big\|f_{n_1}\big\|_{L^2(\mathbb{R}^d)} + \sum_{k=1}^{\infty} \big\|f_{n_{k+1}} - f_{n_k}\big\|_{L^2(\mathbb{R}^d)} \\ &\leqslant \big\|f_{n_1}\big\|_{L^2(\mathbb{R}^d)} + \sum_{k=1}^{\infty} \frac{1}{2^k} \\ &\leqslant \big\|f_{n_1}\big\|_{L^2(\mathbb{R}^d)} + 1 \\ &< \infty,\end{aligned}$$

ce qui signifie que :

$$F \in L^2(\mathbb{R}^d).$$

Autrement dit, l'intégrale de son carré est finie :

$$\int_{\mathbb{R}^d} F(x)^2\, dx < \infty,$$

et ceci force ses valeurs à être finies en presque tout $x \in \mathbb{R}^d$:

$$0 \leqslant F(x) < \infty.$$

Ensuite, pour tout entier $\mathrm{K} \geqslant 1$, la fonction issue d'une simplification téléscopique :

$$f_{n_{\mathrm{K}}}(x) = f_{n_1}(x) + \sum_{k=1}^{\mathrm{K}-1} \big(f_{n_{k+1}}(x) - f_{n_k}(x)\big)$$

satisfait alors uniformément par rapport à K :

$$\begin{aligned}\big|f_{n_{\mathrm{K}}}(x)\big| &\leqslant \big|f_{n_1}(x)\big| + \sum_{k=1}^{\mathrm{K}-1} \big|f_{n_{k+1}}(x) - f_{n_k}(x)\big| \\ &\leqslant F(x),\end{aligned}$$

et donc, en faisant $\mathrm{K} \longrightarrow \infty$, nous voyons que la série infinie $\sum_{k=1}^{\infty}$ converge absolument pour définir, en presque tout $x \in \mathbb{R}^d$, une certaine fonction mesurable finie :

$$\lim_{\mathrm{K}\to\infty} f_{n_{\mathrm{K}}}(x) =: f(x) \in \mathbb{R}. \tag{4.4}$$

C'est elle, notre fonction désirée, nous l'avons façonnée ! Car de ce qui précède, l'inégalité valable presque partout :

$$|f| \leqslant F,$$

d'où $|f|^2 \leqslant F^2$ — souvenons-nous que f est à valeurs dans $\mathbb{C}$, ce qui justifie ici la notation module $|f|$ — nous assure crucialement que :

$$\int_{\mathbb{R}^d} |f(x)|^2\, dx \leqslant \int_{\mathbb{R}^d} F(x)^2\, dx < \infty,$$

ce qui montre que :

$$f \in L^2(\mathbb{R}^d).$$

Naturellement, on s'attend à ce que :

$$0 = \lim_{k\to\infty} \big\|f_{n_k} - f\big\|_{L^2(\mathbb{R}^d)},$$

et pour vérifier cela, on remarque que l'inégalité valable par construction presque partout :

$$\big|f_{n_k} - f\big| \leqslant 2\,F$$

assure la domination uniforme :

$$\begin{aligned}\big\|f_{n_k} - f\big\|_{L^2(\mathbb{R}^d)} &\leqslant 2\,\|F\|_{L^2(\mathbb{R}^d)} \\ &< \infty,\end{aligned}$$

donc grâce au théorème de la convergence dominée de Lebesgue, on conclut que la sous-suite $\big(f_{n_k}\big)_{k=1}^{\infty}$ de la suite $\big(f_n\big)_{n=1}^{\infty}$ converge en norme L^2 vers la fonction $f \in L^2(\mathbb{R}^d)$:

$$\begin{aligned}\lim_{k\to\infty} \int_{\mathbb{R}^d} \big|f_{n_k}(x) - f(x)\big|^2\,dx &= \int_{\mathbb{R}^d} \lim_{k\to\infty} \big|f_{n_k}(x) - f(x)\big|^2\,dx \\ \text{[Utiliser (4.4)]}\qquad &= 0.\end{aligned}$$

Pour achever la démonstration du théorème de complétude de $L^2(\mathbb{R}^d)$, il reste encore à faire voir que la *suite entière* $\big(f_n\big)_{n=1}^{\infty}$ converge en norme L^2 vers la même fonction-limite f que la sous-suite $\big(f_{n_k}\big)_{k=1}^{\infty}$.

Or cela découle logiquement de tout ce qui a été dit, puisque, si $\varepsilon > 0$ est arbitrairement petit, en revenant à l'expression (4.3) de la Cauchycité de $(f_n)_{n=1}^{\infty}$, si un indice $\mathrm{K} \gg 1$ de la sous-suite est choisi assez grand pour que :

$$n_{\mathrm{K}} \geqslant \mathrm{N}(\varepsilon),$$

et pour que, grâce à la convergence qui vient d'être établie :

$$\big\|f_{n_{\mathrm{K}}} - f\big\|_{L^2(\mathbb{R}^d)} \leqslant \varepsilon,$$

alors pour tout $n \geqslant \mathrm{N}(\varepsilon)$, une simple inégalité du triangle :

$$\begin{aligned}\big\|f_n - f\big\|_{L^2(\mathbb{R}^d)} &\leqslant \big\|f_n - f_{n_{\mathrm{K}}}\big\|_{L^2(\mathbb{R}^d)} + \big\|f_{n_{\mathrm{K}}} - f\big\|_{L^2(\mathbb{R}^d)} \\ &\leqslant \varepsilon + \varepsilon,\end{aligned}$$

offre la délivrance. □

Une propriété additionnelle de $L^2(\mathbb{R}^d)$, surprenante au premier abord mais très proche du théorème de Weierstrass d'approximation des fonctions continues par des polynômes, stipule qu'il suffit d'une famille *dénombrable* de fonctions pour approximer toutes les fonctions possibles.

Théorème 4.5. *Dans l'espace $L^2(\mathbb{R}^d)$, il existe une* suite *de fonctions :*

$$\big(f_n\big)_{n=1}^{\infty} \in L^2(\mathbb{R}^d)$$

qui est dense *:*

$$\forall\, f \in L^2(\mathbb{R}^d) \quad \forall\, \varepsilon > 0 \quad \exists\; f_{\mathrm{N}(\varepsilon)} \in \big(f_n\big)_{n=1}^{\infty} \quad \big\|f - f_{\mathrm{N}(\varepsilon)}\big\|_{L^2(\mathbb{R}^d)} \leqslant \varepsilon.$$

Démonstration. Introduisons la famille de toutes les fonctions combinaisons linéaires finies :

$$\sum_{\text{finie}} \lambda_{\mathbb{Q}} \cdot \mathbf{1}_{R_{\mathbb{Q}}}(x),$$

de fonctions indicatrices de rectangles d-dimensionnels compacts :

$$R_{\mathbb{Q}} = [a_{1,\mathbb{Q}}, b_{1,\mathbb{Q}}] \times \cdots \times [a_{d,\mathbb{Q}}, b_{d,\mathbb{Q}}] \qquad {\scriptstyle (-\infty < a_{i,\mathbb{Q}} < b_{i,\mathbb{Q}} < \infty)}$$

à coordonnées *rationnelles :*

$$a_{i,\mathbb{Q}},\ b_{i,\mathbb{Q}}\ \in\ \mathbb{Q} \qquad\qquad (1\leqslant i\leqslant d),$$

et à coefficients complexes :

$$\lambda_{\mathbb{Q}}\ \in\ \mathbb{Q}+i\,\mathbb{Q},$$

dont les parties réelles et imaginaires sont elles aussi rationnelles. Comme $\mathbb{Q}^d$ et $\mathbb{Q}^2$ sont dénombrables, et comme une réunion dénombrable d'ensembles dénombrables est toujours dénombrable, cet ensemble de fonctions que l'on pourrait appeler *fonctions en escalier rationnelles* a la propriété mirifique d'être *dénombrable*. Après rénumérotation, nous pouvons donc le considérer comme étant une suite :

$$\bigg\{\sum_{\mathsf{finie}}\lambda_{\mathbb{Q}}\cdot\mathbf{1}_{\mathbb{R}_{\mathbb{Q}}}\bigg\}\ =\ \big(f_n\big)_{n=1}^{\infty},$$

et nous allons démontrer que cette suite est *dense* dans $L^1(\mathbb{R}^d)$.

Soit en effet une fonction quelconque $f\in L^2(\mathbb{R}^d)$ et soit $\varepsilon>0$ arbitrairement petit. Pour tout entier $K\geqslant 1$, la fonction doublement tronquée définie par :

$$T_K f := \begin{cases} f(x) & \text{lorsque } \max\limits_{1\leqslant i\leqslant d}|x_i|\leqslant K \text{ et lorsque } \big|f(x)\big|\leqslant K,\\ 0 & \text{sinon}\end{cases}$$

satisfait au moins ponctuellement :

$$\big|f-T_K f\big|^2\ \leqslant\ 4\,|f|^2,$$

et comme les valeurs $f(x)$ sont nécessairement finies presque partout puisque $\int|f|^2<\infty$, on a (exercice mental) :

$$f(x)=\lim_{K\to\infty}T_K f(x) \qquad\qquad (\text{pour presque tout } x\in\mathbb{R}^d).$$

Le théorème de la convergence dominée assure alors que :

$$0=\lim_{K\to\infty}\big|\big|f-T_K f\big|\big|_{L^2(\mathbb{R}^d)}.$$

Étant donné $\varepsilon>0$ arbitrairement petit, il existe donc $K=K(\varepsilon)\gg 1$ assez grand pour que :

$$\big|\big|f-T_K f\big|\big|_{L^2(\mathbb{R}^d)}\ \leqslant\ \varepsilon/2.$$

Or la fonction doublement tronquée $T_K f$ est mesurable, elle est bornée et elle a comme support un ensemble borné. Par conséquent, elle est intégrable :

$$T_K f\ \in\ L^1(\mathbb{R}^d).$$

On pense alors spontanément à un théorème censé être déjà connu au niveau de ce cours.

Théorème 4.6. [Densité de fonctions-types concrètes] *Les trois familles de fonctions suivantes sont chacune denses dans l'espace $L^1(\mathbb{R}^d)$ pour la norme L^1 :*

• *les fonctions dites étagées, qui sont sommes finies d'ensembles caractéristiques de sous-ensembles mesurables quelconques $E\subset\mathbb{R}^d$:*

$$\sum_{\mathsf{finie}}\lambda\cdot\mathbf{1}_E \qquad\qquad (\lambda\in\mathbb{C});$$

• *les fonctions dites* en escalier, *qui sont combinaisons linéaires finies de fonctions caractéristiques :*

$$\sum_{\text{finie}} \lambda \cdot \mathbf{1}_R \qquad\qquad (\lambda \in \mathbb{C}),$$

de rectangles d-dimensionnels bornés :

$$R = [a_1, b_1] \times \cdots \times [a_d, b_d] \qquad\qquad (a_i, b_i \in \mathbb{R});$$

• *les fonctions continues à support compact :*

$$\Big\{ f \in \mathscr{C}^0(\mathbb{R}^d, \mathbb{C}) \colon\ f \equiv 0 \ \text{ sur } \ \mathbb{R}^d \backslash \big\{ \max_{1 \leqslant i \leqslant d} |x_i| \geqslant K \big\} \ \text{ pour } \ K \gg 1 \Big\}. \qquad \square$$

En appliquant la deuxième assertion de densité à la fonction bornée $T_K f$ de support borné donc intégrable :

$$T_K f \,\in\, L^1(\mathbb{R}^d),$$

on trouve donc une fonction étagée :

$$\sum_{\text{finie}} \lambda \cdot \mathbf{1}_R,$$

telle que :

$$\Big\| T_K f - \sum_{\text{finie}} \lambda \cdot \mathbf{1}_R \Big\|_{L^1(\mathbb{R}^d)} \,\leqslant\, \frac{\varepsilon^2}{24\,K},$$

en norme L^1, mais pas pour l'instant en norme L^2, et c'est pourquoi on ajuste la contrainte d'approximation à être à l'avance aussi petite que $\frac{\varepsilon^2}{24\,K}$ afin que tout se passe comme il le faudra à la fin, notamment afin d'obtenir un majorant de la forme $\varepsilon/2 + \varepsilon/2 = \varepsilon$.

Or un examen de la construction de cette approximante dans la démonstration du théorème de densité montre qu'on peut aisément supposer que sa taille et son support ne débordent pas trop de celui de $T_K f$:

• $\big| \sum \lambda \cdot \mathbf{1}_R \big| \leqslant 2\,K$,

• $R \subset \{ |x| \leqslant 2\,K \}$,

en abrégeant $|x| := \max\big\{ |x_1|, \ldots, |x_d| \big\}$.

Ensuite, on perturbe très légèrement ces coefficients $\lambda \in \mathbb{C}$ en des coefficients rationnels :

$$\lambda_{\mathbb{Q}} \,\in\, \mathbb{Q} + i\,\mathbb{Q},$$

et on perturbe aussi les extrémités des segments qui définissent les rectangles d-dimensionnels R pour obtenir de nouveaux rectangles $R_{\mathbb{Q}}$ à segments-extrémités dans $\mathbb{Q}$ de telle sorte que :

$$\Big\| \sum_{\text{finie}} \lambda \cdot \mathbf{1}_R - \sum_{\text{finie}} \lambda_{\mathbb{Q}} \cdot \mathbf{1}_{R_{\mathbb{Q}}} \Big\|_{L^1(\mathbb{R}^d)} \,\leqslant\, \frac{\varepsilon^2}{24\,K},$$

d'où par inégalité triangulaire :

$$\Big\| T_K f - \sum_{\text{finie}} \lambda_{\mathbb{Q}} \cdot \mathbf{1}_{R_{\mathbb{Q}}} \Big\|_{L^1(\mathbb{R}^d)} \,\leqslant\, \frac{\varepsilon^2}{24\,K} + \frac{\varepsilon^2}{24\,K} = \frac{\varepsilon^2}{12\,K},$$

avec de plus :

• $\big| \sum \lambda_{\mathbb{Q}} \cdot \mathbf{1}_{R_{\mathbb{Q}}} \big| \leqslant 2\,K$,

• $R_{\mathbb{Q}} \subset \{|x| \leqslant 2\,K\}$.

En particulier, on a la majoration aisée utile :

$$\max_{|x|\leqslant 2K} \bigg|T_K f - \sum_{\text{finie}} \lambda_{\mathbb{Q}} \cdot \mathbf{1}_{R_{\mathbb{Q}}}\bigg| \leqslant 3\,K,$$

Pour vérifier que la fonction ainsi obtenue $\sum \lambda_{\mathbb{Q}} \cdot \mathbf{1}_{R_{\mathbb{Q}}}$ approxime alors f en norme L^2 à ε près, il suffit de vérifier qu'elle approxime $T_K f$ à $\varepsilon/2$ près, ce qui repose sur des majorations-engrenages maintenant parfaitement huilées :

$$\begin{aligned}
\bigg(\bigg\|T_K f - \sum_{\text{finie}} \lambda_{\mathbb{Q}} \cdot \mathbf{1}_{R_{\mathbb{Q}}}\bigg\|_{L^2(\mathbb{R}^d)}\bigg)^2 &= \int_{|x|\leqslant 2K} \bigg|T_K f - \sum_{\text{finie}} \lambda_{\mathbb{Q}} \cdot \mathbf{1}_{R_{\mathbb{Q}}}\bigg|^{1+1} \\
&\leqslant \max_{|x|\leqslant 2K} \bigg|T_K f - \sum_{\text{finie}} \lambda_{\mathbb{Q}} \cdot \mathbf{1}_{R_{\mathbb{Q}}}\bigg|^1 \int_{|x|\leqslant K} \bigg|T_K f - \sum_{\text{finie}} \lambda_{\mathbb{Q}} \cdot \mathbf{1}_{R_{\mathbb{Q}}}\bigg|^1 \\
&\leqslant 3\,K \bigg\|T_K f - \sum_{\text{finie}} \lambda_{\mathbb{Q}} \cdot \mathbf{1}_{R_{\mathbb{Q}}}\bigg\|_{L^1(\mathbb{R}^d)} \\
&\leqslant 3\,K\,\frac{\varepsilon^2}{12\,K} \\
&= \Big(\frac{\varepsilon}{2}\Big)^2,
\end{aligned}$$

d'où au final par simple inégalité triangulaire :

$$\begin{aligned}
\bigg\|f - \sum_{\text{finie}} \lambda_{\mathbb{Q}} \cdot \mathbf{1}_{R_{\mathbb{Q}}}\bigg\|_{L^2(\mathbb{R}^d)} &\leqslant \big\|f - T_K f\big\|_{L^2(\mathbb{R}^d)} + \bigg\|T_K f - \sum_{\text{finie}} \lambda_{\mathbb{Q}} \cdot \mathbf{1}_{R_{\mathbb{Q}}}\bigg\|_{L^2(\mathbb{R}^d)} \\
&\leqslant \varepsilon/2 + \varepsilon/2,
\end{aligned}$$

ce qui conclut. □

5. Définition des espaces de Hilbert abstraits et des bases hilbertiennes

Nous pouvons maintenant définir les espaces de Hilbert abstraits et généraux, sans égard à leur dimension, *i.e.* sans mentionner qu'il soient ou bien de dimension finie, ou bien de dimension infinie (dénombrable ou non-dénombrable).

Résumons notre parcours. Nous avons établi que ℓ^2 est un espace vectoriel hermitien complet en tout point analogue à $\mathbb{C}^n$, afin de présenter et d'étudier un exemple concret et tangible d'espace de Hilbert *avant* toute définition abstraite. Cette stratégie de type inductif qui consiste à faire passer le concret avant l'abstrait n'est pas dénuée de sens, puisque nous verrons plus loin que *tout* espace de Hilbert abstrait de dimension infinie dénombrable est *isométriquement isomorphe* à ℓ^2 (Théorème dit de Riesz-Fischer qui sera vu ultérieurement). Nous utiliserons à présent les lettres f, g, h pour désigner des éléments d'un espace de Hilbert, en ayant à l'idée que les applications principales de la théorie ont cours dans certains espaces de *fonctions*.

Définition 5.1. [Espace de Hilbert] Un espace vectoriel normé $(H, \|\cdot\|)$ sur $\mathbb{C}$ est un *espace de Hilbert* s'il est *complet* et si sa norme :

$$\|f\| = \sqrt{\langle f, f\rangle} \qquad (f \in H)$$

provient d'un produit scalaire hermitien $\langle \cdot, \cdot \rangle$ satisfaisant les cinq axiomes suivants :

(i) $\langle g, f \rangle = \overline{\langle f, g \rangle}$;

(ii) $\langle f' + f'', g \rangle = \langle f', g \rangle + \langle f'', g \rangle$,
$\langle f, g' + g'' \rangle = \langle f, g' \rangle + \langle f, g'' \rangle$;

(iii) $\langle \lambda f, g \rangle = \lambda \langle f, g \rangle$,
$\langle f, \mu g \rangle = \overline{\mu} \langle f, g \rangle$;

(iv) $\langle f, f \rangle \in \mathbb{R}_+$;

(v) $\langle f, f \rangle = 0 \Rightarrow f = 0$.

En des termes équivalents, un *espace de Hilbert* est un $\mathbb{C}$-espace vectoriel hermitien $\big(H, \langle \cdot, \cdot \rangle\big)$ qui est *complet*, ou encore un espace préhilbertien *complet*.

Observons sans attendre que les conditions (iv) et (v) assurent que $\sqrt{\langle f, f \rangle}$ a un sens et ne s'annule que si $f = 0$, et donc :

$$\|f\| := \sqrt{\langle f, f \rangle}$$

est un bon candidat pour une norme, mais deux axiomes doivent encore être vérifiés. Par (iii), on a l'homogénéité :

$$\|\lambda f\| = (\lambda \overline{\lambda})^{1/2} \, \|f\| = |\lambda| \, \|f\|.$$

En développant $\|f + tg\|^2$ et en raisonnant exactement comme dans la preuve du Théorème 2.10, on établit l'inégalité de Cauchy-Schwarz :

$$\forall \, f, g \in H \qquad \big|\langle f, g \rangle\big| \leqslant \|f\| \, \|g\|,$$

l'égalité n'étant satisfaite que lorsque f et g sont colinéaires. De cette inégalité en majorant les deux termes $\langle f, g \rangle$ et $\langle g, f \rangle$ dans :

$$\|f + g\|^2 = \langle f + g, \, f + g \rangle = \|f\|^2 + \langle f, g \rangle + \langle g, f \rangle + \|g\|^2,$$

on déduit ensuite l'inégalité triangulaire :

$$\|f + g\| \leqslant \|f\| + \|g\|,$$

ce qui achève de montrer que $\sqrt{\langle f, f \rangle}$ fournit une norme sur H.

La notion générale d'*espace de Hilbert* abstrait (du verbe « abstraire ») ces deux propriétés et les pose comme définition. Les espaces de Hilbert permettent de résoudre de nombreux problèmes d'analyse : équations intégro-différentielles, séries de Fourier, sans compter que les axiomes de la Mécanique Quantique postulent qu'il existe toujours un espace de Hilbert correspondant à une expérience de microphysique donnée.

Si un espace préhilbertien n'est pas complet, on peut le compléter d'une manière essentiellement unique en un espace de Hilbert (complet), comme le propose de l'établir l'Exercice 22.

Théorème 5.2. *Étant donné un espace préhilbertien $\big(H_0, \langle \cdot, \cdot \rangle_{H_0}\big)$, il existe un espace de Hilbert $\big(H, \langle \cdot, \cdot \rangle\big)$ — complet ! — unique à isomorphisme unitaire près satisfaisant :*

(i) *$H \supset H_0$;*

(ii) *H_0 est dense dans H ;*

(iii) *$\langle f, g \rangle_H = \langle f, g \rangle_{H_0}$ pour tous $f, g \in H_0$.* □

5.3. Cinq concepts abstraits fondamentaux. Maintenant, introduisons les notions de *densité*, de *séparabilité*, de *totalité*, de *familles orthonormées* et de *bases hilbertiennes*.

Nous montons ici d'un cran en abstraction. Il est important de méditer la signification des définitions qui vont suivre. Le moment de leur première apparition exige de se forger des intuitions mathématiques rigoureuses. Le moment de leur compréhension exige aussi de revenir sur leur signification lorsque nous aurons démontré quelques théorèmes qui les utilisent.

Définition 5.4. [Densité] Une partie D de H est dite *dense dans* H si :

$$\forall\, h \in H \quad \forall\, \varepsilon > 0 \quad \exists\, f \in D \quad \text{tel que} \quad \|f - h\| < \varepsilon.$$

Autrement dit, les éléments de D sont arbitrairement proches de tous les élément de H. Ou encore : Pour la topologie sur H naturellement issue de la distance $d(f,g) := \|f - g\|$ associée à la norme, la *fermeture* $\overline{D} = H$ de D remplit H tout entier.

Cette notion de densité a aussi un sens dans les espaces vectoriels normés quelconques (qui ne sont pas forcément de Hilbert), et même dans les espaces métriques sur lesquels seule une distance existe.

Définition 5.5. [Séparabilité] Un espace de Hilbert est *séparable* s'il possède un sous-ensemble dénombrable dense. Autrement dit, H est séparable s'il existe une *suite* $(f_k)_{k=1}^{\infty}$ d'éléments de H satisfaisant :

$$\forall\, h \in H \quad \forall\, \varepsilon > 0 \quad \exists\, f_k \quad \text{tel que} \quad \|h - f_k\| < \varepsilon.$$

À nouveau, cette notion a aussi un sens dans n'importe quel espace métrique, par exemple dans tout espace vectoriel normé $(E, \|\cdot\|_E)$. La plupart des espaces de Hilbert que l'on rencontre dans la théorie des fonctions sont en fait séparables, par exemple pour $\mathsf{p} \geqslant 1$ les espaces $L^{\mathsf{p}}(\mathbb{R}^n)$ de fonctions telles que $\int |f|^{\mathsf{p}} < \infty$, puisque l'on sait que la famille des fonctions qui sont égales à une certaine constante rationnelle sur un hypercube à coordonnées rationnelles et nulles ailleurs sont denses dans $L^{\mathsf{p}}(\mathbb{R}^n)$; le Théorème 4.5 a au moins démontré cela en détail dans le cas $\mathsf{p} = 2$, et la démonstration pour $1 \leqslant \mathsf{p} < \infty$ quelconque est similaire.

Quelques-uns des énoncés abstraits que nous démontrerons ci-dessous ne feront pas d'hypothèse de séparabilité, mais on peut fort bien les comprendre en supposant mentalement qu'ils sont séparables pour s'en faire une idée utile et adéquate.

On veillera à ne pas confondre le terme « séparable » avec la notion d'espace topologique X « séparé », qui par définition a la propriété que pour toute paire de points $x \neq y$ distincts, il existe deux voisinages ouverts non vides $\mathsf{U} \ni x$ et $\mathsf{V} \ni y$ tels que $\mathsf{U} \cap \mathsf{V} = \emptyset$.

Toutefois, cette notion de séparabilité ainsi définie ne tient pas encore compte de la structure d'espace vectoriel de H, car on pourrait naturellement penser qu'avec une suite $(f_k)_{k=1}^{\infty}$, toutes les combinaisons linéaires des f_k sont aussi données. Voici donc une définition plus appropriée.

Définition 5.6. [Totalité] Une partie F d'un espace de Hilbert H est dite *totale* si l'ensemble des combinaisons linéaires finies d'éléments de F est dense dans H. Autrement dit, la famille $F = (f_i)_{i \in \mathscr{I}}$ indexée par un ensemble $\mathscr{I}$ quelconque est totale si :

$$\forall\, h \in H \quad \forall\, \varepsilon > 0 \quad \exists\, k \in \mathbb{N}^* \quad \exists\, f_{i_1}, \dots, f_{i_k} \in F \quad \exists\, a_1, \dots, a_k \in \mathbb{C}$$
$$\text{tels que} \quad \|h - a_1\, f_{i_1} - \cdots - a_k\, f_{i_k}\| < \varepsilon.$$

Enfin, la notion de base d'un espace vectoriel se généralise aux espaces de Hilbert, qui sont le plus souvent de dimension infinie. Mais comme on dispose d'un produit scalaire, on peut même parler d'orthogonalité entre des vecteurs.

En géométrie euclidienne (dimension finie), toute famille de vecteurs qui est à la fois orthogonale et de norme 1 est dite *orthonormée*, concept qui se généralise à la dimension infinie.

Définition 5.7. [Famille orthonormée] Une famille $\{e_i\}_{i\in\mathscr{I}}$ d'éléments d'un espace de Hilbert H est dite *orthonormée* si :

$$\forall\, i_1 \in \mathscr{I} \quad \forall\, i_2 \in \mathscr{I} \quad \langle e_{i_1}, e_{i_2}\rangle = \delta_{i_1,i_2} := \begin{cases} 1 & \text{si } \ i_1 = i_2; \\ 0 & \text{si } \ i_1 \neq i_2. \end{cases}$$

Mais parce qu'il n'est pas clair *a priori* que tout vecteur d'un espace de Hilbert H de dimension infinie puisse être décomposé comme somme *nécessairement infinie* de multiples de vecteurs d'une certaine base (puisqu'une telle somme pourrait fort bien ne pas converger), on définit un premier concept de base dans H en ne considérant que des combinaisons linéaires qui sont *finies*.

Définition 5.8. [Base hilbertienne] Une *base hilbertienne* de H, ou *base orthonormée*, est une famille $\{e_i\}_{i\in\mathscr{I}}$ de vecteurs de H qui est orthonormée, et qui est de plus totale.

En résumé, nous avons présenté cinq concepts fondamentaux au sujet des espaces de Hilbert qui culminent en une définition prudente de la notion de base par laquelle l'infini est rejeté dans la densité, à travers l'hypothèse que la famille $\{e_i\}_{i\in\mathscr{I}}$ est totale. Comme cela arrive très souvent en mathématiques, un champ définitionnel présenté au début d'une théorie se pose *en attente* de théorèmes plus informatifs qui vont métamorphoser l'appréhension des concepts en les rendant plus clairement concrets et manipulables. La suite des événements va donc rendre plus compréhensibles ces définitions.

5.9. Bases hilbertiennes et séparabilité. Maintenant, que peut-on dire lorsque H est séparable ? Naturellement, on peut s'attendre à ce que tous les concepts se ramènent à des collections dénombrables d'objets. Deux résultats fondamentaux confirment cette intuition.

Proposition 5.10. *Tout espace de Hilbert possédant une famille totale qui est une* suite $(f_k)_{k=1}^{\infty}$ *(dénombrable) est séparable.*

Démonstration. Voici le problème : dans la définition de famille totale, les combinaisons linéaires finies autorisées sont à coefficients complexes arbitraires, et l'on sait que $\mathbb{R}$ et $\mathbb{C}$ ne sont *pas* dénombrables, d'après le fameux argument diagonal de Cantor. Toutefois, on sait aussi que le corps des nombres *rationnels* $\mathbb{Q}$, qui est dénombrable, est dense dans $\mathbb{R}$, d'où il découle immédiatement que $\mathbb{Q} + i\,\mathbb{Q}$ est dense dans $\mathbb{C}$. Fixons maintenant un entier $k \geqslant 1$. L'ensemble $\mathscr{C}_k'(f)$ des combinaisons linéaires à coefficients rationnels des k premiers éléments de la suite :

$$a_1' f_1 + \cdots + a_k'\, f_k \qquad\qquad (a_i' \in \mathbb{Q} + i\,\mathbb{Q})$$

est égal à $(\mathbb{Q} + i\,\mathbb{Q}) \times \cdots \times (\mathbb{Q} + i\,\mathbb{Q})$ avec exactement k facteurs, et comme le produit d'un nombre fini d'ensembles dénombrables est dénombrable, $\mathscr{C}_k'(f)$ est dénombrable. Puisqu'une réunion dénombrable d'ensembles dénombrables est elle aussi dénombrable, l'ensemble :

$$\mathscr{C}'(f) := \mathscr{C}_1'(f) \cup \mathscr{C}_2'(f) \cup \cdots \cup \mathscr{C}_k'(f) \cup \cdots$$

est lui aussi dénombrable. Cet argument nous permet donc de remplacer les combinaisons linéaires finies à coefficients complexes par des combinaisons linéaires finies à coefficients rationnels. Mais il nous reste à vérifier scrupuleusement la densité.

Revenons donc à notre suite $(f_k)_{k=1}^{\infty}$, qui est totale par hypothèse. Ainsi :

$$\begin{aligned}\forall\, h \in H \quad \forall\, \varepsilon > 0 \quad \exists\, k \geqslant 1 \quad \exists\, a_1, \dots, a_k \in \mathbb{C}\\ \text{tels que} \quad \|h - a_1\, f_1 - \cdots - a_k\, f_k\| < \varepsilon.\end{aligned}$$

Pour tout $j = 1, \dots, k$, choisissons alors un $a_j' \in \mathbb{Q} + i\,\mathbb{Q}$ tel que :

$$|a_j' - a_j| \leqslant \tfrac{\varepsilon}{k \max\,(\|f_1\|, \dots, \|f_k\|)},$$

ce qui est possible (observons que nous avons ajusté à l'avance la très grande proximité de a_j' à a_j, *cf.* ce qui va suivre). Maintenant, majorons la différence entre h et la nouvelle combinaison linéaire à coefficients rationnels en faisant apparaître (artificiellement) la combinaison linéaire initiale que l'on contrôle par hypothèse :

$$\begin{aligned}\|h - a_1'\, f_1 - \cdots - a_k'\, f_k\| &= \|h - a_1 f_1 - \cdots - a_k f_k - (a_1' - a_1)\, f_1 - \cdots - (a_k' - a_k)\, f_k\|\\ &\leqslant \|h - a_1 f_1 - \cdots - a_k f_k\| + |a_1' - a_1|\, \|f_1\| + \cdots + |a_k' - a_k|\, \|f_k\|\\ &\leqslant \varepsilon + \tfrac{\varepsilon}{k \max\,(\|f_1\|, \dots, \|f_k\|)} \big(\|f_1\| + \cdots + \|f_k\|\big)\\ &\leqslant \varepsilon + \tfrac{\varepsilon}{\underline{k}_{\circ}\, \underline{\max\,(\|f_1\|, \dots, \|f_k\|)}_{\circ\circ}} \big(\underline{k}_{\circ}\, \underline{\max\,(\|f_1\|, \dots, \|f_k\|}_{\circ\circ}\big)\\ &\leqslant 2\, \varepsilon.\end{aligned}$$

On voit donc que tout élément $h \in H$ peut être approximé à 2ε près par un élément de $\mathscr{C}_k'(f)$. En conclusion, l'ensemble $\mathscr{C}'(f)$ des combinaisons linéaires finies à coefficients rationnels de la suite $(f_k)_{k=1}^{\infty}$, ensemble qui est dénombrable, est bien dense dans H, ce qui achève cette preuve très détaillée. □

Commentons cette démonstration. Un premier moment nous a convaincu pour des raisons de théorie des ensembles que les combinaisons linéaires finies à coefficients complexes pouvaient être remplacées par des combinaisons à coefficients rationnels, ces dernières étant dénombrables. La densité de $\mathbb{Q} + i\mathbb{Q}$ dans $\mathbb{C}$ est donc la cause principale de la vérité de cette proposition. Mais l'argument complet de preuve n'est pas terminé, et la seconde partie de la démonstration repose sur la manipulation d'inégalités. Toutefois, avec l'expérience, le mathématicien se forge l'intuition que normes et valeurs absolues sont faites pour que la densité passe au travers sans problème, et donc il se convainc que la deuxième partie de la démonstration va de soi.

Corollaire 5.11. *L'espace* $\ell^2_{\mathbb{C}} = \big\{z = (z_i)_{i=1}^{\infty} \colon z_i \in \mathbb{C},\ \sum |z_i|^2 < \infty\big\}$ *est séparable.*

Démonstration. Pour $i \geqslant 1$ entier, soit $e_i := (0, \dots, 0, 1, 0, \dots)$ le i-ème vecteur de base, avec 1 à la i-ème position, et 0 ailleurs, de telle sorte que l'on peut écrire :

$$z \,=\, \big(z_1, \dots, z_i, \dots\big) \,=\, \sum_{i=1}^{\infty} z_i\, e_i.$$

La convergence $\sum |z_i|^2 < \infty$ assure que, pour tout $\varepsilon > 0$, il existe $\mathrm{I} = \mathrm{I}(\varepsilon) \gg 1$ assez grand pour que :

$$\sum_{i=\mathrm{I}+1}^{\infty} |z_i|^2 \,\leqslant\, \varepsilon^2,$$

d'où l'approximation :

$$\bigg\| z - \sum_{i=1}^{\mathrm{I}} z_i\, e_i \bigg\|_{\ell_2} \leqslant \varepsilon,$$

qui montre que la famille dénombrable de vecteurs $(e_i)_{i=1}^{\infty}$ est totale dans $\ell^2_{\mathbb{C}}$. La Proposition 5.10 qui précède achève alors le travail. □

Voici maintenant le premier résultat fondamental de la théorie des espaces de Hilbert.

Théorème 5.12. *Tout espace de Hilbert H de dimension infinie qui est* séparable *possède une base hilbertienne qui est* dénombrable.

Extrayons le sens de cet énoncé : la séparabilité de H (dénombrabilité d'un sous-ensemble dense) est *cause* de la dénombrabilité d'au moins une base hilbertienne. Est-ce vrai pour *toute* base hilbertienne ? Oui, et nous le verrons juste après.

Démonstration. Soit donc $(g_n)_{n=1}^{\infty}$ une suite dense dans H. Pour commencer, en inspectant ces éléments les uns à la suite des autres :

$$g_1,\ g_2,\ g_3,\ g_4,\ g_5,\ \dots,$$

on décide de ne conserver un élément g_{n+1} que s'il n'appartient pas à l'espace vectoriel engendré par les n précédents éléments $g_1, g_2, \dots, g_n$. De la sorte, on extrait facilement une sous-suite $(g_{n_k})_{k=1}^{\infty}$ telle que les k vecteurs $g_{n_1}, g_{n_2}, \dots, g_{n_k}$ sont toujours linéairement indépendants entre eux, quel que soit $k \geqslant 1$, en se souvenant que par construction, on a aussi :

$$\mathsf{Vect}\big(g_{n_1},\ g_{n_2},\ \dots,\ g_{n_{k-1}},\ g_{n_k}\big) = \mathsf{Vect}\big(g_1, g_2, g_3, \dots, g_{n_k-1}, g_{n_k}\big).$$

Maintenant, en appliquant le procédé d'orthonormalisation de Gram-Schmidt, on peut construire une nouvelle suite $(\widetilde{g}_k)_{k=1}^{\infty}$ telle que :

- $\mathsf{Vect}\big(\widetilde{g}_1, \dots, \widetilde{g}_k\big) = \mathsf{Vect}\big(g_{n_1}, \dots, g_{n_k}\big)$,
- Les k vecteurs $\big(\widetilde{g}_1, \dots, \widetilde{g}_k\big)$ forment une base orthonormée,

et ce, pour tout $k \geqslant 1$. Sachant qu'une base hilbertienne est une famille orthonormée qui est de plus totale, l'assertion suivante conclura la preuve du théorème.

Assertion 5.13. *Alors la suite infinie dénombrable $\big(\widetilde{g}_k\big)_{k=1}^{\infty}$ est totale dans H.*

En effet, soit $h \in H$ quelconque et soit $\varepsilon > 0$ arbitraire. Comme par hypothèse la suite $(g_n)_{n=1}^{\infty}$ est dense dans H, il existe un entier $N \geqslant 1$ tel que $\|h - g_N\| < \varepsilon$. Mais par construction, il existe un certain autre entier inférieur ou égal $n_K \leqslant N$ tel que :

$$g_N \in \mathsf{Vect}\big(g_1, \dots, g_N\big) = \mathsf{Vect}\big(g_{n_1}, \dots, g_{n_K}\big) = \mathsf{Vect}\big(\widetilde{g}_1, \dots, \widetilde{g}_K\big).$$

Donc il existe des constantes $a_1, \dots, a_K$ telles que ce vecteur g_N s'écrive comme la combinaison linéaire finie :

$$g_N = a_1\, \widetilde{g}_1 + \cdots + a_K\, \widetilde{g}_K,$$

On a donc démontré que pour tout $h \in H$, quel que soit $\varepsilon > 0$, il existe une combinaison linéaire finie $a_1\, \widetilde{g}_1 + \cdots + a_K \widetilde{g}_K$ de la suite $\big(\widetilde{g}_k\big)_{k=1}^{\infty}$ telle que :

$$\|h - a_1\, \widetilde{g}_1 - \cdots - a_K\, \widetilde{g}_K\| < \varepsilon.$$

Ceci exprime précisément que la suite orthonormée $\big(\widetilde{g}_k\big)_{k=1}^{\infty}$ est totale et la preuve du théorème s'achève. □

Mentalisons à nouveau la démonstration en essayant de la reparcourir pour embrasser ses causalités et surtout, la résumer adéquatement. Comme on part d'une suite dénombrable dense et que l'on veut construire une base, on peut supprimer les vecteurs inutiles. Bien entendu, il faut orthonormaliser. Mais cette opération change complètement les vecteurs. Heureusement, tout reste dans une suite grandissante d'espaces vectoriels, et l'on ne perd pas les éléments de H atteints par densité pourvu que l'on s'autorise des combinaisons linéaires quelconques (notion de totalité).

Proposition 5.14. *Si H est un espace de Hilbert séparable, toute famille orthonormée de H est de cardinal au plus dénombrable.*

Démonstration. Soit donc $\{e_i\}_{i \in \mathscr{I}}$ une famille orthonormée dans H. Pour prouver que $\mathscr{I}$ est au plus dénombrable, il suffit de démontrer l'existence d'une *injection* de $\mathscr{I}$ *dans* $\mathbb{N}^*$.

Remarquons tout d'abord que si i_1 et i_2 sont deux éléments distincts de $\mathscr{I}$, on a :

$$\begin{aligned} \|e_{i_1} - e_{i_2}\|^2 &= \langle e_{i_1} - e_{i_2},\, e_{i_1} - e_{i_2} \rangle \\ &= \langle e_{i_1},\, e_{i_1} \rangle - \underline{\langle e_{i_2},\, e_{i_1} \rangle}_{\circ} - \underline{\langle e_{i_1},\, e_{i_2} \rangle}_{\circ} + \langle e_{i_2},\, e_{i_2} \rangle \\ &= 2, \end{aligned}$$

c'est-à-dire $\|e_{i_1} - e_{i_2}\| = \sqrt{2}$.

Par séparabilité de H, il existe une suite dénombrable $(g_n)_{n=1}^{\infty}$ dense dans H. Pour chaque indice $i \in \mathscr{I}$, il existe donc un élément $g_{n(i)}$ de cette suite qui approche assez bien e_i :

$$\|e_i - g_{n(i)}\| \leqslant \tfrac{\sqrt{2}}{3},$$

et ici, $\frac{\sqrt{2}}{3}$ qui semble une marge d'erreur assez grossière sera en fait une quantité assez petite.

En choisissant pour chaque $i \in \mathscr{I}$ un seul entier $n(i)$ parmi tous ceux qui vérifient cette inégalité, on définit alors une application $\mathscr{I} \longrightarrow \mathbb{N}^*$, $i \longmapsto n(i)$.

En insérant $-g_{n(i_1)} + g_{n(i_1)} - g_{n(i_2)} + g_{n(i_2)}$ et en utilisant l'inégalité triangulaire, on a alors pour tous $i_1 \neq i_2$:

$$\begin{aligned} \sqrt{2} = \|e_{i_1} - e_{i_2}\| &\leqslant \|e_{i_1} - g_{n(i_1)}\| + \|g_{n(i_1)} - g_{n(i_2)}\| + \|g_{n(i_2)} - e_{i_2}\| \\ &\leqslant \|g_{n(i_1)} - g_{n(i_2)}\| + 2\,\tfrac{\sqrt{2}}{3}, \end{aligned}$$

d'où :

$$\tfrac{\sqrt{2}}{3} \leqslant \|g_{n(i_1)} - g_{n(i_2)}\|.$$

On en déduit immédiatement que $g_{n(i_1)} \neq g_{n(i_2)}$ et donc que $n(i_1) \neq n(i_2)$ pour tous $i_1 \neq i_2$ distincts. Ainsi, l'application $i \longmapsto n(i)$ de $\mathscr{I}$ à valeurs dans $\mathbb{N}^*$ est injective, ce qui implique $\mathsf{Card}(\mathscr{I}) \leqslant \mathsf{Card}(\mathbb{N}^*)$ et termine la preuve. □

Corollaire 5.15. *Si H est séparable, toutes ses bases hilbertiennes sont de cardinal fini ou infini dénombrable.* □

Proposition 5.16. *Soit H un espace de Hilbert ayant une base hilbertienne $\{e_i\}_{i \in \mathscr{I}}$. Si $f \in H$ est orthogonal à tous les e_i, $i \in \mathscr{I}$, alors $f = 0$.*

Cet énoncé, valable sans aucune hypothèse sur le cardinal de $\mathscr{I}$, implique qu'à une base hilbertienne, on ne peut jamais ajouter un vecteur de norme 1 qui est orthogonal à tous les vecteurs de ladite base. En termes plus abstraits : une base hilbertienne est une famille orthonormée qui est toujours maximale pour l'inclusion.

Dans la démonstration de cette proposition et aussi dans les énoncés qui suivront, nous seront amenés à calculer la norme au carré de la somme de deux ou plusieurs éléments, comme nous l'avons déjà fait quelques fois auparavant. Observons donc à l'avance que l'on a la formule :

$$\begin{aligned}\|f_1 + f_2 + \cdots + f_k\|^2 &= \langle f_1 + f_2 + \cdots + f_k,\ f_1 + f_2 + \cdots + f_k\rangle \\ &= \|f_1\|^2 + \|f_2\|^2 + \cdots + \|f_k\|^2 + \\ &\quad + \sum_{1 \leqslant i_1 < i_2 \leqslant k} \langle f_{i_1},\, f_{i_2}\rangle + \sum_{1 \leqslant i_1 < i_2 \leqslant k} \langle f_{i_2},\, f_{i_1}\rangle.\end{aligned}$$

Comme $\langle f_{i_2}, f_{i_1}\rangle = \overline{\langle f_{i_1}, f_{i_2}\rangle}$, le résultat peut être contracté en :

$$\|f_1 + \cdots + f_k\|^2 = \|f_1\|^2 + \cdots + \|f_k\|^2 + 2\,\mathrm{Re} \sum_{1 \leqslant i_1 < i_2 \leqslant k} \langle f_{i_1},\, f_{i_2}\rangle.$$

Définition 5.17. Deux vecteurs z et w d'un espace hermitien $\big(E, \langle\cdot,\cdot\rangle\big)$ sont dits *orthogonaux* lorsque leur produit scalaire est nul :

$$0 = \langle z, w\rangle.$$

Lorsque les f_i sont orthogonaux entre eux, cette formule se réduit à une simple expression du théorème de Pythagore :

$$\|f_1 + \cdots + f_k\|^2 = \|f_1\|^2 + \cdots + \|f_k\|^2.$$

De plus, si chaque f_i est dilaté d'un facteur $\lambda_i \in \mathbb{C}$, on a aussi :

$$\|\lambda_1\, f_1 + \cdots + \lambda_k\, f_k\|^2 = |\lambda_1|^2\, \|f_1\|^2 + \cdots + |\lambda_k|^2\, \|f_k\|^2.$$

Dans le cas de deux éléments $(k = 2)$, la formule se réduit bien entendu à :

$$\|f + g\|^2 = \|f\|^2 + 2\,\mathrm{Re}\,\langle f,\, g\rangle + \|g\|^2,$$

et lorsque f est orthogonal à g, on a le théorème de Pythagore :

$$\|f + g\|^2 = \|f\|^2 + \|g\|^2.$$

C'est quand on élève les normes au carré que le formalisme se prête le mieux aux calculs : souvenons-nous en !

Démonstration. Comme la famille $\{e_i\}_{i \in \mathscr{I}}$ est totale dans H, pour tout $\varepsilon > 0$, il existe une combinaison linéaire finie $\sum_{j \in \mathscr{J}} a_j\, e_j$ pour j variant dans un sous-ensemble *fini* $\mathscr{J} \subset \mathscr{I}$, telle que :

$$\Big\| f - \sum_{j \in \mathscr{J}} a_j\, e_j \Big\| < \varepsilon.$$

Élevons les deux membres de cette inégalité au carré. L'orthogonalité de f à tous les e_j et l'orthonormalité de la sous-famille $\{e_j\}_{j \in \mathscr{J}}$ nous permettent alors d'éliminer les termes croisés :

$$\begin{aligned}
\Big\| f - \sum_{j\in\mathscr{J}} a_j\, e_j \Big\|^2 &= \|f\|^2 - 2\mathsf{Re}\,\Big\langle f,\, \sum_{j\in\mathscr{J}} a_j\, e_j \Big\rangle + \sum_{j\in\mathscr{J}} \|a_j\, e_j\|^2 \\
&= \|f\|^2 - \sum_{j\in\mathscr{J}} 2\mathsf{Re}\,\big[\overline{a}_j \underline{\langle f,\, e_j\rangle}\big] + \sum_{j\in\mathscr{J}} |a_j|^2\, \|e_j\|^2 \\
&= \|f\|^2 + \sum_{j\in\mathscr{J}} |a_j|^2.
\end{aligned}$$

Mais comme par hypothèse, cette quantité est $< \varepsilon^2$, et que les termes obtenus $\|f\|^2$ et $\sum_{j\in\mathscr{J}} |a_j|^2$ sont tous deux positifs, on en déduit que $\|f\|^2 < \varepsilon^2$ (et aussi que $\sum_{j\in\mathscr{J}} |a_j|^2 < \varepsilon^2$, mais cela servira pas). L'inégalité $\|f\| < \varepsilon$ étant valable pour tout $\varepsilon > 0$, on en déduit $\|f\| = 0$, puis $f = 0$, comme annoncé. □

6. Meilleure approximation et projection orthogonale

Comme nous l'avions dit lorsque nous anticipions avec ℓ^2 l'étude des espaces de Hilbert H abstraits et généraux, l'intérêt principal du produit scalaire hermitien, c'est la possibilité de *faire de la géométrie* quasiment comme en dimension finie. En particulier, on peut conceptualiser la notion de projection orthogonale sur un sous-espace vectoriel de dimension finie de H, en s'imaginant intuitivement que tout se passe comme, par exemple, dans le plan ou dans l'espace à trois dimensions (réelles ou complexes). Mais l'intérêt principal de la vision géométrique appliquée aux espaces de Hilbert, lesquels sont la plupart du temps des espaces de *fonctions*, c'est que la géométrie va apporter des informations nouvelles et très intéressantes sur les fonctions.

Dans l'étude de la géométrie euclidienne en dimensions deux et trois, on utilise régulièrement les projections orthogonales sur une droite ou sur un plan. Il en va de même pour la géométrie dans un espace vectoriel complexe de dimension finie muni d'un produit scalaire hermitien. Qu'en est-il en dimension infinie ? Comme on le sait, les calculs et la géométrie sont simples lorsqu'on rapporte toutes les quantités à une certaine base orthonormée fixée à l'avance.

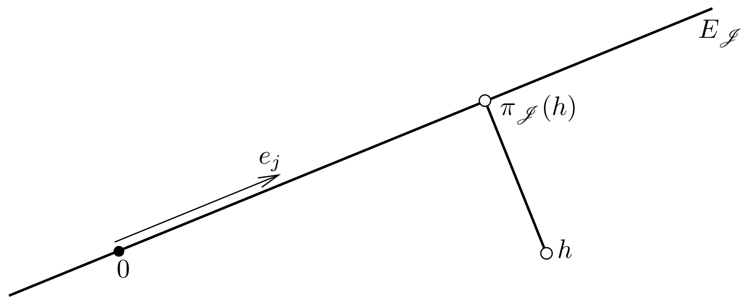
Pour raisonner avec un peu plus de généralité, considérons une famille orthonormée quelconque $\{e_i\}_{i\in\mathscr{I}}$ dans un espace de Hilbert, sans aucune hypothèse sur le cardinal de $\mathscr{I}$. Bien entendu, on pourra s'imaginer mentalement pour fixer les idées que H est séparable (puisque c'est dans ce cas-là que la théorie possède des applications riches et nombreuses aux espaces de fonctions), que $\mathscr{I}$ est dénombrable et même que $\{e_i\}_{i\in\mathscr{I}}$ est une base hilbertienne.

Pour déchiffrer le sens d'un théorème mathématique, et à un niveau supérieur, pour appréhender plus globalement une théorie mathématique donnée, il est recommandé (gymnastique intellectuelle nécessaire) de lire régulièrement les énoncés qui sont un peu trop généraux en les reformulant mentalement avec des hypothèses plus concrètes.

En tout cas, soit $\{e_i\}_{i\in\mathscr{I}}$ une famille orthonormée de H. Considérons alors un sous-espace vectoriel E de dimension finie de H. On peut, sans réelle perte de généralité, supposer que E est engendré par un certain nombre fini des e_i, car cela revient à s'imaginer que la famille $\{e_i\}_{i\in\mathscr{I}}$ a été adaptée à l'avance à E. Soit donc $\mathscr{J} \subset \mathscr{I}$ un sous-ensemble fini et :

$$E = E_{\mathscr{J}} := \mathsf{Vect}_{\mathbb{C}}\big(e_j \colon j \in \mathscr{J}\big).$$

Soit h un vecteur quelconque de H. Comment projette-t-on h orthogonalement sur E ? Y a-t-il une formule ? Certainement oui en dimension finie, mais rappelons les idées géométriques et vérifions que tout se généralise bien à la dimension infinie.



Le projeté orthogonal $\pi_{\mathscr{J}}(h)$ de h sur $E_{\mathscr{J}}$ doit vérifier deux conditions suivantes :

• $\pi_{\mathscr{J}}(h) \in \mathsf{Vect}\big(e_j \colon j \in \mathscr{J}\big)$, c'est-à-dire qu'il existe des constantes $a_j \in \mathbb{C}$, $j \in \mathscr{J}$, telles que :

$$\pi_{\mathscr{J}}(h) = \sum_{j \in \mathscr{J}} a_j\, e_j;$$

• c'est un projeté *orthogonal*.

Mais que veut dire orthogonal, ici ? Élaborons une figure, H étant représenté par un plan, E étant représenté par une droite (comme H et E sont tous deux des espaces vectoriels, on doit spécifier l'origine sur la figure). La seule manière de s'imaginer que $\pi_{\mathscr{J}}(h)$ est un projeté orthogonal, c'est de dire, comme le fait voir la figure, que :

$$\text{le vecteur } h - \pi_{\mathscr{J}}(h) \text{ est orthogonal à tous les vecteurs } e_k,\ k \in \mathscr{J}.$$

Et cette simple condition géométrique impose une équation sur les coefficients inconnus a_j :

$$\begin{aligned} 0 &= \Big\langle h - \sum_{j \in \mathscr{J}} a_j\, e_j,\ e_k \Big\rangle \\ &= \langle h,\, e_k \rangle - \sum_{j \in \mathscr{J}} a_j\, \underbrace{\langle e_j,\, e_k \rangle}_{\delta_{j,k}} \end{aligned}$$

qui permet de les déterminer tous uniquement :

$$a_k = \langle h, e_k \rangle \qquad\qquad (k \in \mathscr{J}).$$

Proposition 6.1. [Projection sur un sous-espace vectoriel de dimension finie] *Soit* $\{e_i\}_{i \in \mathscr{I}}$ *une famille orthonormée d'un espace de Hilbert* H *et soit* $\mathscr{J} \subset \mathscr{I}$ *un sous-ensemble* fini *quelconque. Alors le projeté orthogonal* $\pi_{\mathscr{J}}(h)$ *de* h *sur le sous-espace vectoriel de dimension finie :*

$$E_{\mathscr{J}} := \mathsf{Vect}_{\mathbb{C}}\big(e_j \colon j \in \mathscr{J}\big)$$

est donné par :

$$\pi_{\mathscr{J}}(h) = \sum_{j \in \mathscr{J}} \langle h, e_j \rangle\, e_j.$$

Cette formule est exactement la même que celle connue en géométrie euclidienne : les coefficients de la combinaison linéaire sont de simples produits scalaires. Il est important d'observer que c'est l'orthonormalité de la famille qui a rendu aussi facile la détermination des coefficients a_l, car sinon, si les produit scalaires $\langle e_j, e_k\rangle$ avaient été quelconques, il aurait fallu inverser la matrice $\big(\langle e_j, e_k\rangle\big)_{j\in\mathscr{J}}^{k\in\mathscr{J}}$ pour déterminer tous les a_l.

N'oublions pas de noter au passage que le théorème de Pythagore s'applique :

$$\|h\|^2 = \|\pi_{\mathscr{J}}(h)\|^2 + \|h - \pi_{\mathscr{J}}(h)\|^2,$$

puisque $h - \pi_{\mathscr{J}}(h)$ est orthogonal (d'après la définition) à $\pi_{\mathscr{J}}(h)$.

Ainsi par anticipation, on pourrait dire que le vecteur $h - \pi_{\mathscr{J}}(h)$ est la projection orthogonale de h, parallèlement à $E_{\mathscr{J}}$, dans l'orthogonal :

$$E_{\mathscr{J}}^{\perp} := \big\{f \in H \colon \langle f, g\rangle = 0 \quad \forall\, g \in E_{\mathscr{J}}\big\},$$

mais il y a une petite subtilité de la théorie : lorsqu'un sous-espace vectoriel $E \subset H$ est de dimension *infinie* (ce qui n'est pas le cas ici !), on ne peut pas en général parler de supplémentaire orthogonal $E^{\perp}$, car il faut supposer de plus que E est *fermé* pour que la formule :

$$H = E \oplus E^{\perp}$$

soit vraie, nous en dirons plus ultérieurement. Pour l'instant, nos sous-espaces $E_{\mathscr{J}}$ sont de dimension *finie*, car nous avançons (très) prudemment vers la dimension infinie.

Maintenant que nous avons bien conceptualisé la notion géométrique de projeté orthogonal, nous pouvons le caractériser en termes analytiques de la manière suivante.

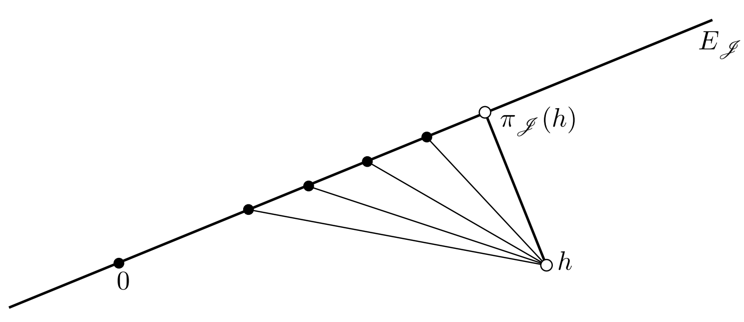
Théorème 6.2. [Meilleure approximation] *Soit $\{e_i\}_{i\in\mathscr{I}}$ une famille orthonormée quelconque dans espace de Hilbert H. Pour toute partie finie non vide $\mathscr{J} \neq \emptyset$ de $\mathscr{I}$, si l'on introduit une notation abrégée pour le produit scalaire de f avec chaque e_i :*

$$c_i(f) := \langle f,\, e_i\rangle \qquad\qquad (i \in \mathscr{I}),$$

alors pour tout vecteur $f \in H$ et pour toutes constantes quelconques $a_j \in \mathbb{C}$, $j \in \mathscr{J}$, on a l'inégalité :

$$\Big\| f - \sum_{j\in\mathscr{J}} c_j(f)\, e_j \Big\| \leqslant \Big\| f - \sum_{j\in\mathscr{J}} a_j\, e_j \Big\|.$$

De plus, cette inégalité est une égalité lorsque et seulement lorsque $a_j = c_j(f)$ pour tout $j \in \mathscr{J}$.



Géométriquement parlant, l'inégalité exprime que la distance de f à sa projection orthogonale est toujours inférieure à la distance de f à un élément *quelconque* de l'espace vectoriel $E_{\mathscr{J}}$, à savoir :

$$\|f - \pi_{\mathscr{J}}(f)\| \leqslant \|f - g\|$$

pour tout $g \in E_{\mathscr{J}}$ de la forme $g = \sum_{j \in \mathscr{J}} a_j \, e_j$. Par ailleurs, la notation $c_i(f)$ pour dire « i-ème coefficient de f » est quelque peu inspirée de la théorie des séries de Fourier, nous en reparlerons.

Avant la démonstration, voici une interprétation en termes de théorie des fonctions de ce théorème. Si l'on définit la distance de f à $E_{\mathscr{J}}$ par :

$$d(f, E_{\mathscr{J}}) := \inf_{g \in E_{\mathscr{J}}} \|f - g\|,$$

le théorème précédent nous dit alors que cet infimum est un *minimum* et que ce minimum est atteint en l'unique point $g := \pi_{\mathscr{J}}(f)$. Autrement dit, la projection orthogonale de f sur $E_{\mathscr{J}}$ est *la* meilleure approximation de f par un élément de l'espace vectoriel $E_{\mathscr{J}}$. Ainsi, lorsque H est un espace de fonctions « compliquées » et lorsque $\mathscr{J}$ indexe un ensemble fini de fonctions-types « simples », le théorème dit que $\pi_{\mathscr{J}}(f)$ est la meilleure approximation de la fonction f comme combinaison linéaire de ces fonctions-types simples.

Démonstration. Comme nous l'avons déjà signalé, il est avantageux d'élever au carré toute inégalité entre normes, car les normes au carré se prêtent mieux aux calculs grâce à leur expression $\|f\|^2 = \langle f, f \rangle$ en fonction du produit scalaire hermitien. Nous avons d'ailleurs déjà établi à ce sujet la formule élémentaire :

$$\|z + w\|^2 = \|z\|^2 + \|w\|^2 + 2\mathrm{Re}\,\langle z, w \rangle.$$

Ici, dans le terme à étudier $\|f - \sum_j a_j e_j\|^2$, faisons apparaître $-\pi_{\mathscr{J}}(f) + \pi_{\mathscr{J}}(f)$ comme terme central, appliquons la formule élémentaire et changeons en $\sum_{k \in \mathscr{J}}$ l'indice de sommation pour la deuxième somme :

$$\begin{aligned}
\Big\| f - \sum_{j \in \mathscr{J}} a_j \, e_j \Big\|^2 &= \Big\| \underbrace{f - \sum_{j \in \mathscr{J}} c_j(f) \, e_j}_{z = f - \pi_{\mathscr{J}}(f)} + \underbrace{\sum_{j \in \mathscr{J}} \big(c_j(f) - a_j\big) \, e_j}_{w} \Big\|^2 \\
&= \Big\| f - \pi_{\mathscr{J}}(f) \Big\|^2 + \Big\| \sum_{k \in \mathscr{J}} \big(c_k(f) - a_k\big) \, e_k \Big\|^2 + \\
&\quad + 2\mathrm{Re} \, \underline{\Big\langle f - \pi_{\mathscr{J}}(f), \ \textstyle\sum_{k \in \mathscr{J}} \big(c_k(f) - a_k\big) \, e_k \Big\rangle}.
\end{aligned}$$

Or nous avons déjà vu que le terme souligné à la dernière ligne s'annule, puisque $f - \pi_{\mathscr{J}}(f)$ est, par définition, orthogonal à tous les e_k, $k \in \mathscr{J}$. D'autre part, grâce à l'orthonormalité des e_k (on a $\|e_k\| = 1$), on en déduit que :

$$\Big\| f - \sum_{j \in \mathscr{J}} a_j \, e_j \Big\|^2 = \Big\| f - \pi_{\mathscr{J}}(f) \Big\|^2 + \sum_{k \in \mathscr{J}} \big| c_k(f) - a_k \big|^2.$$

Le dernier terme à droite étant positif, on a bien l'inégalité annoncée :

$$\Big\| f - \sum_{j \in \mathscr{J}} a_j \, e_j \Big\|^2 \geqslant \Big\| f - \pi_{\mathscr{J}}(f) \Big\|^2,$$

et cette inégalité n'est une égalité que lorsque la somme $\sum_{k\in\mathscr{J}} |c_k(f) - a_k|^2$ à termes positifs que nous avons fait disparaître s'annule, ce qui implique $c_k(f) = a_k$ pour tout $k \in \mathscr{J}$. □

Corollaire 6.3. *Soit $\{e_i\}_{i\in\mathscr{I}}$ une famille orthonormée de H. Alors pour tout $f \in H$ et toutes parties finies $\mathscr{K} \subset \mathscr{J}$ non vides de $\mathscr{I}$, on a :*

$$\Big\| f - \sum_{j\in\mathscr{J}} c_j(f)\, e_j \Big\| \leqslant \Big\| f - \sum_{j\in\mathscr{K}} c_j(f)\, e_j \Big\|$$

Démonstration. Il suffit d'appliquer *logiquement* le théorème avec $a_j := c_j(f)$ pour $j \in \mathscr{K}$ et avec $a_j := 0$ pour $j \in \mathscr{J} \backslash \mathscr{K}$. □

Puisque la logique mécanique n'éclaire rien, dévoilons la signification géométrique de cette inégalité :

$$\| f - \pi_{\mathscr{J}}(f) \| \leqslant \| f - \pi_{\mathscr{K}}(f) \|.$$

Autrement dit, plus un espace vectoriel grossit par inclusions croissantes, plus par projection orthogonale, l'approximation offerte s'améliore.

7. Bessel, Parseval, Plancherel, Développement dans une base hilbertienne

Dans le théorème ci-dessous, il faut observer que la famille $\{e_i\}_{i\in\mathscr{I}}$ n'est pas supposée totale et qu'elle peut même être finie. En fait, lorsque $\{e_i\}_{i\in\mathscr{I}}$ est une base hilbertienne, l'inégalité de Bessel deviendra une *égalité* que l'on appellera l'identité de Plancherel, *voir* la suite.

Théorème 7.1. [Inégalité de Bessel] *Si $\{e_i\}_{i\in\mathscr{I}}$ est une famille orthonormée au plus dénombrable de H, alors pour tout vecteur $f \in H$:*

$$\sum_{i\in\mathscr{I}} \big|\langle f,\, e_i\rangle\big|^2 \leqslant \|f\|^2 < \infty.$$

Démonstration. Soit $\mathscr{I}' \subset \mathscr{I}$ un sous-ensemble d'indices *fini* quelconque. Nous avons déjà vu que la projection orthogonale $\pi_{\mathscr{I}'}(f)$ de f sur $E_{\mathscr{I}'} = \mathrm{Vect}_{\mathbb{C}}(e_i \colon i \in \mathscr{I}')$ est orthogonale au vecteur $f - \pi_{\mathscr{I}'}(f)$, donc le théorème de Pythagore appliqué deux fois nous donne :

$$\begin{aligned}
\|f\|^2 &= \Big\| \sum_{i\in\mathscr{I}'} \langle f, e_i\rangle\, e_i \Big\|^2 + \Big\| f - \sum_{i\in\mathscr{I}'} \langle f, e_i\rangle\, e_i \Big\|^2 \\
&\geqslant \Big\| \sum_{i\in\mathscr{I}'} \langle f, e_i\rangle\, e_i \Big\|^2 \\
&= \sum_{i\in\mathscr{I}'} \big|\langle f, e_i\rangle\big|^2,
\end{aligned}$$

ce qui donne l'inégalité de Bessel pour toute famille *finie* $\mathscr{I}' \subset \mathscr{I}$. Or $\mathscr{I}$ est dénombrable, donc à la limite on obtient l'inégalité de Bessel (les sommes partielles finies d'une série numérique à termes $\geqslant 0$ qui est bornée convergent vers une limite majorée par la même borne). □

Proposition 7.2. [Convergence commutative] *Soit* $\{e_i\}_{i=1}^{\infty}$ *une* suite *orthonormée dans un espace de Hilbert. Si l'on note* $c_i(f) := \langle f, e_i \rangle$ *pour tout* $f \in H$*, alors la série :*

$$\sum_{i=1}^{\infty} c_i(f)\, e_i$$

converge commutativement dans H*, à savoir, les deux conditions suivantes sont satisfaites :*

- *il existe un unique élément* $g \in H$ *tel que*

$$\lim_{n \to \infty} \Big\| g - \sum_{i=1}^{n} c_i(f)\, e_i \Big\| = 0\,;$$

- *pour toute bijection* $\tau \colon \mathbb{N}^* \to \mathbb{N}^*$*, on a aussi, avec le même élément* g *:*

$$\lim_{n \to \infty} \Big\| g - \sum_{i=1}^{n} c_{\tau(i)}(f)\, e_{\tau(i)} \Big\| = 0.$$

À nouveau, la suite orthonormée $(e_i)_{i=1}^{\infty}$ n'est pas supposée totale afin de gagner en généralité.

Démonstration. Grâce à l'inégalité de Bessel, on sait déjà que :

$$\sum_{i=1}^{\infty} |c_i(f)|^2 \leqslant \|f\|^2 < \infty.$$

Donc cette série de réels positifs converge dans $\mathbb{R}_+$. Par conséquent, la suite de ses sommes partielles vérifie le critère de Cauchy dans $\mathbb{R}_+$, à savoir :

$$\forall\, \varepsilon > 0 \quad \exists\, I \geqslant 1 \quad \Big(i_1,\, i_2 \geqslant I \Rightarrow \sum_{i_1 \leqslant i \leqslant i_2} |c_i(f)|^2 \leqslant \varepsilon \Big).$$

Mais en raison de l'orthonormalité de la suite $(e_i)_{i=1}^{\infty}$, on a :

$$\Big\| \sum_{i_1 \leqslant i \leqslant i_2} c_i(f)\, e_i \Big\|^2 = \sum_{i_1 \leqslant i \leqslant i_2} |c_i(f)|^2 \leqslant \varepsilon.$$

Cette dernière inégalité montre donc que la série *vectorielle* $\sum_i c_i(f)\, e_i$ constituée d'éléments de H vérifie le critère de Cauchy. Comme tout espace de Hilbert est supposé complet par définition, cette série converge donc (pour la distance de H) vers un certain élément g de H, qui est bien sûr unique.

La convergence commutative est un phénomène général partagé par toutes les séries à termes positifs ou nuls, en particulier par $\sum_{i=1}^{\infty} |c_i(f)|^2$.

Lemme 7.3. *Si* $\sum_{k=1}^{\infty} a_k$ *est une série à termes réels* $a_k \geqslant 0$ *qui converge vers un élément* $\ell \in \mathbb{R}_+$*, alors pour toute bijection quelconque* $\tau \colon \mathbb{N}^* \to \mathbb{N}^*$*, la série renumérotée* $\sum_{k=1}^{\infty} a_{\tau(k)}$ *converge à nouveau vers le même élément* $\ell \in \mathbb{R}_+$.

Démonstration. Afin de montrer que la limite est la même, considérons tout naturellement la différence des sommes partielles jusqu'à un (grand) entier n :

$$\Big| \sum_{k_1=1}^{n} a_{k_1} - \sum_{k_2=1}^{n} a_{\tau(k_2)} \Big|.$$

Si, pour deux indices k_1 et k_2, on a $k_1 = \tau(k_2)$, le terme $a_{k_1} - a_{\tau(k_2)}$ s'annihile. Dans la différence entre les deux sommes, il reste donc, à un signe $\pm$ près, seulement les termes qui n'apparaissent que dans une seule somme.

Rappelons que la *différence symétrique* $\mathscr{A} \Delta \mathscr{B}$ entre deux sous-ensembles $\mathscr{A}$ et $\mathscr{B}$ d'un ensemble $\mathscr{E}$ est constituée de l'ensemble des éléments de $\mathscr{E}$ qui n'appartiennent qu'à *un seul* des deux ensembles $\mathscr{A}$ et $\mathscr{B}$. En utilisant la notation $\mathscr{D}^c := \mathscr{E} \backslash \mathscr{D}$ pour désigner le complémentaire d'un sous-ensemble $\mathscr{D} \subset \mathscr{E}$, on peut aussi écrire cette différence symétrique comme :

$$\mathscr{A} \Delta \mathscr{B} := (\mathscr{A} \cap \mathscr{B}^c) \bigcup (\mathscr{A}^c \cap \mathscr{B}).$$

En revenant à notre différence entre deux sommes, si nous notons :

$$\mathscr{K}_n := \big\{1, 2, 3, \dots, n\big\} \Delta \big\{\tau(1), \tau(2), \tau(3), \dots, \tau(n)\big\},$$

nous pouvons donc la majorer, en tenant compte du fait que $|\pm a_k| = a_k$:

$$\begin{aligned} \Big| \sum_{k_1=1}^{n} a_{k_1} - \sum_{k_2=1}^{n} a_{\tau(k_2)} \Big| &\leqslant \sum_{k \in \mathscr{K}_n} a_k \\ &\leqslant \sum_{k=\mathsf{min}(\mathscr{K}_n)}^{\infty} a_k. \end{aligned}$$

Pour démontrer que cette dernière série majorante tend vers zéro avec n, il suffit donc de vérifier que $\mathsf{min}\mathscr{K}_n$ devient arbitrairement grand quand n tend vers l'infini.

Fixons alors un entier N quelconque. La condition $\mathsf{min}(\mathscr{K}_n) \geqslant N$ est réalisée dès que tous les entiers $j = 1, 2, 3, \dots, N-1$ sont aussi dans l'ensemble $\{\tau(1), \tau(2), \dots, \tau(n)\}$. Mais pour garantir cela, il suffit que :

$$n \geqslant \mathsf{max} \big\{\tau^{-1}(j) \colon 1 \leqslant j \leqslant N-1\big\},$$

et cette dernière quantité est finie. Ceci démontre que $\mathsf{min}(\mathscr{K}_n)$ tend bien vers l'infini avec n, et conclut la preuve. □

Avec les notations du lemme, l'orthonormalité de la suite $(e_k)_{k=1}^{\infty}$ nous permet aussi de majorer pour tout $n \in \mathbb{N}$:

$$\Big\| \sum_{k=1}^{\infty} c_k(f)\, e_k - \sum_{k=1}^{n} c_{\tau(k)}(f)\, e_{\tau(k)} \Big\|^2 = \sum_{k \in \mathscr{K}_n} |c_k(f)|^2 \leqslant \sum_{k=\mathsf{min}\mathscr{K}_n}^{\infty} |c_k(f)|^2$$

par un majorant qui tend tout aussi bien vers zéro quand n tend vers l'infini. Ceci achève la preuve du théorème. □

Voici enfin l'énoncé qui présente les espaces de Hilbert séparables comme des espaces munis de vraies bases (infinies dénombrables) avec lesquelles on peut faire des calculs essentiellement comme en dimension finie.

Théorème 7.4. [Décomposition selon une base hilbertienne] *Soit H un espace de Hilbert de dimension infinie qui est* séparable. *Si $\{e_i\}_{i=1}^{\infty}$ est une base hilbertienne (dénombrable) de H, alors tout vecteur $f \in H$ se décompose de manière unique comme combinaison linéaire (infinie) des e_i :*

$$f = \sum_{i=1}^{\infty} \langle f, e_i \rangle\, e_i,$$

où la série converge (commutativement) pour la topologie de H, à savoir : pour toute bijection $\tau\colon \mathbb{N}^* \to \mathbb{N}^*$, on a :

$$0 = \lim_{n\to\infty} \Big\| f - \sum_{i=1}^{n} \langle f, e_{\tau(i)}\rangle\, e_{\tau(i)} \Big\|.$$

Réciproquement, étant donné une famille orthonormée $\{e_i\}_{i\in\mathscr{I}}$ de H (nécessairement de cardinal infini dénombrable), si l'on a $f = \sum_{i\in\mathscr{I}} \langle f, e_i\rangle e_i$ pour tout $f \in H$, alors la famille $\{e_i\colon i \in \mathscr{I}\}$ est en fait totale, et c'est donc une base hilbertienne.

Démonstration. Nous savons qu'il suffit d'établir la convergence vers f de la série lorsque $\tau\colon \mathbb{N}^* \to \mathbb{N}^*$ est l'identité. Démontrons donc que la quantité :

$$\delta_n := \Big\| f - \sum_{i=1}^{n} \langle f, e_i\rangle\, e_i \Big\|$$

tend vers zéro lorsque n tend vers l'infini. Observons d'ores et déjà que le Corollaire 6.3 nous informe que la suite δ_n est décroissante.

Puisque par hypothèse, la collection $\{e_i\}_{i=1}^{\infty}$ est totale dans H, pour tout $\varepsilon > 0$ de la forme $\varepsilon = \frac{1}{k} > 0$ avec $k \in \mathbb{N}^*$ grand, il existe une combinaison linéaire finie :

$$\sum_{i=1}^{n_k} a_{k,i}\, e_i$$

des e_i jusqu'à un certain entier n_k (qui dépend de k) à coefficients complexes $a_{k,i}$ (qui dépendent aussi de k) telle que :

$$\Big\| f - \sum_{i=1}^{n_k} a_{k,i}\, e_i \Big\| \leqslant \frac{1}{k}.$$

Mais grâce à la propriété de meilleure approximation énoncée par le Théorème 6.2, on a automatiquement :

$$\delta_{n_k} = \Big\| f - \sum_{i=1}^{n_k} \langle f, e_i\rangle\, e_i \Big\| \leqslant \Big\| f - \sum_{i=1}^{n_k} a_{k,i}\, e_i \Big\| \underset{k\to\infty}{\longrightarrow} 0.$$

Ainsi, la sous-suite $(\delta_{n_k})_{k=1}^{\infty}$ des δ_n tend vers zéro lorsque $k \longrightarrow \infty$, et comme $(\delta_n)_{n\geqslant 1}$ est décroissante, c'est donc la suite $(\delta_n)_{n=1}^{\infty}$ *tout entière* qui converge vers zéro. Ceci prouve la convergence dans H de la série $\sum_{i=1}^{\infty} \langle f, e_i\rangle\, e_i$ vers f.

Pour ce qui est de la réciproque, si l'on indexe l'ensemble dénombrable $\mathscr{I}$ par des entiers $i \geqslant 1$ la convergence de $\sum_{i=1}^{\infty} \langle f, e_i\rangle e_i$ vers f signifie précisément que les sommes partielles $\sum_{i=1}^{n} \langle f, e_i\rangle e_i$ peuvent être rendues arbitrairement proches de f, donc la famille $\{e_i\}_{i=1}^{\infty}$ est totale. □

Lemme 7.5. [Continuité de la norme] *Si $(f_k)_{k=1}^{\infty}$ est une suite dans un espace de Hilbert quelconque H qui converge vers un élément $f_\infty \in H$, à savoir :*

$$\lim_{k\to\infty} \|f_k - f_\infty\| = 0,$$

alors les normes de f_k convergent aussi vers la norme de f_∞ :

$$\lim_{k\to\infty} \|f_k\| = \|f_\infty\|.$$

Démonstration. En élevant au carré la norme de f_k, en faisant apparaître $-f_\infty + f_\infty$, en développant :

$$\|f_k\|^2 = \|f_k - f_\infty + f_\infty\|^2 = \|f_k - f_\infty\|^2 + \|f_\infty\|^2 + 2\mathsf{Re}\,\langle f_k - f_\infty,\, f_\infty\rangle,$$

et en appliquant l'inégalité de Cauchy-Schwarz après une inégalité triangulaire :

$$\Big|\|f_k\|^2 - \|f_\infty\|^2\Big| \leqslant \|f_k - f_\infty\|^2 + 2\,\|f_k - f_\infty\|\,\|f_\infty\|,$$

on majore la différence $\|f_k\|^2 - \|f_\infty\|^2$ par une quantité qui tend par hypothèse vers zéro lorsque $k \to +\infty$. □

Théorème 7.6. [Identité de Plancherel] *Si H est un espace de Hilbert séparable de dimension infinie (dénombrable) muni d'une base hilbertienne $\{e_i\}_{i=1}^{\infty}$, alors la norme au carré de tout vecteur $f \in H$:*

$$f = \sum_{i=1}^{\infty} \langle f, e_i\rangle\, e_i$$

qui est décomposé selon cette base s'obtient simplement en sommant les modules au carré de ses coefficients :

$$\|f\|^2 = \sum_{i=1}^{\infty} \big|\langle f, e_i\rangle\big|^2.$$

[Identité de Parseval] *Plus généralement, le produit scalaire $\langle f, g\rangle$ de f avec un autre élément quelconque g de H :*

$$g = \sum_{i=1}^{\infty} \langle g, e_i\rangle\, e_i$$

qui est lui aussi décomposé selon cette même base, s'obtient en sommant les produits (conjugués) de leurs coefficients :

$$\langle f, g\rangle = \sum_{i=1}^{\infty} \langle f, e_i\rangle\, \overline{\langle g, e_i\rangle}.$$

Ce qu'on appelle « identité de Plancherel » n'est donc autre que la version du Théorème de Pythagore en dimension infinie : immortalité et renouveau d'un théorème vieux de vingt-cinq siècles ! Et bien entendu aussi, l'identité de Parseval est la généralisation aux espaces de Hilbert de la règle bien connue de calcul du produit scalaire $x_1y_1 + \cdots + x_ny_n$ entre deux vecteurs de $\mathbb{R}^n$ ayant les coordonnées $(x_1, \ldots, x_n)$ et $(y_1, \ldots, y_n)$ dans la base orthogonale canonique.

Démonstration. Si $\{e_i\}_{i=1}^{\infty}$ est une base hilbertienne de H, on sait que : pour tout $f \in H$, la somme partielle :

$$\sum_{i=1}^{n} \langle f, e_i\rangle\, e_i$$

tend vers f lorsque n tend vers l'infini. Par continuité de la norme sur H, on en déduit aussi la convergence dans $\mathbb{R}^+$:

$$\Big\|\sum_{i=1}^{n} \langle f, e_i\rangle e_i\Big\|^2 \underset{n\to\infty}{\longrightarrow} \|f\|^2,$$

ce qui, grâce à l'orthonormalité des e_i, signifie précisément que l'identité de Plancherel est satisfaite :

$$\sum_{i=1}^{n} |\langle f, e_i\rangle|^2 \underset{n\to\infty}{\longrightarrow} \|f\|^2.$$

Pour la preuve de l'identité de Parseval, nous avons besoin d'un énoncé sur les opérateurs continus entre espaces vectoriels normés qui nous permettra de réinterpréter la fameuse inégalité de Cauchy-Schwarz. Le théorème suivant, que nous rappelons avec sa démonstration, ramène les problèmes de continuité à des majorations sur les normes.

Théorème 7.7. . *Soient $(F, \|\cdot\|_F)$ et $(G, \|\cdot\|_G)$ deux espaces vectoriels normés et soit L une application linéaire de F dans G. Les trois propriétés suivantes sont équivalentes :*

(i) *l'application L est continue à l'origine ;*

(ii) *l'application L est uniformément continue ;*

(iii) *il existe une constante $C \geqslant 0$ telle que l'on ait :*

$$\forall\, f \in F \qquad \|L(f)\|_G \leqslant C\,\|f\|_F.$$

La dernière inégalité **(iii)** exprime que la *norme d'opérateur* de L, définie par :

$$|\!|\!|L|\!|\!|_{\mathsf{Lin}(F,G)} := \sup_{f\neq 0} \frac{\|L(f)\|_G}{\|f\|_F} = \sup_{\|f\|_F\leqslant 1} \|L(f)\|_G$$

est finie et qu'elle est majorée par :

$$|\!|\!|L|\!|\!|_{\mathsf{Lin}(F,G)} \leqslant C.$$

Dans les applications concrètes aux espaces de fonctions, on est souvent amené à essayer de majorer $\|L(f)\|_G$ en faisant apparaître $\|f\|_F$ à une constante près dans le majorant. Bien entendu, lorsqu'on manipule des inégalités strictes, on perd de l'information, et la constante que l'on obtient est presque toujours strictement supérieure à la norme d'opérateur $|\!|\!|L|\!|\!|$. Raffiner les calculs afin d'obtenir la meilleure constante C, c'est-à-dire $|\!|\!|L|\!|\!|$, ouvre souvent sur des questions mathématiques délicates.

Démonstration. Montrons que **(i)** implique **(iii)**. La continuité à l'origine dit que pour tout $\varepsilon > 0$, il existe $\eta > 0$ tel que $\|f\|_F \leqslant \eta$ entraîne $\|L(f)\|_G \leqslant \varepsilon$. Prenons $\varepsilon = 1$. Soit maintenant $f \in F$ quelconque, mais non nul. Nous pouvons alors appliquer l'inégalité $\|L(f')\|_G \leqslant 1$ au vecteur $f' := \eta\, \frac{f}{\|f\|_F}$ qui satisfait bien sûr $\|f'\| \leqslant \eta$, ce qui nous donne :

$$\big\|L\big(\eta\, \tfrac{f}{\|f\|_F}\big)\big\|_G \leqslant 1 \Longleftrightarrow \|L(f)\|_G \leqslant \tfrac{1}{\eta}\, \|f\|_F \qquad (\forall\, f \in F\backslash\{0\}),$$

mais cette dernière égalité est aussi valable pour $f = 0$, puisque L est linéaire.

L'implication **(ii)** $\Rightarrow$ **(i)** étant triviale, il reste à montrer que **(iii)** entraîne **(ii)**. Mais en appliquant **(iii)** à $f - g$ avec $f, g \in H$, étant donné $\varepsilon > 0$ arbitrairement petit, il suffit de choisir $\eta = \varepsilon/C$ pour garantir que :

$$\big(\|f - g\|_F \leqslant \eta\big) \Longrightarrow \big(\|Lf - Lg\|_G \leqslant \varepsilon\big),$$

ce qui achève la preuve. □

Maintenant, l'inégalité de Cauchy-Schwarz :

$$\big|\langle f, g\rangle\big| \leqslant \|f\|\,\|g\|$$

montre *simultanément* que les deux applications suivantes, linéaire et sesquilinéaire, de H à valeurs dans $\mathbb{C}$ et appelées « produit scalaire avec un vecteur fixé » :

$$\begin{aligned} L_g \colon f &\longmapsto \langle f, g\rangle \\ H &\longrightarrow \mathbb{C} \end{aligned} \qquad \text{et} \qquad \begin{aligned} L_f \colon g &\longmapsto \langle f, g\rangle \\ H &\longrightarrow \mathbb{C} \end{aligned}$$

sont continues.

Nous pouvons revenir maintenant à l'identité de Parseval. Tout d'abord, si on applique l'inégalité élémentaire déjà vue en (2.4) :

$$2\,|zz'| \leqslant |z|^2 + |z'|^2$$

satisfaite par deux nombres complexes arbitraires $z, z' \in \mathbb{C}$ au terme général de la série $\sum_i \langle f, e_i\rangle\,\overline{\langle g, e_i\rangle}$, on déduit que cette série est absolument convergente, puisque :

$$\begin{aligned} 2\sum_{i=1}^{\infty} \big|\langle f, e_i\rangle\,\overline{\langle g, e_i\rangle}\big| &\leqslant \sum_{i=1}^{\infty} |\langle f, e_i\rangle|^2 + \sum_{i=1}^{\infty} |\langle g, e_i\rangle|^2 \\ &= \|f\|^2 + \|g\|^2 < \infty \end{aligned}$$

grâce à l'identité de Plancherel. Comme $\{e_i\}_{i=1}^{\infty}$ est une base hilbertienne, le Théorème 7.4 nous assure que :

$$\sum_{i=1}^{n} \langle f, e_i\rangle e_i \underset{n\to\infty}{\longrightarrow} f \qquad \text{et que :} \qquad \sum_{i=1}^{n} \langle g, e_i\rangle e_i \underset{n\to\infty}{\longrightarrow} g$$

pour la topologie définie par la norme de H. Mais comme nous venons de voir que le produit scalaire est une forme bilinéaire *continue* sur H (en raison de l'inégalité de Cauchy-Schwarz), on en déduit aussi que :

$$\Big\langle \sum_{i=1}^{n} \langle f, e_i\rangle e_i,\ \sum_{i=1}^{n} \langle g, e_i\rangle e_i \Big\rangle \underset{n\to\infty}{\longrightarrow} \langle f, g\rangle,$$

ce qui, grâce à l'orthonormalité des e_i et à la bilinéarité du produit scalaire :

$$\sum_{i_1=1}^{n}\sum_{i_2=1}^{n} \langle f, e_{i_1}\rangle\,\overline{\langle g, e_{i_2}\rangle}\,\underbrace{\langle e_{i_1}, e_{i_2}\rangle}_{\delta_{i_1,i_2}} = \sum_{i=1}^{n} \langle f, e_i\rangle\,\overline{\langle g, e_i\rangle}$$

signifie précisément que l'identité de Parseval est satisfaite :

$$\sum_{i=1}^{\infty} \langle f, e_i\rangle\,\overline{\langle g, e_i\rangle} = \langle f, g\rangle.$$

La démonstration détaillée est terminée. □

Proposition 7.8. [Unicité du développement] *Soit $\{e_i\}_{i=1}^{\infty}$ une base hilbertienne dénombrable d'un espace de Hilbert séparable H de dimension infinie, et soit $(a_i)_{i=1}^{\infty}$ une famille de scalaires telle que la série $\sum_{i=1}^{\infty} a_i\, e_i$ converge au sens de la topologie de H vers un élément $f = \sum a_i e_i$. Alors $a_i = \langle f, e_i\rangle$ pour tout $i \geqslant 1$.*

Démonstration. Par hypothèse, nous avons donc :

$$f_k := \sum_{i=1}^{k} a_i\, e_i \underset{k\to\infty}{\longrightarrow} f$$

lorsque k tend vers l'infini. D'autre part, pour tous entiers $i \geqslant 1$ et $k \geqslant 1$, on a par orthonormalité des e_i :

$$\langle f_k,\, e_i\rangle = \begin{cases} a_i & \text{si } i \leqslant k \\ 0 & \text{si } i \geqslant k+1. \end{cases}$$

Pour tout i fixé, la suite de nombres complexes $\big(\langle f_k, e_i\rangle\big)_{k=1}^{\infty}$ est donc constante égale à a_i à partir du rang $k \geqslant i$. Ainsi, elle converge (trivialement) vers a_i.

Par ailleurs, toujours avec i fixé, nous savons grâce à l'inégalité de Cauchy-Schwarz que l'application $f \mapsto \langle f, e_i\rangle$ de H à valeurs dans $\mathbb{C}$ est une forme linéaire *continue*. Comme f_k tend vers f pour la topologie de H, on en déduit que :

$$\langle f_k,\, e_i\rangle \underset{k\to\infty}{\longrightarrow} \langle f, e_i\rangle.$$

Par unicité de la limite dans $\mathbb{C}$, on en déduit finalement que $a_i = \langle f, e_i\rangle$ pour tout $i \geqslant 1$. $\square$

Une application linéaire L entre deux espace vectoriels normés $(F, \|\cdot\|_F)$ et $(G, \|\cdot\|_G)$ est une *isométrie* si on a :

$$\|L(f)\|_G = \|f\|_F$$

pour tout élément $f \in F$. Il en découle immédiatement que L est injective. Si elle est de plus surjective, on dit que L est un isomorphisme isométrique.

Théorème 7.9. [Riesz-Fischer] *Tout espace de Hilbert H séparable de dimension infinie est isométriquement isomorphe à ℓ^2.*

Cet énoncé est la généralisation, en dimension infinie dénombrable, du théorème d'après lequel tout espace vectoriel hermitien de dimension finie n est isomorphe à $\mathbb{C}^n$ muni du produit scalaire hermitien standard $z_1\overline{w}_1 + \cdots + z_n\overline{w}_n$. Il est important d'ajouter qu'il y a autant d'isomorphismes isométriques entre H et ℓ^2 qu'il y a de bases hilbertiennes, et donc qu'aucun tel isomorphisme isométrique n'est canonique.

Démonstration. Rappelons que ℓ^2 est l'espace des suites infinies dénombrables $z = (z_1, z_2, \dots, z_i, \dots)$ de nombres complexes $z_i \in \mathbb{C}$ telles que :

$$\|z\|_{\ell^2}^2 := \sum_{i=1}^{\infty} |z_i|^2 < \infty,$$

et que c'est un espace de Hilbert pour le produit scalaire hermitien :

$$\langle z, w\rangle_{\ell^2} := \sum_{i=1}^{\infty} z_i\, \overline{w_i},$$

de manière entièrement analogue à ce qui vaut en dimension finie.

Soit donc H un espace de Hilbert de dimension infinie qui est séparable. Bien entendu, nous savons grâce au Théorème 5.12 que H possède une base hilbertienne dénombrable ; notons-la comme auparavant $\{e_i\}_{i=1}^{\infty}$.

L'application qui va nous fournir l'isomorphisme isométrique est alors tout simplement celle qui, à un élément quelconque $f \in H$, associe tous les coefficients de son développement dans une telle base hilbertienne :

$$\Psi\colon\quad H \longrightarrow \ell^2, \qquad f \longmapsto \big(\langle f, e_i\rangle\big)_{i=1}^{\infty}.$$

En effet, l'identité de Plancherel montre que $\Psi(f)$ est bien un élément de ℓ^2 et que Ψ est une isométrie :

$$\|\Psi(f)\|_{\ell^2} = \|f\|_H. \qquad (\forall\, f \in H).$$

L'injectivité de Ψ en découle immédiatement. Par ailleurs, Ψ étant clairement linéaire, cette égalité interprétée comme *in*égalité $\|\Psi(f)\|_{\ell^2} \leqslant \|f\|_H$ montre que Ψ est une application linéaire *continue*.

Il ne reste plus qu'à établir la surjectivité de Ψ. Soit donc $z = (z_i)_{i=1}^{\infty}$ un élément quelconque de ℓ^2. Pour montrer l'existence d'un $f \in H$ tel que $\Psi(f) = z$, définissons une suite $(f_k)_{k=1}^{\infty}$ d'éléments de H par :

$$f_k := \sum_{i=1}^{k} z_i\, e_i \qquad (k \geqslant 1).$$

Grâce à l'orthonormalité des e_i, on peut tester si le critère de Cauchy est satisfait :

$$\begin{aligned} \forall\, k_1, \quad \forall\, k_2 \geqslant k_1 + 1, \qquad \|f_{k_2} - f_{k_1}\|^2 &= \sum_{k_1+1 \leqslant i \leqslant k_2} |z_i|^2\, \|e_i\|^2 \\ &\leqslant \sum_{i=k_1+1}^{\infty} |z_i|^2 \underset{k_1 \to \infty}{\longrightarrow} 0, \end{aligned}$$

et cette dernière quantité tend visiblement vers zéro lorsque k_1 tend vers l'infini. Donc la suite $(f_k)_{k=1}^{\infty}$ est bien de Cauchy, et comme H est complet par définition, cette suite possède une (unique) limite f dans H, ce qui s'écrit encore :

$$f = \sum_{i=1}^{\infty} z_i\, e_i,$$

la série étant convergente pour la topologie de H. Mais d'après la Proposition 7.8, on a alors nécessairement $z_i = \langle f, e_i \rangle$ pour tout $i \geqslant 1$, et donc $\Psi(f) = z$, comme annoncé. Fin de la preuve ! □

Comme nous l'avions dit par anticipation au début de ces considérations, la notion purement abstraite d'espace de Hilbert séparable se ramène, *via* le choix d'une base hilbertienne (orthonormée), à celle de l'espace ℓ^2 concret. Toutefois, comme les bases hilbertiennes sont nombreuses et variables, l'isomorphisme Ψ ci-dessus n'est en rien canonique ; il indique seulement que la combinatoire du calcul en coordonnées dans une base hilbertienne fixée est exactement la même que dans l'espace ℓ^2, ce qui est néanmoins très satisfaisant pour l'intuition mathématique.

8. Polynômes orthogonaux

Soit un intervalle ouvert $]a, b[$ non nécessairement borné avec :

$$-\infty \leqslant a < b \leqslant \infty.$$

Définition 8.1. Un *poids* sur $]a, b[$ est une fonction continue strictement positive :

$$w\colon \quad]a, b[\longrightarrow]0, \infty[,$$

dont tous les *moments d'ordre* $n \in \mathbb{N}$, à savoir les intégrales :

$$\int_a^b |x|^n\, w(x)\, dx \;<\; \infty$$

sont supposés *finis*.

Avec une telle fonction-poids w, lettre-initiale du mot anglais correspondant *weight*, introduisons l'espace vectoriel :

$$\mathscr{H}_w := \left\{ f \in \mathscr{C}^0\big(]a,b[, \mathbb{C}\big) \colon \int_a^b |f(x)|^2\, w(x)\, dx \,<\, \infty \right\}.$$

On vérifie aisément que $\mathscr{H}_w$ est un espace préhilbertien muni du produit scalaire naturel :

$$\langle f, g \rangle_w := \int_a^b f(x)\, \overline{g(x)}\, w(x)\, dx,$$

et de la norme naturelle :

$$\| f \|_w := \left(\int_a^b |f(x)|^2\, w(x)\, dx \right)^{\frac{1}{2}}.$$

Toutefois, $\mathscr{H}_w$ n'est pas forcément complet en général, et afin que ce soit un vrai espace de Hilbert, il faut étudier des conditions spécifiques qui assureraient sa complétude.

Avant d'entamer une telle étude, discutons le cas simple où $-\infty < a < b < \infty$ où l'intervalle considéré $[a,b]$ est compact, et où le poids $w \in \mathscr{C}^0([a,b], \mathbb{R}_+^*)$ est continu sur $[a,b]$, de telle sorte que :

$$0 \,<\, c := \min_{a \leqslant x \leqslant b} w(x) \,\leqslant\, \max_{a \leqslant x \leqslant b} w(x) =: C \,<\, \infty.$$

Dans ce cas, nous allons argumenter que l'on a l'isomorphisme :

$$\big(\mathscr{H}_w,\ \| \cdot \|_w\big) \ \cong\ \big(L^2\big([a,b], \mathbb{C}\big),\ \| \cdot \|_{L^2}\big).$$

Définition 8.2. Une application $\mathbb{C}$-linéaire $U \colon H \longrightarrow H'$ entre deux espaces de Hilbert — ou simplement préhilbertiens — $(H, \langle \cdot, \cdot \rangle)$ et $(H', \langle \cdot, \cdot \rangle)$ est dite *unitaire* lorsque :

(i) U est bijective ;

(ii) $\big\| U(f) \big\|_{H'} = \| f \|_H$ pour tout $f \in H$.

En particulier, U est une application linéaire *continue*.

De plus, $U^{-1} \colon H' \longrightarrow H$ existe, est une application $\mathbb{C}$-linéaire (exercice), et pour tout $f' \in H'$, on a :

$$\begin{aligned} \big\| U^{-1}(f') \big\|_H &= \big\| U\big(U^{-1}(f')\big) \big\|_{H'} \\ &= \| f' \|_{H'}, \end{aligned}$$

donc l'inverse :

$$U^{-1} \colon \quad H' \longrightarrow H$$

est aussi unitaire.

Ensuite, grâce aux identités de polarisation énoncées dans le Théorème 3.11, on démontre (exercice) que $U \colon H \longrightarrow H'$ est unitaire si et seulement si :

$$\big\langle U(f), U(g) \big\rangle_{H'} = \langle f, g \rangle_H \qquad\qquad (\forall\, f \in H,\ \forall\, g \in H).$$

Terminologie 8.3. Deux espaces de Hilbert — ou simplement préhilbertiens — H et H' sont dits *unitairement équivalents* lorsqu'il existe une application unitaire $U \colon H \longrightarrow H'$.

On se convainc alors aisément (exercice) que H est complet si et seulement si H' l'est.

Avec tous ces concepts, on s'aperçoit (exercice) que l'application :

$$\begin{aligned} I\colon \quad L^2([a,b]) &\longrightarrow \mathscr{H}_w \\ f &\longmapsto w^{-\frac{1}{2}}\, f \end{aligned}$$

est une équivalence unitaire, ce qui offre le fait que $\mathscr{H}_w$ est complet. Une autre manière de procéder (solution de l'exercice) est de regarder l'application identité (!) :

$$\begin{aligned} I\colon \quad L^2([a,b]) &\longrightarrow \mathscr{H}_w \\ f &\longmapsto f, \end{aligned}$$

qui n'est pas en général une équivalence unitaire entre espaces préhilbertiens, mais qui satisfait quand même une première inégalité :

$$\big\| I(f) \big\|_w^2 \ = \ \int_a^b \big| f(x) \big|^2\, w(x)\, dx \ \leqslant\ C\, \| f \|_{L^2}^2,$$

et une seconde inégalité :

$$c\, \| f \|_{L^2}^2 \ \leqslant\ \int_a^b \big| f(x) \big|^2\, w(x)\, dx \ = \ \| f \|_w^2$$

qui remplissent le même rôle argumentatif que l'égalité des normes dans le contexte unitaire en établissant un isomorphisme, et donc, puisque $L^2([a,b])$ est de Hilbert, nous déduisons que $\mathscr{H}_w$ est bien un espace de Hilbert lorsque $[a,b]$ est compact, et lorsque w est continue sur $[a,b]$.

Revenons maintenant au cas général d'un intervalle ouvert $]a,b[$ non nécessairement borné. Au moins pour tout $n \in \mathbb{N}$, l'espace $\mathscr{H}_w$ contient manifestement le sous-espace :

$$\mathscr{P}_n \subset \mathscr{H}_w$$

constitué des polynômes à coefficients complexes de degré $\leqslant n$.

Proposition 8.4. *Il existe une unique famille :*

$$\big(P_n\big)_{n=0}^{\infty}$$

de polynômes unitaires $P_n = P_n(x)$ de degré n qui sont orthogonaux deux à deux :

$$\begin{aligned} \langle P_{n_1},\, P_{n_2} \rangle_w &= 0 \quad \text{ lorsque } \ n_1 \neq n_2, \\ \langle P_{n_1},\, P_{n_2} \rangle_w &> 0 \quad \text{ lorsque } \ n_1 = n_2. \end{aligned}$$

Démonstration. Puisque P_0 est unitaire, nécessairement $P_0(x) := 1$, puis :

$$\langle P_0, P_0 \rangle_w \ = \ \int_a^b 1^2\, w(x)\, dx \ > \ 0.$$

Construisons P_n par récurrence sur l'entier $n \geqslant 1$, en supposant déjà connus $P_0, \dots, P_{n-1}$. Comme $\deg(P_k) = k$, ces polynômes $P_0, \dots, P_{n-1}$ forment une base de l'espace vectoriel $\mathscr{P}_{n-1}$. On peut donc chercher P_n sous la forme :

$$P_n(x) = x^n - \lambda_{n-1}\, P_{n-1}(x) - \cdots - \lambda_0\, P_0(x),$$

où les constantes $\lambda_{n-1}, \dots, \lambda_0 \in \mathbb{C}$ sont inconnues. Mais les n conditions d'orthogonalité supposées :

$$
\begin{aligned}
0 &= \langle P_n,\, P_0 \rangle_w \\
&\cdots\cdots\cdots\cdots \\
0 &= \langle P_n,\, P_{n-1} \rangle_w
\end{aligned}
$$

se lisent comme le système linéaire :

$$
\begin{aligned}
0 &= \langle x^n,\, P_0 \rangle_w - \lambda_{n-1} \underline{\langle P_{n-1}, P_0 \rangle_w} - \cdots - \lambda_0 \langle P_0, P_0 \rangle_w, \\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\
0 &= \langle x^n,\, P_{n-1} \rangle - \lambda_{n-1} \langle P_{n-1}, P_{n-1} \rangle_w - \cdots - \lambda_0 \underline{\langle P_0, P_{n-1} \rangle_w},
\end{aligned}
$$

lequel se réduit manifestement à un système diagonal par orthogonalité (hypothèse de récurrence) :

$$
\begin{aligned}
0 &= \langle x^n,\, P_0 \rangle_w - \lambda_0 \underbrace{\langle P_0,\, P_0 \rangle_w}_{\neq 0}, \\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\
0 &= \langle x^n,\, P_{n-1} \rangle_w - \lambda_{n-1} \underbrace{\langle P_{n-1},\, P_{n-1} \rangle_w}_{\neq 0},
\end{aligned}
$$

la solution en les λ_k étant visiblement unique. □

Exemple 8.5. Quatre familles de polynômes orthogonaux interviennent fréquemment :

- Polynômes de Laguerre : poids $w(x) = e^{-x}$ sur l'intervalle $[0, \infty[$.
- Polynômes de Hermite : poids $w(x) = e^{-x^2}$ sur l'intervalle $[-\infty, \infty[$.
- Polynômes de Legendre : poids $w(x) = 1$ sur l'intervalle $[-1, +1]$.
- Polynômes de Tchebychev : poids $w(x) = \frac{1}{\sqrt{1-x^2}}$ sur l'intervalle $[-1, +1[$.

En revenant à un poids $w(x)$ général dont les moments d'ordre $n \in \mathbb{N}$ quelconque sont finis, un calcul simple montre (exercice) que :

$$
P_0(x) = 1,
$$

et que :

$$
P_1(x) = x - \frac{\int_a^b x\, w(x)\, dx}{\int_a^b w(x)\, dx}.
$$

Des formules générales simples permettent, dans les applications, de calculer rapidement les polynômes orthogonaux uniques de $P_n(x)$.

Proposition 8.6. *Les polynômes orthogonaux associés à un poids w comme ci-dessus vérifient les relations de récurrence :*

$$
P_{n+1}(x) = \big(x - \alpha_n\big)\, P_n(x) - \beta_n\, P_{n-1}(x) \qquad\qquad (\forall\, n \geqslant 1),
$$

avec les constantes :

$$
\begin{aligned}
\alpha_n &:= \frac{\int_a^b x\, P_n(x)^2\, w(x)\, dx}{\int_a^b P_n(x)^2\, w(x)\, dx}, \\
\beta_n &:= \frac{\int_a^b P_n(x)^2\, w(x)\, dx}{\int_a^b P_{n-1}(x)^2\, w(x)\, dx},
\end{aligned}
$$

Démonstration. Fixons $n \geqslant 1$. On sait que $P_0, \dots, P_{n+1}$ forment une base de l'espace vectoriel $\mathscr{P}_{n+1}$ des polynômes de degré $\leqslant n+1$. Comme les polynômes sont unitaires, la différence :

$$x\,P_n(x) - P_{n+1}(x)$$

n'a plus de monôme en x^{n+1}, donc elle appartient à $\mathscr{P}_n$, à savoir elle s'exprime sous la forme :

$$x\,P_n(x) - P_{n+1}(x) = \sum_{k=0}^{n} c_k\,P_k(x),$$

avec certaines constantes $c_k \in \mathbb{C}$. On doit donc démontrer que :

$$\begin{aligned} c_n &= \alpha_n, \\ c_{n-1} &= \beta_n, \\ c_k &= 0 \quad \text{pour } \ k = n-2, \dots, 1, 0. \end{aligned}$$

À cette fin, pour un entier $l \leqslant n$ quelconque, multiplions l'équation précédente par $P_l(x)\,w(x)$, intégrons sur $]a, b[$:

$$\langle xP_n,\, P_l\rangle_w - \underline{\langle P_{n+1},\, P_l\rangle_w} = \sum_{k=0}^{n} c_k\, \underbrace{\langle P_k,\, P_l\rangle_w}_{=\,\delta_{k,l}}$$

ce qui, en tenant compte des orthogonalités, se simplifie agréablement comme :

$$\langle xP_n,\, P_l\rangle_w = c_l\, \underbrace{\langle P_l,\, P_l\rangle_w}_{\neq\,0},$$

et se résout en :

$$c_l = \frac{\langle xP_n,\, P_l\rangle_w}{\langle P_l,\, P_l\rangle_w}.$$

Lorsque $l = n$, on obtient l'expression annoncée :

$$\begin{aligned} c_n &= \frac{\langle xP_n,\, P_n\rangle_w}{\langle P_n,\, P_n\rangle_w} \\ &= \alpha_n. \end{aligned}$$

Ensuite, lorsque $l = n-1$, on obtient :

$$c_{n-1} = \frac{\langle xP_n,\, P_{n-1}\rangle_w}{\langle P_{n-1},\, P_{n-1}\rangle_w},$$

mais le numérateur n'est pas encore celui de la constante annoncée β_n. Qu'à cela ne tienne, l'astuce consiste à faire passer x de l'autre côté :

$$\langle xP_n,\, P_{n-1}\rangle_w = \langle P_n,\, xP_{n-1}\rangle_w,$$

puis à faire observer que :

$$P_n - x\,P_{n-1} \,\in\, \mathscr{P}_{n-1}$$

est de degré $\leqslant n-1$ par unitarité des polynômes, d'où l'orthogonalité :

$$\langle P_n,\, P_n - xP_{n-1}\rangle_w = 0,$$

c'est-à-dire :

$$\langle P_n,\, xP_{n-1}\rangle_w = \langle P_n,\, P_n\rangle_w$$

et au final, on a bien transformé le numérateur de c_{n-1} pour faire voir qu'il est égal à celui annoncé de $\beta_n = c_{n-1}$.

Enfin et très facilement, lorsque $l \leqslant n-2$, il est clair que le numérateur de :

$$c_l = \frac{\langle P_n,\, xP_l\rangle_w}{\langle P_l,\, P_l\rangle_w}$$

est nul, puisque $x\, P_l \in \mathscr{P}_{n-1}$, et puisque P_n est orthogonal à ce sous-espace. □

L'intérêt principal des polynômes orthogonaux est leur application aux équations différentielles, en analyse numérique, en analyse complexe, en analyse harmonique, en probabilités. En se restreignant ici au cadre des espaces de Hilbert élémentaires, on trouve un résultat de meilleure approximation naturelle, conséquence directe d'un théorème déjà vu, et qui a le mérite de s'appliquer dans des contextes très variés.

Théorème 8.7. *Pour toute fonction $f \in \mathscr{H}_w$ appartenant à l'espace L^2 à poids :*

$$\mathscr{H}_w = L^2\big(]a,b[,w\big),$$

il existe un unique polynôme $R_n(f)$ de degré $\leqslant n$ qui minimise la distance à f dans l'espace $\mathscr{P}_n$ des polynômes de degré $\leqslant n$:

$$\big\|f - R_n(f)\big\|_w = \inf_{g \in \mathscr{P}_n} \big\|f - g\big\|_w.$$

Plus précisément, $R_n(f)$ est le projeté orthogonal de f sur $\mathscr{P}_n$. □

Ce polynôme est appelé *polynôme de meilleure approximation quadratique de f à l'ordre* n.

Démonstration. Comme le système $\big(P_n\big)_{n=0}^{\infty}$ de polynômes précédemment construits n'est pas nécessairement orthonormé — il est juste orthogonal —, une légère précision concernant l'expression du projeté orthogonal $R_n(f)$ sera bienvenue.

Si on cherche en effet $R_n(f)$ sous la forme :

$$R_n(f) = \sum_{k=0}^{n} \mu_k\, P_k,$$

alors les conditions :

$$\begin{aligned} 0 &= \langle R_n(f),\, P_0\rangle_w, \\ &\cdots\cdots\cdots\cdots\cdots\cdots \\ 0 &= \langle R_n(f),\, P_n\rangle_w, \end{aligned}$$

permettent aisément (exercice) de déterminer les $(n+1)$ constantes $\mu_0, \dots, \mu_n$, et l'on trouve :

$$R_n(f)(x) = \sum_{k=0}^{n} \frac{\langle f,\, P_k\rangle_w}{\langle P_k,\, P_k\rangle_w}\, P_k(x),$$

ce qui conclut. □

Maintenant, qu'en est-il de la *complétude* de ces espaces L^2 à fonction-poids w ? Le cas où l'intervalle d'intégration est borné :

$$-\infty < a < b < \infty$$

est très accessible, tandis que le cas non borné, non traité ici, est plus délicat.

Théorème 8.8. *Si l'intervalle ouvert* $]a, b[$ *est* borné*, alors pour un poids continu quelconque :*

$$w\colon \quad]a, b[\longrightarrow]0, \infty[$$

dont tous les moments d'ordre $n \in \mathbb{N}$ *sont finis :*

$$\int_a^b |x|^n \, w(x)\, dx \;<\; \infty,$$

la suite unique de polynômes orthogonaux unitaires $\big(P_n\big)_{n=0}^{\infty}$ *de degré* n *pour le produit scalaire* $\langle \cdot, \cdot \rangle_w$ *est toujours une* base hilbertienne *de* $\mathscr{H}_w$.

Notons que la notion de base hilbertienne a un sens dans tout espace préhilbertien, sans en supposer la complétude.

Démonstration. Il s'agit d'établir que la famille des polygones orthogonaux $\big(P_n\big)_{n=0}^{\infty}$ est *totale*.

Supposons d'abord que $f \in \mathscr{C}^0([a, b])$ est continue sur l'intervalle *fermé* $[a, b]$. Notons $Q_n(f)$ le polynôme de meilleure approximation de degré $\leqslant n$ de f, pas pour la norme $\|\cdot\|_w$, mais pour la norme du sup :

$$\big\|f\big\|_{\mathscr{C}^0([a,b])} := \sup_{x \in [a,b]} \big|f(x)\big|.$$

Le théorème de Weierstrass stipule que :

$$0 = \lim_{n \to \infty} \big\|f - Q_n(f)\big\|_{\mathscr{C}^0([a,b])}.$$

Alors par l'inégalité évidente :

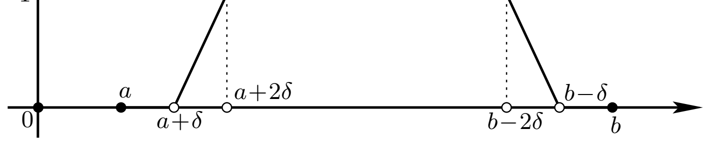
$$\begin{aligned}\bigg(\int_a^b \big|f(x) - Q_n(f)(x)\big|^2 \, w(x)\, dx\bigg)^{\frac{1}{2}} &\leqslant \max_{x \in [a,b]} \big|f(x) - Q_n(f)(x)\big| \,\bigg(\int_a^b w(x)\, dx\bigg)^{\frac{1}{2}} \\ &= \big\|f - Q_n(f)\big\|_{\mathscr{C}^0([a,b])} \underbrace{\bigg(\int_a^b w(x)\, dx\bigg)^{\frac{1}{2}}}_{\text{quantité finie}},\end{aligned}$$

on déduit instantanément que l'on a aussi :

$$0 = \lim_{n \to \infty} \big\|f - Q_n(f)\big\|_w.$$

Dans le cas général, rappelons que f est seulement supposée continue sur l'intervalle *ouvert* $]a, b[$. Soit alors pour $\delta > 0$ petit avec $\delta \ll b - a$, la fonction affine par morceaux

$$\chi_\delta(x)\colon \quad \mathbb{R} \to [0, 1]$$



dont on trouvera l'expression par formules (exercice). La différence :

$$f - f\,\chi_\delta$$

vaut 0 près de a et près de b. En utilisant :

$$0 \leqslant 1 - \chi_\delta \leqslant 1,$$

On estime alors :

$$\begin{aligned}\Big(\big\|f - f\,\chi_\delta\big\|_w\Big)^2 &= \int_{a+\delta}^{a+2\delta} \big|f(1-\chi_\delta)(x)\big|^2\,w(x)\,dx + \int_{b-2\delta}^{b-\delta} \big|f(1-\chi_\delta)(x)\big|^2\,w(x)\,dx \\ &\leqslant \int_{a+\delta}^{a+2\delta} \big|f\big|^2\,w + \int_{b-2\delta}^{b-\delta} \big|f\big|^2\,w,\end{aligned}$$

et ce dernier majorant tend vers 0 lorsque $\delta \to 0$, parce que :

$$\int_a^b \big|f\big|^2\,w \,<\, \infty,$$

et parce que :

$$\mathsf{mesure}\bigg([a+\delta,\ a+2\delta] \bigcup [b-2\delta,\ b-\delta]\bigg) = 2\,\delta \,\longrightarrow\, 0;$$

Par conséquent :

$$\forall\,\varepsilon > 0, \qquad \exists\,\delta > 0, \qquad \big\|f - f\chi_\delta\big\|_w \,\leqslant\, \varepsilon.$$

Ensuite, avec $\varepsilon > 0$ arbitrairement petit, puisque la fonction $f\chi_\delta$ est continue sur l'intervalle fermé $[a,b]$, la première partie de la démonstration assure qu'il existe un polynôme :

$$R \,\in\, \mathscr{P}_{N(\varepsilon)}$$

de degré $\leqslant N(\varepsilon)$ assez grand pour que :

$$\big\|f\chi_\delta - R\big\|_w \,\leqslant\, \varepsilon.$$

Alors l'inégalité du triangle :

$$\begin{aligned}\big\|f - R\big\|_w &\leqslant \big\|f - f\chi_\delta\big\|_w + \big\|f\chi_\delta - R\big\|_w \\ &\leqslant \varepsilon + \varepsilon,\end{aligned}$$

montre que l'on peut toujours approximer f en norme $\|\cdot\|_w$ avec une précision arbitraire par des polynômes, ce qui veut dire que la famille :

$$\big(1,\ x,\ x^2,\ \dots\big) \,=\, \big(P_0,\ P_1,\ P_2,\ \dots\big)$$

est bien *totale* dans $\big(\mathscr{H}_w,\ \|\cdot\|_w\big)$. □

Toutefois, malgré la beauté du Théorème 8.8, lorsque l'intervalle $]a,b[$ n'est *pas* borné, la suite de polynômes orthogonaux $(P_n)_{n=0}^{\infty}$ construits généralement n'est pas toujours totale.

Exemple 8.9. Sur l'intervalle $]0,\infty[$, soit le poids :

$$w(x) \,:=\, x^{-\log x} \,=\, e^{-\log x\,\log x}.$$

Nous affirmons que la fonction non identiquement nulle :

$$f(x) \,:=\, \sin\big(2\pi\log x\big)$$

est orthogonale à tous les monômes x^n avec $n \geqslant 0$ entier, donc à tous les P_n : elle n'est alors approximable à $\varepsilon > 0$ arbitrairement petit près par aucune combinaison linéaire des P_n.

En effet, le changement de variable $y = \log x$, d'où $e^y\,dy = dx$, permet tout d'abord de vérifier que tous les moments d'ordre $n \geqslant 0$ de la fonction f :

$$\int_0^\infty x^n\,x^{-\log x}\,dx \;=\; \int_{-\infty}^\infty e^{ny}\,e^{-y^2}\,e^y\,dy \;<\; \infty$$

sont finis, grâce à la petitesse écrasante de e^{-y^2} lorsque $|y| \longrightarrow \infty$.

Ensuite, ce même changement de variable permet de calculer comme suit le produit scalaire :

$$\begin{aligned}
\langle x^n, f\rangle_w &= \int_{-\infty}^\infty e^{ny}\,\sin\big(2\pi\,y\big)\,e^{-y^2}\,e^y\,dy \\
&= \int_{-\infty}^\infty e^{(n+1)y-y^2}\,\sin\big(2\pi\,y\big)\,dy \\
&= \int_{-\infty}^\infty e^{-(y+\frac{n+1}{2})^2+\frac{(n+1)^2}{4}}\sin\big(2\pi\,y\big)\,dy \\
&= e^{\frac{(n+1)^2}{4}}\int_{-\infty}^\infty e^{-(y+\frac{n+1}{2})^2}\,\sin\big(2\pi\,y\big)\,dy \\
\text{[Changer } t := y+\tfrac{n+1}{2}\text{]} \qquad &= e^{\frac{(n+1)^2}{4}}\,(-1)^{n+1}\int_{-\infty}^\infty e^{-t^2}\,\sin\big(2\pi\,t\big)\,dt \\
&= 0,
\end{aligned}$$

pour confirmer, par simple imparité de $t \longmapsto \sin(2\pi\,t)$, toutes les orthogonalités affirmées.

En utilisant des résultats d'analyse complexe (non admis dans ce cours), on peut démontrer le résultat suivant.

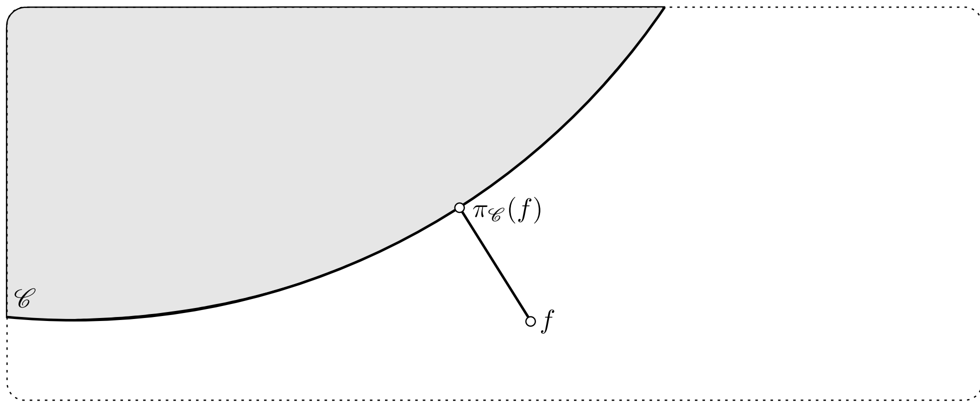
Théorème 8.10. *Sur un intervalle non borné* $]a,b[\, \subset \mathbb{R}$, *si la fonction-poids* $w\colon\,]a,b[\, \longrightarrow \mathbb{R}^+$ *satisfait pour une certaine constante* $\alpha > 0$ *:*

$$\int_a^b w(x)\,e^{\alpha\,|x|}\,dx \;<\; \infty,$$

alors la suite de polynômes orthogonaux $(P_n)_{n=0}^\infty$ *construits dans la Proposition 8.4 constitue une base hilbertienne de l'espace préhilbertien* $\mathscr{H}_w$, *qui est alors un espace de Hilbert en tant que tel.* □

9. Projection sur un convexe fermé et théorème de représentation de Riesz

Le théorème suivant joue un rôle très important dans les applications aux espaces fonctionnels : tout comme le théorème du point fixe, il affirme l'existence d'un unique élément vérifiant une certaine (in)égalité. Il sert notamment à établir l'existence de solutions pour des équations ou inéquations fonctionnelles.



Définition 9.1. Un sous-ensemble $\mathscr{C} \subset E$ d'un $\mathbb{R}$- ou $\mathbb{C}$-espace vectoriel E est dit *convexe* si le segment fermé défini par deux points quelconques lui appartenant :

$$\forall\, c_1 \in \mathscr{C} \quad \forall\, c_2 \in \mathscr{C} \quad \forall\, 0 \leqslant t \leqslant 1 \quad t\, c_1 + (1-t)\, c_2 \in \mathscr{C}$$

est encore entièrement contenu en lui.

Le théorème suivant, qui requiert réellement que $\mathscr{C}$ soit *fermé*, sera principalement appliqué au cas où $\mathscr{C}$ est un sous-espace vectoriel fermé.

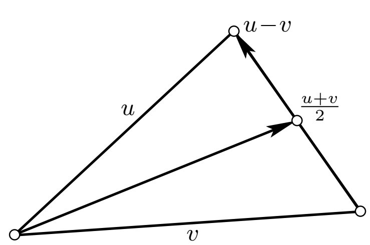
Théorème 9.2. [Projection orthogonale sur un convexe fermé] *Soit H un espace de Hilbert quelconque (pas nécessairement séparable) et soit $\mathscr{C}$ une partie* convexe *non vide de H qui est de plus* fermée. *Alors pour tout $f \in H$, il existe un unique point de $\mathscr{C}$, appelé* projection de f sur $\mathscr{C}$ *et noté $\pi_{\mathscr{C}}(f)$, dont la distance à f soit minimale :*

$$\|f - \pi_{\mathscr{C}}(f)\| = \min_{g \in \mathscr{C}} \|f - g\|.$$

Démonstration. En se souvenant du Théorème 3.11, ou en développant les carrés scalaires du membre de droite, le lecteur vérifiera facilement la *formule de la médiane :*

$$\|u\|^2 + \|v\|^2 = 2\Big\|\tfrac{u+v}{2}\Big\|^2 + \tfrac{1}{2}\,\|u - v\|^2 \qquad\qquad (u,\, v \in H),$$

qui remonte à Pythagore et à Euclide



Posons maintenant :

$$d := \inf\big\{\|f - g\| \colon g \in \mathscr{C}\big\},$$

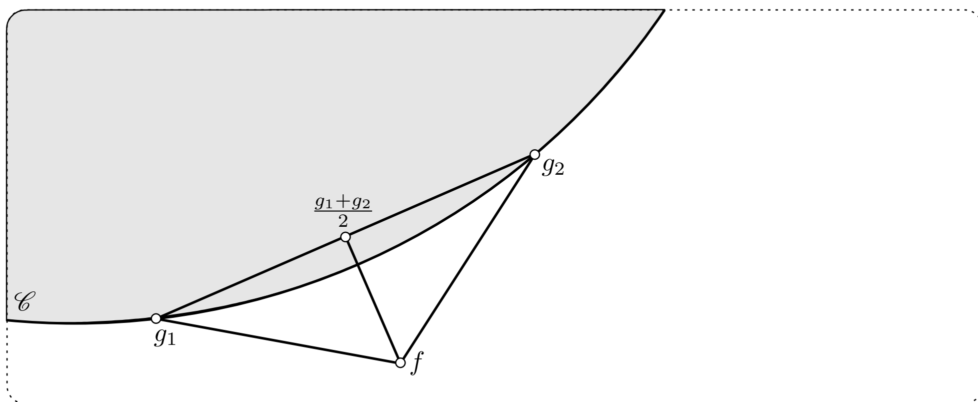
et établissons l'*unicité* de $\pi_{\mathscr{C}}(f)$ (l'existence viendra ensuite). Par contradiction, s'il existait *deux* éléments *distincts* g_1 et g_2 de $\mathscr{C}$ réalisant cette borne inférieure :

$$d = \|f - g_1\| = \|f - g_2\|,$$

alors leur milieu $\frac{g_1+g_2}{2}$, qui appartiendrait aussi à l'ensemble *convexe* $\mathscr{C}$, satisferait, *via* l'identité de la médiane appliquée à $u := f - g_1$ et à $v := f - g_2$:

$$\begin{aligned} 2\Big\| f - \tfrac{g_1+g_2}{2} \Big\|^2 &= \|f - g_1\|^2 + \|f - g_2\|^2 - \tfrac{1}{2}\,\|g_2 - g_1\|^2 \\ &= 2d^2 - \tfrac{1}{2}\,\|g_2 - g_1\|^2 < 2d^2, \end{aligned}$$

et donc le vecteur médian $\frac{g_1+g_2}{2} \in \mathscr{C}$ satisferait $\|f - \frac{g_1+g_2}{2}\| < d$, ce qui contredirait manifestement la définition de l'infimum d.



Géométriquement parlant, nous avons seulement utilisé le fait que, dans un triangle ABC isocèle en A, la hauteur AH issue de A est strictement inférieure à la longueur isocèle $AB = AC$.

À présent, établissons l'existence. Par définition de toute borne inférieure telle que d, il existe une suite $(g_k)_{k=1}^{\infty}$ d'éléments de $\mathscr{C}$ telle que $\|f - g_k\|$ tende vers d lorsque k tend vers ∞. Alors nous allons voir que la convexité de $\mathscr{C}$ produit une espèce de « miracle » qui impose à toute telle suite d'être automatiquement de Cauchy.

En effet, si l'on introduit à nouveau le milieu $\frac{g_{k_1}+g_{k_2}}{2}$ entre deux éléments quelconques g_{k_1} et g_{k_2} d'une telle suite, on peut à nouveau appliquer l'identité de la médiane à $u := f - g_{k_1}$ et à $v := f - g_{k_2}$:

$$\tfrac{1}{2}\,\|g_{k_1} - g_{k_2}\|^2 = \|f - g_{k_1}\|^2 + \|f - g_{k_2}\|^2 - 2\Big\|\tfrac{g_{k_1}+g_{k_2}}{2} - f\Big\|^2.$$

Mais puisque ce milieu $\frac{g_{k_1}+g_{k_2}}{2}$ appartient encore au convexe $\mathscr{C}$, sa distance $\|\frac{g_{k_1}+g_{k_2}}{2} - f\|$ à f est automatiquement $\geqslant d$, donc on en déduit une inégalité :

$$\tfrac{1}{2}\,\|g_{k_1} - g_{k_2}\|^2 \leqslant \|f - g_{k_1}\|^2 + \|f - g_{k_2}\|^2 - 2\,d^2$$

qui montre que $\|g_{k_1} - g_{k_2}\|$ peut être rendu arbitrairement petit, pourvu que $k_1, k_2 \geqslant K$ soient assez grands, puisque l'on a par hypothèse :

$$\|f - g_{k_1}\|^2 \underset{k_1\to\infty}{\longrightarrow} d^2 \qquad\qquad \text{et} \qquad\qquad \|f - g_{k_2}\|^2 \underset{k_2\to\infty}{\longrightarrow} d^2.$$

Ainsi, la suite $(g_k)_{k=1}^{\infty}$ est bien de Cauchy. Mais l'espace de Hilbert H est complet, donc il existe un unique élément $g_\infty \in H$ qui est la limite des g_k lorsque k tend vers l'infini. Et comme $\mathscr{C}$ a été supposé fermé, cette limite g_∞ appartient automatiquement à $\mathscr{C}$, ce qu'il fallait démontrer. $\square$

Théorème 9.3. [Hyperplan de support] *Sous la même hypothèse que H est un espace de Hilbert quelconque (pas forcément séparable) et que $\mathscr{C}$ est une partie convexe fermée non vide de H, le projeté $\pi_{\mathscr{C}}(f)$ sur $\mathscr{C}$ d'un élément quelconque $f \in H$ vérifie les inégalités :*

$$\mathsf{Re}\left\langle f - \pi_{\mathscr{C}}(f),\, h - \pi_{\mathscr{C}}(f)\right\rangle \leqslant 0, \qquad \textit{pour tout } \ h \in \mathscr{C}.$$

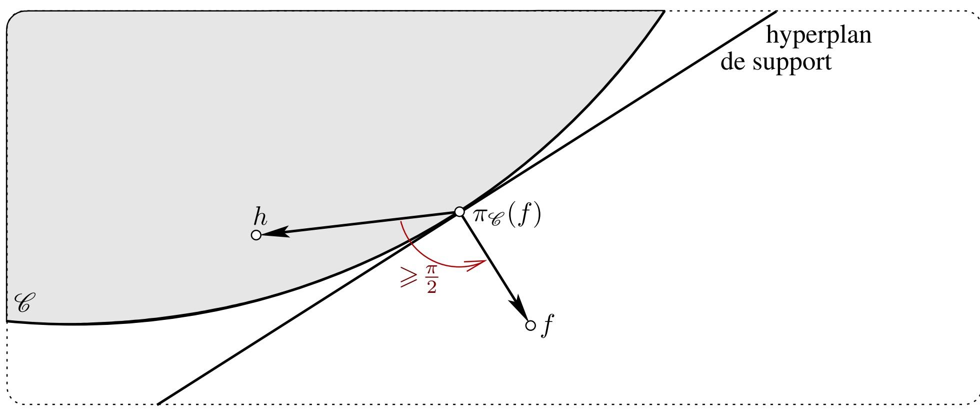
Réciproquement, si $g \in \mathscr{C}$ est un point du convexe fermé qui vérifie les mêmes inégalités :

$$\mathsf{Re}\big\langle f - g,\, h - g\big\rangle \leqslant 0 \qquad \textit{pour tout } \ h \in \mathscr{C},$$

alors nécessairement g est unique et n'est autre que le projeté de f :

$$g = \pi_{\mathscr{C}}(f).$$

Lorsque l'espace de Hilbert H est défini sur $\mathbb{R}$, on enlève les parties réelles 'Re'.



Géométriquement parlant et si on l'interprète dans un espace hilbertien réel, cette dernière caractérisation exprime que l'angle entre le vecteur $f - \pi_{\mathscr{C}}(f)$ et tout autre vecteur $h - \pi_{\mathscr{C}}(f)$ d'extrémité un autre élément quelconque $h \in \mathscr{C}$ est toujours $\geqslant \frac{\pi}{2}$.

Autrement dit, tous les points h du convexe fermé $\mathscr{C}$ se situent dans le demi-espace fermé :

$$\Big\{h \in H \colon \mathsf{Re}\,\langle f - \pi_{\mathscr{C}}(f),\, h - \pi_{\mathscr{C}}(f)\rangle \leqslant 0\Big\}$$

dont le bord, défini par l'*équation* :

$$\mathsf{Re}\,\langle f - \pi_{\mathscr{C}}(f),\, h - \pi_{\mathscr{C}}(f)\rangle = 0,$$

est lui-même un *hyperplan* réel passant par le point projeté $\pi_{\mathscr{C}}(f) \in \mathscr{C}$, hyperplan qui est orthogonal au vecteur $f - \pi_{\mathscr{C}}(f)$. Ainsi, le projeté sur le convexe fermé $\mathscr{C}$ s'identifie à un simple projeté sur cet hyperplan réel. On nomme parfois « *hyperplan de support* » tout hyperplan qui garantit qu'un objet géométrique se situe entièrement dans un, et dans un seul des deux côtés (fermés) qu'il définit.

Démonstration. Soit donc $h \in \mathscr{C}$ quelconque. Pour tout réel t avec $0 \leqslant t \leqslant 1$, le point :

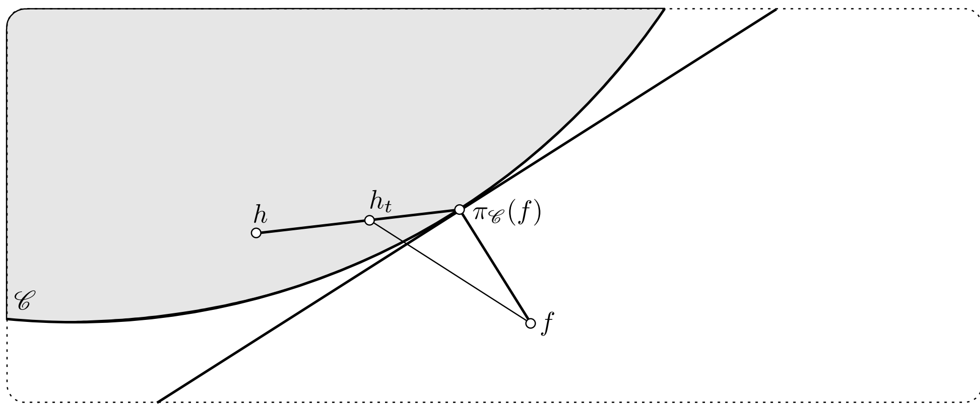
$$h_t := \pi_{\mathscr{C}}(f) + t\big(h - \pi_{\mathscr{C}}(f)\big)$$

appartient aussi au convexe fermé $\mathscr{C}$. On a donc une inégalité automatique que l'on peut développer :

$$\begin{aligned}
\|\pi_{\mathscr{C}}(f) - f\|^2 &\leqslant \|h_t - f\|^2 \\
&= \big\|\pi_{\mathscr{C}}(f) - f + t(h - \pi_{\mathscr{C}}(f))\big\|^2 \\
&= \|\pi_{\mathscr{C}}(f) - f\|^2 + 2t\,\mathsf{Re}\,\langle \pi_{\mathscr{C}}(f) - f,\ h - \pi_{\mathscr{C}}(f)\rangle + t^2\,\|h - \pi_{\mathscr{C}}(f)\|^2,
\end{aligned}$$

et après simplification évidente, cette inégalité se réduit à :

$$0 \leqslant 2t\,\mathsf{Re}\,\langle \pi_{\mathscr{C}}(f) - f,\ h - \pi_{\mathscr{C}}(f)\rangle + t^2\,\|h - \pi_{\mathscr{C}}(f)\|^2.$$



Avec $a, b \in \mathbb{R}$ et $a < 0$, nous savons que $at + bt^2 \sim at < 0$ est négatif pour $0 < t \leqslant \varepsilon$ petit. En faisant aussi $0 < t \leqslant \varepsilon$ ci-dessus, on déduit donc que le coefficient de t ci-dessus ne peut *pas* être < 0, donc est $\geqslant 0$, ce qui établit l'inégalité voulue (noter le changement de signe) :

$$\mathsf{Re}\,\langle f - \pi_{\mathscr{C}}(f),\ h - \pi_{\mathscr{C}}(f)\rangle \leqslant 0,$$

valable pour tout $h \in \mathscr{C}$.

Il reste à établir l'unicité d'un élément $g \in \mathscr{C}$ satisfaisant :

$$\mathsf{Re}\,\langle f - g,\, h - g\rangle \leqslant 0,$$

pour tout $h \in \mathscr{C}$. Or si g satisfait de telles inégalités, nous pouvons estimer la distance à f d'un élément quelconque $h \in \mathscr{C}$ en y insérant $-g + g$:

$$\begin{aligned}
\|h - f\|^2 = \|h - g + g - f\|^2 &= \underbrace{\|h - g\|^2}_{\geqslant 0} + \underbrace{2\mathsf{Re}\,\langle h - g,\, g - f\rangle}_{\geqslant 0} + \|g - f\|^2 \\
&\geqslant \|g - f\|^2,
\end{aligned}$$

et déduire qu'elle est toujours supérieure ou égale à la distance de f à g. Or d'après le Théorème 9.2 qui précède, un tel élément g existe, est unique, et n'est autre que $g = \pi_{\mathscr{C}}(f)$, comme annoncé. □

Définition 9.4. [Orthogonal à une partie quelconque] Dans un espace de Hilbert H quelconque, l'orthogonal d'un sous-ensemble $B \subset H$ arbitraire est défini en toute généralité comme étant constitué des vecteurs qui sont orthogonaux à *tous* les éléments de B :

$$B^\perp := \big\{h \in H\colon\ \langle h, b\rangle = 0\ \ \forall\, b \in B\big\}.$$

C'est un sous-espace vectoriel de H (exercice). On notera que $B_1 \subset B_2$ implique $B_1^\perp \supset B_2^\perp$.

Lemme 9.5. [Trivial mais important] *Soit B une partie arbitraire d'un espace de Hilbert H quelconque. Alors son orthogonal :*

$$B^{\perp} = \big\{h \in H\colon \ \langle h, b\rangle = 0 \ \ \forall\, b \in B\big\}$$

est un sous-espace vectoriel de H qui est toujours fermé *dans H.*

Démonstration. En effet, si $(h_k)_{k=1}^{\infty}$ est une suite de Cauchy quelconque dans $B^{\perp}$ satisfaisant donc $\langle h_k, b\rangle = 0$, pour tout b fixé, l'inégalité de Cauchy-Schwarz assure (exercice direct laissé au lecteur) que, si l'on note h_∞ son unique limite dans H (qui est complet), on a encore $\langle h_\infty, b\rangle = 0$, ce qui montre bien que $h_\infty \in B^{\perp}$. □

Proposition 9.6. [Supplémentaire orthogonal] *Soient H un espace de Hilbert et F un sous-espace vectoriel de H qui est* fermé. *Alors son orthogonal :*

$$F^{\perp} = \big\{g \in H\colon \langle f, g\rangle = 0 \ \ \forall\, f \in F\big\}.$$

est un sous-espace vectoriel de H qui est lui aussi fermé, appelé supplémentaire orthogonal *de F. Surtout, tout élément $h \in H$ se décompose de manière unique sous la forme :*

$$h = f + g \quad \textit{avec} \quad f \in F \quad \textit{et} \quad g \in F^{\perp}.$$

En outre, les éléments f et g de cette décomposition sont orthogonaux entre eux et sont les projections (orthogonales) de h sur les deux ensembles convexes fermés F et sur $F^{\perp}$, respectivement, à savoir l'on peut écrire :

$$H = F \overset{\perp}{\oplus} F^{\perp}, \qquad F \cap F^{\perp} = \{0\}.$$

Démonstration. Observons que tout sous-espace vectoriel fermé de H est, en particulier, un sous-ensemble convexe fermé de H.

Soit donc $h \in H$ et notons comme précédemment $\pi_F(h)$ sa projection sur le fermé convexe F (cette projection existe grâce au Théorème 9.2). Pour tout élément $f' \in F$ et tout scalaire $\lambda \in \mathbb{C}$, le point $\pi_F(h) + \lambda\, f'$ appartient à F et on a donc d'après le Théorème 9.3 ci-dessus :

$$\mathsf{Re}\,\langle h - \pi_F(h),\ \pi_F(h) + \lambda\, f' - \pi_F(h)\rangle \leqslant 0,$$

à savoir :

$$\mathsf{Re}\,\langle h - \pi_F(h),\ \lambda\, f'\rangle \leqslant 0.$$

Or un nombre complexe z qui vérifie $\mathsf{Re}\,\big(\overline{\lambda}\, z\big) \leqslant 0$ pour tout $\lambda \in \mathbb{C}$ est nécessairement nul (exercice élémentaire). Donc il vient :

$$\langle h - \pi_F(h),\ f'\rangle = 0,$$

et ce, pour tout $f' \in F$. Ceci montre que $h - \pi_F(h)$ appartient à $F^{\perp}$ et prouve donc l'existence de la décomposition :

$$h = \underbrace{\pi_F(h)}_{=:\, f \in F} + \underbrace{h - \pi_F(h)}_{=:\, g \in F^{\perp}}.$$

Quant à l'unicité, s'il existait une autre décomposition $h = f' + g'$ du même type avec $f' \in F$ et $g' \in F^{\perp}$, le vecteur $u := f' - f = g' - g$ appartiendrait à la fois à F et à $F^{\perp}$. Or si $u \in F$ appartient aussi à $F^{\perp}$, on doit avoir $\langle u, u\rangle = 0$, donc $u = 0$, et enfin $f' = f$ et $g' = g$. La preuve est terminée. □

Corollaire 9.7. *Comme à l'instant, soit H un espace de Hilbert quelconque et soit F un sous-espace vectoriel de H qui est fermé. Alors les deux conditions suivantes sont équivalentes :*

(i) $F = H$;

(ii) $F^{\perp} = \{0\}$.

Démonstration. Cette équivalence découle par simple logique (exercice mental) de la décomposition orthogonale qui précède $H = F \overset{\perp}{\oplus} F^{\perp}$ et de son unicité. $\square$

Lemme 9.8. *Soit B une partie quelconque d'un espace de Hilbert H. Alors l'orthogonal de B coïncide avec l'orthogonal de son adhérence $\overline{B}$:*

$$B^{\perp} = \overline{B}^{\perp}.$$

Démonstration. En effet, l'inclusion $B \subset \overline{B}$ donne immédiatement par retour à la définition :

$$\overline{B}^{\perp} \subset B^{\perp}.$$

Pour établir l'inclusion inverse $B^{\perp} \subset \overline{B}^{\perp}$, étant donné un vecteur quelconque $h \in B^{\perp}$, à savoir $\langle h, b \rangle = 0$ pour tout $b \in B$, on doit démontrer que $h \in \overline{B}^{\perp}$, à savoir que $\langle h, \overline{b} \rangle = 0$ pour tout $\overline{b} \in \overline{B}$. Si donc $(b_k)_{k=1}^{\infty}$ est une suite de Cauchy d'éléments de B qui converge vers un élément arbitraire $\overline{b} \in \overline{B}$ dans H complet :

$$\forall\, \varepsilon > 0 \quad \exists\, K \geqslant 1 \quad \Big(k \geqslant K \implies \| \overline{b} - b_k \| \leqslant \varepsilon \Big),$$

l'inégalité de Cauchy-Schwarz (à nouveau elle !) nous assure que le nombre réel positif :

$$\begin{aligned} \big| \langle h, \overline{b} \rangle \big| &= \big| \langle h, \overline{b} \rangle - \langle h, b_k \rangle \big| \\ &\leqslant \|h\| \, \big\| \overline{b} - b_k \big\| \\ &\leqslant \|h\| \, \varepsilon \end{aligned}$$

peut être rendu arbitrairement petit, donc h est bien orthogonal aussi à tous les éléments $\overline{b} \in \overline{B}$. $\square$

Corollaire 9.9. [Critère de totalité] *Soit H un espace de Hilbert quelconque et soit $A \subset H$ une partie quelconque. Alors les trois conditions suivantes sont équivalentes :*

(i) *A est totale ;*

(ii) *l'espace vectoriel $\mathsf{Vect}_{\mathsf{fini}}(A)$ constitué des combinaisons linéaires finies d'éléments de A à coefficients complexes possède une adhérence :*

$$\overline{\mathsf{Vect}_{\mathsf{fini}}(A)} = H$$

égale à tout l'espace de Hilbert ambient ;

(iii) *l'ensemble des vecteurs orthogonaux à tous les éléments de A se réduit au vecteur nul, à savoir :*

$$\{0\} = A^{\perp} = \big\{ h \in H \colon \ \langle h, a \rangle = 0 \ \forall\, a \in A \big\}.$$

Démonstration. L'équivalence entre **(i)** et **(ii)** est une reformulation de la notion de totalité telle qu'elle a été définie au début de ce cours sur les espaces de Hilbert (exercice mental).

Montrons l'équivalence entre **(ii)** et **(iii)**. Par linéarité du produit scalaire, il est clair que l'orthogonal de A coïncide avec l'ortogonal de ses combinaisons linéaires finies :

$$A^\perp = \big[\mathsf{Vect}_{\mathsf{fini}}(A)\big]^\perp.$$

Mais nous venons de voir que prendre l'orthogonal d'une partie B quelconque revient à prendre l'orthogonal de sa fermeture $\overline{B}$, donc :

$$A^\perp = \big[\mathsf{Vect}_{\mathsf{fini}}(A)\big]^\perp = \big[\overline{\mathsf{Vect}_{\mathsf{fini}}(A)}\big]^\perp.$$

Le Corollaire 9.7 vu il y a quelques instants appliqué à l'espace vectoriel fermé :

$$F := \overline{\mathsf{Vect}_{\mathsf{fini}}(A)}$$

disait alors que $F = H$ si et seulement si $F^\perp = \{0\}$, ce qui nous donne l'équivalence désirée. □

Théorème 9.10. [Théorème de représentation de Riesz] *Soit H un espace de Hilbert quelconque. À tout vecteur fixé $g_0 \in H$, on peut faire correspondre la forme linéaire continue :*

$$\begin{aligned} L_{g_0}\colon \quad f &\longmapsto \langle f, g_0\rangle \\ H &\longrightarrow \mathbb{C} \end{aligned}$$

« produit scalaire avec g_0 ».

Réciproquement, étant donné une forme linéaire continue arbitraire L sur H, il existe un et un seul vecteur $g_0 \in H$ tel que l'on ait :

$$L = L_{g_0} = \langle \cdot, g_0\rangle.$$

Démonstration. Vérifions tout d'abord que L_{g_0} est une forme linéaire continue. Pour tout $f \in H$, Cauchy-Schwarz donne :

$$\big|L_{g_0}(f)\big| = \big|\langle f, g_0\rangle\big| \leqslant \underbrace{\|g_0\|_H}_{\text{constante}} \, \|f\|_H.$$

Ceci prouve que L_{g_0} est continue, et de plus, que sa norme d'opérateur satisfait :

$$|\!|\!| L_{g_0} |\!|\!| = \sup_{f\neq 0} \frac{|L_{g_0}(f)|}{\|f\|_H} \leqslant \|g_0\|_H.$$

Qui plus est, en considérant $L_{g_0}(g_0)$, on constate facilement que $|\!|\!| L_{g_0} |\!|\!| = \|g_0\|_H$.

Soit maintenant L une forme linéaire continue sur H non identiquement nulle. Le sous-espace vectoriel noyau :

$$F := L^{-1}(0)$$

est fermé (puisque L est continue) et distinct de H (puisque $L \not\equiv 0$). Grâce au Corollaire 9.7, l'orthogonal $F^\perp$ n'est donc pas réduit à $\{0\}$ et il existe par conséquent un élément non nul :

$$g \in F^\perp \backslash \{0\}.$$

Cet élément n'appartient alors pas à F, donc le nombre complexe :

$$\lambda := L(g)$$

est non nul.

Ensuite, pour tout $f \in H$ on peut poser :

$$f = \tfrac{L(f)}{L(g)}\, g + \Big(f - \tfrac{L(f)}{L(g)}\, g\Big) =: f_1 + f_2.$$

On remarque que le second terme f_2, qui vérifie $L(f_2) = 0$, appartient à $F = L^{-1}(0)$ tandis que f_1, qui est un multiple de g, appartient à $F^{\perp}$. En effectuant le produit scalaire avec g, on obtient donc :

$$\langle f,\, g\rangle = \langle f_1, g\rangle + \underbrace{\langle f_2, g\rangle}_{\circ} = \tfrac{L(f)}{L(g)}\, \|g\|^2 + 0.$$

Il suffit maintenant de poser $g_0 := \frac{\overline{L(g)}}{\|g\|^2}\, g$ pour obtenir $L(f) = \langle f, g_0\rangle$ quel que soit f. Cela montre que L est égale à L_{g_0}.

Enfin, l'unicité (facile) est laissée en exercice à tout(e) étudiant(e) consciencieux(se) qui souhaite réussir à l'examen. □

Pour terminer ce chapitre, discutons brièvement les notions d'opérateurs linéaires et de leurs adjoints. Soient H_1 et H_2 deux $\mathbb{C}$-espaces de Hilbert.

Définition 9.11. Une application $T\colon H_1 \longrightarrow H_2$ est appelée un *opérateur linéaire*, ou une *transformation linéaire*, lorsqu'elle est $\mathbb{C}$-linéaire :

$$T\big(\lambda\, f + \mu\, g\big) \;=\; \lambda\, T(f) + \mu\, T(g) \qquad\qquad (\forall\, \lambda, \mu \in \mathbb{C},\, \forall\, f, g \in H_1),$$

d'où $T(0) = 0$.

Définition 9.12. On dit qu'un opérateur linéaire $T\colon H_1 \longrightarrow H_2$ est *borné* lorsqu'il existe une constante $0 \leqslant C < \infty$ telle que :

$$\big\|T(f)\big\|_{H_2} \;\leqslant\; C\, \|f\|_{H_1} \qquad\qquad (\forall\, f \in H_1).$$

La *norme d'opérateur* de T, l'infimum de telles constantes C, vaut alors :

$$|\!|\!|T|\!|\!| \;:=\; \sup_{\|f\|_{H_1} \leqslant 1} \big\|T(f)\big\|_{H_2}.$$

On a alors une majoration fondamentale :

$$\big\|T(f)\big\|_{H_2} \;\leqslant\; |\!|\!|T|\!|\!| \cdot \|f\|_{H_1} \qquad\qquad (\forall\, f \in H_1).$$

Avec $H_1 = H_2 =: H$, un exemple trivial est l'opérateur identité Id de norme $|\!|\!|\mathsf{Id}|\!|\!| = 1$.

Lemme 9.13. *La norme d'un opérateur linéaire borné $T\colon H_1 \longrightarrow H_2$ vaut aussi :*

$$|\!|\!|T|\!|\!| \;=\; \sup\Big\{\big|\langle T(f), g\rangle_{H_2}\big| \colon\; \|f\|_{H_1} \leqslant 1,\; \|g\|_{H_2} \leqslant 1\Big\}.$$

Démonstration. Si $|\!|\!|T|\!|\!| \leqslant C$ pour une constante $0 \leqslant C < \infty$, l'inégalité de Cauchy-Schwarz donne, pour tout $\|f\|_{H_1} \leqslant 1$ et tout $\|g\|_{H_2} \leqslant 1$:

$$\big|\langle T(f), g\rangle_{H_2}\big| \;\leqslant\; \|T(f)\|_{H_2} \cdot \|g\|_{H_2} \;\leqslant\; C\, \|f\|_{H_1}\, \|g\|_{H_2} \;\leqslant\; C,$$

et ainsi, en faisant tendre $|\!|\!|T|\!|\!| \xleftarrow{<} C$:

$$\sup\Big\{\big|\langle T(f), g\rangle_{H_2}\big| \colon\; \|f\|_{H_1} \leqslant 1,\; \|g\|_{H_2} \leqslant 1\Big\} \;\leqslant\; |\!|\!|T|\!|\!|.$$

Pour l'inégalité inverse, faisons l'hypothèse, pour une constante $0 \leqslant C < \infty$, que :

$$\sup\Big\{\big|\langle T(f), g\rangle_{H_2}\big| \colon\; \|f\|_{H_1} \leqslant 1,\; \|g\|_{H_2} \leqslant 1\Big\} \;\leqslant\; C,$$

et montrons que $\|T(f)\|_{H_2} \leqslant C\, \|f\|_{H_1}$ pour tout $f \in H_1$, d'où découlera $|\!|\!|T|\!|\!| \leqslant C$, ce qui donnera bien en faisant tendre C vers ce supremum :

$$|\!|\!|T|\!|\!| \;\leqslant\; \sup\Big\{\big|\langle T(f), g\rangle_{H_2}\big| \colon\; \|f\|_{H_1} \leqslant 1,\; \|g\|_{H_2} \leqslant 1\Big\}.$$

Lorsque $f = 0$ ou $T(f) = 0$, il n'y a rien à faire.

Supposons donc que $f \neq 0$ et que $T(f) \neq 0$, introduisons alors les deux éléments de norme 1 :

$$f' := \frac{f}{\|f\|_{H_1}} \qquad \text{et} \qquad g' := \frac{T(f)}{\|T(f)\|_{H_2}},$$

et appliquons-leur l'hypothèse :

$$\begin{aligned} C &\geqslant \big|\langle T(f'), g'\rangle_{H_2}\big| \\ &= \frac{|\langle T(f), T(f)\rangle_{H_2}|}{\|f\|_{H_1}\,\|T(f)\|_{H_2}}, \end{aligned}$$

ce qui donne $\|T(f)\|_{H_2} \leqslant C\,\|f\|_{H_1}$, conclusion ! □

Définition 9.14. Une transformation linéaire $T\colon H_1 \longrightarrow H_2$ entre espaces de Hilbert est dite *continue* si, pour toute suite $(f_n)_{n=1}^{\infty}$ d'éléments $f_n \in H_1$ qui converge vers un élément $f_\infty \in H_1$, on a aussi convergence dans H_2 :

$$T(f_n) \underset{n\to\infty}{\longrightarrow} T(f_\infty).$$

Proposition 9.15. *On a équivalence entre :*

$$\Big(T\colon\ H_1 \longrightarrow H_2 \text{ est borné}\Big) \iff \Big(T\colon\ H_1 \longrightarrow H_2 \text{ est continu}\Big).$$

Démonstration. Si T est borné, et si $f_n \underset{n\to\infty}{\longrightarrow} f_\infty$, il est clair que :

$$\big\|T(f_n) - T(f_\infty)\big\|_{H_2} \leqslant \|T\| \cdot \|f_n - f_\infty\|_{H_1} \underset{n\to\infty}{\longrightarrow} 0,$$

donc T est continu.

Réciproquement, si T est continu, supposons par l'absurde qu'il ne soit pas borné, c'est-à-dire que pour tout $n \geqslant 1$, il existe $f_n \in H_1\backslash\{0\}$ avec :

$$\big\|T(f_n)\big\|_{H_2} \geqslant n\,\|f_n\|_{H_1}.$$

Alors le vecteur :

$$f_n' := \frac{f_n}{n\,\|f_n\|_{H_1}}$$

est de norme $\|f_n'\|_{H_1} = \frac{1}{n}$, donc $f_n' \underset{n\to\infty}{\longrightarrow} 0$, et comme T est continu (en 0), on doit aussi avoir $T(f_n') \underset{n\to\infty}{\longrightarrow} 0$, ce qui n'est manifestement pas cohérent avec :

$$\big\|T(f_n')\big\|_{H_2} = \frac{\|T(f_n)\|_{H_2}}{n\,\|f_n\|_{H_1}} \geqslant 1.$$

□

L'application la plus immédiate du Théorème 9.10 de représentation de Riesz est la construction, pour tout opérateur linéaire borné $T\colon H_1 \longrightarrow H_2$, d'un *adjoint* $H_1 \longleftarrow H_2 : T^*$.

Théorème 9.16. [Adjoint] *Si $T\colon H_1 \longrightarrow H_2$ est un opérateur linéaire borné (continu) entre espaces de Hilbert, il existe un unique opérateur linéaire borné (continu) :*

$$H_1 \longleftarrow H_2 : \quad T^*,$$

appelé l'adjoint de T, satisfaisant :

$$\langle T(f), g\rangle_{H_2} = \langle f, T^*(g)\rangle_{H_1} \qquad (\forall\, f \in H_1,\ \forall\, g \in H_2).$$

De plus :

(1) $\|T^*\| = \|T\|$;
(2) $\big(T^*\big)^* = T$.

Démonstration. Pour montrer (simultanément) l'existence et l'unicité de T^*, observons que pour tout $g \in H_2$ fixé, l'application linéaire :

$$\begin{aligned} \Lambda_g : \quad H_1 &\longrightarrow \mathbb{C} \\ f &\longmapsto \langle T(f), g\rangle_{H_2} \end{aligned}$$

est *bornée*, parce que, comme T est borné, on peut estimer :

$$\begin{aligned} \big|\Lambda_g(f)\big| &= \big|\langle T(f), g\rangle_{H_2}\big| \\ \text{[Cauchy-Schwarz]} \qquad &\leqslant \big\|T(f)\big\|_{H_2} \cdot \|g\|_{H_2} \\ &\leqslant \|T\| \cdot \|g\|_{H_2} \cdot \|f\|_{H_1} \\ &= \mathsf{constante} \cdot \|f\|_{H_1}. \end{aligned}$$

Le Théorème 9.10 (magique !) de représentation de Riesz fournit alors un unique élément $h \in H_2$ tel que $\Lambda_g(\cdot)$ coïncide avec le produit scalaire avec h :

$$\Lambda_g(f) = \langle f, h\rangle_{H_1} \qquad (\forall\, f \in H_1).$$

On se convainc aisément (exercice) que la dépendance de h vis-à-vis de g est alors $\mathbb{C}$-linéaire, ce qui justifie de noter $h = T^*(g)$, avec un certain opérateur *adjoint* T^* qui est $\mathbb{C}$-linéaire et qui satisfait bien au final l'identité fondamentale :

$$\langle T(f), g\rangle_{H_2} = \langle f, T^*(g)\rangle_{H_1} \qquad (\forall\, f \in H_1,\ \forall\, g \in H_2),$$

qu'il est avisé de lire en voyant T « *passer de l'autre côté du produit scalaire en cueillant une étoile* ».

La propriété **(1)** s'obtient instantanément comme suit grâce au Lemme 9.13 :

$$\begin{aligned} \|T\| &= \sup\Big\{\big|\langle T(f), g\rangle_{H_2}\big| :\ \|f\|_{H_1} \leqslant 1,\ \|g\|_{H_2} \leqslant 1\Big\} \\ &= \sup\Big\{\big|\langle f, T^*(g)\rangle_{H_2}\big| :\ \|f\|_{H_1} \leqslant 1,\ \|g\|_{H_2} \leqslant 1\Big\} \\ \text{[Conjuguer]} \qquad &= \sup\Big\{\big|\langle T^*(g), f\rangle_{H_2}\big| :\ \|f\|_{H_1} \leqslant 1,\ \|g\|_{H_2} \leqslant 1\Big\} \\ &= \|T^*\|. \end{aligned}$$

Enfin pour la propriété **(2)**, le bi-adjoint $(T^*)^*$ de T, *i.e.* l'adjoint de T^*, étant défini par :

$$\langle T^*(g), h\rangle_{H_1} = \langle g, (T^*)^*(h)\rangle_{H_2} \qquad (\forall\, g \in H_2,\ \forall\, h \in H_1),$$

ou, de manière équivalente après une simple conjugaison complexe, par :

$$\langle h, T^*(g)\rangle_{H_1} = \langle (T^*)^*(h), g\rangle_{H_2} \qquad (\forall\, g \in H_2,\ \forall\, h \in H_1),$$

une comparaison avec les identités qui définissent T^* donne :

$$\langle T(h), g\rangle_{H_2} = \langle (T^*)^*(h), g\rangle_{H_2} \qquad (\forall\, g \in H_2,\ \forall\, h \in H_1),$$

ce qui force $T(h) = (T^*)^*(h)$, d'où comme annoncé $T = (T^*)^*$. □

Une *identité de polarisation*, que l'on peut vérifier par un calcul direct, va s'avèrer utile dans un instant :

$$\begin{aligned} 4\,\langle T(f), g\rangle = \big\langle T(f+g), f+g\big\rangle - \big\langle T(f-g), f-g\big\rangle + \\ + i\,\big\langle T(f+ig), f+ig\big\rangle - i\,\big\langle T(f-ig), f-ig\big\rangle. \end{aligned}$$

Définition 9.17. Un opérateur linéaire $T\colon H \longrightarrow H$ d'un $\mathbb{C}$-espace de Hilbert H à valeurs dans lui-même est dit *auto-adjoint* lorsque :

$$T^* = T.$$

Proposition 9.18. *La norme d'opérateur $\|T\|$ d'un opérateur linéaire $T\colon H \longrightarrow H$ auto-adjoint $T^* = T$ peut être calculée comme :*

$$\|T\| = \sup\big\{\big|\langle T(f), f\rangle\big|\colon \|f\| = 1\big\}.$$

Ce résultat doit être comparé à la formule du Lemme 9.13, valable pour tout opérateur linéaire.

Démonstration. D'après ce lemme, on a tout d'abord une inégalité dans un sens facile :

$$\begin{aligned} N &:= \sup\big\{\big|\langle T(f), f\rangle\big|\colon \|f\|_H = 1\big\} \\ &\leqslant \sup\big\{\big|\langle T(f), g\rangle\big|\colon \|f\|_H = 1, \|g\|_H = 1\big\} \\ &= \|T\|. \end{aligned}$$

Pour l'inégalité inverse $N \geqslant \|T\|$ — qui requiert plus de travail —, observons tout d'abord que pour tout vecteur $h \in H$, grâce à $T^* = T$, la quantité :

$$\langle T(h), h\rangle = \langle h, T^*(h)\rangle = \langle h, T(h)\rangle = \overline{\langle T(h), h\rangle}$$

est *réelle*, et donc lorsqu'on prend la partie réelle dans l'identité de polarisation ci-dessus :

$$\mathrm{Re}\,\big[\langle T(f), g\rangle\big] = \frac{1}{4}\Big[\big\langle T(f+g), f+g\big\rangle - \big\langle T(f-g), f-g\big\rangle\Big] + 0 + 0,$$

deux termes s'évanouissent pour notre plus grand bien.

Car ensuite, l'inégalité $|\langle T(h), h\rangle| \leqslant N\,\|h\|^2$ découlant de la définition de N donne :

$$\begin{aligned} \Big|\mathrm{Re}\,\big[\langle T(f), g\rangle\big]\Big| &\leqslant \frac{N}{4}\Big[\|f+g\|^2 + \|f-g\|^2\Big] \\ \text{[Identité du parallélogramme]} \qquad &= \frac{N}{2}\big[\|f\|^2 + \|g\|^2\big]. \end{aligned}$$

En supposant $\|f\| \leqslant 1$ et $\|g\| \leqslant 1$, nous obtenons :

$$\Big|\mathrm{Re}\,\langle T(f), g\rangle\Big| \leqslant N.$$

Enfin, pour nous débarasser de cette partie réelle, il suffit de remplacer g par $e^{i\theta}g$ avec θ égal à l'argument de $\langle T(f), g\rangle$, et nous obtenons :

$$\big|\langle T(f), g\rangle\big| \leqslant N \qquad\qquad (\forall\, \|f\| = 1,\ \forall\, \|g\| = 1),$$

donc en prenant le supremum comme dans Lemme 9.13, l'inégalité inverse visée $\|T\| \leqslant N$ arrive — et clôt ce chapitre ! $\square$

10. Exercices

Exercice 1. Sur le $\mathbb{Q}$-espace vectoriel de dimension finie 2 :

$$\mathbb{Q}[\sqrt{2}] := \big\{a + b\sqrt{2}\colon\ a, b \in \mathbb{Q}\big\},$$

on introduit les deux normes :

$$\big\|a + b\sqrt{2}\big\|_1 := |a| + |b| \qquad \text{et} \qquad \big\|a + b\sqrt{2}\big\|_2 := \big|a + b\sqrt{2}\big|,$$

et on se propose d'établir qu'elles ne sont *pas* équivalentes.

(a) Commencer par vérifier que $\mathbb{Q}[\sqrt{2}]$ est bien un $\mathbb{Q}$-espace vectoriel de dimension 2, et que $\|\cdot\|_1$ et $\|\cdot\|_2$ sont bien des normes.
(b) Trouver une constante $0 < c < \infty$ assurant que :
$$c\,\|z\|_2 \leqslant \|z\|_1 \qquad (\forall\, z \in \mathbb{Q}[\sqrt{2}]).$$
(c) Pour $n \geqslant 1$ entier, montrer que les deux éléments :
$$x_n := \big(1-\sqrt{2}\big)^n \qquad \text{et} \qquad y_n := \big(1+\sqrt{2}\big)^n$$
appartiennent à $\mathbb{Q}[\sqrt{2}]$.
(d) Montrer qu'il n'existe pas de constante $0 < C < \infty$ telle que :
$$\|z\|_1 \leqslant C\,\|z\|_2 \qquad (\forall\, z \in \mathbb{Q}[\sqrt{2}]).$$
Indication: Si $x_n = a_n - b_n\sqrt{2}$ et si $y_n = a_n + b_n\sqrt{2}$, avec a_n, b_n entiers, montrer que $a_n \longrightarrow \infty$ lorsque $n \longrightarrow \infty$, puis, réaliser que $\|x_n\|_1 \longrightarrow \infty$ tandis que $\|x_n\|_2 \longrightarrow 0$.

Exercice 2. Montrer qu'une forme sesquilinéaire quelconque φ sur un espace vectoriel complexe E est hermitienne si et seulement si l'on a $\varphi(z,z) \in \mathbb{R}$ pour tout $z \in E$. Indication: Considérer $\varphi(z,w) + \varphi(w,z)$ et aussi $\varphi(\sqrt{-1}\,z, w) + \varphi(w, \sqrt{-1}\,z)$.

Exercice 3. Montrer qu'il n'existe aucun produit scalaire dans $\mathbb{R}^2$ dont la norme associée serait :
$$\big\|(x,y)\big\| := \max\big(|x|,\,|y|\big).$$
Exercice 4. Soit un intervalle fermé $[a,b] \subset \mathbb{R}$ avec $-\infty < a < b < \infty$.
(a) Montrer que l'application :
$$(f,g) \longmapsto \langle f,g\rangle := \int_a^b f(x)\,\overline{g(x)}\,dx$$
définit un produit scalaire hermitien sur l'espace :
$$\mathscr{C}^0\big([a,b],\,\mathbb{C}\big)$$
des fonctions continues sur $[a,b]$ à valeurs complexes.
(b) Cet espace $\mathscr{C}^0\big([a,b],\mathbb{C}\big)$ est-il complet pour la norme associée $\big(\int_a^b \big|f(x)\big|^2\,dx\big)^{\frac{1}{2}}$?

Exercice 5. **(a)** Soit $(E, \|\cdot\|)$ un espace vectoriel normé sur $\mathbb{C}$. Démontrer que sa norme $\|\cdot\|$ provient d'un produit scalaire hermitien si et seulement si elle satisfait l'identité du parallélogramme :
$$\forall\,(z,w) \in E^2 \qquad \|z+w\|^2 + \|z-w\|^2 = 2\big(\|z\|^2 + \|w\|^2\big),$$
et que dans ce cas, le produit scalaire hermitien qui définit $\|\cdot\|$ est nécessairement de la forme suivante :
$$\langle z,w\rangle = \tfrac{1}{4}\,\big(\|z+w\|^2 - \|z-w\|^2 + \sqrt{-1}\,\|z+\sqrt{-1}\,w\|^2 - \sqrt{-1}\,\|z-\sqrt{-1}\,w\|^2\big).$$
Indication: Pour démontrer que la condition est suffisante, on pourra considérer l'application $\langle\cdot,\cdot\rangle$ ainsi définie et démontrer successivement qu'elle satisfait les cinq propriétés suivantes :
- pour tout $z \in E$, on a $\langle z,z\rangle = \|z\|^2$;
- pour tout $(z,w) \in E^2$, on a $\langle z,w\rangle = \overline{\langle w,z\rangle}$;
- pour tout $(z,w,t) \in E^3$, on a $\langle z+w,t\rangle = 2\langle z,t/2\rangle + 2\langle w,t/2\rangle$;
- pour tout $(z,w,t) \in E^3$, on a $\langle z+w,t\rangle = \langle z,t\rangle + \langle w,t\rangle$;
- pour tout $(z,w) \in E^2$ et tout $\lambda \in \mathbb{C}$, on a $\langle \lambda z,w\rangle = \lambda\langle z,w\rangle$.

(b) En déduire que pour tout $\mathsf{p} \in [1,\infty[$ tel que $\mathsf{p} \neq 2$, il n'existe aucun produit scalaire hermitien sur l'espace $L^{\mathsf{p}}([0,1],dx)$ des fonctions $f\colon [0,1] \to \mathbb{C}$ de p-ème puissance intégrable dont dériverait sa norme :
$$\|f\|_{L^{\mathsf{p}}} := \left(\int_0^1 |f(x)|^{\mathsf{p}}\,dx\right)^{1/\mathsf{p}} < \infty.$$
Indication: On pourra examiner ce que donne l'identité du parallélogramme appliquée aux deux fonctions :
$$f\colon x \mapsto \begin{cases} 1 & \text{si } x \in [0,1/2] \\ 0 & \text{si } x \in \,]1/2,1] \end{cases} \qquad \text{et} \qquad g := 1-f.$$

Exercice 6. Soient E et F deux espaces vectoriels réels munis chacun d'un produit scalaire $\langle \cdot, \cdot \rangle_E$ et $\langle \cdot, \cdot \rangle_F$. Soit $f \colon E \to F$ une application avec $f(0_E) = 0_F$ telle que :

$$\|f(x) - f(y)\|_F = \|x - y\|_E,$$

pour tous $x, y \in E$. Montrer que f est nécessairement linéaire. Indication: Montrer tout d'abord que f préserve le produit scalaire.

Exercice 7. Soit $f \colon E \longrightarrow F$ une application linéaire entre deux espaces vectoriels réels (ou complexes) chacun munis d'un produit scalaire (hermitien) qui préserve la norme. Vérifier que f préserve alors aussi le produit scalaire (hermitien).

Exercice 8. On note $\ell^2_{\mathbb{R}}$ l'espace des suites infinies $x = (x_1, x_2, \dots, x_i, \dots)$ de nombres réels $x_i \in \mathbb{R}$ telles que $\sum_{i=1}^{\infty} |x_i|^2 < \infty$ et on le munit du produit scalaire habituel $\langle x, y \rangle_{\ell^2_{\mathbb{R}}} := \sum_{i=1}^{\infty} x_i y_i$. On considère aussi la collection de suites infinies :

$$E := \Big\{ (\alpha_n)_{n=1}^{\infty} \colon \alpha_n \in \mathbb{R}, \quad \sum_{n=1}^{\infty} \frac{\alpha_n^2}{n^2} < \infty \Big\}.$$

(a) Vérifier que $\ell^2_{\mathbb{R}}$ est contenu dans E.

(b) Montrer que E est un espace vectoriel sur $\mathbb{R}$ pour la loi de linéarité composante-par-composante :

$$\lambda\, (\alpha_n)_{n=1}^{\infty} + \mu\, (\beta_n)_{n=1}^{\infty} := \big(\lambda\, \alpha_n + \mu\, \beta_n \big)_{n=1}^{\infty}.$$

(c) Montrer que :

$$\langle \alpha, \beta \rangle_E := \sum_{n \geqslant 1} \frac{\alpha_n\, \beta_n}{n^2}$$

définit aussi un produit scalaire sur E.

(d) Montrer que l'application :

$$\varphi \colon \left[\begin{array}{c} \ell^2_{\mathbb{R}} \longrightarrow E \\ (x_i)_{i=1}^{\infty} \longmapsto (i\, x_i)_{i=1}^{\infty} =: (\alpha_n)_{n=1}^{\infty} \end{array} \right.$$

est un isomorphisme isométrique, *i.e.* un isomorphisme linéaire satisfaisant $\|\varphi(x)\|_E = \|x\|_{\ell^2_{\mathbb{R}}}$ pour tout $x \in \ell^2_{\mathbb{R}}$.

Exercice 9. Soit $\ell^{\infty} = \ell^{\infty}_{\mathbb{C}}$ l'espace vectoriel des suites $z = (z_i)_{i=1}^{\infty} \in \mathbb{C}^{\infty}$ vérifiant :

$$\sup_{i \geqslant 1} |z_i| \, < \, \infty.$$

(a) Montrer que la formule $\|z\|_{\ell^{\infty}} := \sup_{i \geqslant 1} |z_i|$ définit une *norme* sur ℓ^{∞}.

(b) Montrer que $\big(\ell^{\infty}, \| \cdot \|_{\ell^{\infty}} \big)$ n'est pas séparable. Indication: Utiliser le fait que $\{0, 1\}^{\mathbb{N}^*}$ n'est pas dénombrable, et, pour deux éléments $z = (z_i)_{i=1}^{\infty} \in \ell^{\infty}$ et $z' = (z'_i)_{i=1}^{\infty} \in \ell^{\infty}$ de composantes $z_i, z'_i \in \{0, 1\}$ égales seulement aux deux lettres dont le langage infra-primitif des ordinateurs est constitué, utiliser aussi le fait que les deux boules ouvertes :

$$\big\{ w \in \ell^{\infty} \colon \, \|w - z\|_{\ell^{\infty}} < \tfrac{1}{2} \big\} \qquad \text{et} \qquad \big\{ w' \in \ell^{\infty} \colon \, \|w' - z'\|_{\ell^{\infty}} < \tfrac{1}{2} \big\}$$

sont disjointes.

Exercice 10. Soit E un $\mathbb{R}$-espace vectoriel muni d'une norme $\| \cdot \|$.

(a) Montrer que deux vecteurs $x, y \in E$ quelconques satisfont toujours l'inégalité :

$$\|x\|, \, \|y\| \, \leqslant \, \max \big\{ \|x + y\|, \, \|x - y\| \big\}.$$

(b) Montrer qu'une boule fermée entièrement contenue dans une autre boule fermée a toujours un rayon qui est inférieur ou égal à celui de la première ; précisément, si $x_1, x_2 \in E$, si $r_1, r_2 > 0$ et si :

$$\big\{ x \in E \colon \, \|x - x_1\| \leqslant r_1 \big\} =: B_1 \, \subset \, B_2 := \big\{ x \in E \colon \, \|x - x_2\| \leqslant r_2 \big\},$$

montrer que $r_1 \leqslant r_2$.

Exercice 11. Montrer qu'aucune des deux normes suivantes sur $\mathbb{C}^n$:

$$\|z\|_\infty := \max\big\{|z_1|, \ldots, |z_n|\big\} \qquad \text{et} \qquad \|z\|_1 := |z_1| + \cdots + |z_n|$$

ne dérive d'un produit scalaire.

Exercice 12. Soit un nombre $1 \leqslant \mathsf{p} < \infty$. On note $\ell^{\mathsf{p}}_{\mathbb{C}}$ l'ensemble des suites infinies de nombres complexes :

$$z = \big(z_1, z_2, \ldots, z_i, \ldots\big) \in \mathbb{C}^\infty,$$

dont la somme des puissances p-èmes des composantes converge :

$$\sum_{i=1}^{\infty} \big|z_i\big|^{\mathsf{p}} < \infty.$$

(a) En utilisant le fait que la fonction $\mathbb{R}_+ \ni t \longmapsto t^{\mathsf{p}} \in \mathbb{R}_+$ est convexe (lorsque $\mathsf{p} \geqslant 1$!), montrer pour tout $0 < \alpha < 1$ l'inégalité :

$$\big|z + w\big|^{\mathsf{p}} \leqslant \alpha \,\big|\tfrac{z}{\alpha}\big|^{\mathsf{p}} + (1-\alpha)\,\big|\tfrac{w}{1-\alpha}\big|^{\mathsf{p}}.$$

(b) Montrer que la quantité :

$$\|z\|_{\mathsf{p}} := \bigg(\sum_{i=1}^{\infty} \big|z_i\big|^{\mathsf{p}}\bigg)^{1/\mathsf{p}}$$

définit une norme sur ℓ^{p} et que ℓ^{p} est un $\mathbb{C}$-espace vectoriel. Indication: Pour établir l'inégalité triangulaire, poser $\alpha := \frac{\|z\|_{\mathsf{p}}}{\|z\|_{\mathsf{p}} + \|w\|_{\mathsf{p}}}$ ci-dessus.

(c) Montrer que ℓ^{p} est complet pour $\|\cdot\|_{\mathsf{p}}$.

(d) Lorsque $\mathsf{p} < 1$, a-t-on une norme avec cette définition ?

(e) Soit à nouveau $\mathsf{p} \geqslant 1$. On considère le sous-espace :

$$M_{\mathsf{p}} := \Big\{z = (z_i)_{i=1}^{\infty} \in \ell^{\mathsf{p}} \colon \textstyle\sum_{i=1}^{\infty} z_i = 0\Big\}.$$

Pour quelles valeurs de $\mathsf{p} \geqslant 1$ ce sous-espace est-il dense dans ℓ^{p} ?

Exercice 13. On note X l'espace des fonctions continues $f\colon \mathbb{R} \longrightarrow \mathbb{C}$ qui sont sommes finies quelconques d'exponentielles trigonométriques :

$$f(t) = \sum_{k=1}^{n} a_k\, e^{i\theta_k t},$$

avec $n \in \mathbb{N}$, $a_k \in \mathbb{C}$, $\theta_k \in \mathbb{R}$ quelconques. Étant donné deux fonctions $f, g \in X$ dans cet espace, on pose :

$$\langle f,\, g\rangle := \lim_{T \to \infty} \frac{1}{2\,T} \int_{-T}^{T} f(t)\,\overline{g(t)}\,dt.$$

(a) Montrer que $\langle\cdot,\,\cdot\rangle$ définit un produit scalaire sur X.

(b) Montrer que la famille $\big(t \longmapsto e^{i\theta t}\big)_{\theta \in \mathbb{R}}$ est orthonormée.

(c) L'espace X est-il de Hilbert ?

(d) L'espace X est-il séparable ?

Exercice 14. Soit $\big(H, \langle\cdot,\cdot\rangle\big)$ un $\mathbb{C}$-espace de Hilbert.

(a) Si $z, w \in H$ satisfont $\mathrm{Re}\,\langle z,\, w\rangle = \|z\|^2 = \|w\|^2$, montrer que $z = w$.

(b) Soient maintenant deux suites $(z_n)_{n=1}^{\infty}$ et $(w_n)_{n=1}^{\infty}$ satisfaisant $\|z_n\| \leqslant 1$ et $\|w_n\| \leqslant 1$. Si $\lim_{n\to\infty} \langle z_n,\, w_n\rangle = 1$, montrer que $\lim_{n\to\infty} \|z_n - w_n\| = 0$.

(c) Si $\lim_{n\to\infty} \|z_n + w_n\| = 2$, montrer que $\lim_{n\to\infty} \|z_n - w_n\| = 0$.

Exercice 15. On munit le $\mathbb{R}$-espace vectoriel $\mathbb{R}[x] \subset \mathscr{C}^0(\mathbb{R}, \mathbb{R})$ des polynômes réels d'un produit scalaire quelconque $\langle\cdot,\cdot\rangle$, dont on note $\|\cdot\|$ la norme associée. À partir de la famille génératrice libre $\big(x^k\big)_{k=0}^{\infty}$ de tous les monômes, le procédé de Gram-Schmidt construit une famille orthonormée $\big(P_k(x)\big)_{k=0}^{\infty}$ avec $\deg P_k = k$ pour tout $k \geqslant 0$. Soit alors $\big(Q_k(x)\big)_{k=0}^{\infty}$ la famille :

$$Q_k(x) := \sum_{0 \leqslant i \leqslant k} \frac{1}{i+1}\, P_i(x).$$

(a) Montrer que la suite $\big(Q_k(x)\big)_{k=0}^{\infty}$ est de Cauchy dans $\big(\mathbb{R}[x], \|\cdot\|\big)$.

(b) Montrer que $\big(\mathbb{R}[x], \|\cdot\|\big)$ n'est pas un espace de Hilbert.

Exercice 16. On considère $\mathscr{C}^0\big([-1,1],\mathbb{R}\big)$ muni du produit scalaire L^2 :

$$\langle f, g\rangle_{L^2} \ := \int_{-1}^{1} f(x)\, g(x)\, dx \qquad\qquad (f, g \in \mathscr{C}^0[-1,1]),$$

et on note $\|\cdot\|_{L^2}$ la norme associée, distincte de celle intrinsèque aux fonctions continues :

$$\|f\|_{\mathscr{C}^0} \ := \max_{-1\leqslant x\leqslant 1} \big|f(x)\big| \qquad\qquad (f \in \mathscr{C}^0[-1,1]).$$

(a) Montrer que la suite, définie pour $n \geqslant 1$, de fonctions :

$$f_n(x) \ := \begin{cases} -1 & \text{lorsque} \ \ -1 \leqslant x \leqslant -\frac{1}{n}, \\ n\,x & \text{lorsque} \ \ -\frac{1}{n} < x < \frac{1}{n}, \\ 1 & \text{lorsque} \ \ \frac{1}{n} \leqslant x \leqslant 1, \end{cases}$$

est de Cauchy dans $\big(\mathscr{C}^0[-1,1], \|\cdot\|_{L^2}\big)$.

(b) Montrer que $(f_n)_{n=1}^{\infty}$ converge en norme L^2 vers la fonction :

$$f_\infty(x) \ := \begin{cases} -1 & \text{lorsque} \ \ -1 \leqslant x < 0, \\ 0 & \text{pour} \ \ x = 0, \\ 1 & \text{lorsque} \ \ 0 < x \leqslant 1. \end{cases}$$

(c) Qu'en est-t-il en norme $\|\cdot\|_{\mathscr{C}^0}$? Interpréter.

Exercice 17. Une fonctionnelle linéaire :

$$L\colon \quad L^2\big([0,1],\mathbb{R}\big) \longrightarrow \mathbb{R}$$

est dite *positive* si :

$$\Big(f \ \geqslant \ 0 \text{ presque partout}\Big) \ \Longrightarrow \ L(f) \ \geqslant \ 0.$$

L'objectif est de démontrer que toute fonctionnelle linéaire positive est continue.

(a) Montrer que $L(g) \geqslant L(f)$ dès que $f, g \in L^2[0,1]$ satisfont $g \geqslant f$ presque partout.

(b) Montrer que $L(|f|) \geqslant \big|L(f)\big|$, pour toute $f \in L^2[0,1]$.

(c) On raisonne par l'absurde, à savoir on suppose qu'il existe une fonctionnelle linéaire positive L sur $L^2[0,1]$ qui n'est pas continue. Montrer qu'il existe une suite $(u_k)_{k=1}^{\infty}$ de fonctions $u_k \in L^2[0,1]$ avec $u_k \geqslant 0$ presque partout satisfaisant :

$$\|u_k\|_{L^2} \ = \ 1 \qquad \text{et} \qquad L(u_k) \ \geqslant \ k\,2^k \qquad\qquad (\forall\, k \geqslant 1).$$

(d) Montrer que la série $\sum_{k=1}^{\infty} \frac{u_k}{2^k}$ est normalement convergente dans $L^2[0,1]$.

(e) Soit donc $\sum_{k=1}^{\infty} \frac{u_k}{2^k} =: u \in L^2[0,1]$. Montrer que l'on a presque partout :

$$\frac{u_k}{2^k} \ \leqslant \ u \qquad\qquad (\forall\, k \geqslant 1).$$

Indication: Si une suite converge dans L^2 en norme L^2 vers une certaine fonction-limite de carré intégrable, on peut toujours en extraire une sous-suite qui converge ponctuellement presque partout vers cette même fonction-limite.

(f) Montrer que $L(u) \geqslant k$ pour tout $k \geqslant 1$, et conclure.

(g) Lorsqu'on change la norme sur $L^2[0,1]$, l'énonce :

$$L \text{ positive} \ \Longrightarrow \ L \text{ continue},$$

peut devenir faux. Après avoir vérifié que $L^2[0,1] \subset L^1[0,1]$, donner un exemple de fonctionnelle linéaire positive sur :

$$\big(L^2[0,1],\ \|\cdot\|_{L^1}\big),$$

qui n'est *pas* continue.

Exercice 18. Dans l'espace vectoriel $\mathscr{C}^0\big([0,1],\mathbb{R}\big)$ des fonctions continues sur l'intervalle $[0,1]$, muni de la norme :

$$\|f\|_{\mathscr{C}^0} := \max_{0\leqslant x\leqslant 1} \big|f(x)\big| \qquad\qquad (f\in\mathscr{C}^0[0,1]),$$

trouver une famille $(e_k)_{k=1}^\infty$ qui est libre et totale, ainsi qu'une fonction $g\in\mathscr{C}^0[0,1]$, telles que l'on ait deux représentations :

$$\begin{aligned} g &= \sum_{k=1}^\infty \lambda_k\, e_k \\ &= \sum_{k=1}^\infty \mu_k\, e_k, \end{aligned}$$

pour deux collections *distinctes* de coefficients réels $(\lambda_k)_{k=1}^\infty$ et $(\mu_k)_{k=1}^\infty$. Cela serait-il possible dans un espace de Hilbert ? Indication: Penser à écrire $e^x - 1 = \sum_{k=1}^\infty \frac{x^k}{k!}$.

Exercice 19. Déterminer les trois constantes réelles $a,b,c\in\mathbb{R}$ qui minimisent la valeur de :

$$\int_{-1}^1 \big|x^3 - a\,x^2 - b\,x - c\big|^2\,dx.$$

Exercice 20. On munit l'espace $\mathscr{C}^1\big([0,1],\mathbb{R}\big)$ des fonctions continûment dérivables sur $[0,1]$ du produit scalaire :

$$\langle f,g\rangle := \int_0^1 \big(f(x)\,g(x) + f'(x)\,g'(x)\big)\,dx \qquad\qquad (f,g\in\mathscr{C}^1[0,1]),$$

et on note $\|\cdot\|$ la norme associée.

(a) Montrer que $\langle\cdot,\cdot\rangle$ est effectivement un produit scalaire.

(b) Montrer que les deux sous-espaces suivants de $\mathscr{C}^1[0,1]$ — noter que $\mathscr{C}^2\subset\mathscr{C}^1$ — :

$$V := \big\{g\in\mathscr{C}^1[0,1]\colon\ g(0)=g(1)=0\big\} \qquad \text{et} \qquad W := \big\{h\in\mathscr{C}^2[0,1]\colon\ h''=h\big\},$$

sont orthogonaux entre eux.

(c) Vérifier que :

$$W = \big\{x\longmapsto \lambda\,e^x + \mu\,e^{-x}\colon\ (\lambda,\mu)\in\mathbb{R}^2\big\},$$

et exprimer la norme d'un tel élément général de W.

(d) Montrer, pour tout $f\in\mathscr{C}^1[0,1]$, qu'il existe deux réels $\lambda_f,\mu_f\in\mathbb{R}$ tels que :

$$\big(x\longmapsto f(x) - \lambda_f\,e^x - \mu_f\,e^{-x}\big) \in V.$$

(e) Établir que V et W sont supplémentaires orthogonaux dans $\mathscr{C}^1[0,1]$.

(f) Exprimer l'opérateur de projection orthogonale :

$$\pi_W\colon\quad \mathscr{C}^1[0,1]\longrightarrow W.$$

(g) Maintenant, pour deux constantes réelles $\alpha,\beta\in\mathbb{R}$ fixées, on regarde :

$$E_{\alpha,\beta} := \big\{f\in\mathscr{C}^1[0,1]\colon\ f(0)=\alpha,\ f(1)=\beta\big\}.$$

Déterminer :

$$\inf_{f\in E_{\alpha,\beta}} \|f\|^2.$$

Indication: Appliquer le théorème de Pythagore à la décomposition d'éléments $f\in E_{\alpha,\beta}$ sur $V\overset{\perp}{\oplus}W$.

Exercice 21. Soit $(H,\|\cdot\|_H)$ un espace de Hilbert. Montrer que sa boule unité fermée $\{f\in H\colon \|f\|_H\leqslant 1\}$ est compacte si et seulement si $\dim H<\infty$.

Exercice 22. [Complétion d'un espace préhilbertien] Soit H_0 un $\mathbb{C}$-espace vectoriel préhilbertien (pas nécessairement complet), de produit scalaire hermitien $\langle \cdot, \cdot \rangle_{H_0}$ défini positif. On considère la collection de toutes les suites $(f_n)_{n=1}^{\infty}$ d'éléments $f_n \in H_0$ qui sont de Cauchy pour $\| \cdot \|_{H_0}$, on introduit la relation :

$$(f_n)_{n=1}^{\infty} \sim (f'_n)_{n=1}^{\infty} \iff 0 = \lim_{n \to \infty} \big\| f_n - f'_n \big\|_{H_0},$$

et on définit le quotient :

$$H := \Big\{ (f_n)_{n=1}^{\infty} \in H_0 \Big\} \Big/ \sim .$$

(a) Vérifier que $\sim$ est une relation d'équivalence. On notera F la classe d'équivalence d'une suite $(f_n)_{n=1}^{\infty} \in H_0$.

(b) Montrer que H hérite d'une structure de $\mathbb{C}$-espace vectoriel.

(c) Montrer que H hérite du produit scalaire $\langle \cdot, \cdot \rangle_H$ défini par :

$$\langle F, G \rangle_H := \lim_{n \to \infty} \langle f_n, g_n \rangle_{H_0},$$

où F et G sont représentés pas $(f_n)_{n=1}^{\infty}$ et $(g_n)_{n=1}^{\infty}$.

(d) Avec des suites constantes, vérifier que $H \supset H_0$, et que $\langle f, g \rangle_H = \langle f, g \rangle_{H_0}$ pour tous $f, g \in H_0$.

(e) Montrer que H_0 est dense dans H.

(f) Montrer que H est complet. Indication: Étant donné une suite de Cauchy $(F^k)_{k=1}^{\infty}$ dans $(H, \| \cdot \|_H)$, chaque F^k étant représenté par une suite de Cauchy $(f_n^k)_{n=1}^{\infty}$ d'éléments $f_n^k \in H_0$, montrer que l'élément $F \in H$ représenté par la suite — dont on vérifiera qu'elle est bien de Cauchy — :

$$\big(f_{\mathrm{N}(n)}^n \big)_{n=1}^{\infty},$$

où $N(k) \gg 1$ est choisi suffisamment grand pour que :

$$\big\| f_n^k - f_{\mathrm{N}(k)}^k \big\|_{H_0} \leqslant \frac{1}{k} \qquad\qquad (\forall\, n \geqslant \mathrm{N}(k)),$$

satisfait :

$$0 = \lim_{k \to \infty} \big\| F - F^k \big\|_H.$$

(g) Enfin, soient H et H' deux espaces de Hilbert (complets) qui sont des *complétions* de H_0 au sens où :

$$H_0 \subset H, \qquad \overline{H_0}^{\| \cdot \|_H} = H, \qquad \langle \cdot, \cdot \rangle_H \big|_{H_0} = \langle \cdot, \cdot \rangle_{H_0},$$
$$H_0 \subset H', \qquad \overline{H_0}^{\| \cdot \|_{H'}} = H', \qquad \langle \cdot, \cdot \rangle_{H'} \big|_{H_0} = \langle \cdot, \cdot \rangle_{H_0}.$$

Montrer qu'il existe un isomorphisme $\mathbb{C}$-linéaire $\Phi \colon H \longrightarrow H'$ avec $\Phi\big|_{H_0} = \mathrm{Id}_{H_0}$ qui est *unitaire* au sens où :

$$\big\| \Phi(F) \big\|_{H'} = \| F \|_H \qquad\qquad (\forall\, F \in H).$$

Indication: Si $F \in H$ est représenté par une suite $(f_n)_{n=1}^{\infty}$ de Cauchy dans H_0, montrer que $(f_n)_{n=1}^{\infty}$ définit aussi un élément $F' \in H'$.

Exercice 23. Montrer que les deux ensembles suivants de fonctions sont denses dans $L^1(\mathbb{R}^d)$.

(a) Les fonctions étagées.

(b) Les fonctions continues à support compact.

Exercice 24. On note E le $\mathbb{C}$-sous-espace vectoriel de $L^2\big([0,1]\big)$ constitué des fonctions $f \colon [0,1] \longrightarrow \mathbb{C}$ de carré intégrable pour lesquelles il existe une fonction notée $\Lambda_f \in L^2\big([0,1]\big)$ vérifiant :

$$\int_0^1 f(x)\, \varphi'(x)\, dx = - \int_0^1 \Lambda_f(x)\, \varphi(x)\, dx,$$

pour toute fonction $\varphi \in \mathscr{C}_c^{\infty}\big([0,1]\big)$ de classe $\mathscr{C}^{\infty}$ sur $[0,1]$ et dont le support :

$$\mathrm{supp}(\varphi) \overset{\mathrm{def}}{:=} \overline{\big\{ t \in [0,1] \colon \varphi(t) \neq 0 \big\}}$$

est un sous-ensemble *compact* de $[0,1]$.

(a) Montrer que E est un sous-ensemble dense de $L^2\big([0,1]\big)$.

(b) Montrer que :

$$\langle f,\, g\rangle := \int_0^1 f(x)\,\overline{g(x)}\,dx + \int_0^1 \Lambda_f(x)\,\overline{\Lambda_g(x)}\,dx$$

définit un produit scalaire sur E.

(c) Montrer que E, muni de ce produit scalaire, est un espace de Hilbert.

Exercice 25. [Polynômes de Hermite] On appelle n-ème *polynôme de Hermite* le résultat de :

$$H_n(x) := (-1)^n\, e^{x^2}\, \frac{d^n}{dx^n}\big(e^{-x^2}\big).$$

Obtient-on bien des polynômes ? Pourquoi ?

(a) Calculer H_0, H_1, H_2, H_3, H_4, H_5. Deviner des régularités qui semblent générales.

(b) Montrer que l'on a, pour tout $n \geqslant 1$:

$$H_n(x) = 2\,x\,H_{n-1}(x) - H'_{n-1}(x).$$

(c) Raisonner par récurrence pour obtenir l'expression explicite :

$$H_n(x) = n! \sum_{0\leqslant j\leqslant \frac{n}{2}} (-1)^j\, \frac{1}{j!\;(n-2j)!}\,\big(2\,x\big)^{n-2j},$$

et confirmer ce qui avait été deviné en **(a)**.

(d) Montrer que les polynômes de Hermite forment un système orthogonal par rapport au poids $w(x) = e^{-x^2}$ et que leur norme au carré vaut :

$$\big|\big|H_n\big|\big|_w^2 = 2^n\, n!\, \sqrt{\pi}.$$

Exercice 26. [Polynômes de Legendre] Soit $L^2\big([-1,1],\mathbb{C}\big)$ l'espace des fonctions mesurables sur $[-1,1]$ dont le module au carré est d'intégrale $< \infty$, que l'on munit du produit scalaire usuel. On appelle *polynôme de Legendre* de degré $n \geqslant 0$ le polynôme :

$$L_n(x) := \frac{1}{2^n\, n!}\, \frac{d^n}{dx^n}\,\big[(x^2-1)^n\big].$$

(a) Pour deux entiers $m > n \geqslant 0$, intégrer n fois par partie l'intégrale :

$$\int_{-1}^1 \frac{d^m}{dx^m}\big[(x^2-1)^m\big]\;\frac{d^n}{dx^n}\big[(x^2-1)^n\big]\,dx,$$

en prenant itérativement la primitive du premier facteur et en déduire que L_m est orthogonal à L_n.

(b) Par la même méthode, calculer $||L_n||^2_{L^2[-1,1]}$.

(c) En déduire qu'il existe une constante strictement positive $c_n > 0$ — que l'on précisera ! — telle que $\big(c_n\, L_n\big)_{n\in\mathbb{N}}$ est un système *orthonormal* de $L^2\big([-1,1],\mathbb{C}\big)$.

Exercice 27. On dit qu'un sous-ensemble $\mathscr{A}$ d'un espace métrique est *précompact* si pour tout $\varepsilon > 0$, il existe un nombre fini de parties $\mathscr{A}_1,\dots,\mathscr{A}_k$ de $\mathscr{A}$ toutes de diamètre $\leqslant \varepsilon$ dont la réunion $\cup_{1\leqslant i\leqslant k}\,\mathscr{A}_i = \mathscr{A}$ est égale à $\mathscr{A}$. Soit alors $\mathscr{A}$ un sous-ensemble quelconque d'un espace vectoriel normé $(E,||\cdot||)$. Démontrer que $\mathscr{A}$ est précompact si et seulement si $\mathscr{A}$ est borné et si pour tout $\varepsilon > 0$, il existe un sous-espace vectoriel F_ε de E de dimension finie tel que pour tout $a \in \mathscr{A}$, on a $\mathsf{dist}(a, F_\varepsilon) \leqslant \varepsilon$.

Exercice 28. Soit $a = (a_n)_{n=1}^\infty$ une suite quelconque de nombres réels strictement positifs. On note ℓ_a^2 la collection des suites de nombres complexes $z = (z_n)_{n=1}^\infty$ telles que la série $\sum_{n=1}^\infty a_n\,|z_n|^2$ soit convergente.

(a) Vérifier que ℓ_a^2 est un espace vectoriel sur $\mathbb{C}$.

(b) Démontrer que l'expression :

$$\langle z_n, w_n\rangle_{\ell_a^2} := \sum_{n=1}^\infty a_n\, z_n \overline{w}_n$$

définit un produit scalaire sur ℓ_a^2.

(c) Démontrer que l'application :

$$\iota_a\colon (z_n)_{n=1}^\infty \longmapsto \big(\sqrt{a_n}\,z_n\big)_{n=1}^\infty$$

est un isomorphisme isométrique de ℓ_a^2 sur $\ell_{\mathbb{C}}^2$. En déduire que ℓ_a^2 est aussi un espace de Hilbert.

(d) Soient $(a_n)_{n=1}^{\infty}$ et $(b_n)_{n=1}^{\infty}$ deux suites de nombres réels strictement positifs. Démontrer que, si $0 = \lim_{n\to\infty} \frac{a_n}{b_n}$, alors la boule unité fermée de ℓ_b^2 est un sous-ensemble compact de ℓ_a^2. Indication: Utiliser l'Exercice 27.

Exercice 29. Montrer qu'un opérateur linéaire $T\colon H_1 \longrightarrow H_2$ est automatiquement borné lorsque H_1 est de dimension finie, et que tel n'est pas le cas autrement.

Exercice 30. Soit $(\varphi_j)_{j=1}^{\infty}$ une base hilbertienne de $L^2(\mathbb{R}^d)$. Montrer que la collection doublement infinie :

$$\big(\varphi_j(x)\,\varphi_k(y)\big)_{j,k=1}^{\infty}$$

constitue une base hilbertienne de $L^2(\mathbb{R}^d \times \mathbb{R}^d)$. Indication: Commencer par vérifier l'orthonormalité. Ensuite, avec $F \in L^2(\mathbb{R}^d \times \mathbb{R}^d)$ et pour tout $k \geqslant 1$ fixé, introduire $F_k(x) := \int_{\mathbb{R}^d} F(x,y)\,\overline{\varphi_k(y)}\,dy$. En supposant $0 = \langle F, \varphi_j \varphi_k\rangle_{L^2(\mathbb{R}^d\times\mathbb{R}^d)}$ pour tout j, obtenir $0 = \int_{\mathbb{R}^d} F_k(x)\,\overline{\varphi_j(x)}\,dx$.

Exercice 31. Soit $E \subset H$ un sous-espace vectoriel d'un $\mathbb{C}$-espace de Hilbert H. Montrer que son bi-orthogonal $(E^{\perp})^{\perp}$ est le plus petit sous-espace *fermé* de H qui contient E.

Exercice 32. Soit H un espace de Hilbert et soit $C \subset H$ un sous-ensemble convexe fermé non vide. Montrer, en utilisant un théorème du cours, que l'opérateur de projection π_C sur C est 1-*lipschitzien*, au sens où il satisfait :

$$\|\pi_C(x_2) - \pi_C(x_1)\| \leqslant \|x_2 - x_1\|,$$

pour tous $x_1, x_2 \in H$.

Exercice 33. Soit π_F la projection orthogonale sur un sous-espace vectoriel fermé $F \subset H$ d'un $\mathbb{C}$-espace de Hilbert H :

$$\pi_F\big|_F \,=\, \mathrm{Id}_F \qquad\qquad \text{et} \qquad\qquad \pi_F\big|_{F^{\perp}} \,=\, 0.$$

(a) Vérifier que $\pi_F \circ \pi_F = \pi_F$ et montrer que $\pi_F^* = \pi_F$.

(b) Réciproquement, si $P\colon H \longrightarrow H$ est un opérateur linéaire borné satisfaisant $P \circ P = P$ ainsi que $P^* = P$, montrer que P est la projection orthogonale sur un certain sous-espace fermé de H.

(c) Démontrer que tout sous-espace fermé $F \subset H$ d'un espace de Hilbert séparable H est aussi un espace de Hilbert séparable.

Exercice 34. Soit $E \subset \mathbb{R}^d$ un sous-ensemble mesurable, et soit $F \subset L^2(\mathbb{R}^d)$ le sous-espace des fonctions qui s'annulent presque partout dans le complémentaire $\mathbb{R}^d \backslash E$. Montrer que la projection orthogonale $\pi_F\colon L^2(\mathbb{R}^d) \longrightarrow F$ est donnée par :

$$\pi_F(f) \,=\, f \cdot \mathbf{1}_E \qquad\qquad (f \in L^2(\mathbb{R}^d)).$$

Exercice 35. Soient π_{F_1} et π_{F_2} deux projections orthogonales sur deux sous-espaces fermés $F_1, F_2 \subset H$ d'un $\mathbb{C}$-espace de Hilbert H.

(a) Montrer que leur composition $\pi_{F_1} \circ \pi_{F_2}$ est encore une projection orthogonale si et seulement si $\pi_{F_1} \circ \pi_{F_2} = \pi_{F_2} \circ \pi_{F_1}$ commutent.

(b) Dans ce cas, montrer que $\pi_{F_1} \circ \pi_{F_2}$ est la projection orthogonale sur $F_1 \cap F_2$.

Exercice 36. Soit H un $\mathbb{C}$-espace de Hilbert. On note $\mathscr{L}(H)$ l'espace vectoriel de tous les opérateurs $T\colon H \longrightarrow H$ (linéaires) bornés, muni de la norme d'opérateur :

$$|\!|\!|T|\!|\!| \,:=\, \inf\big\{0 \leqslant C < \infty\colon\ \|T(f)\| \leqslant C\,\|f\|,\ \forall\, f \in H\big\}.$$

(a) Montrer que $|\!|\!|T_1 + T_2|\!|\!| \leqslant |\!|\!|T_1|\!|\!| + |\!|\!|T_2|\!|\!|$ pour tous $T_1, T_2 \in \mathscr{L}(H)$.

(b) Établir que la quantité :

$$d(T_1, T_2) \,:=\, |\!|\!|T_1 - T_2|\!|\!|$$

définit une métrique sur $\mathscr{L}(H)$.

(c) Montrer que $\mathscr{L}(H)$ est complet pour cette métrique.

Exercice 37. [Déterminants de Gram] Soit $\mathbb{K} = \mathbb{R}$ ou $= \mathbb{C}$ le corps des nombres réels ou le corps des nombres complexes. Soit E un $\mathbb{K}$-espace vectoriel préhilbertien, *i.e.* muni d'un produit scalaire $\langle \cdot, \cdot \rangle$ mais qui n'est pas nécessairement complet (dans le cas où E est de dimension infinie !). Soient $e_1, \dots, e_n$ des vecteurs de E. On appelle *matrice de Gram* de $e_1, \dots, e_n$ la matrice :

$$\mathsf{Gram}\big(e_1, \dots, e_n\big) := \Big(\langle e_i,\, e_j \rangle \Big)_{1 \leqslant i \leqslant n}^{1 \leqslant j \leqslant n},$$

l'indice i numérotant ici les lignes. Son déterminant :

$$\det \mathsf{Gram}\big(e_1, \dots, e_n\big) = \begin{vmatrix} \langle e_1, e_1 \rangle & \cdots & \langle e_1, e_n \rangle \\ \vdots & \cdots & \vdots \\ \langle e_n, e_1 \rangle & \cdots & \langle e_n, e_n \rangle \end{vmatrix}$$

est naturellement appelé *déterminant de Gram* de $e_1, \dots, e_n$.

(a) Vérifier que cette matrice est *symétrique* dans le cas $\mathbb{K} = \mathbb{R}$, ou *hermitienne* dans le cas $\mathbb{K} = \mathbb{C}$, à savoir par définition égale à la conjuguée de sa transposée.

(b) Dans le cas de $n = 2$ vecteurs, que peut-on dire de $\mathsf{Gram}(e_1, e_2)$? Quel théorème fondamental le Sherlock Holmes reconnaît-il ici sans aucune aide de son cher Watson ? Et qu'est-on tenté alors de conjecturer pour tout $n \geqslant 2$?

(c) Soit maintenant $X = (x_1, \dots, x_n) \in \mathbb{K}^n$ un vecteur-ligne quelconque et soit :

$$\big(\mathsf{Gram}(e) \cdot X\big)_i$$

la i-ème composante du vecteur-colonne $\mathsf{Gram}(e) \cdot X$. Montrer par le calcul que l'on a :

$$\big\langle X,\, \mathsf{Gram}(e) \cdot X \big\rangle = \bigg\| \sum_{i=1}^{n} \overline{x}_i\, e_i \bigg\|^2.$$

(d) Montrer que la matrice $\mathsf{Gram}(e_1, \dots, e_n)$ est toujours positive, et qu'elle est *définie* positive lorsque et seulement lorsque $e_1, \dots, e_n$ sont $\mathbb{K}$-linéairement indépendants.

(e) Le but principal de cet exercice est de montrer que si $F \subset E$ est un K-sous-espace vectoriel de dimension finie $n \geqslant 1$ muni d'une base $(e_1, \dots, e_n)$ qui n'est pas nécessairement orthogonale, alors pour tout $x \in E$, la distance au carré de x à F est donnée comme le quotient suivant de déterminants de Gram :

$$\big[\mathsf{dist}(x, F)\big]^2 = \frac{\det \mathsf{Gram}(e_1, \dots, e_n, x)}{\det \mathsf{Gram}(e_1, \dots, e_n)}.$$

Cette formule vient donc complémenter le cours, dans lequel on a systématiquement supposé que les vecteurs étaient orthogonaux entre eux afin de simplifier grandement les formules. Commencer alors par justifier l'existence d'un unique vecteur $\pi_F(x) \in F$ qui réalise la distance de x à F :

$$\begin{aligned} \mathsf{dist}\big(x, F\big) &\overset{\mathsf{def}}{=} \inf_{y \in F} \|x - y\| \\ &= \big\|x - \pi_F(x)\big\|. \end{aligned}$$

(f) Montrer par le calcul que :

$$\det \mathsf{Gram}\big(e_1, \dots, e_n, x\big) = \begin{vmatrix} \langle e_1, e_1 \rangle & \cdots & \langle e_1, e_n \rangle & \overline{\langle \pi_F(x), e_1 \rangle} \\ \vdots & \cdots & \vdots & \vdots \\ \langle e_n, e_1 \rangle & \cdots & \langle e_n, e_n \rangle & \overline{\langle \pi_F(x), e_n \rangle} \\ \langle \pi_F(x), e_1 \rangle & \cdots & \langle \pi_F(x), e_n \rangle & \big\|\pi_F(x)\big\|^2 \end{vmatrix} + \\ + \begin{vmatrix} \langle e_1, e_1 \rangle & \cdots & \langle e_1, e_n \rangle & 0 \\ \vdots & \cdots & \vdots & \vdots \\ \langle e_n, e_1 \rangle & \cdots & \langle e_n, e_n \rangle & 0 \\ \langle \pi_F(x), e_1 \rangle & \cdots & \langle \pi_F(x), e_n \rangle & \big\|x - \pi_F(x)\big\|^2 \end{vmatrix},$$

et conclure.

Exercice 38. On considère ici comme connu le théorème de Weierstrass d'après lequel la suite des monômes standard $\big(x^n\big)_{n\in\mathbb{N}}$ est totale dans $\mathscr{C}^0\big([0,1],\mathbb{R}\big)$ pour la norme $\|\cdot\|_{\mathscr{C}^0}$ de la borne supérieure. Puisque l'on a (exercice rapide) :

$$\|f\|_{L^2} \leqslant \|f\|_{\mathscr{C}^0},$$

pour toute $f \in L^2\big([0,1]\big)$, la totalité de cette suite est instantanément *héritée* par L^2.

Étant alors donnée une suite $\big(\alpha_n\big)_{n\geqslant 1}$ de nombres strictement positifs $\alpha_n \in \mathbb{R}_+^*$ qui est strictement croissante : $\alpha_n < \alpha_{n+1}$, on considère la famille, indexée par $n \in \mathbb{N}^*$, des fonctions-monômes :

$$[0,1] \ni x \longmapsto x^{\alpha_n} \in \mathbb{R},$$

qui appartiennent tous au sous-espace vectoriel dense :

$$\mathscr{C}^0\big([0,1],\mathbb{R}\big) \subset L^2\big([0,1],\mathbb{R}\big)$$

des fonctions continues sur le segment $[0,1]$ dans l'espace de Hilbert des fonctions de carré intégrable sur $[0,1]$ muni de la norme standard $\int_0^1 f(x)^2\,dx$ notée $\|f\|_{L^2}$, et on se demande : *quand donc une telle suite est-elle (aussi) totale dans $L^2\big([0,1]\big)$?*

Pour tout $N \in \mathbb{N}^*$, on introduit l'espace vectoriel :

$$E_N := \mathsf{Vect}_{\mathbb{R}}\big(x^{\alpha_1}, x^{\alpha_2}, \dots, x^{\alpha_N}\big).$$

(a) Soit maintenant une fonction-monôme x^m avec $m \in \mathbb{R}_+$ quelconque. Montrer que :

$$\inf_{f\in E_N} \big\|x^m - f\big\|_{L^2} = \frac{1}{\sqrt{2m+1}} \prod_{i=1}^N \left|\frac{\alpha_i - m}{\alpha_i + m + 1}\right|.$$

Indication: En appliquant le résultat de l'exercice précédent, on se ramènera à deux déterminants de Gram que l'on calculera par des techniques d'algèbres linéaire ; notamment, on démontrera, en commençant au besoin par examiner les petites valeurs $N = 1, 2, 3$, que :

$$\begin{vmatrix} \frac{1}{\alpha_1+\alpha_1+1} & \cdots & \frac{1}{\alpha_1+\alpha_N+1} & \frac{1}{\alpha_1+m+1} \\ \cdot\cdot & \cdots & \cdot\cdot & \cdot\cdot \\ \frac{1}{\alpha_N+\alpha_1+1} & \cdots & \frac{1}{\alpha_N+\alpha_N+1} & \frac{1}{\alpha_N+m+1} \\ \frac{1}{m+\alpha_1+1} & \cdots & \frac{1}{m+\alpha_N+1} & \frac{1}{m+m+1} \end{vmatrix} = \frac{\prod_{1\leqslant i<j\leqslant N}\big(\alpha_i-\alpha_j\big)^2 \prod_{1\leqslant i\leqslant N}\big(\alpha_i - m\big)^2}{(2m+1)\prod_{1\leqslant i,j\leqslant N}\big(\alpha_i+\alpha_j+1\big)\prod_{1\leqslant i\leqslant N}\big(\alpha_i+m+1\big)^2}.$$

(b) Montrer que la famille des fonctions-monômes $\big(x^{\alpha_n}\big)_{n\geqslant 1}$ est totale dans l'espace $\mathscr{C}^0\big([0,1],\mathbb{R}\big)$ muni de la norme $\|\cdot\|_{L^2}$ *si et seulement si* — joli Théorème dû à Müntz et à Szász — la série suivante diverge :

$$\sum_{n=1}^{\infty} \frac{1}{\alpha_n} = \infty.$$

Indication: Premièrement, dans la circonstance spéciale (la moins significative) où la suite $\big(\alpha_n\big)_{n\geqslant 1}$ est majorée :

$$\sup_{n\to\infty} \alpha_n = \lim_{n\to\infty} \alpha_n =: A < \infty,$$

de telle sorte que l'on a automatiquement $\sum_n \frac{1}{\alpha_n} = \infty$, montrer que pour tout $m \in \mathbb{R}^+$, on a :

$$\lim_{N\to\infty} \prod_{i=1}^N \left|\frac{\alpha_i - m}{\alpha_i + m + 1}\right| = 0.$$

Deuxièmement, dans la circonstance (plus fréquente et plus naturelle) où $\lim_{n\to\infty} \frac{1}{\alpha_n} = \infty$, montrer tout d'abord que l'on a aussi pour toute constante positive :

$$\lim_{n\to\infty} \frac{1}{\alpha_n + \mathsf{constante}} = \infty.$$

Ensuite, pour tout $m \in \mathbb{R}^+$ fixé quelconque, choisir un entier $N_1 \gg 1$ assez grand pour que $\alpha_i > m$ pour tout $i \geqslant N_1$, prendre $N_2 \geqslant N_1$ arbitraire, et montrer que :

$$\left(\lim_{N_2\to\infty} \prod_{N_1\leqslant i\leqslant N_2} \frac{\alpha_i - m}{\alpha_i + m + 1} = 0\right) \iff \left(\lim_{N_2\to\infty} \sum_{i=N_1}^{N_2} \frac{1}{\alpha_i + m + 1} = \infty\right)$$

(On pourra prendre le logarithme de ce produit et utiliser la majoration $\log(1-x) < -x$ valable pour tout $0 < x < 1$.) Pour conclure, effectuer une synthèse lumineuse qui répond de manière rigoureuse et complète à la question **(b)**, en remplissant tous les raisonnements qui n'ont pas été explicitement indiqués.