Compléments sur les espaces L^p

Quelques sous-espaces denses dans $\mathrm{L}^p(\mathbb{R}^d)$, pour $1 \le p < +\infty$

Rappel : on a traité un exercice qui donnait, dans un cas particulier, le théorème d'approximation suivant : *soit* (X, d) *un espace métrique, soient* $\mathcal{B}_\mathrm{X}$ *sa tribu borélienne et* μ *une mesure sur* $(\mathrm{X}, \mathcal{B}_\mathrm{X})$ *; si* A *est un borélien de* X*, contenu dans un ouvert* W *de* **mesure finie**, $\mu(\mathrm{W}) < +\infty$*, il existe pour tout* $\varepsilon > 0$ *un fermé* F *et un ouvert* V *de* X*, tels que* $\mu(\mathrm{V}) < +\infty$,

$$\mathrm{F} \subset \mathrm{A} \subset \mathrm{V} \quad et \quad \mu(\mathrm{V}) - \mu(\mathrm{F}) = \mu(\mathrm{V} \setminus \mathrm{F}) < \varepsilon.$$

Il en résulte qu'il existe une fonction φ *réelle continue sur* X*, nulle hors de* W*, telle que*

$$0 \le \varphi \le 1 \quad et \quad \int_\mathrm{X} |\mathbf{1}_\mathrm{A} - \varphi| \, d\mu < \varepsilon.$$

Définition. Le *support* d'une fonction continue f définie sur $\mathbb{R}^d$ est l'adhérence de l'ensemble des points où f est non nulle,

$$\mathrm{supp}(f) = \overline{\{x \in \mathbb{R}^d : f(x) \ne 0\}}.$$

On notera $\mathcal{K}(\mathbb{R}^d)$ l'espace vectoriel des fonctions réelles (ou complexes, selon le contexte) définies sur $\mathbb{R}^d$ et dont le support est compact ; dire que $f \in \mathcal{K}(\mathbb{R}^d)$ revient à dire que la fonction f est nulle en dehors d'un certain sous-ensemble borné B de $\mathbb{R}^d$.

Théorème. *L'espace* $\mathcal{K}(\mathbb{R}^d)$ *des fonctions continues à support compact est dense dans* $\mathrm{L}^p(\mathbb{R}^d)$*, pour tout* $p \in [1, +\infty[$.

Le cas de tous les espaces $\mathrm{L}^p(\mathbb{R}^d)$, pour $1 \le p < +\infty$ est essentiellement le même et on écrira la preuve dans le cas $p = 1$.

Démonstration. Il est clair que les fonctions étagées sont denses dans L^1 : pour toute fonction $f \in \mathrm{L}^1(\mathbb{R}^d)$, il existe une suite de fonctions étagées (f_n) telle que $|f_n| \le |f|$ et $f_n \to f$ simplement ; par convergence dominée, f est limite dans L^1 des f_n. Pour montrer le théorème, il suffit de voir que toute indicatrice $\mathbf{1}_\mathrm{A}$ d'un borélien A *de mesure finie* (ce qui équivaut à dire que $\mathbf{1}_\mathrm{A} \in \mathrm{L}^1$) peut être approchée par des fonctions continues à support compact : par linéarité on obtiendra que toute fonction étagée peut être approchée par une fonction continue à support compact, d'où le résultat puisque les fonctions étagées sont denses dans $\mathrm{L}^1(\mathbb{R}^d)$.

Soit A un borélien de $\mathbb{R}^d$, de mesure de Lebesgue $\lambda(\mathrm{A})$ finie ; l'ensemble A est réunion de la suite croissante (A_n), où A_n désigne l'intersection de A avec la boule de rayon n ; on sait alors que $\lambda(\mathrm{A}_n)$ tend vers $\lambda(\mathrm{A})$, ce qui signifie que pour n assez grand on aura

$$\lambda(\mathrm{A}) - \lambda(\mathrm{A}_n) = \int_{\mathbb{R}^d} |\mathbf{1}_\mathrm{A}(x) - \mathbf{1}_{\mathrm{A}_n}(x)| \, dx < \varepsilon.$$

Finalement, il suffit pour montrer le théorème d'approcher, pour la norme de $\mathrm{L}^1(\mathbb{R}^d)$, toute indicatrice de borélien borné par une fonction continue à support compact. Si A est un borélien borné, on peut le placer dans un ouvert borné de la forme $\mathrm{W} =]-a, a[^d$, qui est de mesure finie pour la mesure de Lebesgue. D'après le théorème d'approximation appliqué à $\mathrm{X} = \mathbb{R}^d$, il existe une fonction continue φ sur $\mathbb{R}^d$, nulle hors de W et telle que

$$\int_{\mathbb{R}^d} \bigl|\mathbf{1}_{\mathrm{A}}(x) - \varphi(x)\bigr| \, dx < \varepsilon,$$

ce qui termine la preuve, car le support de φ est contenu dans $[-a, a]^d$, qui est compact.

Remarque. Les fonctions *en escalier* sont denses aussi (en plusieurs dimensions, on appellera ainsi toute combinaison linéaire de fonctions indicatrices $\mathbf{1}_{\mathrm{P}}$ de produits d'intervalles $\mathrm{P} = \prod_{i=1}^d [a_i, b_i] \subset \mathbb{R}^d$).

Pour montrer cette densité, on peut constater qu'une fonction continue à support compact peut être approchée dans L^p $(1 \le p < \infty)$ par des fonctions en escalier (utiliser la continuité uniforme) ; cette méthode n'est pas la plus naturelle, mais elle évite d'avoir à refaire pour les fonctions en escalier un raisonnement analogue à celui qui a donné le *théorème d'approximation* qui commence ce paragraphe. L'exercice suivant est une tentative d'obtenir un énoncé qui soit assez général pour pouvoir s'appliquer à plusieurs situations semblables, mais du coup sa formulation est un peu «tordue».

Exercice. On donne un espace mesuré $(\Omega, \mathcal{A}, \mu)$ et un sous-espace vectoriel E de l'intersection $\mathrm{L}^\infty(\Omega, \mathcal{A}, \mu) \cap \mathrm{L}^1(\Omega, \mathcal{A}, \mu)$; on suppose que E est stable par produit des fonctions. On introduit la classe $\mathcal{D}$ des parties $\mathrm{D} \in \mathcal{A}$ telles que $\mu(\mathrm{D}) < \infty$ et

$$\forall p \in [1, \infty[, \ \forall \varepsilon > 0, \ \exists \varphi \in \mathrm{E}, \quad \|\mathbf{1}_{\mathrm{D}} - \varphi\|_p < \varepsilon.$$

a. On fixe $\mathrm{D}_0 \in \mathcal{D}$. Montrer que la classe $\mathcal{B}(\mathrm{D}_0)$ des parties $\mathrm{B} \in \mathcal{A}$ telles que $\mathrm{B} \cap \mathrm{D}_0 \in \mathcal{D}$ est une tribu de parties de Ω.

Indication : le traitement de l'intersection de deux ensembles amène naturellement au produit de deux fonctions ; on utilisera le fait que l'approximation des indicatrices a lieu dans tous les L^p et on appliquera, par exemple, l'inégalité $\|fg\|_p \le \|f\|_{2p} \, \|g\|_{2p}$ qui résulte de Cauchy-Schwarz).

b. On prend $(\Omega, \mathcal{A}, \mu) = (\mathbb{R}^d, \mathcal{B}_{\mathbb{R}^d}, \lambda)$ et pour E, on prend l'un des espaces suivants :

– l'espace des fonctions continues à support compact sur $\mathbb{R}^d$;

– l'espace des combinaisons linéaires d'indicatrices $\mathbf{1}_{\mathrm{P}}$ de pavés $\mathrm{P} = \prod_{i=1}^d [a_i, b_i]$;

– l'espace des combinaisons linéaires de fonctions de la forme $\prod_{i=1}^d f_i(x_i)$, avec f_i continue à support compact sur $\mathbb{R}$.

Montrer dans chaque cas que $\mathrm{D}_0 = [-a, a]^d$ est dans $\mathcal{D}$ pour tout $a > 0$, et en déduire la densité dans tous les $\mathrm{L}^p(\mathbb{R}^d)$ $(\ 1 \le p < +\infty)$ de l'espace E correspondant.

Inégalité de Hölder

Voir les notes de 2004-2005.

Espaces de Hilbert

Considérons un espace vectoriel E sur $\mathbb{R}$ ou sur $\mathbb{C}$, et posons $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$ pour pouvoir parler des deux cas ensemble. Un *produit scalaire* sur le $\mathbb{K}$-espace vectoriel E est une application de $\mathrm{E} \times \mathrm{E}$ dans $\mathbb{K}$ qui vérifie les trois propriétés suivantes :

– pour tout $y \in \mathrm{E}$, l'application $x \in \mathrm{E} \to \langle x, y \rangle$ est $\mathbb{K}$-linéaire([a]) sur E ;
– on a $\langle y, x \rangle = \overline{\langle x, y \rangle}$ pour tous $x, y \in \mathrm{E}$;
– et enfin $\langle x, x \rangle$ est réel ≥ 0 pour tout $x \in \mathrm{E}$.

La deuxième propriété implique que $\langle x, x \rangle$ est réel pour tout vecteur $x \in \mathrm{E}$. On notera que

$$\begin{aligned} &\langle x+y, x+y \rangle = \langle x, x \rangle + \langle x, y \rangle + \langle y, x \rangle + \langle y, y \rangle = \\ (1)\qquad &= \langle x, x \rangle + \langle x, y \rangle + \overline{\langle x, y \rangle} + \langle y, y \rangle = \langle x, x \rangle + 2\,\mathrm{Re}\,\langle x, y \rangle + \langle y, y \rangle. \end{aligned}$$

L'étude du trinôme $t \in \mathbb{R} \to \langle x + ty, x + ty \rangle \ge 0$ permet de montrer([b]) l'inégalité de Cauchy-Schwarz :

$$\bigl|\langle x, y \rangle\bigr|^2 \le \langle x, x \rangle\, \langle y, y \rangle.$$

Il en résulte en particulier que $x \to \langle x, x \rangle^{1/2}$ est une semi-norme sur E, car

$$\langle x+y, x+y \rangle = \langle x, x \rangle + 2\,\mathrm{Re}\,\langle x, y \rangle + \langle y, y \rangle \le \langle x, x \rangle + 2\,|\langle x, y \rangle| + \langle y, y \rangle \le$$

$$\le \langle x, x \rangle + 2\,\langle x, x \rangle^{1/2} \langle y, y \rangle^{1/2} + \langle y, y \rangle = \bigl(\langle x, x \rangle^{1/2} + \langle y, y \rangle^{1/2}\bigr)^2.$$

Définition. On appelle *espace de Hilbert* un espace vectoriel H sur $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$ muni d'un produit scalaire $(x, y) \to \langle x, y \rangle \in \mathbb{K}$, tel que la semi-norme $x \to \sqrt{\langle x, x \rangle}$ soit une norme sur H, qui rende cet espace **complet**.

Si H est un espace de Hilbert, on note $\|x\| = \sqrt{\langle x, x \rangle}$ pour tout $x \in \mathrm{H}$. L'inégalité de Cauchy-Schwarz dit que $|\langle x, y \rangle| \le \|x\|\, \|y\|$.

Exemple. L'espace $\mathrm{L}^2(\Omega, \mathcal{A}, \mu)$ est un espace de Hilbert pour le produit scalaire

$$\langle f, g \rangle = \int_\Omega f(s) \overline{g(s)}\, d\mu(s).$$

L'espace ℓ^2 est un cas particulier, obtenu lorsque $\Omega = \mathbb{N}$ est muni de la mesure de comptage (définie par $\mu(\{n\}) = 1$ pour tout $n \in \mathbb{N}$).

Si x, y sont deux vecteurs de l'espace H, on a d'après (1)

$$\|x+y\|^2 = \langle x+y, x+y \rangle = \|x\|^2 + 2\,\mathrm{Re}\,\langle x, y \rangle + \|y\|^2.$$

On a de même $\|x-y\|^2 = \|x\|^2 - 2\,\mathrm{Re}\,\langle x, y \rangle + \|y\|^2$, ce qui donne, en additionnant, la *relation du parallélogramme*

$$\|x+y\|^2 + \|x-y\|^2 = 2\,\|x\|^2 + 2\,\|y\|^2.$$

Si on pose $u = x + y$ et $v = x - y$, la relation prend la forme([c])

$$\frac{\|u\|^2 + \|v\|^2}{2} = \Bigl\|\frac{u+v}{2}\Bigr\|^2 + \Bigl\|\frac{u-v}{2}\Bigr\|^2.$$

Lemme 1. *Soient* $(u_1, \ldots, u_n)$ *des vecteurs deux à deux orthogonaux d'un espace de Hilbert* H *; on a*

$$\Big\|\sum_{k=1}^{n} u_k\Big\|^2 = \sum_{k=1}^{n} \|u_k\|^2.$$

En particulier, des vecteurs orthogonaux non nuls *sont linéairement indépendants.*

Démonstration. Le premier point est facile : il suffit de développer le carré scalaire $\langle \sum_{k=1}^n u_k, \sum_{k=1}^n u_k \rangle$. Pour l'indépendance linéaire, supposons que les vecteurs (u_k) soient non nuls, deux à deux orthogonaux, et que $\sum_{k=1}^n c_k u_k = 0$; alors

$$0 = \Big\|\sum_{k=1}^{n} c_k u_k\Big\|^2 = \sum_{k=1}^{n} |c_k|^2 \, \|u_k\|^2$$

donc $|c_k| \, \|u_k\| = 0$ pour tout k, et $c_k = 0$ puisque $\|u_k\| \neq 0$; ainsi, la seule combinaison linéaire nulle est celle dont tous les coefficients (c_k) sont nuls, donc les vecteurs (u_k) sont linéairement indépendants.

Lemme 2. *Soit* $(e_1, \ldots, e_n)$ *une suite orthonormée finie dans un espace de Hilbert* H, *et posons* $\mathrm{F} = \mathrm{Vect}(e_1, \ldots, e_n)$ *; pour tout vecteur* $x \in \mathrm{H}$, *le vecteur*

$$\mathrm{P_F}\, x = \sum_{i=1}^{n} \langle x, e_i \rangle \, e_i$$

est la projection orthogonale de x *sur* F, *c'est-à-dire que* $\mathrm{P_F}\, x \in \mathrm{F}$ *et que le vecteur* $x - \mathrm{P_F}\, x$ *est orthogonal à* F. *On a* $\|x - \mathrm{P_F}\, x\| \leq \|x - z\|$ *pour tout* $z \in \mathrm{F}$ *(le point* $\mathrm{P_F}\, x$ *est le point de* F *le plus proche de* x*) ; on a également* $\|\mathrm{P_F}\, x\| \leq \|x\|$ *(la projection orthogonale diminue la norme des vecteurs).*

Démonstration. Il est évident que $y = \mathrm{P_F}\, x \in \mathrm{F}$, et il est clair que $\langle y, e_j \rangle = \langle x, e_j \rangle$ pour tout $j = 1, \ldots, n$: en effet

$$\langle y, e_j \rangle = \Big\langle \sum_{i=1}^{n} \langle x, e_i \rangle \, e_i, e_j \Big\rangle = \sum_{i=1}^{n} \langle x, e_i \rangle \, \delta_{i,j} = \langle x, e_j \rangle$$

donc $x - y$ est orthogonal à tous les (e_j), $j = 1, \ldots, n$, ce qui implique que $x - y$ est orthogonal à F. Pour la deuxième affirmation, si $z \in \mathrm{F}$ on écrit $x - z = (x - \mathrm{P_F}\, x) + (\mathrm{P_F}\, x - z)$ et on utilise l'orthogonalité de $x - \mathrm{P_F}\, x$ et de $\mathrm{P_F}\, x - z \in \mathrm{F}$,

$$\|x - z\|^2 = \|x - \mathrm{P_F}\, x\|^2 + \|\mathrm{P_F}\, x - z\|^2 \geq \|x - \mathrm{P_F}\, x\|^2.$$

Si on choisit $z = 0 \in \mathrm{F}$, la ligne précédente indique que

$$\|x\|^2 = \|x - \mathrm{P_F}\, x\|^2 + \|\mathrm{P_F}\, x\|^2 \geq \|\mathrm{P_F}\, x\|^2.$$

Lemme. *Soit $(u_n)_{n\in\mathbb{N}}$ une famille de vecteurs deux à deux orthogonaux dans un espace de Hilbert* H *; la série $\sum u_n$ converge dans* H *si et seulement si $\sum_{n\in\mathbb{N}} \|u_n\|^2 < +\infty$, et dans ce cas*
$$\Bigl\|\sum_{n\in\mathbb{N}} u_n\Bigr\|^2 = \sum_{n\in\mathbb{N}} \|u_n\|^2.$$
Si $(e_n)_{n\in\mathbb{N}}$ est une famille orthonormée, la série de vecteurs $\sum c_n e_n$ converge dans H *si et seulement si $\sum_n |c_n|^2 < +\infty$, et dans ce cas on a $\bigl\|\sum_{n\in\mathbb{N}} c_n e_n\bigr\|^2 = \sum_n |c_n|^2$.*

Démonstration. La deuxième partie de l'énoncé résulte immédiatement de la première. Supposons $\sum_{n\geq 0} \|u_n\|^2 < +\infty$, et posons $s_n = \sum_{k=0}^n u_k \in \mathrm{H}$; pour $\mathrm{N} \leq m < n$ on peut écrire
$$s_n - s_m = \sum_{k=m+1}^{n} u_k,$$
et comme les vecteurs (u_k) sont deux à deux orthogonaux, on a
$$\|s_n - s_m\|^2 = \sum_{k=m+1}^{n} \|u_k\|^2 \leq \sum_{k>\mathrm{N}} \|u_k\|^2 = \varepsilon_{\mathrm{N}}^2$$
qui tend vers 0 quand $\mathrm{N} \to +\infty$ (reste d'une série numérique convergente). La suite (s_n) est donc de Cauchy dans l'espace complet H, ce qui entraîne qu'elle converge vers un vecteur $s \in \mathrm{H}$, qui est par définition la somme de la série de vecteurs $\sum_{k\geq 0} u_k$. Par continuité de la norme on obtient
$$\Bigl\|\sum_{k\geq 0} u_k\Bigr\|^2 = \sum_{k\geq 0} \|u_k\|^2.$$
Inversement, la convergence de la série de vecteurs implique que $\|s_n\|^2 = \sum_{k=0}^n \|u_k\|^2$ tend vers $\|\sum_{k\geq 0} u_k\|^2$, c'est-à-dire que la série numérique $\sum_{k\geq 0} \|u_k\|^2$ converge.

Définition. On appelle *base hilbertienne* d'un espace de Hilbert H une famille orthonormée $(e_i)_{i\in\mathrm{I}}$ qui est de plus *totale* dans H, c'est-à-dire que l'espace vectoriel engendré $\mathrm{Vect}(e_i : i \in \mathrm{I})$ est dense dans H.

Proposition. *Supposons que* H *admette une base hilbertienne* dénombrable*. Pour toute énumération $(e_n)_{n\in\mathbb{N}}$ de cette base, et pour tout vecteur x de* H*, on a*
$$x = \sum_{k=0}^{+\infty} \langle x, e_k\rangle\, e_k \quad \text{et} \quad \|x\|^2 = \sum_{k=0}^{+\infty} |\langle x, e_k\rangle|^2.$$

Démonstration. D'après le lemme 2, on sait que la projection orthogonale de x sur $\mathrm{F}_n = \mathrm{Vect}(e_0, \ldots, e_n)$ est égale à $x_n = \sum_{k=0}^n \langle x, e_k\rangle\, e_k$ pour tout $n \geq 0$. Soit $\varepsilon > 0$ donné ; puisque la famille (e_i) est totale, il existe un vecteur $z \in \mathrm{Vect}(e_i)$ tel que $\|x - z\| < \varepsilon$; on peut trouver un entier N_0 tel que $z \in \mathrm{F}_{\mathrm{N}_0}$. Pour tout entier $n \geq \mathrm{N}_0$, on a aussi $z \in \mathrm{F}_n$, et on déduit du lemme 2 que $\|x - x_n\| \leq \|x - z\| < \varepsilon$. Il en résulte que la suite (x_n) tend vers x ; comme les (x_n) sont les sommes partielles de la série $\sum_{k=0}^{+\infty} \langle x, e_k\rangle\, e_k$, le résultat voulu est établi.

On voit qu'une base hilbertienne $(e_n)_{n\geq 0}$ de H est une suite orthonormée qui vérifie pour tout $x \in \mathrm{H}$ la première propriété indiquée dans la proposition précédente. En effet, cette propriété implique clairement que la suite $(e_n)_{n\geq 0}$ doit être totale dans H.

Exercice. Déterminants de Gram. Soient $(x_1, \ldots, x_{n+1})$ des vecteurs d'un espace de Hilbert, tels que $\mathrm{F} = \mathrm{Vect}(x_1, \ldots, x_n)$ soit de dimension n ; montrer que

$$\mathrm{dist}^2(x_{n+1}, \mathrm{F}) = \frac{\det\bigl(\langle x_i, x_j\rangle\bigr)_{i,j=1,\ldots,n+1}}{\det\bigl(\langle x_i, x_j\rangle\bigr)_{i,j=1,\ldots,n}}\,.$$

Exercice : le système de Haar. On définit une fonction «fondamentale» h sur $\mathbb{R}$ par la formule $h = \mathbf{1}_{(0,1/2)} - \mathbf{1}_{(1/2,1)}$. On définit ensuite des fonctions sur $[0,1]$ en posant $h_{0,0}(t) = h(t)$ pour $t \in [0,1]$, puis pour tout $k \geq 0$ et tout $j = 0, \ldots, 2^k - 1$

$$\forall t \in [0,1], \quad h_{k,j}(t) = 2^{k/2}\, h(2^k t - j).$$

Montrer que le système formé de la fonction constante $\mathbf{1}$ et des fonctions $(h_{k,j})$, $k \geq 0$ et $j = 0, \ldots, 2^k - 1$, constitue une base hilbertienne de $\mathrm{L}^2(0,1)$.

Gram-Schmidt

Proposition. *Soit $(v_n)_{n\in\mathbb{N}}$ une suite de vecteurs linéairement indépendants dans un espace de Hilbert* H *; il existe une suite orthonormée $(e_n)_{n\in\mathbb{N}}$ telle que*

$$\mathrm{Vect}(e_0, \ldots, e_n) = \mathrm{Vect}(v_0, \ldots, v_n)$$

pour tout entier $n \geq 0$.

Démonstration : par récurrence sur $n \geq 0$; on commence avec $e_0 = \|v_0\|^{-1} v_0$ (le vecteur v_0 n'est pas nul puisqu'il fait partie d'un système libre). Si $e_0, \ldots, e_n$ sont déjà déterminés, posons $\mathrm{V}_n = \mathrm{Vect}(e_0, \ldots, e_n)$; on a aussi $\mathrm{V}_n = \mathrm{Vect}(v_0, \ldots, v_n)$ par hypothèse de récurrence ; la suite $e_0, \ldots, e_n$ est une base orthonormée de V_n, ce qui permet d'exprimer la projection orthogonale P_n de H sur V_n par

$$\forall x \in \mathrm{H}, \quad \mathrm{P}_n\, x = \sum_{k=0}^{n} \langle x, e_k\rangle\, e_k.$$

Puisque la suite (v_k) est libre, le vecteur v_{n+1} n'est pas dans $\mathrm{V}_n = \mathrm{Vect}(v_0, \ldots, v_n)$, donc le vecteur $x_{n+1} = v_{n+1} - \mathrm{P}_n v_{n+1}$ n'est pas nul, et ce vecteur est orthogonal à V_n, donc à chacun des vecteurs $e_0, \ldots, e_n$. On pose ([d])

$$e_{n+1} = \|x_{n+1}\|^{-1} x_{n+1}.$$

Ce vecteur de norme un est, comme x_{n+1}, orthogonal à chacun des vecteurs $e_0, \ldots, e_n$. Comme $e_{n+1} = a\, v_{n+1} + w_n$, $a \neq 0$, $w_n \in \mathrm{V}_n$ et $\mathrm{Vect}(e_0, \ldots, e_n) = \mathrm{Vect}(v_0, \ldots, v_n)$, on vérifie que $\mathrm{Vect}(e_0, \ldots, e_n, e_{n+1}) = \mathrm{Vect}(v_0, \ldots, v_n, v_{n+1})$.

Lemme. *Si* E *est un espace normé séparable et de dimension infinie, on peut trouver dans* E *une suite de vecteurs* $(v_n)_{n\in\mathbb{N}}$ *indépendants telle que* $\mathrm{Vect}(v_n : n \ge 0)$ *soit dense dans* E.

Démonstration. Puisque E est séparable on peut trouver un ensemble dénombrable dense, qu'on énumérera dans une suite $(x_n)_{n\in\mathbb{N}}$; on définit par récurrence une sous-suite d'entiers (n_k) telle que pour tout $k \ge 0$, la propriété suivante soit satisfaite :

— les vecteurs $x_{n_0}, \ldots, x_{n_k}$ sont libres, et l'espace vectoriel V_{k+1} qu'ils engendrent contient tous les x_j pour lesquels $0 \le j \le n_k$.

On commence avec $\mathrm{V}_0 = \{0\}$. Si $n_0 < \ldots < n_k$ sont déjà définis et possèdent la propriété ci-dessus, on note que V_{k+1} est fermé, comme sous-espace de dimension finie de E, et différent de E qui est supposé de dimension infinie ; il est donc impossible que V_{k+1} contienne tous les termes de la suite (x_n), qui est dense dans E. On désignera par n_{k+1} le plus petit entier n tel que $x_n \notin \mathrm{V}_{k+1}$; évidemment, $n_{k+1} > n_k$. Il est clair que le nouveau système est libre, puisque les premiers vecteurs étaient libres et que le dernier n'est pas dans l'espace V_{k+1} engendré par les premiers ; par définition de n_{k+1}, on a $x_n \in \mathrm{V}_{k+1}$ pour tout $n < n_{k+1}$, et $x_n \in \mathrm{V}_{k+2}$ pour tout $n \le n_{k+1}$.

On considère $v_k = x_{n_k}$; on voit que $\mathrm{V} = \mathrm{Vect}(v_k : k \ge 0)$ contient tous les vecteurs x_n, donc V est dense dans E.

Proposition. *Soit* H *un espace de Hilbert séparable et de dimension infinie ; il existe une base hilbertienne* $(e_n)_{n\in\mathbb{N}}$ *de* H *indexée par* $\mathbb{N}$.

Démonstration. Il résulte du lemme précédent qu'il existe une suite (v_n) libre telle que $\mathrm{Vect}(v_n : n \ge 0)$ soit dense dans H. La suite orthonormée $(e_n)_{n\in\mathbb{N}}$ construite par Gram-Schmidt à partir de la suite $(v_n)_{n\in\mathbb{N}}$ est telle que $\mathrm{Vect}(e_n : n \ge 0)$ est dense dans H : c'est la définition d'une base hilbertienne.

Si $(e_n)_{n\in\mathbb{N}}$ est une base hilbertienne de H, on a vu que pour tout vecteur $x \in \mathrm{H}$,

$$x = \sum_{n=0}^{+\infty} \langle x, e_n\rangle\, e_n \quad \text{et} \quad \|x\|^2 = \sum_{n=0}^{+\infty} |\langle x, e_n\rangle|^2.$$

Les *polynômes orthogonaux* sont introduits de cette façon. Par exemple, les *polynômes d'Hermite* sont obtenus par le procédé de Gram-Schmidt appliqué à la suite des fonctions monôme $t \to t^n$, $n \ge 0$, dans l'espace de Hilbert $\mathrm{L}^2(\mathbb{R}, \mathrm{e}^{-x^2}\, dx)$. En réalité, les polynômes d'Hermite ne sont pas les fonctions de norme un données par Gram-Schmidt, mais plutôt les polynômes unitaires (dont le coefficient de plus haut degré est égal à 1) qui leur sont proportionnels.

Théorème de projection. *Si* C *est un convexe fermé non vide d'un espace de Hilbert* H*, il existe pour tout* $x \in \mathrm{H}$ *un point* $y \in \mathrm{C}$ *qui réalise la distance de* x *à* C,

$$\|x - y\| = \mathrm{dist}(x, \mathrm{C}) = \inf\{\|x - c\| : c \in \mathrm{C}\}.$$

Il n'existe qu'un seul point y *avec cette propriété. On l'appelle la* projection de x sur C.

Démonstration. On peut trouver une suite (c_n) de points de C telle que $\|x - c_n\|$ tende vers $d = \mathrm{dist}(x, \mathrm{C})$. Avec la relation du parallélogramme, on verra que la suite (c_n) est de Cauchy. Puisque H est complet, la suite (c_n) converge vers un vecteur y, et $y \in \mathrm{C}$ parce que C est fermé. Comme la distance est une fonction continue, on obtient $\|x-y\| = d$ à la limite.

Il reste à montrer le caractère Cauchy de la suite (c_n). Soit $\varepsilon > 0$ donné et soit n un entier assez grand pour que $\|x-c_j\|^2 < d^2+\varepsilon^2$ pour tout $j \ge n$; si $j,k \ge n$, désignons par m le milieu de c_j et c_k ; on a $m \in \mathrm{C}$ puisque l'ensemble C est convexe, donc $\|x-m\| \ge d$; appliquons la relation du parallélogramme avec $u = x-c_j$, $v = x-c_k$; alors $(u+v)/2 = x-m$ et $(u-v)/2 = (c_k-c_j)/2$,

$$d^2+\varepsilon^2 \ge \frac{\|x-c_j\|^2+\|x-c_k\|^2}{2} = \|x-m\|^2 + \Bigl\|\frac{c_k-c_j}{2}\Bigr\|^2 \ge d^2 + \Bigl\|\frac{c_k-c_j}{2}\Bigr\|^2$$

ce qui montre que $\|c_k-c_j\| \le 2\varepsilon$. Pour l'unicité, on reprend le calcul de la ligne précédente en remplaçant c_j, c_k par deux candidats projection y et y'.

Corollaire 1. *Si* F *est un sous-espace vectoriel fermé d'un espace de Hilbert* H*, il existe pour tout* $x \in \mathrm{H}$ *un point* $y \in \mathrm{F}$ *tel que* $x-y \perp \mathrm{F}$ *; le point* y *est appelé* projection orthogonale *de* x *sur* F*. L'application* $\mathrm{P_F}$*, qui associe à chaque* $x \in \mathrm{H}$ *sa projection* $\mathrm{P_F}\, x = y$ *sur* F*, est linéaire et de norme* 1*, sauf si* $\mathrm{F} = \{0\}$ *(dans ce cas* $\mathrm{P_F} = 0$*).*

Démonstration. Si $y \in \mathrm{F}$ et $x-y \perp \mathrm{F}$, il est clair que y est le point de F le plus proche de x (si $z \in \mathrm{F}$, $\|x-z\|^2 = \|x-y\|^2+\|y-z\|^2 \ge \|x-y\|^2$) donc y est la projection de x sur le convexe fermé F. Inversement, si y est le point de F le plus proche de x et si $v \in \mathrm{F}$, on a $y+tv \in \mathrm{F}$ pour tout réel t, donc

$$\|x-y-tv\|^2 \ge \|x-y\|^2\,;$$

il en résulte que

$$-2t\ \mathrm{Re}\,\langle x-y, v\rangle + t^2\|v\|^2 \ge 0$$

pour tout réel t, ce qui n'est possible que si $\mathrm{Re}\,\langle x-y,v\rangle = 0$. Si l'espace H est réel, on sait ainsi que $x-y$ est orthogonal à tous les vecteurs v de F ; dans le cas complexe, on peut ($^{\mathrm{e}}$) appliquer ce qui précède au vecteur $\mathrm{i}v \in \mathrm{F}$ pour obtenir aussi $\mathrm{Im}\,\langle x-y,v\rangle = 0$, et finalement $\langle x-y,v\rangle = 0$ pour tout $v \in \mathrm{F}$.

Si $x_1, x_2 \in \mathrm{H}$ ont pour projections y_1 et $y_2 \in \mathrm{F}$, le vecteur $\lambda_1 y_1 + \lambda_2 y_2$ est dans F et on voit facilement que $(\lambda_1x_1+\lambda_2x_2)-(\lambda_1y_1+\lambda_2y_2)$ est orthogonal à F, ce qui montre que $\mathrm{P_F}(\lambda_1x_1+\lambda_2x_2) = \lambda_1\mathrm{P_F}(x_1)+\lambda_2\mathrm{P_F}(x_2)$. On sait que la projection diminue la norme, car $\|x\|^2 = \|x-\mathrm{P_F}\,x\|^2+\|\mathrm{P_F}\,x\|^2 \ge \|\mathrm{P_F}\,x\|^2$; on a donc $\|\mathrm{P_F}\| \le 1$; si $\mathrm{F} \ne \{0\}$, on applique $\mathrm{P_F}$ à un vecteur v non nul, élément de F : on doit avoir $\|v\| = \|\mathrm{P_F}v\| \le \|\mathrm{P_F}\|\,\|v\|$, ce qui implique $\|\mathrm{P_F}\| \ge 1$.

Corollaire 2. *Si* A *est un sous-ensemble d'un espace de Hilbert* H*, le sous-espace vectoriel* Vect(A) *engendré par* A *est dense dans* H *si et seulement si* 0 *est le seul vecteur de* H *qui soit orthogonal à la partie* A.

Démonstration. Si v est non nul et orthogonal à A, l'ensemble

$$v^\perp = \{w \in \mathrm{H} : w \perp v\}$$

est un sous-espace vectoriel fermé distinct de H (il ne contient pas v) qui contient A ; puisque le sous-espace vectoriel $v^\perp$ contient A, il contient aussi le sous-espace engendré Vect(A), et comme $v^\perp$ est fermé, il contient l'adhérence de Vect(A) ; par conséquent Vect(A) n'est pas dense dans H, puisque $\overline{\mathrm{Vect(A)}} \subset v^\perp \ne \mathrm{H}$.

Si Vect(A) n'est pas dense, son adhérence F est un sous-espace vectoriel (petit exercice), fermé et distinct de H. Si x est choisi hors de F et si y est la projection de x sur F, le vecteur $v = x-y$ est non nul et il est orthogonal à F, donc à A.

On a déjà introduit la terminologie *famille sommable* dans le cas de réels positifs au chapitre Intégration ; une famille de *vecteurs* $(v_i)_{i\in \mathrm{I}}$ d'un espace normé est dite *sommable* s'il existe un vecteur x possédant la propriété suivante : pour tout $\varepsilon > 0$, on peut trouver un ensemble fini $\mathrm{J}_0 \subset \mathrm{I}$ tel que pour tout sous-ensemble fini $\mathrm{J} \supset \mathrm{J}_0$, on ait

$$\Bigl\| x - \sum_{j\in \mathrm{J}} v_j \Bigr\| < \varepsilon.$$

On note dans ce cas

$$x = \sum_{i\in \mathrm{I}} v_i$$

et on dit que le vecteur x est la *somme* de la famille sommable $(v_i)_{i\in \mathrm{I}}$.

Lemme 3 : inégalité de Bessel. *Soient* H *un espace de Hilbert et* $(e_i)_{i\in \mathrm{I}}$ *une famille orthonormée dans* H *; pour tout* $x \in \mathrm{H}$ *la famille* $(|\langle x, e_i\rangle|^2)_{i\in \mathrm{I}}$ *est sommable et*

$$\sum_{i\in \mathrm{I}} |\langle x, e_i\rangle|^2 \le \|x\|^2.$$

Démonstration. Pour qu'une famille $(u_i)_{i\in \mathrm{I}}$ de réels ≥ 0 soit sommable, il faut et il suffit qu'il existe un majorant M pour l'ensemble des sommes finies $\sum_{i\in \mathrm{J}} u_i$, $\mathrm{J} \subset \mathrm{I}$, J fini, et on a alors $\sum_{i\in \mathrm{I}} u_i \le \mathrm{M}$. Il suffit donc de montrer le résultat du lemme pour une suite finie $e_1, \dots, e_n$. On a vu que si on pose $y = \sum_{i=1}^n \langle x, e_i\rangle\, e_i$, le vecteur $x - y$ est orthogonal au sous-espace $\mathrm{F} = \mathrm{Vect}(e_1, \dots, e_n)$, et

$$\|x\|^2 \ge \|y\|^2 = \sum_{i=1}^n |\langle x, e_i\rangle|^2$$

(lemme 2) d'où le résultat.

Lemme 4. *Soit* $(u_i)_{i\in \mathrm{I}}$ *une famille orthogonale dans un espace de Hilbert* H *; la famille est sommable dans* H *si et seulement si* $\sum_{i\in \mathrm{I}} \|u_i\|^2 < +\infty$*, et dans ce cas*

$$\Bigl\|\sum_{i\in \mathrm{I}} u_i\Bigr\|^2 = \sum_{i\in \mathrm{I}} \|u_i\|^2.$$

Si $(e_i)_{i\in \mathrm{I}}$ *est une famille orthonormée, la famille de vecteurs* $(c_i e_i)_{i\in \mathrm{I}}$ *est sommable dans* H *si et seulement si* $\sum_{i\in \mathrm{I}} |c_i|^2 < +\infty$*, et dans ce cas on a* $\bigl\|\sum_{i\in \mathrm{I}} c_i e_i\bigr\|^2 = \sum_{i\in \mathrm{I}} |c_i|^2$.

Démonstration. Supposons $\sum_{i\in \mathrm{I}} \|u_i\|^2 < +\infty$. On peut trouver une suite croissante d'ensembles finis $\mathrm{I}_n \subset \mathrm{I}$ telle que la suite croissante $s_n = \sum_{i\in \mathrm{I}_n} \|u_i\|^2$ tende vers la somme $s = \sum_{i\in \mathrm{I}} \|u_i\|^2$ de la famille $(\|u_i\|^2)_{i\in \mathrm{I}}$. Posons $\mathrm{U}_n = \sum_{i\in \mathrm{I}_n} u_i$ pour tout $n \ge 0$. Si $m < n$ on a par orthogonalité

$$\|\mathrm{U}_n - \mathrm{U}_m\|^2 = \sum_{i\in \mathrm{I}_n \setminus \mathrm{I}_m} \|u_i\|^2 = s_n - s_m. \tag{2}$$

A partir de là, il est clair que la suite (U_n) est de Cauchy dans H, donc converge vers un vecteur $x \in \mathrm{H}$. Ce vecteur x est la *somme* de la famille $(u_i)_{i\in\mathrm{I}}$: pour tout $\varepsilon > 0$ donné, on peut trouver un sous-ensemble J_0 fini dans I tel que $\|x - \sum_{i\in\mathrm{J}} u_i\| < \varepsilon$ pour tout J fini contenant J_0 ; il suffit en effet de choisir $\mathrm{J}_0 = \mathrm{I}_m$ pour un entier m tel que $s - s_m < \varepsilon^2/4$. En passant à la limite dans (2) quand $\mathrm{U}_n \to x$, on obtient

$$\|x - \mathrm{U}_m\|^2 = s - s_m < \varepsilon^2/4,$$

c'est-à-dire que $\big\|x - \sum_{j\in\mathrm{J}_0} u_j\big\| < \varepsilon/2$. Si J est un sous-ensemble fini de I contenant J_0, on aura par orthogonalité

$$s \ge \Big\|\sum_{j\in\mathrm{J}} u_j\Big\|^2 = \Big\|\sum_{j\in\mathrm{J}_0} u_j\Big\|^2 + \Big\|\sum_{j\in\mathrm{J}\setminus\mathrm{J}_0} u_j\Big\|^2 \ge s - \varepsilon^2/4 + \Big\|\sum_{j\in\mathrm{J}\setminus\mathrm{J}_0} u_j\Big\|^2 ;$$

ceci implique $\|\sum_{j\in\mathrm{J}\setminus\mathrm{J}_0} u_j\| < \varepsilon/2$ et $\|x - \sum_{j\in\mathrm{J}} u_j\| < \varepsilon$.

La norme de la somme de la famille s'obtient en passant à la limite dans l'égalité du lemme 1, qui donne $\|\mathrm{U}_n\|^2 = s_n$ pour tout $n \ge 0$. Si on suppose réciproquement que la famille est sommable, il existe un vecteur $x \in \mathrm{H}$ et un ensemble fini J_0 tels que $\|x - \sum_{j\in\mathrm{J}} u_j\| \le 1$ pour tout J fini contenant J_0, ce qui montre que

$$\sum_{j\in\mathrm{J}} \|u_j\|^2 = \Big\|\sum_{j\in\mathrm{J}} u_j\Big\|^2 \le \big(\|x\| + 1\big)^2$$

pour tout J fini contenant J_0, et ceci implique $\sum_{i\in\mathrm{I}} \|u_i\|^2 \le \big(\|x\| + 1\big)^2 < +\infty$.

L'un des intérêts (peut-être mineur) de la notion de *famille sommable* est de permettre d'utiliser la notion de base hilbertienne $(e_i)_{i\in\mathrm{I}}$ dans le cas le plus général (l'ensemble d'indices I peut être non dénombrable) pour représenter les vecteurs de l'espace, sans faire de périphrases : tout vecteur x de H est la *somme* $x = \sum_{i\in\mathrm{I}} \langle x, e_i\rangle\, e_i$ de la famille sommable $(\langle x, e_i\rangle\, e_i)_{i\in\mathrm{I}}$. En effet, la famille est sommable par la conjonction des lemmes 3 et 4 ; si y désigne la somme de la famille, on voit que $x - y$ est orthogonal à tous les e_i, donc à leurs combinaisons linéaires, donc à H entier puisque (e_i) est totale, donc $x - y$ est le vecteur nul. Ainsi, on a bien $x = y$.

Pour tout ensemble d'indices I, on a un espace de Hilbert $\ell^2(\mathrm{I})$ des familles $(c_i)_{i\in\mathrm{I}}$ de scalaires telles que $\sum_{i\in\mathrm{I}} |c_i|^2 < +\infty$, qui a une base hilbertienne canonique qu'on devine ; tout espace de Hilbert H est isomorphe à l'un de ces espaces $\ell^2(\mathrm{I})$ d'après le théorème qui suit. En effet, si H admet une base hilbertienne $(e_i)_{i\in\mathrm{I}}$, l'égalité de Bessel-Parseval montre que H est isométriquement isomorphe à $\ell^2(\mathrm{I})$, par l'application linéaire qui associe à chaque vecteur $x \in \mathrm{H}$ la famille $(\langle x, e_i\rangle)_{i\in\mathrm{I}} \in \ell^2(\mathrm{I})$ formée par les coordonnées de x dans la base $(e_i)_{i\in\mathrm{I}}$.

Pour la culture, démontrons le théorème hors-programme qui suit.

Théorème. *Tout espace de Hilbert admet des bases hilbertiennes.*

Démonstration. Considérons un espace de Hilbert H, et désignons par $\mathcal{U}$ l'ensemble de toutes les parties $\mathrm{U} \subset \mathrm{H}$ qui vérifient les deux propriétés suivantes :

– si $x \in \mathrm{U}$, $\|x\| = 1$;
– si $x, y \in \mathrm{U}$ et $x \ne y$, alors $x \perp y$.

On ordonne $\mathcal{U}$ par l'inclusion des parties de H ; muni de cet ordre, l'ensemble $\mathcal{U}$ est inductif : pour toute famille $(\mathrm{U}_i)_{i\in\mathrm{I}}$ d'éléments de $\mathcal{U}$ qui est totalement ordonnée par inclusion, on peut trouver dans $\mathcal{U}$ un majorant pour cette famille, à savoir le sous-ensemble

$$\mathrm{U} = \bigcup_{i\in\mathrm{I}} \mathrm{U}_i \subset \mathrm{H},$$

pour lequel on montre facilement que $\mathrm{U} \in \mathcal{U}$. D'après le lemme de Zorn, l'ensemble $\mathcal{U}$ admet des éléments maximaux ; soit U_0 un tel élément maximal ; les deux propriétés qui définissent la classe $\mathcal{U}$ montrent que les éléments de U_0 constituent un système orthonormé ; on va de plus montrer que U_0 engendre un sous-espace dense dans H, c'est-à-dire que U_0 est une base hilbertienne de H.

Considérons le sous-espace vectoriel $\mathrm{Vect}(\mathrm{U}_0)$; s'il n'est pas dense dans H, on peut trouver un vecteur v_1 de norme un, orthogonal à U_0 (prendre v non nul, orthogonal à U_0, donné par le corollaire 2, puis $v_1 = \|v\|^{-1}v$). Alors $\mathrm{U}_0 \cup \{v_1\}$ est un élément de $\mathcal{U}$ qui contredit la maximalité de U_0. On en déduit que $\mathrm{H}_0 = \mathrm{H}$, et que U_0 fournit ([f]) une base hilbertienne de H.

Notes.

([a]) Il n'y a pas accord universel sur le choix du côté qui doit donner la $\mathbb{C}$-linéarité (l'autre côté étant alors *anti*linéaire). Les mathématiciens proches de la physique adoptent souvent la convention opposée de celle que nous avons choisie (histoire de *bra* et de *ket*, pour ceux qui connaissent).

([b]) Pour prouver cette inégalité de Cauchy-Schwarz, écrivons d'abord

$$|\langle x, y\rangle| = a\,\langle x, y\rangle$$

où a est un scalaire tel que $|a| = 1$, de sorte que $\langle ax, y\rangle = a\,\langle x, y\rangle = |\langle x, y\rangle|$; il suffit ensuite d'écrire $\langle \lambda ax - \mu y, \lambda ax - \mu y\rangle \ge 0$, en choisissant deux nombres λ, μ réels tels que $\lambda > \langle y, y\rangle^{1/2}$ et $\mu > \langle x, x\rangle^{1/2}$. On obtient en développant le carré scalaire

$$0 \le \lambda^2 |a|^2 \langle x, x\rangle - 2\lambda\mu\,\mathrm{Re}\,\langle ax, y\rangle + \mu^2\langle y, y\rangle < 2\,\lambda^2\,\mu^2 - 2\lambda\mu\big|\langle x, y\rangle\big|.$$

Ainsi $\big|\langle x, y\rangle\big| \le \lambda\,\mu$, et il ne reste qu'à faire décroître λ et μ vers $\langle y, y\rangle^{1/2}$ et $\langle x, x\rangle^{1/2}$, pour obtenir à la limite l'inégalité de Cauchy-Schwarz.

([c]) On peut aussi donner pour la relation du parallélogramme la forme

$$\|u\|^2 + \|v\|^2 = \Big\|\frac{u+v}{\sqrt{2}}\Big\|^2 + \Big\|\frac{u-v}{\sqrt{2}}\Big\|^2$$

qui indique que la « rotation d'angle $-\pi/4$ »

$$\begin{pmatrix} \frac{1}{\sqrt{2}}\,\mathrm{Id}_\mathrm{H} & \frac{1}{\sqrt{2}}\,\mathrm{Id}_\mathrm{H} \\ -\frac{1}{\sqrt{2}}\,\mathrm{Id}_\mathrm{H} & \frac{1}{\sqrt{2}}\,\mathrm{Id}_\mathrm{H} \end{pmatrix}$$

définit une isométrie dans l'espace de Hilbert $\mathrm{H} \times \mathrm{H}$. Plus généralement toute matrice orthogonale $\mathrm{A} \in \mathrm{O}(n)$ donne une isométrie de H^n, qu'on munit de la norme hilbertienne définie ainsi : si $\mathbf{x} = (x_1, \ldots, x_n) \in \mathrm{H}^n$, on pose

$$\|\mathbf{x}\| = \Big(\sum_{i=1}^{n} \|x_i\|^2\Big)^{1/2}.$$

([d]) On pourra noter le caractère très algorithmique du procédé de Gram-Schmidt.

([e]) Dans le cas complexe, sous-espace vectoriel signifie bien entendu $\mathbb{C}$-sous-espace : si $x \in \mathrm{F}$, alors $\mathrm{i}x \in \mathrm{F}$.

([f]) Si H est réduit à $\{0\}$, la classe $\mathcal{U}$ est réduite à la partie vide, qui est aussi l'élément maximal U_0 de la preuve ; on peut considérer que la base vide est une base orthonormée de $\{0\}$... Si on n'aime pas ces considérations tordues, il est préférable d'exiger que H soit différent de $\{0\}$.