2. Espaces de Hilbert

2.1. Produits scalaires

Définition 2.1.1. Soient X et Y deux espaces vectoriels complexes ; une application $f : \mathrm{X} \to \mathrm{Y}$ est dite *antilinéaire* si, pour tous $x, y \in \mathrm{X}$ et tout $\lambda \in \mathbb{C}$ on a $f(x+y) = f(x) + f(y)$ et $f(\lambda x) = \overline{\lambda} f(x)$.

On notera que la composition de deux (ou d'un nombre pair) d'applications antilinéaires est une application linéaire ; de plus, la composition d'une linéaire et d'une antilinéaire est antilinéaire.

Définition 2.1.2. Soit X un espace vectoriel complexe ; on appelle *forme sesquilinéaire* sur X une application $\mathrm{B} : \mathrm{X} \times \mathrm{X} \to \mathbb{C}$ telle que, pour tout $y \in \mathrm{X}$, l'application $x \to \mathrm{B}(x, y)$ soit linéaire et telle que pour tout $x \in \mathrm{X}$, l'application $y \to \mathrm{B}(x, y)$ soit antilinéaire (de X dans $\mathbb{C}$).

Rappelons qu'une forme bilinéaire B sur un espace vectoriel réel X est dite *symétrique* si, pour tous $x, y \in \mathrm{X}$, on a $\mathrm{B}(y, x) = \mathrm{B}(x, y)$.

Proposition 2.1.1 : identité de polarisation.

(i) *Soient* X *un espace vectoriel complexe et* B *une forme sesquilinéaire sur* X *; pour tous* $x, y \in \mathrm{X}$ *on a*

$$4\,\mathrm{B}(x, y) = \mathrm{B}(x+y, x+y) - \mathrm{B}(x-y, x-y) + i\mathrm{B}(x+iy, x+iy) - i\mathrm{B}(x-iy, x-iy).$$

(ii) *Soient* X *un espace vectoriel réel et* B *une forme bilinéaire symétrique sur* X *; pour tous* $x, y \in \mathrm{X}$ *on a*

$$4\,\mathrm{B}(x, y) = \mathrm{B}(x+y, x+y) - \mathrm{B}(x-y, x-y).$$

En particulier, pour connaître une forme sesquilinéaire ou une forme bilinéaire symétrique B sur X, il suffit de connaître $\mathrm{B}(x, x)$ pour tout $x \in \mathrm{X}$.

Corollaire 2.1.2. *Soient* X *un espace vectoriel complexe et* B *une forme sesquilinéaire sur* X *; les conditions suivantes sont équivalentes :*

(i) *pour tous* $x, y \in \mathrm{X}$ *on a* $\mathrm{B}(y, x) = \overline{\mathrm{B}(x, y)}$ *;*
(ii) *pour tout* $x \in \mathrm{X}$*, on a* $\mathrm{B}(x, x) \in \mathbb{R}$*.*

Démonstration. Posons $\mathrm{S}(x, y) = \mathrm{B}(x, y) - \overline{\mathrm{B}(y, x)}$; c'est une forme sesquilinéaire. Par la proposition 1, S est nulle si et seulement si, pour tout $x \in \mathrm{X}$, on a $\mathrm{S}(x, x) = 0$, ce qui est bien le cas.

//

Soit X un espace vectoriel complexe ; on appelle *forme hermitienne* sur X une forme sesquilinéaire vérifiant les conditions équivalentes du corollaire 2. On peut résumer ces conditions ainsi : la forme φ sur $\mathrm{X} \times \mathrm{X}$ est hermitienne si elle vérifie les deux conditions suivantes :

- pour tout $y \in \mathrm{X}$, l'application $x \to \varphi(x, y)$ est $\mathbb{C}$-linéaire sur X ;
- pour tous $x, y \in \mathrm{X}$, on a $\varphi(y, x) = \overline{\varphi(x, y)}$.

Une forme hermitienne B sur un espace vectoriel complexe X est dite *positive* si, pour tout $x \in \mathrm{X}$, le nombre $\mathrm{B}(x, x)$ est réel ≥ 0. Rappelons qu'une forme bilinéaire symétrique B sur un espace vectoriel réel X est dite *positive* si, pour tout $x \in \mathrm{X}$, on a $\mathrm{B}(x, x) \geq 0$.

Convenons d'appeler *produit scalaire* une forme symétrique positive sur un espace réel ou une forme hermitienne positive sur un espace complexe. Le plus souvent, nous noterons les produits scalaires $(x, y) \to \langle x, y \rangle$.

Proposition 2.1.3 : inégalité de Cauchy-Schwarz. *Soit* X *un espace vectoriel muni d'un produit scalaire ; pour tous* $x, y \in \mathrm{X}$ *on a*

$$|\langle x, y \rangle|^2 \leq \langle x, x \rangle \, \langle y, y \rangle.$$

Démonstration. Soit $u \in \mathbb{K}$ de module 1 tel que $u \, \langle x, y \rangle = |\langle x, y \rangle|$; pour $t \in \mathbb{R}$, le produit scalaire $\langle ux + ty, ux + ty \rangle$ est positif. Or

$$\langle ux + ty, ux + ty \rangle = \langle ux, ux \rangle + 2t \operatorname{Re}\langle ux, y \rangle + t^2 \langle y, y \rangle = \langle x, x \rangle + 2t |\langle x, y \rangle| + t^2 \langle y, y \rangle.$$

Ce polynôme du deuxième degré en t est positif pour tout $t \in \mathbb{R}$, donc son discriminant est négatif ou nul. Cela donne $|\langle x, y \rangle|^2 - \langle x, x \rangle \, \langle y, y \rangle \leq 0$.

//

Corollaire 2.1.4. *Soit* X *un espace vectoriel muni d'un produit scalaire ; l'application* $x \to \langle x, x \rangle^{1/2}$ *est une semi-norme sur* X.

Démonstration. Pour tous $x, y \in \mathrm{X}$, on a

$$\begin{aligned} \langle x + y, x + y \rangle &= \langle x, x \rangle + \langle y, y \rangle + \langle x, y \rangle + \overline{\langle x, y \rangle} \\ &\leq \langle x, x \rangle + \langle y, y \rangle + 2|\langle x, y \rangle| \leq (\langle x, x \rangle^{1/2} + \langle y, y \rangle^{1/2})^2 \end{aligned}$$

par la proposition 3.

//

Notons encore une relation utile, appelée la *relation du parallélogramme*,

$$\langle x + y, x + y \rangle + \langle x - y, x - y \rangle = 2 \big(\langle x, x \rangle + \langle y, y \rangle \big).$$

2.2. Espaces de Hilbert, orthogonalité, bases

Un *produit scalaire* sur un espace vectoriel réel ou complexe X est donc une application $(x, y) \to \langle x, y\rangle$ de $\mathrm{X} \times \mathrm{X}$ dans $\mathbb{K}$ telle que

$x \in \mathrm{X} \to \langle x, y\rangle$ est $\mathbb{K}$-linéaire pour tout $y \in \mathrm{X}$ fixé,
$\langle y, x\rangle = \overline{\langle x, y\rangle}$ pour tous $x, y \in \mathrm{X}$,
$\langle x, x\rangle \geq 0$ pour tout $x \in \mathrm{X}$.

Certains auteurs exigent qu'un produit scalaire vérifie $\langle x, x\rangle > 0$ pour tout $x \neq 0_{\mathrm{X}}$. Dans ce cas la semi-norme $x \to \sqrt{\langle x, x\rangle}$ du corollaire précédent est une norme sur l'espace X.

On appelle *espace préhilbertien* un espace vectoriel X (réel ou complexe) muni d'un produit scalaire tel que la semi-norme $p(x) = \sqrt{\langle x, x\rangle}$ soit une norme sur X. Tout espace préhilbertien sera considéré comme espace normé, muni de la norme ci-dessus, qui sera notée simplement $\|x\|$ désormais.

Proposition 2.2.1. *Soit* X *un espace préhilbertien ; pour tout vecteur* $y \in \mathrm{X}$ *la forme linéaire* $\ell_y : x \to \langle x, y\rangle$ *est continue de* X *dans* $\mathbb{K}$. *L'application* $y \to \ell_y$ *est antilinéaire et isométrique de* X *dans* X^*.

Démonstration. Pour $x \in \mathrm{X}$ on a $|\ell_y(x)| \leq \|x\|\,\|y\|$ par la proposition 1.3, donc l'application linéaire ℓ_y est continue et $\|\ell_y\| \leq \|y\|$. Or $\|y\|^2 = \ell_y(y) \leq \|\ell_y\|\,\|y\|$, d'où l'on déduit que $\|y\| = \|\ell_y\|$. On vérifie sans peine que l'application $y \to \ell_y$ est antilinéaire.

//

Définition 2.2.1. On appelle *espace de Hilbert* un espace vectoriel H (réel ou complexe) muni d'un produit scalaire $(x, y) \to \langle x, y\rangle$ 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 notera $\|x\| = \sqrt{\langle x, x\rangle}$ pour tout $x \in \mathrm{H}$. L'inégalité de Cauchy-Schwarz s'écrit alors $|\langle x, y\rangle| \leq \|x\|\,\|y\|$.

Exemple 2.2.2. L'espace $\mathrm{L}_2(\Omega, \mu)$ 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 $n \in \mathbb{N}$).

Définition 2.2.3. Soit H un espace de Hilbert ; on dit que les vecteurs x et y de H sont *orthogonaux* si $\langle x, y\rangle = 0$. Soit $(x_n)_{n\geq 0}$ une suite infinie de vecteurs de H ou bien $(x_1, \ldots, x_{\mathrm{N}})$ une suite finie ; on dit que la suite est *orthogonale* si les x_n sont deux à deux orthogonaux, c'est à dire si $\langle x_m, x_n\rangle = 0$ lorsque $m \neq n$; on dit que c'est une *suite orthonormée* si de plus, pour tout n, on a $\|x_n\| = 1$.

Si x est orthogonal à $y_1, \ldots, y_n$, alors x est orthogonal à toutes les combinaisons linéaires de $y_1, \ldots, y_n$ d'après la linéarité du produit scalaire par rapport à sa première variable. Le vecteur x est donc orthogonal au sous-espace vectoriel engendré

$$\mathrm{F} = \mathrm{Vect}(y_1, \ldots, y_n).$$

Si x est orthogonal à tous les vecteurs d'un ensemble A, alors x est aussi orthogonal à l'adhérence de A (parce que l'application $a \to \langle a, x\rangle$ est continue).

Lemme 2.2.3. *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. Facile, en développant le carré scalaire $\langle \sum_{k=1}^n u_k, \sum_{k=1}^n u_k \rangle$.

//

Lemme 2.2.4. *Soit* $(e_1, \ldots, e_n)$ *une suite orthonormée finie dans un espace de Hilbert* H *; posons* $\mathrm{F} = \mathrm{Vect}(e_1, \ldots, e_n)$ *; pour tout vecteur* $x \in \mathrm{H}$*, le vecteur*

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

est la projection orthogonale de x *sur* F*, c'est à dire que* $y \in \mathrm{F}$ *et que le vecteur* $x - y$ *est orthogonal* à F*.*

Démonstration. Il est évident que $y \in \mathrm{F}$, et il est clair que $\langle y, e_j \rangle = \langle x, e_j \rangle$ pour tout $j = 1, \ldots, n$, donc $x - y$ est orthogonal à tous les (e_j), ce qui implique que $x - y$ est orthogonal à F.

//

Lemme 2.2.5 : inégalité de Bessel. *Soient* H *un espace de Hilbert et* $(e_n)_{n \ge 0}$ *une suite orthonormée dans* H *; pour tout* $x \in \mathrm{H}$ *la série numérique* $\sum_k |\langle x, e_k \rangle|^2$ *est convergente et*

$$\sum_{k \ge 0} |\langle x, e_k \rangle|^2 \le \|x\|^2.$$

Démonstration. Il suffit de montrer le résultat pour une suite finie $e_1, \ldots, e_n$. 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 $\mathrm{F} = \mathrm{Vect}(e_1, \ldots, e_n)$, donc $x - y$ est orthogonal à $y \in \mathrm{F}$. On aura puisque $x = y + (x - y)$

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

d'où le résultat.

//

Lemme 2.2.6. *Soit* $(u_n)_{n\geq 0}$ *une suite orthogonale dans un espace de Hilbert* H *; la série de vecteurs* $\sum_k u_k$ *converge dans* H *si et seulement si* $\sum \|u_k\|^2 < +\infty$*, et dans ce cas*

$$\Big\|\sum_{k=0}^{+\infty} u_k\Big\|^2 = \sum_{k=0}^{+\infty} \|u_k\|^2.$$

Si $(e_n)_{n\geq 0}$ *est une suite orthonormée, la série de vecteurs* $\sum_k c_k e_k$ *converge si et seulement si* $\sum |c_k|^2 < +\infty$*, et dans ce cas on a* $\big\|\sum_{k=0}^{+\infty} c_k e_k\big\|^2 = \sum_{k=0}^{+\infty} |c_k|^2$.

Démonstration. Posons $\mathrm{U}_n = \sum_{i=0}^{n} u_i$. Si $m < n$ on a par orthogonalité

$$\|\mathrm{U}_n - \mathrm{U}_m\|^2 = \sum_{k=m+1}^{n} \|u_k\|^2.$$

A partir de là, il est clair que la suite (U_n) est de Cauchy dans H si et seulement si la série numérique $\sum_k \|u_k\|^2$ vérifie le critère de convergence de Cauchy. La norme de la somme de la série s'obtient en passant à la limite dans l'égalité du lemme 3.

//

Lemme 2.2.7. *Soit* $(e_n)_{n\geq 0}$ *une suite orthonormée dans* H *et soit* F *le sous-espace vectoriel fermé engendré par la suite* $(e_n)_{n\geq 0}$ *; pour tout vecteur* $y \in \mathrm{F}$*, on a*

$$y = \sum_{k=0}^{+\infty} \langle y, e_k\rangle\, e_k.$$

Démonstration. Posons $c_j = \langle y, e_j\rangle$ pour tout $j \geq 0$, et $z = \sum_{k=0}^{+\infty} c_k\, e_k$. Cette série converge d'après le lemme 5 et le lemme précédent, et $z \in \mathrm{F}$. Pour tout $j \geq 0$, on voit en passant à la limite grâce à la continuité de l'application $x \to \langle x, e_j\rangle$

$$\langle z, e_j\rangle = \lim_n \langle \sum_{i=0}^{n} c_i e_i, e_j\rangle = c_j = \langle y, e_j\rangle$$

ce qui montre que $y - z$ est orthogonal à chacun des vecteurs e_j, donc $y - z$ est orthogonal à F. Puisque $y - z \in \mathrm{F}$, il en résulte que $y - z = 0_{\mathrm{H}}$, d'où le résultat.

//

Définition 2.2.4. On appelle *base hilbertienne* d'un espace de Hilbert séparable H de dimension infinie une suite orthonormée $(e_n)_{n\geq 0}$ qui est de plus *totale* dans H. On dit aussi *base orthonormée* de H.

Certaines bases hilbertiennes sont naturellement indexées par un ensemble dénombrable spécifique, par exemple $\mathrm{I} = \mathbb{Z}$, plutôt que par l'ensemble $\mathbb{N}$. Du point de vue théorique, il n'y a pas de différence et nous écrirons les preuves avec $\mathrm{I} = \mathbb{N}$.

Proposition 2.2.8. *Supposons que* $(e_n)_{n\ge0}$ *soit une base orthonormée de l'espace de Hilbert séparable* H *de dimension infinie. Pour tout vecteur* x *de* H, *on a*

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

On voit qu'une base hilbertienne $(e_n)_{n\ge0}$ 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 $(e_n)_{n\ge0}$ doit être totale dans H.

Démonstration. Par définition d'une base orthonormée, la suite $(e_n)_{n\ge0}$ est totale dans H, ce qui signifie que le sous-espace vectoriel fermé F engendré par cette suite est égal à H. Il suffit d'appliquer le lemme 7 pour obtenir la première partie de la conclusion, et le lemme 6 pour la seconde.

//

Théorème 2.2.9. *Pour tout espace de Hilbert séparable* H *de dimension infinie, il existe une base orthonormée* $(e_n)_{n\ge0}$.

Si H est de dimension finie, l'existence de base orthonormée (bien entendu *finie*) a été vue en DEUG. Le cas des espaces de Hilbert non séparables sera examiné au chapitre 6.

Démonstration. Soit H un espace de Hilbert séparable de dimension infinie ; on peut trouver une suite croissante $(\mathrm{E}_n)_{n\ge0}$ de sous-espaces de dimension finie de H, telle que $\dim \mathrm{E}_n = n$ pour tout $n \ge 0$ et telle que $\bigcup_n \mathrm{E}_n$ soit dense dans H (proposition 1.7.1). On construit la suite orthonormée par récurrence de façon que pour tout $n \ge 1$, la suite $(e_1, \ldots, e_n)$ soit une base orthonormée de E_n. On commence en prenant pour e_1 un vecteur de norme un dans E_1. Supposons $e_1, \ldots, e_n$ définis, de façon que $(e_1, \ldots, e_n)$ soit une base orthonormée de E_n. Puisque $\mathrm{E}_{n+1} \ne \mathrm{E}_n$, on peut choisir un vecteur $x_{n+1} \in \mathrm{E}_{n+1}$ qui n'est pas dans E_n. Soit y la projection orthogonale de x_{n+1} sur E_n. On a $x_{n+1} \ne y$ puisque $x_{n+1} \notin \mathrm{E}_n$. Le vecteur $z = x_{n+1} - y$ est non nul et orthogonal à E_n. On prend pour e_{n+1} un multiple de norme un du vecteur z. Par construction $(e_1, \ldots, e_n, e_{n+1})$ est une suite orthonormée dans E_{n+1}, donc une base de E_{n+1} (puisque $\dim \mathrm{E}_{n+1} = n+1$).

La suite $(e_n)_{n\ge0}$ est totale dans H puisque l'espace vectoriel qu'elle engendre contient la réunion $\bigcup_{n\ge0} \mathrm{E}_n$ qui est dense dans H.

//

Exemples 2.2.5.

1. La suite canonique $(\mathbf{e}_n)_{n\ge0}$ de l'espace ℓ_2 est évidemment une base orthonormée de ℓ_2.

2. Considérons l'espace de Hilbert $\mathrm{L}_2(0, 2\pi)$ des fonctions complexes de carré sommable pour la mesure $dx/2\pi$. Pour chaque entier relatif $n \in \mathbb{Z}$, soit f_n la fonction définie par

$$f_n(s) = \mathrm{e}^{ins}$$

pour tout $s \in [0, 2\pi]$. Il est facile de vérifier que les fonctions $(f_n)_{n\in\mathbb{Z}}$ forment une suite orthonormée dans $\mathrm{L}_2(0, 2\pi)$. En revanche, il faut une petite démonstration pour voir que ce système est total (voir la section 3.6). Il s'agit donc d'une base orthonormée de $\mathrm{L}_2(0, 2\pi)$.

2.3. Théorème de projection

Théorème 2.3.1 : théorème de projection. *Soient* H *un espace de Hilbert et* C *une partie convexe fermée non vide de* H *; pour tout* $x \in \mathrm{H}$, *il existe un et un seul point* y_0 *de* C *en lequel la fonction* $y \to \|y - x\|$ *atteint son minimum sur* C. *On a de plus*

$$\forall y \in \mathrm{C}, \quad \mathrm{Re}\langle x - y_0, y - y_0 \rangle \le 0.$$

Démonstration. En translatant le convexe C, on peut se ramener au cas où $x = 0_{\mathrm{H}}$. Notons alors

$$d = \inf\{d(y, 0_{\mathrm{H}}) : y \in \mathrm{C}\} = \inf\{\|y\| : y \in \mathrm{C}\}$$

la distance de 0_{H} à C. Si y et z sont deux points de C, on a $(y + z)/2 \in \mathrm{C}$ puisque C est convexe, donc $\|(y + z)/2\| \ge d$; de plus la relation du parallélogramme

$$\|(y + z)/2\|^2 + \|(y - z)/2\|^2 = (\|y\|^2 + \|z\|^2)/2$$

implique pour tous $y, z \in \mathrm{C}$

$$(*) \qquad 0 \le \|(y - z)/2\|^2 \le (\|y\|^2 + \|z\|^2)/2 - d^2.$$

Pour tout entier $n \ge 1$, posons

$$\mathrm{C}_n = \{y \in \mathrm{C} : \|y\|^2 \le d^2 + 1/n\}.$$

L'ensemble C_n est une partie fermée non vide de H ; d'après la relation $(*)$, on a $\|(y - z)/2\|^2 \le 1/n$ pour tous $y, z \in \mathrm{C}_n$. Le diamètre de C_n est donc inférieur ou égal à $2/\sqrt{n}$, et il tend donc vers 0. Comme l'espace H est complet, l'intersection des fermés emboîtés C_n qui est égale à $\{y \in \mathrm{C} : \|y\| = d\}$, contient un et un seul point, qui est le point y_0 cherché.

Compte tenu de notre translation simplificatrice, la relation à démontrer ensuite devient $\mathrm{Re}(\langle -y_0, y - y_0 \rangle) \le 0$ pour tout $y \in \mathrm{C}$; pour $t \in [0, 1]$, on a $y_0 + t(y - y_0) \in \mathrm{C}$, donc $\|y_0 + t(y - y_0)\| \ge \|y_0\|$, ce qui donne en développant le carré de la norme

$$2t\,\mathrm{Re}(\langle y_0, y - y_0 \rangle) + t^2\|y - y_0\|^2 \ge 0$$

pour $0 \le t \le 1$; pour finir on divise par $t > 0$ que l'on fait ensuite tendre vers 0, et on obtient $\mathrm{Re}(\langle y_0, y - y_0 \rangle) \ge 0$.

//

Un cas particulier important est celui où C est un sous-espace vectoriel fermé F de H. Dans ce cas on a $\langle x - y_0, z \rangle = 0$ pour tout vecteur $z \in \mathrm{F}$, c'est à dire que $x - y_0 \perp \mathrm{F}$. Dans le cas de la projection sur un sous-espace vectoriel fermé F, la projection y_0 de x sur F est entièrement caractérisée par les deux conditions suivantes

– le vecteur y_0 appartient à F ;
– le vecteur $x - y_0$ est orthogonal à F.

En effet, si ces conditions sont vérifiées et si y est un élément quelconque de F, on aura

$$(*) \qquad \|x - y\|^2 = \|(x - y_0) + (y_0 - y)\|^2 = \|x - y_0\|^2 + \|y_0 - y\|^2$$

parce que $y_0 - y \in \mathrm{F}$ est orthogonal à $x - y_0$. Cette relation montre que $\|x-y\|^2 \geq \|x-y_0\|^2$ pour tout $y \in \mathrm{F}$, c'est à dire que y_0 est bien le point de F le plus proche du point x.

On notera $\mathrm{P_F}(x) = y_0$ la projection orthogonale de x sur F. La caractérisation ci-dessus montre que $\mu \mathrm{P_F}(x) + \mu' \mathrm{P_F}(x')$ est la projection de $\mu x + \mu' x'$, autrement dit l'application $\mathrm{P_F}$ est une application linéaire. L'égalité $(*)$ ci-dessus donne aussi $\|x-y\| \geq \|\mathrm{P_F}(x) - y\|$ pour tout $y \in \mathrm{F}$, donc $\|x\| \geq \|\mathrm{P_F}(x)\|$ en prenant $y = 0$; on a donc $\|\mathrm{P_F}\| \leq 1$.

Si F est un sous-espace vectoriel fermé d'un espace hilbertien H on appelle *projecteur orthogonal* sur F l'opérateur borné $\mathrm{P_F} : \mathrm{H} \to \mathrm{H}$ qui associe à tout vecteur $x \in \mathrm{H}$ sa projection sur F.

Exemple 2.3.1. Espérance conditionnelle. Si $(\Omega, \mathcal{A}, \mu)$ est un espace de probabilité et si $\mathcal{F}$ est une sous-tribu de $\mathcal{A}$, on peut considérer le sous-espace vectoriel F de L_2 formé de toutes les fonctions qui sont $\mathcal{F}$-mesurables. Le sous-espace F est fermé, et la projection orthogonale de L_2 sur F s'appelle *l'espérance conditionnelle*. Par exemple, si $\Omega = [0,1]^2$ est muni de sa tribu borélienne et de la mesure de Lebesgue, si $\mathcal{F}$ est la sous-tribu formé de tous les ensembles de la forme $\mathrm{A} \times [0,1]$, où A varie parmi les boréliens de $[0,1]$, le sous-espace F est formé des fonctions qui ne dépendent que de la première variable et la projection $\mathrm{P_F} f = \mathrm{E}(f|\mathcal{F})$ d'une fonction $f \in \mathrm{L}_2$ est donnée par

$$\mathrm{E}(f|\mathcal{F})(x,y) = \int_0^1 f(x,u)\,du.$$

Corollaire 2.3.2. *Soient* H *un espace de Hilbert,* F *un sous-espace vectoriel fermé séparable de* H*, et* $(e_n)_{n \geq 0}$ *une base hilbertienne du sous-espace* F*. Pour tout vecteur* $x \in \mathrm{H}$*, la projection orthogonale de* x *sur* F *est donnée par*

$$\mathrm{P_F}(x) = \sum_{k=0}^{+\infty} \langle x, e_k \rangle \, e_k.$$

Définition 2.3.2. On dit que des parties A et B d'un espace de Hilbert H sont *orthogonales* si tout élément de A est orthogonal à tout élément de B. Soit A une partie de H ; on appelle orthogonal de A l'ensemble $\mathrm{A}^\perp$ des éléments de H orthogonaux à A.

Il est clair que $\mathrm{A}^\perp$ est un sous-espace vectoriel fermé de H.

Proposition 2.3.3. *Soient* H *un espace de Hilbert et* F *un sous-espace vectoriel fermé de* H *; on a* $\mathrm{P_F} + \mathrm{P_{F^\perp}} = \mathrm{Id_H}$*. Il en résulte que* $\mathrm{F} \oplus \mathrm{F}^\perp = \mathrm{H}$ *et* $\mathrm{F}^{\perp\perp} = \mathrm{F}$.

Démonstration. Commençons par une évidence : par définition, tout vecteur de F est orthogonal à $\mathrm{F}^\perp$, donc $\mathrm{F} \subset \mathrm{F}^{\perp\perp}$. Soit maintenant $x \in \mathrm{H}$ quelconque et écrivons $x = \mathrm{P_F}(x) + (x - \mathrm{P_F}(x))$; d'après les propriétés de la projection orthogonale sur le sous-espace vectoriel F, on a bien que $x - \mathrm{P_F}(x) \in \mathrm{F}^\perp$, et de plus la différence $x - (x - \mathrm{P_F}(x)) = \mathrm{P_F}(x) \in \mathrm{F}$ est orthogonale à $\mathrm{F}^\perp$; cela montre que $x - \mathrm{P_F}(x)$ est la projection orthogonale de x sur $\mathrm{F}^\perp$, c'est à dire que $\mathrm{P_{F^\perp}} = \mathrm{Id_H} - \mathrm{P_F}$. La relation $\mathrm{Id_H} = \mathrm{P_F} + \mathrm{P_{F^\perp}}$ implique évidemment que H est la somme de F et $\mathrm{F}^\perp$. On vérifie ensuite que la somme est directe : si $x \in \mathrm{F} \cap \mathrm{F}^\perp$ alors $\langle x, x \rangle = 0$ donc $x = 0_\mathrm{H}$.

Pour finir, si on a un vecteur $x \in \mathrm{F}^{\perp\perp}$, il est orthogonal à $\mathrm{F}^\perp$ par définition, donc 0_H est sa projection orthogonale sur $\mathrm{F}^\perp$ et la relation $\mathrm{P_F}(x) = (\mathrm{Id_H} - \mathrm{P_{F^\perp}})(x) = x$ montre que $x \in \mathrm{F}$.

//

Corollaire 2.3.4. *Soit* H *un espace de Hilbert ;*

(i) *pour toute partie* A *de* H, *l'ensemble* $(\mathrm{A}^\perp)^\perp$ *est le plus petit sous-espace vectoriel fermé de* H *contenant* A *;*

(ii) *si* Y *est un sous-espace vectoriel de* H, *on a* $(\mathrm{Y}^\perp)^\perp = \overline{\mathrm{Y}}$.

Démonstration. Montrons le point (i). Soit F le plus petit sous-espace vectoriel fermé de H contenant A ; on sait que tout vecteur y orthogonal à A est aussi orthogonal à l'espace vectoriel Y engendré par A (par linéarité du produit scalaire), puis à l'adhérence $\mathrm{F} = \overline{\mathrm{Y}}$ de ce sous-espace (par continuité du produit scalaire). Inversement tout vecteur orthogonal à F est évidemment orthogonal à A. On a donc $\mathrm{A}^\perp = \mathrm{F}^\perp$, donc $(\mathrm{A}^\perp)^\perp = \mathrm{F}^{\perp\perp} = \mathrm{F}$.

Le point (ii) découle de (i), puisque le plus petit sous-espace fermé de H contenant Y est l'adhérence $\overline{\mathrm{Y}}$.

//

A tout vecteur $y \in \mathrm{H}$ on a associé la forme linéaire continue ℓ_y définie par

$$(*) \qquad \forall x \in \mathrm{H}, \quad \ell_y(x) = \langle x, y \rangle$$

et on a vu que $\|\ell_y\| = \|y\|$ (proposition 2.1).

Proposition 2.3.5. *Soit* H *un espace de Hilbert ; l'application isométrique antilinéaire* $y \to \ell_y$ *de l'équation* $(*)$ *est une bijection de* H *sur le dual* H^*. *En d'autres termes, pour toute forme linéaire continue* ℓ *sur* H, *il existe un vecteur* $y_\ell \in \mathrm{H}$ *unique qui représente la forme linéaire* ℓ *au sens suivant :*

$$\forall x \in \mathrm{H}, \quad \ell(x) = \langle x, y_\ell \rangle.$$

Démonstration. Soit $\ell \in \mathrm{H}^*$; si $\ell = 0$ il suffit de (et il faut) prendre $y_\ell = 0_{\mathrm{H}}$. Si $\ell \neq 0$, notons F son noyau (fermé). Puisque $\mathrm{F} \neq \mathrm{H}$, on peut choisir un vecteur z orthogonal à F et tel que $\ell(z) = 1$. Tout vecteur $x \in \mathrm{H}$ peut s'écrire

$$x = (x - \ell(x)\, z) + \ell(x)\, z = x' + \ell(x)\, z$$

avec $x' = x - \ell(x)\, z$ qui est dans F puisque $\ell(x') = \ell(x) - \ell(x)\ell(z) = 0$. On a pour tout $x \in \mathrm{H}$, puisque $x' \perp z$

$$\langle x, z \rangle = \langle \ell(x)\, z, z \rangle = \langle z, z \rangle\, \ell(x)$$

ce qui montre que $\ell = \|z\|^{-2}\, \ell_z$. Il suffit de prendre $y_\ell = \|z\|^{-2}\, z$ pour obtenir le résultat voulu.

//

Exemple 2.3.3. Pour toute forme linéaire continue ℓ sur $\mathrm{L}_2(\Omega, \mu)$, il existe une fonction $g \in \mathrm{L}_2$ telle que

$$\forall f \in \mathrm{L}_2, \quad \ell(f) = \int_\Omega f(s) \overline{g(s)}\, d\mu(s).$$

Une application très utile est le "petit" théorème de Radon-Nikodym. Si μ, ν sont deux mesures positives sur un espace mesurable $(\Omega, \mathcal{A})$, avec ν finie et μ σ-finie, et si $\nu(\mathrm{A}) \le \mu(\mathrm{A})$ pour tout $\mathrm{A} \in \mathcal{A}$, il existe une fonction mesurable bornée f telle que

$$\nu(\mathrm{A}) = \int_{\mathrm{A}} f(s)\, d\mu(s) = \int_{\Omega} \mathbf{1}_{\mathrm{A}}(s) f(s)\, d\mu(s)$$

pour tout ensemble $\mathrm{A} \in \mathcal{A}$, c'est à dire que la mesure ν peut se représenter comme la mesure de densité f par rapport à μ.

La démonstration fonctionne ainsi : il résulte de l'hypothèse que $\int h\, d\nu \le \int h\, d\mu$ pour toute fonction mesurable positive h, et ceci implique que $\|g\|_{\mathrm{L}_2(\nu)} \le \|g\|_{\mathrm{L}_2(\mu)}$ pour toute fonction $g \in \mathrm{L}_2(\mu)$, ce qui montre que $\mathrm{L}_2(\mu) \subset \mathrm{L}_2(\nu)$. Comme ν est finie, la forme linéaire $g \to \int g\, d\nu$ est définie et continue sur $\mathrm{L}_2(\nu)$, donc sur $\mathrm{L}_2(\mu)$. On peut donc la représenter par une fonction $f \in \mathrm{L}_2(\mu)$, c'est à dire que

$$\int g\, d\nu = \int g f\, d\mu$$

pour toute fonction $g \in \mathrm{L}_2(\mu)$. En appliquant avec $g = \mathbf{1}_{\mathrm{A}}$ on obtient le résultat annoncé.

Somme hilbertienne de sous-espaces orthogonaux

Proposition 2.3.6. *Si* F_1 *et* F_2 *sont deux sous-espaces vectoriels fermés orthogonaux de l'espace de Hilbert* H, *le sous-espace* $\mathrm{F} = \mathrm{F}_1 + \mathrm{F}_2$ *est fermé et la projection orthogonale de* H *sur* F *est donnée par* $\mathrm{P}_{\mathrm{F}} = \mathrm{P}_{\mathrm{F}_1} + \mathrm{P}_{\mathrm{F}_2}$.

Supposons donnée dans un espace de Hilbert H une suite $(\mathrm{F}_n)_{n\ge 0}$ de sous-espaces vectoriels fermés, deux à deux orthogonaux. On sait que pour toute famille $(x_n)_{n\ge 0}$ de vecteurs telle que $x_n \in \mathrm{F}_n$ pour tout $n \ge 0$ et $\sum \|x_k\|^2 < +\infty$, la série $\sum x_k$ converge dans H ; le vecteur $x = \sum_{k=0}^{+\infty} x_k$, qui est limite de $y_{\mathrm{N}} = \sum_{k=0}^{\mathrm{N}} x_k$, appartient à l'espace vectoriel fermé F engendré par la famille $(\mathrm{F}_n)_{n\ge 0}$: en effet, le sous-espace F contient chaque y_{N} puisqu'il contient $\mathrm{F}_0, \ldots, \mathrm{F}_{\mathrm{N}}$ et il contient la limite x puisqu'il est fermé. Inversement

Proposition 2.3.7. *Le sous-espace vectoriel fermé* F *engendré par une famille* $(\mathrm{F}_n)_{n\ge 0}$ *de sous-espaces vectoriels fermés de* H *deux à deux orthogonaux coïncide avec*

$$\{x = \sum_{k=0}^{+\infty} x_k : \forall n \ge 0,\ x_n \in \mathrm{F}_n \quad et \quad \sum \|x_k\|^2 < +\infty\}.$$

Démonstration. Désignons par G l'ensemble ci-dessus ; on a déjà expliqué que l'espace fermé F engendré par les (F_n) doit contenir G. Inversement, soit $x \in \mathrm{F}$ et désignons par x_n, pour tout entier $n \ge 0$, la projection orthogonale de x sur F_n ; on vérifie facilement que $s_n = x_0 + \cdots + x_n$ est la projection orthogonale de x sur le sous-espace fermé $\mathrm{F}_0 + \cdots + \mathrm{F}_n$. Il en résulte que $\|s_n\| \le \|x\|$ pour tout n, ce qui implique que $\sum \|x_k\|^2 \le \|x\|$ et permet de définir $x' = \sum_{k=0}^{+\infty} x_k \in \mathrm{G} = \lim_n s_n$. On va montrer que $x' = x$; puisque x appartient au sous-espace fermé engendré par les F_n, on peut trouver pour tout $\varepsilon > 0$ un entier N et un vecteur $y \in \mathrm{F}_0 + \cdots + \mathrm{F}_{\mathrm{N}}$ tels que $\|x - y\| < \varepsilon$. Mais par définition de la projection orthogonale, on a $\|x - s_{\mathrm{N}}\| \le \|x - y\|$, ce qui montre que $\lim_n s_n = x$ et $x = x'$.

//