Chapitre 2. Espaces de Hilbert

2.1. *Produit scalaire, orthogonalité, bases hilbertiennes*

On considère un espace vectoriel E sur $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$. On appelle *produit scalaire* sur E une application $(x, y) \in \mathrm{E} \times \mathrm{E} \to \langle x, y\rangle \in \mathbb{K}$ telle que

– l'application $x \in \mathrm{E} \to \langle x, y\rangle$ est $\mathbb{K}$-linéaire pour tout $y \in \mathrm{E}$ fixé ;

– pour tous $x, y \in \mathrm{E}$, on a $\langle y, x\rangle = \overline{\langle x, y\rangle}$;

– le nombre $\langle x, x\rangle$, qui est réel d'après la ligne précédente, vérifie $\langle x, x\rangle \geq 0$ pour tout $x \in \mathrm{E}$, et de plus $\langle x, x\rangle > 0$ pour tout vecteur x non nul.

On dira *semi-produit scalaire* ([a]) lorsqu'on supposera seulement que $\langle x, x\rangle \geq 0$ pour tout vecteur x, mais en permettant que $\langle x, x\rangle = 0$ pour des vecteurs x non nuls. L'application $y \to \langle x, y\rangle$ est *additive*,

$$\langle x, y_1 + y_2\rangle = \overline{\langle y_1 + y_2, x\rangle} = \overline{\langle y_1, x\rangle} + \overline{\langle y_2, x\rangle} = \langle x, y_1\rangle + \langle x, y_2\rangle.$$

Si x, y sont deux vecteurs de E et λ un scalaire, on a ([b])

$$\langle \lambda x, y\rangle = \lambda \langle x, y\rangle \quad \text{mais} \quad \langle x, \lambda y\rangle = \overline{\langle \lambda y, x\rangle} = \overline{\lambda} \langle x, y\rangle.$$

L'application $y \to \langle x, y\rangle$ n'est donc pas linéaire dans le cas complexe, car l'image du vecteur λy est $\overline{\lambda} \langle x, y\rangle$ (et pas $\lambda \langle x, y\rangle$). On dit qu'une telle application f du $\mathbb{K}$-espace vectoriel E dans un autre $\mathbb{K}$-espace vectoriel, qui vérifie $f(\lambda v + w) = \overline{\lambda} f(v) + f(w)$ pour tout scalaire λ et tous vecteurs $v, w \in \mathrm{E}$, est une *application antilinéaire*. Ainsi, pour tout $x \in \mathrm{E}$ fixé,

l'application $y \to \langle x, y\rangle$ est antilinéaire de E *dans* $\mathbb{K}$.

Lorsque $\mathbb{K} = \mathbb{R}$, toutes les barres de conjugaison sont inutiles, et il n'y a pas de différence entre linéaire et antilinéaire ; toutefois, considérer les deux cas $\mathbb{R}$ et $\mathbb{C}$ ensemble permet d'éviter de devoir tout répéter. Si x et y sont deux vecteurs de E, on obtient que

$$\langle x + y, x + y\rangle = \langle x, x\rangle + 2 \,\mathrm{Re}\langle x, y\rangle + \langle y, y\rangle. \tag{1}$$

En remplaçant y par $-y$, on obtient $\langle x - y, x - y\rangle = \langle x, x\rangle - 2 \,\mathrm{Re}\langle x, y\rangle + \langle y, y\rangle$ et en additionnant les deux on obtient *l'identité du parallélogramme*, appelée ([c]) aussi *relation de la médiane*

$$\langle x + y, x + y\rangle + \langle x - y, x - y\rangle = 2 \langle x, x\rangle + 2 \langle y, y\rangle. \tag{2}$$

Proposition 1 (Cauchy-Schwarz). *Soit* E *un* $\mathbb{K}$*-espace vectoriel muni d'un semi-produit scalaire ; pour tous les vecteurs* x, y *de* E *on a*

$$|\langle x, y\rangle| \leq \sqrt{\langle x, x\rangle}\, \sqrt{\langle y, y\rangle}.$$

Preuve.— Écrivons le nombre complexe $\langle x, y\rangle$ sous forme polaire $\langle x, y\rangle = |\langle x, y\rangle|\,\mathrm{e}^{\mathrm{i}\theta}$, pour un certain nombre réel θ, de sorte que $\langle \mathrm{e}^{-\mathrm{i}\theta}\, x, y\rangle = \mathrm{e}^{-\mathrm{i}\theta}\langle x, y\rangle = |\langle x, y\rangle|$. Choisissons deux nombres réels $\lambda > \sqrt{\langle y, y\rangle}$ et $\mu > \sqrt{\langle x, x\rangle}$. On écrit pour le vecteur $z = \lambda\,\mathrm{e}^{-\mathrm{i}\theta}\, x - \mu y$

$$0 \le \langle z, z\rangle = \langle \lambda\,\mathrm{e}^{-\mathrm{i}\theta}\, x - \mu y, \lambda\,\mathrm{e}^{-\mathrm{i}\theta}\, x - \mu y\rangle$$

$$= \lambda^2\langle x, x\rangle - 2\lambda\mu\,\mathrm{Re}\langle \mathrm{e}^{-\mathrm{i}\theta}\, x, y\rangle + \mu^2\langle y, y\rangle = \lambda^2\langle x, x\rangle - 2\lambda\mu\,|\langle x, y\rangle| + \mu^2\langle y, y\rangle.$$

Il en résulte que

$$2\lambda\mu\,|\langle x, y\rangle| \le \lambda^2\langle x, x\rangle + \mu^2\langle y, y\rangle \le 2\lambda^2\mu^2,$$

d'où $|\langle x, y\rangle| \le \lambda\,\mu$ par simplification, puisque $\lambda\mu > 0$. Pour finir, on fait tendre λ vers $\sqrt{\langle y, y\rangle}$ et μ vers $\sqrt{\langle x, x\rangle}$.

Semi-norme déduite d'un semi-produit scalaire

Soit E un $\mathbb{K}$-espace vectoriel muni d'un semi-produit scalaire ; l'application

$$x \in \mathrm{E} \to \|x\| = \sqrt{\langle x, x\rangle}$$

est une *semi-norme* sur E, c'est-à-dire qu'elle vérifie

$$\|x\| \ge 0, \quad \|\lambda x\| = |\lambda|\,\|x\| \quad \text{et} \quad \|x + y\| \le \|x\| + \|y\|$$

pour tous $x, y \in \mathrm{E}$ et $\lambda \in \mathbb{K}$. Le dernier point *(l'inégalité triangulaire)* résulte de Cauchy-Schwarz,

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

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

Lorsque E est muni d'un produit scalaire, l'hypothèse $\|x\| = 0$ implique $\langle x, x\rangle = 0$, donc $x = 0$ par définition d'un produit scalaire, et la semi-norme est en fait une *norme* sur l'espace vectoriel E.

Exemple. Sur l'espace $\mathrm{E} = \mathcal{L}^2(\Omega, \mu)$ on peut définir le semi-produit scalaire

$$\langle f, g\rangle = \int_\Omega f(\omega)\overline{g(\omega)}\,\mathrm{d}\mu(\omega)\,;$$

il lui correspond la semi-norme

$$\|f\|_2 = \Bigl(\int_\Omega |f(\omega)|^2\,\mathrm{d}\mu(\omega)\Bigr)^{1/2}.$$

Si f est négligeable sans être la fonction nulle, on aura $\|f\|_2 = 0$ sans que le vecteur f soit le vecteur nul de E : cette semi-norme n'est en général pas une norme.

Soit E un $\mathbb{K}$-espace vectoriel muni d'un semi-produit scalaire ; on dit que deux vecteurs $u, v \in \mathrm{E}$ sont *orthogonaux* quand $\langle u, v\rangle = 0$; on note que cette relation est symétrique : $\langle v, u\rangle = \overline{\langle u, v\rangle} = \overline{0} = 0$.

Proposition 2 (Pythagore). *Soit* E *un espace muni d'un semi-produit scalaire ; si* $u_1, \ldots u_n$ *sont des vecteurs de l'espace* E*, deux à deux orthogonaux, on a pour la semi-norme associée*

$$\Bigl\|\sum_{j=1}^{n} u_j\Bigr\|^2 = \sum_{j=1}^{n} \|u_j\|^2.$$

Preuve. — Si u et v sont orthogonaux,

$$\|u+v\|^2 = \langle u+v, u+v\rangle = \langle u, u\rangle + 2\,\mathrm{Re}\langle u, v\rangle + \langle v, v\rangle = \langle u, u\rangle + \langle v, v\rangle = \|u\|^2 + \|v\|^2.$$

Ensuite, on montre la proposition par récurrence sur $n \ge 2$: on pose $v = \sum_{j=1}^{n} u_j$, on remarque que u_{n+1} est orthogonal à v (linéarité du produit scalaire) et

$$\Bigl\|\sum_{j=1}^{n+1} u_j\Bigr\|^2 = \|v + u_{n+1}\|^2 = \|v\|^2 + \|u_{n+1}\|^2 = \Bigl(\sum_{j=1}^{n} \|u_j\|^2\Bigr) + \|u_{n+1}\|^2.$$

Définition. Un *espace de Hilbert* est un $\mathbb{K}$-espace vectoriel E muni d'un produit scalaire, et tel que E soit complet pour la norme associée à ce produit scalaire.

Exemple. Rappelons la définition de l'espace $\mathrm{L}^2(\Omega, \mathcal{A}, \mu)$ (bien noter que la lettre « L » dans le symbole L^2 est droite maintenant). Un élément $\widetilde{f} \in \mathrm{L}^2(\Omega, \mu)$ n'est pas une fonction, mais une *classe de fonctions* : si f est une fonction mesurable sur $(\Omega, \mathcal{A})$ à valeurs dans $\mathbb{K}$, on lui associe l'ensemble $\widetilde{f}$ de toutes les fonctions mesurables f_1 à valeurs dans $\mathbb{K}$ telles que $f_1 = f$ μ-presque partout. Si l'un des éléments d'une classe $\widetilde{f}$ est intégrable, alors tous les autres éléments de la classe sont intégrables et ont la même intégrale, ce qui permet d'employer la notation $\int_\Omega \widetilde{f}\,\mathrm{d}\mu$ et de parler de *classe intégrable*. Les carrés f^2 des éléments d'une classe $\widetilde{f}$ sont dans une même classe, qu'on peut raisonnablement noter $\widetilde{f^2}$ et appeler le carré de la classe $\widetilde{f}$.

L'espace $\mathrm{L}^2(\Omega, \mathcal{A}, \mu)$ est l'espace vectoriel des classes de carré intégrable. Le vecteur nul $\widetilde{0}$ de cet espace vectoriel est la *classe nulle*, c'est-à-dire l'ensemble des fonctions μ-négligeables à valeurs dans $\mathbb{K}$. L'espace $\mathrm{E} = \mathrm{L}^2(\Omega, \mathcal{A}, \mu)$, muni du produit scalaire ([d])

$$\langle \widetilde{f}, \widetilde{g}\rangle = \int_\Omega f(\omega)\overline{g(\omega)}\,\mathrm{d}\mu(\omega)$$

où f, g sont des représentants quelconques des classes $\widetilde{f}, \widetilde{g} \in \mathrm{L}^2$, et muni de la norme $\|\widetilde{f}\|_2 = \langle \widetilde{f}, \widetilde{f}\rangle^{1/2}$, est un espace de Hilbert. Dans la suite on ne mentionnera plus les classes ; on fera « comme si » $\widetilde{f}$ et $\widetilde{g}$ étaient des vraies fonctions, pour toute question ne dépendant pas du représentant choisi, comme le calcul des intégrales par exemple.

Corollaire. *Soit* E *un espace de Hilbert ; des vecteurs* (u_k), $k = 1, \ldots, n$, *orthogonaux et* non nuls *sont linéairement indépendants.*

Preuve. — En effet, si $\sum_{k=1}^n c_k u_k = 0$, on aura par la proposition 2

$$\Bigl\|\sum_{k=1}^n c_k u_k\Bigr\|^2 = \sum_{k=1}^n \|c_k u_k\|^2 = \sum_{k=1}^n |c_k|^2 \|u_k\|^2 = 0$$

donc $|c_k|\,\|u_k\| = 0$ pour tout k, et $c_k = 0$ puisque $\|u_k\| \neq 0$ pour $k = 1, \ldots, n$. Ainsi, la seule combinaison linéaire de ces vecteurs qui soit nulle est celle dont tous les coefficients c_k sont nuls : le système de vecteurs est libre.

Si les vecteurs $e_1, \ldots, e_n$ sont *orthonormés*, c'est-à-dire que $\langle e_k, e_\ell \rangle = \delta_{k,\ell}$ pour tous $k, \ell = 1, \ldots, n$, on a

$$\Bigl\|\sum_{k=1}^n c_k e_k\Bigr\|^2 = \sum_{k=1}^n |c_k|^2.$$

De plus, si $\mathrm{F} = \mathrm{Vect}(e_1, \ldots, e_n)$ et si z est un vecteur de F, les coordonnées de z dans la base $(e_1, \ldots, e_n)$ de F sont données par les produits scalaires $\langle z, e_k \rangle$, $k = 1, \ldots, n$

$$z = \sum_{k=1}^n \langle z, e_k \rangle\, e_k, \quad \|z\|^2 = \sum_{k=1}^n |\langle z, e_k \rangle|^2.$$

En effet, si $z \in \mathrm{F}$, il existe des coefficients c_k tels que $z = \sum_{k=1}^n c_k e_k$; pour tout indice $j = 1, \ldots, n$ on a

$$\langle z, e_j \rangle = \sum_{k=1}^n c_k \langle e_k, e_j \rangle = c_j.$$

Formes linéaires continues

Soit H un espace de Hilbert ; pour tout vecteur $v \in \mathrm{H}$ fixé, l'application

$$\ell_v : x \in \mathrm{H} \to \langle x, v \rangle$$

est une forme linéaire continue sur H. En effet, par Cauchy-Schwarz, on a

$$|\ell_v(x)| = |\langle x, v \rangle| \le \|v\|\,\|x\|.$$

Comme on le sait, ceci implique la continuité de ℓ, puisqu'alors $|\ell_v(x_1) - \ell_v(x_2)| = |\ell_v(x_1 - x_2)| \le \|v\|\,\|x_1 - x_2\|$ pour $x_1, x_2 \in \mathrm{H}$. L'orthogonal du vecteur v,

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

est un sous-espace vectoriel (noyau de ℓ_v), et il est fermé (continuité de ℓ_v).

Le *dual topologique* E′ d'un espace normé E est l'espace vectoriel des formes linéaires *continues* sur E. Il est normé de la façon suivante : si ℓ est une forme linéaire continue sur E, on pose

$$\|\ell\| = \sup\{|\ell(x)| : x \in \mathrm{E},\ \|x\| \le 1\}.$$

On a alors $|\ell(x)| \le \|\ell\|\,\|x\|$ pour tout $x \in \mathrm{E}$; en fait $\|\ell\|$ est le plus petit nombre C tel que l'on ait $|\ell(x)| \le \mathrm{C}\,\|x\|$ pour tout vecteur $x \in \mathrm{E}$. Muni de cette norme, l'espace E′ est complet [e].

Revenons au cas de l'espace de Hilbert H. On vient de voir par Cauchy-Schwarz que $|\ell_v(x)| \le \|v\| \, \|x\|$ pour tout x, donc $\|\ell_v\| \le \|v\|$; en appliquant ℓ_v au vecteur v lui-même, on obtient

$$\|v\|^2 = \langle v, v \rangle = \ell_v(v) \le \|\ell_v\| \, \|v\|,$$

ce qui implique $\|v\| \le \|\ell_v\|$ quand $v \ne 0$. On en déduit que $\|\ell_v\| = \|v\|$ pour tout vecteur $v \in \mathrm{H}$. On montrera plus loin que *toute forme linéaire continue sur un espace de Hilbert est de la forme* ℓ_v, *pour un certain vecteur* $v \in \mathrm{H}$ (théorème 6).

Projection orthogonale sur un sous-espace de dimension finie

Considérons dans un espace de Hilbert H un sous-espace vectoriel $\mathrm{F} = \mathrm{Vect}(e_1, \ldots, e_n)$ engendré par une suite orthonormée finie $e_1, \ldots, e_n$: pour tout vecteur $x \in \mathrm{H}$, posons

$$\mathrm{P_F} x = \sum_{k=1}^{n} \langle x, e_k \rangle \, e_k.$$

Lemme 1. *Pour tout* $x \in \mathrm{H}$, *le point* $\mathrm{P_F} x$ *est dans* F *et* $x - \mathrm{P_F} x$ *est orthogonal à* F. *Le point* $\mathrm{P_F} x$ *est le point de* F *le plus proche de* x *; en particulier,* $\mathrm{P_F} y = y$ *pour tout* $y \in \mathrm{F}$. *De plus* $\|\mathrm{P_F} x\| \le \|x\|$, *c'est-à-dire que*

$$\forall x \in \mathrm{H}, \quad \sum_{k=1}^{n} |\langle x, e_k \rangle|^2 \le \|x\|^2.$$

L'application $x \to \mathrm{P_F} x$ *est linéaire continue de* H *dans* H.

On dit que $\mathrm{P_F} x$ est la *projection orthogonale* de x sur le sous-espace vectoriel F.

Preuve.— Posons $y = \sum_{k=1}^{n} \langle x, e_k \rangle \, e_k$; pour tout $j = 1, \ldots, n$ on a

$$\langle y, e_j \rangle = \langle \sum_{k=1}^{n} \langle x, e_k \rangle \, e_k, e_j \rangle = \sum_{k=1}^{n} \langle x, e_k \rangle \, \langle e_k, e_j \rangle = \langle x, e_j \rangle.$$

On voit ainsi que $\langle e_j, x - y \rangle = 0$ pour tout $j = 1, \ldots, n$; par linéarité du produit scalaire par rapport au premier vecteur, on obtient $\langle \sum_{j=1}^{n} c_j e_j, x - y \rangle = 0$ pour tous les scalaires (c_j), ce qui montre que que $x - y$ est orthogonal à F : on a bien que $y \in \mathrm{F}$ et $x - y \perp \mathrm{F}$. Si z est un vecteur de F quelconque, la différence $y - z$ est encore dans F, donc orthogonale à $x - y$ et par Pythagore

$$\|x - z\|^2 = \|(x - y) + (y - z)\|^2 = \|x - y\|^2 + \|y - z\|^2 \ge \|x - y\|^2.$$

En reprenant la ligne précédente avec $z = 0$,

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

Il est clair sur les formules de définition que $\mathrm{P_F}$ est linéaire, et l'inégalité $\|\mathrm{P_F} x\| \le \|x\|$ implique la continuité de $\mathrm{P_F}$.

Séries de vecteurs dans un espace vectoriel normé

Si $(u_k)_{k\ge 0}$ est une suite de vecteurs d'un espace normé E, on peut définir les *sommes partielles*

$$\mathrm{U}_n = \sum_{k=0}^{n} u_k \in \mathrm{E}\,;$$

par définition, la série de vecteurs $\sum u_k$ est dite *convergente dans* E quand il existe un vecteur $s \in \mathrm{E}$ tel que

$$\lim_n \|s - \mathrm{U}_n\|_{\mathrm{E}} = 0,$$

c'est-à-dire que $s = \lim_n \mathrm{U}_n$ dans E. Ce vecteur s (qui est unique) est appelé *la somme de la série*. On pose alors

$$\sum_{k=0}^{+\infty} u_k = s = \lim_n \mathrm{U}_n.$$

Proposition 3. *Soit $(u_k)_{k\ge 0}$ une suite de vecteurs deux à deux orthogonaux dans un espace de Hilbert* H *; la série $\sum u_k$ converge dans* H *si et seulement si $\sum \|u_k\|^2 < +\infty$, et dans ce cas*

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

Preuve. Posons pour tout entier $n \ge 0$

$$\mathrm{U}_n = u_0 + \cdots + u_n, \quad \mathrm{V}_n = \|\mathrm{U}_n\|^2 = \|u_0\|^2 + \cdots + \|u_n\|^2.$$

Puisque H est complet, la série de vecteurs $\sum u_k$ converge dans H si et seulement si la suite des sommes partielles (U_n) est de Cauchy dans H. Pour tous les entiers $m < n$ on a $\mathrm{U}_n - \mathrm{U}_m = u_{m+1} + \cdots + u_n$, donc par Pythagore

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

Il en résulte que la suite de vecteurs (U_n) est de Cauchy si et seulement si la suite numérique (V_n) est de Cauchy, c'est-à-dire si et seulement si $\sum \|u_k\|^2 < +\infty$. Dans le cas où la série converge, on obtient

$$\Bigl\|\sum_{k=0}^{+\infty} u_k\Bigr\|^2 = \|\lim_n \mathrm{U}_n\|^2 = \lim_n \|\mathrm{U}_n\|^2 = \lim_n \sum_{k=0}^{n} \|u_k\|^2 = \sum_{k=0}^{+\infty} \|u_k\|^2.$$

Le *sous-espace vectoriel* $\mathrm{Vect}(v_i : i \in \mathrm{I})$ *engendré* par une famille $(v_i)_{i\in \mathrm{I}}$ de vecteurs de H est l'ensemble de toutes les combinaisons linéaires de ces vecteurs ; par définition, une combinaison linéaire $w \in \mathrm{Vect}(v_i : i \in \mathrm{I})$ fait intervenir un ensemble *fini* $\mathrm{J} \subset \mathrm{I}$ et des coefficients scalaires $(c_j)_{j\in \mathrm{J}}$,

$$w = \sum_{j\in \mathrm{J}} c_j v_j.$$

On voit facilement que $\mathrm{Vect}(v_i : i \in \mathrm{I})$ est *le plus petit sous-espace vectoriel* de H contenant tous les vecteurs v_i. Si $\mathrm{I} = \mathbb{N}$, on voit que $\mathrm{Vect}(v_n : n \ge 0)$ est la réunion des espaces de dimension finie $\mathrm{Vect}(v_k : 0 \le k \le \mathrm{N})$, lorsque N varie dans $\mathbb{N}$.

On peut montrer facilement que l'adhérence d'un sous-espace vectoriel V est encore un sous-espace vectoriel. Le *sous-espace vectoriel fermé engendré* par la famille $(v_i)_{i \in \mathrm{I}}$ de vecteurs de H est l'adhérence de $\mathrm{Vect}(v_i : i \in \mathrm{I})$: c'est *le plus petit sous-espace vectoriel fermé* contenant tous les vecteurs v_i, $i \in \mathrm{I}$.

Proposition 4. *Soit $(e_k)_{k \geq 0}$ une suite orthonormée infinie dans un espace de Hilbert H ; pour tout vecteur $x \in \mathrm{H}$, la série de vecteurs*

$$(*) \qquad \sum_{k=0}^{+\infty} \langle x, e_k \rangle \, e_k$$

est convergente dans H. *Désignons par* F *le sous-espace vectoriel fermé engendré par la suite $(e_k)_{k \geq 0}$; la somme $y = \sum_{k=0}^{+\infty} \langle x, e_k \rangle \, e_k$ de la série précédente est un vecteur de* F, *et $x - y$ est orthogonal* à F. *On dit encore que ce vecteur y est* la projection orthogonale de x sur F.

De plus, tout vecteur z du sous-espace vectoriel F *peut s'écrire*

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

Preuve.— Pour tout entier $n \geq 0$, posons $\mathrm{F}_n = \mathrm{Vect}(e_0, \ldots, e_n)$; on sait par le lemme 1 que le vecteur $y_n = \sum_{k=0}^{n} \langle x, e_k \rangle \, e_k \in \mathrm{F}$ est la projection orthogonale de x sur F_n, et que

$$\sum_{k=0}^{n} |\langle x, e_k \rangle|^2 = \|y_n\|^2 \leq \|x\|^2.$$

Il en résulte que la série numérique $\sum |\langle x, e_k \rangle|^2$ converge, ce qui entraîne par la proposition 3 la convergence de la série $(*)$ de vecteurs orthogonaux ; la suite (y_n) des sommes partielles converge vers la somme de la série y ; ce vecteur limite y est dans F, puisque F est fermé. Pour $m \leq n$ on a $e_m \in \mathrm{F}_n$ donc $\langle x - y_n, e_m \rangle = 0$. Le produit scalaire avec e_m étant continu, on obtient

$$\langle x - y, e_m \rangle = \lim_n \langle x - y_n, e_m \rangle = 0.$$

Ceci montre que $x - y$ est orthogonal à tous les vecteurs e_m ; comme l'orthogonal $(x-y)^\perp$ est un sous-espace vectoriel, il contient d'abord toutes les combinaisons linéaires des e_m, et comme il est fermé, il contient aussi les limites de combinaisons linéaires, c'est-à-dire finalement tous les vecteurs de F. Ainsi $\mathrm{F} \subset (x - y)^\perp$, ce qui signifie que $x - y \perp \mathrm{F}$.

Prenons maintenant $z \in \mathrm{F}$ et considérons le vecteur $y = \sum_{k=0}^{+\infty} \langle z, e_k \rangle \, e_k$; d'après ce qui précède, on a $y \in \mathrm{F}$ et $z - y \perp \mathrm{F}$; mais puisque $z - y \in \mathrm{F}$, il en résulte que

$$\langle z - y, z - y \rangle = 0,$$

par conséquent $z = y = \sum_{k=0}^{+\infty} \langle z, e_k \rangle \, e_k$, comme annoncé.

Exemple. Les fonctions $e_n(t) = \mathrm{e}^{\mathrm{i}nt}$, $n \in \mathbb{Z}$, sont une suite orthonormée dans l'espace $\mathrm{H} = \mathrm{L}^2([0, 2\pi], \mathrm{d}t/(2\pi))$. Pour toute fonction $f \in \mathrm{H}$, les deux séries de vecteurs

$$\sum_{n=0}^{+\infty} \langle f, e_n \rangle \, e_n, \quad \sum_{n=1}^{+\infty} \langle f, e_{-n} \rangle \, e_{-n}$$

convergent dans H.

Définition. Une *base hilbertienne* d'un espace de Hilbert H est une famille de vecteurs $(e_i)_{i \in \mathrm{I}}$ indexée par un ensemble I, et telle que

– les vecteurs sont de norme un, et deux à deux orthogonaux ;

– l'espace vectoriel engendré $\mathrm{Vect}(e_i, i \in \mathrm{I})$ est dense dans H.

On ne travaillera pas ici avec les espaces de Hilbert qui ont une base hilbertienne indexée par un ensemble I non dénombrable. Mentionnons que *tout espace de Hilbert admet des bases hilbertiennes*, mais nous ne le prouverons que dans le cas d'un espace de Hilbert *séparable* ; dans ce cas, la base peut être indexée par un ensemble I fini ou dénombrable. Si I est infini dénombrable, si $(e_i)_{i \in \mathrm{I}}$ est une base hilbertienne de H et si $(i_n)_{n \geq 0}$ est une énumération quelconque des éléments de I, on a pour tout $x \in \mathrm{H}$

$$x = \sum_{n=0}^{+\infty} \langle x, e_{i_n} \rangle \, e_{i_n}, \qquad \|x\|^2 = \sum_{n=0}^{+\infty} |\langle x, e_{i_n} \rangle|^2.$$

Pour la première égalité, il suffit d'appliquer la proposition 4, puisque le sous-espace fermé engendré par la suite (e_{i_n}) est ici égal à l'espace H tout entier. La deuxième égalité résulte de la proposition 3.

Expression du produit scalaire dans une base hilbertienne infinie dénombrable

Si l'espace de Hilbert H admet une base hilbertienne infinie dénombrable $(e_n)_{n \geq 0}$, si $x, y \in \mathrm{H}$, le vecteur x est limite dans H de la suite

$$x_{\mathrm{N}} = \sum_{k=0}^{\mathrm{N}} \langle x, e_k \rangle \, e_k \, ;$$

puisque la forme linéaire ℓ_y est continue sur H, on obtient

$$\langle x, y \rangle = \ell_y(x) = \lim_{\mathrm{N}} \langle x_{\mathrm{N}}, y \rangle = \lim_{\mathrm{N}} \sum_{k=0}^{\mathrm{N}} \langle x, e_k \rangle \, \langle e_k, y \rangle = \lim_{\mathrm{N}} \sum_{k=0}^{\mathrm{N}} \langle x, e_k \rangle \, \overline{\langle y, e_k \rangle}$$

c'est-à-dire que

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

Exemple de base hilbertienne : le système de Haar

Pour chaque intervalle borné $\mathrm{I} = [a, b[$ posons $m = (a+b)/2$, puis $\mathrm{I}_- = [a, m[$ et $\mathrm{I}_+ = [m, b[$; on décrit une famille infinie d'intervalles par récurrence : la famille $\mathcal{G}_0$ est formée du seul intervalle $[0, 1[$; si $\mathcal{G}_n$ est définie, la famille $\mathcal{G}_{n+1}$ est formée de tous les intervalles $\mathrm{I}_-, \mathrm{I}_+$ lorsque I varie dans $\mathcal{G}_n$.

On voit que $\mathcal{G}_n$ est formée de 2^n intervalles disjoints qui recouvrent $[0, 1[$. On considère la famille de fonctions formée de $h_\emptyset = 1$, puis de toutes les fonctions

$$h_{\mathrm{I}} = |\mathrm{I}|^{-1/2} \big(\mathbf{1}_{\mathrm{I}_-} - \mathbf{1}_{\mathrm{I}_+} \big)$$

où I varie dans toutes les familles $\mathcal{G}_n$, $n \geq 0$ et $|\mathrm{I}|$ dénote la longueur de l'intervalle I. Cette famille de fonctions est une base hilbertienne de l'espace de Hilbert $\mathrm{L}^2(0, 1)$. À cet effet, on vérifie d'abord que ces fonctions (h_{I}) sont deux à deux orthogonales (exercice). Ensuite, l'espace

$$\mathrm{V}_{n+1} = \mathrm{Vect}(\mathbf{1}_{\mathrm{I}} : \mathrm{I} \in \mathcal{G}_{n+1})$$

est de dimension 2^{n+1}, et il contient les fonctions $h_\emptyset$ et les h_{I}, pour I dans la réunion des $\mathcal{G}_k$, $0 \le k \le n$, qui sont indépendantes car orthogonales ; le nombre de ces fonctions est

$$1 + 1 + 2 + 4 + \cdots + 2^n = 2^{n+1},$$

donc elles forment une base de V_{n+1}, et il en résulte que

$$\mathrm{V}_{n+1} = \mathrm{Vect}(h_\emptyset, h_{\mathrm{I}} : \mathrm{I} \in \mathcal{G}_k,\ 0 \le k \le n).$$

On voit donc que l'espace engendré par la famille de toutes les fonctions h_{I} contient tous les espaces V_n ; il est clair que toute fonction en escalier peut être approchée dans $\mathrm{L}^2([0,1])$ par une fonction de la réunion des V_n. Il en résulte que la famille (h_{I}) est une base hilbertienne de $\mathrm{L}^2([0,1])$.

Gram-Schmidt et bases hilbertiennes

Le procédé de Gram-Schmidt est utile en dimension finie comme en dimension infinie. On va utiliser dans la proposition qui suit une convention de notation qui permet de traiter les deux cas en même temps, au prix d'une petite perte de lisibilité.

Proposition 5. *Désignons par* N *un entier* ≥ 0 *fini ou bien* $\mathrm{N} = +\infty$. *Si on a une suite* $(v_n)_{0 \le n < \mathrm{N}}$ *de vecteurs linéairement indépendants dans un espace de Hilbert* H, *il existe une suite orthonormée* $(f_n)_{0 \le n < \mathrm{N}}$ *telle que*

$$\mathrm{Vect}(f_k : 0 \le k \le n) = \mathrm{Vect}(v_k : 0 \le k \le n)$$

pour tout entier n *tel que* $0 \le n < \mathrm{N}$.

Preuve. — Le vecteur v_0 est non nul, puisqu'il fait partie d'un système libre ; on introduit pour commencer le vecteur de norme 1

$$f_0 = \|v_0\|^{-1} v_0,$$

et on a bien que $\mathrm{Vect}(f_0) = \mathrm{Vect}(v_0)$. Supposons que $f_0, \ldots, f_n$ aient été déterminés et que

$$\mathrm{Vect}(v_0, \ldots, v_n) = \mathrm{F}_n = \mathrm{Vect}(f_0, \ldots, f_n) \,;$$

la projection orthogonale P_n sur F_n est donnée pour tout vecteur $x \in \mathrm{H}$ par

$$\mathrm{P}_n x = \sum_{k=0}^{n} \langle x, f_k \rangle \, f_k.$$

Si $n+1 = \mathrm{N}$, c'est fini : on a traité tous les vecteurs disponibles. Sinon, $n+1 < \mathrm{N}$ et puisque les vecteurs (v_j) sont indépendants, on sait que $v_{n+1} \notin \mathrm{Vect}(v_0, \ldots, v_n) = \mathrm{F}_n$, donc $y_{n+1} = \mathrm{P}_n v_{n+1} \ne v_{n+1}$; posons

$$f_{n+1} = \|v_{n+1} - y_{n+1}\|^{-1} (v_{n+1} - y_{n+1}).$$

Ce vecteur de norme 1 est orthogonal à F_n, donc à $f_0, \ldots, f_n$ ce qui montre que les vecteurs $f_0, \ldots, f_{n+1}$ sont orthonormés. Par ailleurs, $f_{n+1} \in \mathrm{F}_{n+1} = \mathrm{Vect}(v_0, \ldots, v_{n+1})$ puisque $y_{n+1} \in \mathrm{F}_n \subset \mathrm{F}_{n+1}$ et $v_{n+1} \in \mathrm{F}_{n+1}$; les $n+2$ vecteurs $f_0, \ldots, f_{n+1}$ sont libres, et ils sont tous dans l'espace F_{n+1}, qui est de dimension $n+2$. Ils forment donc une base de F_{n+1} et par conséquent

$$\mathrm{Vect}(v_0, \ldots, v_{n+1}) = \mathrm{Vect}(f_0, \ldots, f_{n+1}).$$

On a ainsi démontré la possibilité de la récurrence.

Corollaire 1. *Si* F *est un sous-espace vectoriel de dimension finie de* H, *il admet des bases orthonormées ; le sous-espace* F *est fermé*([f]) *dans* H.

Preuve. — Si $\mathrm{N} = \dim \mathrm{F}$, on peut trouver une base $(v_n)_{0 \le n < \mathrm{N}}$ pour F ; en appliquant la proposition 5, on peut remplacer cette base par une base orthonormée $f_0, \ldots, f_{\mathrm{N}-1}$. La projection orthogonale $\mathrm{P_F}$ définie par

$$\forall x \in \mathrm{H}, \quad \mathrm{P_F} x = \sum_{k=0}^{\mathrm{N}-1} \langle x, f_k \rangle \, f_k$$

est continue, et F est exactement l'ensemble des vecteurs $x \in \mathrm{H}$ tels que $\mathrm{P_F} x = x$, donc $\mathrm{F} = \ker(\mathrm{Id_H} - \mathrm{P_F})$ est fermé.

Définition : espaces séparables. Un espace métrique (X, d) est *séparable* s'il existe une suite $(x_n)_{n \ge 0}$ de points de X qui est dense dans X : pour tout $x \in \mathrm{X}$ et tout $\varepsilon > 0$, il existe un entier n tel que $d(x, x_n) < \varepsilon$.

Il est connu que les espaces $\mathbb{R}$, $\mathbb{C}$ ou $\mathbb{R}^m$, pour tout entier $m \ge 2$, sont séparables, mais on va le revoir plus loin. On peut montrer que l'espace $\mathrm{L}^2([a, b])$ est séparable. On déduit de Gram-Schmidt le résultat qui suit.

Corollaire. *Si* H *est un espace de Hilbert de dimension infinie et séparable, il existe une base hilbertienne* $(e_n)_{n \ge 0}$ *pour* H, *indexée par l'ensemble* $\mathbb{N}$ *des entiers* ≥ 0.

Preuve. — Par définition il existe une suite dense (x_k) dans H ; on peut([g]) construire par récurrence une sous-suite $v_n = x_{k_n}$ formée de vecteurs indépendants et telle que $x_j \in \mathrm{Vect}(v_0, \ldots, v_n)$ pour tout $j \le k_n$. D'après Gram-Schmidt, il existe une suite orthonormée $(e_j)_{j \ge 0}$ telle que $\mathrm{Vect}(e_j, j \ge 0)$ contienne tous les $\mathrm{Vect}(v_0, \ldots, v_n)$ et en particulier tous les vecteurs $(x_k)_{k \ge 0}$. Il en résulte que $\mathrm{Vect}(e_j, j \ge 0)$ est dense dans H, donc $(e_j)_{j \ge 0}$ est une base hilbertienne, par définition.

Réciproque : *si* H *admet une base hilbertienne finie ou dénombrable, il est séparable.*

Preuve. — Supposons pour simplifier que $\mathbb{K} = \mathbb{R}$, et supposons d'abord que la base hilbertienne soit finie (l'espace H est de dimension finie dans ce cas), disons $(e_1, \ldots, e_\mathrm{N})$. On considère pour tout $n \ge 1$ l'ensemble A_n des vecteurs z de la forme

$$z = \sum_{k=1}^{\mathrm{N}} c_k \, e_k$$

où $|c_k| \le n$ pour $1 \le k \le \mathrm{N}$; il est clair que $\mathrm{A}_n \subset \mathrm{A}_{n+1}$ pour tout n et que la réunion des A_n est égale à H. Considérons ensuite, pour chaque entier $n \ge \mathrm{N}$, le sous-ensemble B_n des points z de A_n dont les coordonnées c_k sont astreintes à être des multiples entiers $j\, n^{-2}$, $j = -n^3, -n^3+1, \ldots, n^3$ de n^{-2} ; l'approximation d'un élément quelconque a de A_n par un élément de B_n sera possible avec une erreur $\le \mathrm{N}\, n^{-2} \le 1/n$ (remplacer chaque coordonnée de a par le réel $j\, n^{-2}$ le plus proche, avec une erreur $\le n^{-2}$). L'ensemble B_n est fini, de cardinal $(2n^3+1)^\mathrm{N}$; si on place dans une liste $(d_k)_{k \ge 0}$ tous les éléments de $\mathrm{B_N}$, suivis de tous les éléments de $\mathrm{B_{N+1}}$, $\mathrm{B_{N+2}}$, etc., il est clair que la suite (d_k) sera dense dans H.

Si la dimension de H est infinie, soit $(e_k)_{k \ge 0}$ la base hilbertienne de H, donnée par hypothèse ; d'après la première étape on peut trouver pour tout entier k un ensemble dénombrable $\mathrm{D}_k \subset \mathrm{H}_k = \mathrm{Vect}(e_0, \ldots, e_k)$ qui soit dense dans H_k ; la réunion $\mathrm{D} = \bigcup_k \mathrm{D}_k$

est encore dénombrable, et elle est dense dans H : si x est un vecteur de H, on sait que x est la limite dans H des vecteurs

$$x_k = \sum_{j=0}^{k} \langle x, e_j \rangle \, e_j \in \mathrm{H}_k,$$

lorsque $k \to +\infty$; on peut donc trouver k tel que $\|x - x_k\| < \varepsilon/2$. On peut ensuite trouver un élément d de D_k, donc de D, tel que $\|x_k - d\| < \varepsilon/2$ et finalement $\|x - d\| < \varepsilon$, ce qui montre que D est dense dans H.

2.2. *La base hilbertienne des exponentielles complexes*

Pour chaque entier $n \in \mathbb{Z}$, définissons la fonction e_n par

$$\forall x \in \mathbb{R}, \quad e_n(x) = \mathrm{e}^{\mathrm{i}nx}.$$

Ces fonctions sont 2π-périodiques ; on les considérera aussi comme des fonctions sur l'intervalle $[0, 2\pi]$, qu'on munira de la mesure $\mathrm{d}\mu(x) = (2\pi)^{-1}\mathrm{d}x$, multiple de la mesure de Lebesgue. Pour cette mesure μ, l'intervalle $[0, 2\pi]$ est de mesure 1, et on a pour les produits scalaires dans $\mathrm{L}^2(\mu)$, lorsque $m \neq n$

$$\begin{aligned} \langle e_m, e_n \rangle &= \int_0^{2\pi} e_m(x) \overline{e_n(x)} \, \mathrm{d}\mu(x) = \int_0^{2\pi} \mathrm{e}^{\mathrm{i}mx} \, \mathrm{e}^{-\mathrm{i}nx} \, \mathrm{d}\mu(x) \\ &= \frac{1}{2\pi} \int_0^{2\pi} \mathrm{e}^{\mathrm{i}(m-n)x} \, \mathrm{d}x = \frac{1}{2\pi} \Big[\frac{\mathrm{e}^{\mathrm{i}(m-n)x}}{\mathrm{i}(m-n)} \Big]_{x=0}^{2\pi} = 0. \end{aligned}$$

Lorsque $m = n$,

$$\langle e_m, e_m \rangle = \frac{1}{2\pi} \int_0^{2\pi} \mathrm{e}^{\mathrm{i}(m-m)x} \, \mathrm{d}x = \frac{1}{2\pi} \int_0^{2\pi} \mathrm{d}x = 1.$$

Il s'agit donc d'une famille orthonormée. On va maintenant montrer que cette famille $(e_n)_{n \in \mathbb{Z}}$ est une base hilbertienne de l'espace $\mathrm{L}^2([0, 2\pi], \mathrm{d}x/(2\pi))$. Avant de se lancer dans la preuve, on va faire quelques rappels et introduire de nouveaux éléments utiles.

Norme uniforme, convergence uniforme

Si $\mathrm{C}([a, b])$ désigne l'espace vectoriel des fonctions continues sur l'intervalle fermé borné $[a, b]$, on le munit de la *norme uniforme* $\|f\|_u$ définie pour toute fonction continue f par

$$\|f\|_u = \max\{|f(t)| : t \in [a, b]\}$$

(d'après le théorème A5 de l'annexe du chapitre 1, cette norme est égale à $\|f\|_\infty$, la norme induite par l'espace $\mathrm{L}^\infty([a, b])$, où $[a, b]$ est muni de la mesure de Lebesgue). Dire qu'une suite (f_n) de fonctions continues converge uniformément sur $[a, b]$ vers une fonction f peut s'exprimer au moyen de la norme uniforme,

$$\lim_n \|f_n - f\|_u = 0.$$

Quand f est une fonction continue 2π-périodique sur $\mathbb{R}$, son maximum sur $\mathbb{R}$ est identique à son maximum sur n'importe quel intervalle $[a, b]$ de longueur 2π, par exemple égal au maximum de f sur l'intervalle $[a, b] = [0, 2\pi]$.

Intégrale sur une période. Si f est continue 2π-périodique, la dérivée en x de

$$\Phi : x \to \int_x^{x+2\pi} f(t)\,\mathrm{d}t$$

est nulle, puisqu'elle vaut $f(x+2\pi) - f(x) = 0$, ce qui montre que l'intégrale sur une période ne dépend pas de l'intervalle de longueur 2π choisi. Le même résultat est vrai aussi pour une fonction 2π-périodique intégrable sur $[0, 2\pi]$, en découpant l'intégrale : si $0 < x < 2\pi$ on peut écrire, d'abord pour une fonction mesurable positive f (qui permet de raisonner avec des intégrales infinies)

$$\Phi(x) = \int_x^{x+2\pi} f(t)\,\mathrm{d}t = \int_x^{2\pi} f(t)\,\mathrm{d}t + \int_{2\pi}^{x+2\pi} f(t)\,\mathrm{d}t$$

$$= \int_x^{2\pi} f(t)\,\mathrm{d}t + \int_0^x f(t)\,\mathrm{d}t = \int_0^{2\pi} f(t)\,\mathrm{d}t = \Phi(0)\,;$$

pour x quelconque, on peut trouver x_1 de la forme $x + 2k\pi$, $k \in \mathbb{Z}$, tel que $0 < x_1 < 2\pi$; par la périodicité de f on voit facilement que $\Phi(x) = \Phi(x_1)$, et on sait que $\Phi(x_1) = \Phi(0)$ par ce qui précède. Si f est intégrable, on reprend les calculs, et on se sert, pour justifier l'existence des intégrales, du cas de $|f|$ déjà traité.

La *convolution périodique* de f, g, fonctions 2π-périodiques intégrables sur $[0, 2\pi]$, sera définie par

$$\forall x \in \mathbb{R}, \quad (f *_{\text{per}} g)(x) = \int_a^{a+2\pi} f(x-t)g(t)\,\frac{\mathrm{d}t}{2\pi},$$

où la valeur de l'intégrale ne dépend pas de a, d'après la remarque précédente. Le fait que l'intégrale ait un sens pour presque tout x résulte du cas de la convolution sur $\mathbb{R}$, traité au chapitre 1 : en effet, si on se limite à x dans un intervalle borné $[u, v]$, la formule précédente coïncide, au facteur 2π près, avec la convolution sur $\mathbb{R}$ de deux fonctions intégrables sur $\mathbb{R}$, à savoir $\mathbf{1}_{[u-a-2\pi, v-a]}\, f$ et $\mathbf{1}_{[a, a+2\pi]}\, g$. Le fait que la valeur de la convolution périodique ne dépend pas de l'intervalle de longueur 2π choisi permet de montrer par changement de variable que

$$(f *_{\text{per}} g)(x) = (g *_{\text{per}} f)(x).$$

Théorème 1. *Les fonctions $(e_n)_{n\in\mathbb{Z}}$ définies par $e_n(t) = \mathrm{e}^{\mathrm{i}nt}$ forment une base hilbertienne de l'espace de Hilbert $\mathrm{L}^2([0, 2\pi], \mathrm{d}x/(2\pi))$.*

La preuve sera un peu longue. On sait déjà que $(e_n)_{n\in\mathbb{Z}}$ est une suite orthonormée, il reste à prouver qu'elle engendre un sous-espace vectoriel qui est dense dans $\mathrm{L}^2([0, 2\pi])$. Comme on sait que les fonctions continues qui sont nulles en 0 et en 2π sont denses(**h**) dans l'espace $\mathrm{L}^2([0, 2\pi])$, il suffit de voir que toute fonction continue nulle en 0 et 2π peut être approchée en norme L^2 par un *polynôme trigonométrique*

$$\mathrm{P} = \sum_{k=-\mathrm{N}}^{\mathrm{N}} c_k e_k.$$

Comme la mesure de $[0,2\pi]$ est finie, il suffit de montrer une approximation uniforme ; en effet si g est continue sur $[0,2\pi]$ et si P est un polynôme trigonométrique tel que $\|g-\mathrm{P}\|_\infty < \varepsilon$, on aura

$$\|g-\mathrm{P}\|_2^2 = \int_0^{2\pi} |g(x)-\mathrm{P}(x)|^2 \,\frac{\mathrm{d}x}{2\pi} \le \int_0^{2\pi} \|g-\mathrm{P}\|_\infty^2 \,\frac{\mathrm{d}x}{2\pi} = \|g-\mathrm{P}\|_\infty^2$$

donc $\|g-\mathrm{P}\|_2 \le \|g-\mathrm{P}\|_\infty < \varepsilon$. Comme toute fonction continue sur $[0,2\pi]$, nulle en 0 et 2π peut ([i]) être étendue en fonction 2π-périodique continue sur $\mathbb{R}$, il suffit de prouver le théorème qui suit.

Théorème 2 (Weierstrass périodique). *Si f est une fonction continue 2π-périodique sur $\mathbb{R}$, on peut l'approcher uniformément par des polynômes trigonométriques.*

Preuve. — Pour la preuve on va se servir du *noyau de Poisson* : pour $0 < r < 1$ on pose

$$\forall \theta \in \mathbb{R}, \quad \mathrm{P}_r(\theta) = \sum_{n\in\mathbb{Z}} r^{|n|}\,\mathrm{e}^{\mathrm{i}n\theta} = \sum_{n=0}^{+\infty} r^n\,\mathrm{e}^{\mathrm{i}n\theta} + \sum_{n=1}^{+\infty} r^n\,\mathrm{e}^{-\mathrm{i}n\theta}$$

$$= \frac{1}{1-r\,\mathrm{e}^{\mathrm{i}\theta}} + \frac{r\,\mathrm{e}^{-\mathrm{i}\theta}}{1-r\,\mathrm{e}^{-\mathrm{i}\theta}} = \frac{1-r^2}{|1-r\,\mathrm{e}^{\mathrm{i}\theta}|^2} = \frac{1-r^2}{1+r^2-2r\cos(\theta)}.$$

La série converge normalement, ce qui permet par exemple d'intervertir intégrale et série, et de voir ainsi que

$$\int_{-\pi}^{\pi} \mathrm{P}_r(\theta)\,\frac{\mathrm{d}\theta}{2\pi} = \sum_{n\in\mathbb{Z}} r^{|n|} \int_{-\pi}^{\pi} \mathrm{e}^{\mathrm{i}n\theta}\,\frac{\mathrm{d}\theta}{2\pi} = 1,$$

car les intégrales sont nulles, sauf si $n=0$. Cette fonction continue P_r est positive, paire, décroissante sur $[0,\pi]$ ce qui implique quand $0 < \delta < \pi$

$$\int_{-\pi}^{\pi} \mathbf{1}_{\{|t|\ge\delta\}} \mathrm{P}_r(t)\,\frac{\mathrm{d}t}{2\pi} = 2\int_\delta^\pi \mathrm{P}_r(t)\,\frac{\mathrm{d}t}{2\pi} \le 2\mathrm{P}_r(\delta)\int_\delta^\pi \frac{\mathrm{d}t}{2\pi} \le \mathrm{P}_r(\delta) = \frac{1-r^2}{1+r^2-2r\cos(\delta)}$$

qui tend vers 0 quand $r \to 1$, avec δ fixé. Autrement dit, la « masse » de P_r (limitée à l'intervalle $[-\pi,\pi]$) se concentre à l'origine quand $r \to 1$; c'est un cas particulier du phénomène *d'approximation de l'unité*. On montre alors que

a. La convolée $f *_{\mathrm{per}} \mathrm{P}_r$ tend uniformément vers f quand $r \to 1$.

b. La convolée $f *_{\mathrm{per}} \mathrm{P}_r$ est une série trigonométrique, limite uniforme de polynômes trigonométriques.

Pour le point *a*, on se donne $\varepsilon > 0$ et on trouve, parce que f est ([j]) uniformément continue sur $\mathbb{R}$, un réel $\delta > 0$ tel que $|f(x)-f(y)| < \varepsilon/2$ dès que $|x-y| < \delta$; on peut supposer $0 < \delta < \pi$ et on choisit ensuite r suffisamment proche de 1 pour que

$$\int_\delta^\pi \mathrm{P}_r(t)\,\frac{\mathrm{d}t}{2\pi} < \frac{\varepsilon}{1+8\|f\|_u}\ ;$$

on écrit

$$f(x) - (f*\mathrm{P}_r)(x) = \int_{-\pi}^{\pi} \bigl(f(x)-f(x-t)\bigr)\,\mathrm{P}_r(t)\,\frac{\mathrm{d}t}{2\pi}\ ;$$

on découpe l'intégrale selon que $|t| < \delta$ ou bien $|t| \ge \delta$; dans le premier cas, on a $|f(x) - f(x-t)| < \varepsilon/2$ par le choix de δ, et dans le deuxième cas on majore la différence par $2\,\|f\|_u$; on obtient ainsi, pour tout x réel

$$|f(x) - (f * \mathrm{P}_r)(x)| \le \frac{\varepsilon}{2} \int_{-\pi}^{\pi} \mathrm{P}_r(t)\, \frac{\mathrm{d}t}{2\pi} + 2\,\|f\|_u \int_{\delta \le |t| \le \pi} \mathrm{P}_r(t)\, \frac{\mathrm{d}t}{2\pi}$$

$$\le \frac{\varepsilon}{2} + 4\,\|f\|_u\, \frac{\varepsilon}{1 + 8\|f\|_u} < \varepsilon,$$

donc $\|f - f * \mathrm{P}_r\|_u < \varepsilon$. Pour le point b, on écrit grâce à la convergence normale de la série qui définit P_r

$$(f * \mathrm{P}_r)(x) = \int_{-\pi}^{\pi} f(s) \mathrm{P}_r(x-s)\, \frac{\mathrm{d}s}{2\pi} = \int_{-\pi}^{\pi} f(s) \Bigl(\sum_{n \in \mathbb{Z}} r^{|n|}\, \mathrm{e}^{\mathrm{i}n(x-s)} \Bigr) \frac{\mathrm{d}s}{2\pi}$$

$$(*) \qquad = \sum_{n \in \mathbb{Z}} r^{|n|} \int_{-\pi}^{\pi} f(s)\, \mathrm{e}^{\mathrm{i}n(x-s)}\, \frac{\mathrm{d}s}{2\pi} = \sum_{n \in \mathbb{Z}} r^{|n|} c_n(f)\, \mathrm{e}^{\mathrm{i}nx},$$

où

$$c_n(f) = \int_{-\pi}^{\pi} f(s)\, \mathrm{e}^{-\mathrm{i}ns}\, \frac{\mathrm{d}s}{2\pi}$$

est le n ième *coefficient de Fourier complexe* de la fonction 2π-périodique f. Il est clair que $|c_n(f)| \le \|f\|_u$ pour tout n, donc la série dans l'équation $(*)$ est normalement convergente ; ceci implique que la fonction somme $f *_{\mathrm{per}} \mathrm{P}_r$ est limite uniforme des sommes partielles

$$x \to \sum_{n=-\mathrm{N}}^{\mathrm{N}} r^{|n|} c_n(f)\, \mathrm{e}^{\mathrm{i}nx},$$

comme on voulait le montrer. Ceci termine la démonstration du théorème de Weierstrass périodique.

Il résulte de tout ceci que le système trigonométrique est une base hilbertienne : pour toute fonction $f \in \mathrm{L}^2([0, 2\pi])$, on a dans l'espace L^2

$$f = \sum_{n \in \mathbb{Z}} \langle f, e_n \rangle\, e_n$$

où

$$\langle f, e_n \rangle = \int_0^{2\pi} f(t)\, \mathrm{e}^{-\mathrm{i}nt}\, \frac{\mathrm{d}t}{2\pi} = c_n(f).$$

La série indexée par $\mathbb{Z}$ peut sembler troublante, on va la commenter. Tout d'abord, on sait d'après la proposition 4 que les deux séries « ordinaires »

$$\sum_{n=0}^{+\infty} \langle f, e_n \rangle\, e_n, \qquad \sum_{n=1}^{+\infty} \langle f, e_{-n} \rangle\, e_{-n} = \sum_{-\infty}^{-1} \langle f, e_n \rangle\, e_n$$

convergent dans L^2. De plus, si on énumère tous les indices dans $\mathbb{Z}$, par exemple sous la forme $0, -1, 1, -2, 2, \ldots$ c'est-à-dire que $n_{2k} = k$ et $n_{2k-1} = -k$, on sait aussi que

$$f = \sum_{k=0}^{+\infty} \langle f, e_{n_k} \rangle \, e_{n_k}.$$

Mais la somme partielle d'indice $2p$ est égale à

$$\sum_{k=0}^{2p} \langle f, e_{n_k} \rangle \, e_{n_k} = \sum_{n=-p}^{-1} \langle f, e_n \rangle \, e_n + \sum_{n=0}^{p} \langle f, e_n \rangle \, e_n,$$

qui converge quand $p \to +\infty$ vers la somme des deux séries «ordinaires» ; on peut donc écrire pour se rassurer

$$f = \sum_{n \in \mathbb{Z}} \langle f, e_n \rangle \, e_n = \sum_{-\infty}^{-1} \langle f, e_n \rangle \, e_n + \sum_{n=0}^{+\infty} \langle f, e_n \rangle \, e_n.$$

Le fait que $f = \sum_{n \in \mathbb{Z}} \langle f, e_n \rangle \, e_n$ au sens de l'espace L^2 n'indique pas s'il y a égalité pour des valeurs de x, ni même si la série numérique $\sum_{n \in \mathbb{Z}} c_n(f) \, \mathrm{e}^{\mathrm{i} n x}$ converge pour des valeurs de x. C'est vrai d'après le théorème suivant, qui est un résultat très difficile datant du milieu des années 1960 (l'article du mathématicien suédois Lennart Carleson est paru en 1966).

Théorème de Carleson : *pour toute fonction $f \in \mathrm{L}^2([0, 2\pi])$ et pour presque tout x, la série de Fourier*

$$\sum_{k=-\infty}^{+\infty} c_k(f) \, \mathrm{e}^{\mathrm{i} k x}$$

converge et sa somme est égale à $f(x)$.

Quand nous disons que la série converge, cela signifie que *les deux* séries $\sum_{k=0}^{+\infty} c_k(f) \, \mathrm{e}^{\mathrm{i} k x}$ et $\sum_{k=0}^{+\infty} c_{-k}(f) \, \mathrm{e}^{-\mathrm{i} k x}$ convergent.

Exemple : développement en série de Fourier de la fonction périodique f définie par $f(x) = 1 - |x|/\pi$ lorsque $|x| \le \pi$. Pour $n \ne 0$, on voit en utilisant la parité de f que

$$\begin{aligned} c_n(f) &= \int_{-\pi}^{\pi} f(x) \, \mathrm{e}^{-\mathrm{i} n x} \, \frac{\mathrm{d}x}{2\pi} = \int_{-\pi}^{\pi} f(x) \cos(nx) \, \frac{\mathrm{d}x}{2\pi} = \frac{1}{\pi} \int_0^{\pi} (1 - x/\pi) \cos(nx) \, \mathrm{d}x \\ &= \frac{1}{\pi} \Bigl(\frac{1}{n} \Bigl[(1 - x/\pi) \sin(nx) \Bigr]_0^{\pi} + \frac{1}{\pi n} \int_0^{\pi} \sin(nx) \, \mathrm{d}x \Bigr) = \frac{1 - \cos(n\pi)}{\pi^2 n^2}, \end{aligned}$$

qui est égal à 0 pour n pair non nul, et à $2/(\pi^2 n^2)$ pour n impair. On voit aussi que $c_0(f) = 1/2$. La série de Fourier de f est donc

$$\frac{1}{2} + \sum_{n \in \mathbb{Z}} \frac{1 - \cos(n\pi)}{\pi^2 n^2} \, \mathrm{e}^{\mathrm{i} n x} = \frac{1}{2} + \frac{4}{\pi^2} \sum_{k=0}^{+\infty} \frac{\cos((2k+1)x)}{(2k+1)^2}.$$

La série précédente est normalement convergente, et sa somme $g(x)$ définit donc une fonction continue ; comme f et g sont continues et doivent être égales presque partout (puisqu'elles représentent la même classe dans L^2) il en résulte(**k**) que $f(x) = g(x)$ pour tout x. La valeur en $x = 0$ ou $x = \pi$ donne $\sum_{k=0}^{+\infty} (2k+1)^{-2} = \pi^2/8$, ce qui implique l'égalité classique

$$\sum_{n=1}^{+\infty} \frac{1}{n^2} = \frac{\pi^2}{6} \, .$$

2.3. *Projection orthogonale et applications*

Lemme 2. *Soient* H *un espace de Hilbert sur* $\mathbb{K}$, F *un* $\mathbb{K}$*-sous-espace vectoriel de* H, $x \in \mathrm{H}$ *et* $y \in \mathrm{F}$ *; alors* $x - y$ *est orthogonal à* F *si et seulement si* y *est le point de* F *le plus proche de* x,

$$\|x - y\| = d(x, \mathrm{F}) = \min\{\|x - z\| : z \in \mathrm{F}\}.$$

Le point $y \in \mathrm{F}$ *qui minimise la distance de* x *aux points de* F *est unique (s'il existe ; dans ce cas, il est appelé* projection orthogonale *de* x *sur* F*).*

Preuve. — Supposons d'abord que $x - y \perp \mathrm{F}$; pour tout vecteur $z \in \mathrm{F}$, on voit que $y - z \in \mathrm{F}$, donc $y - z$ est orthogonal à $x - y$ et

$$\|x - z\|^2 = \|(x - y) + (y - z)\|^2 = \|x - y\|^2 + \|y - z\|^2 \geq \|x - y\|^2,$$

donc y est le point de F le plus proche de x. Supposons inversement que y soit le point de F le plus proche de x ; si v est un vecteur de F, le vecteur $y + tv$ est dans F pour tout réel t, donc

$$\|x - (y + tv)\|^2 \geq \|x - y\|^2$$

pour tout t ; la fonction

$$t \to \|x - (y + tv)\|^2 = \|x - y\|^2 - 2t \operatorname{Re}\langle x - y, v\rangle + t^2\|v\|^2$$

atteint son minimum en $t = 0$, donc sa dérivée en $t = 0$ est nulle, ce qui donne

$$\operatorname{Re}\langle x - y, v\rangle = 0$$

pour tout vecteur $v \in \mathrm{F}$; ceci est suffisant quand $\mathbb{K} = \mathbb{R}$, mais quand $\mathbb{K} = \mathbb{C}$ il faut faire un pas de plus : comme F est alors un $\mathbb{C}$-sous-espace vectoriel de H, on peut choisir θ réel de façon que si $v_1 = \mathrm{e}^{\mathrm{i}\theta} v$, on ait $\langle x - y, v_1\rangle \in \mathbb{R}$. On a encore $v_1 \in \mathrm{F}$: la première partie du raisonnement montre que $\langle x - y, v_1\rangle = 0$, donc $\langle x - y, v\rangle = 0$, pour tout vecteur $v \in \mathrm{F}$, ce qui signifie que $x - y$ est orthogonal à F.

Le point $y \in \mathrm{F}$ qui minimise la distance est unique : si y' était un autre point de F tel que $x - y'$ soit orthogonal à F, on aurait par différence $y - y' \perp \mathrm{F}$, et comme $y - y' \in \mathrm{F}$, le vecteur $y - y'$, orthogonal à lui-même, serait nul.

Prélude : considérons d'abord un sous-espace vectoriel F fermé, séparable et de dimension infinie. D'après le corollaire de Gram-Schmidt, on peut trouver une base hilbertienne $(f_n)_{n \geq 0}$ pour F. La proposition 4 donne l'existence de la projection orthogonale sur F dans ce cas. On peut ensuite prouver par un petit bricolage l'existence de la projection de x sur un sous-espace vectoriel fermé F quelconque (c'est-à-dire même si F n'est pas séparable) : on peut trouver une suite $(y_n) \subset \mathrm{F}$ telle que $\|x - y_n\|$ tende vers la *distance de* x à F

$$d = d(x, \mathrm{F}) = \inf\{\|x - z\| : z \in \mathrm{F}\} ;$$

le sous-espace vectoriel fermé $\mathrm{G} \subset \mathrm{F}$ engendré par la suite (y_n) vérifie $d(x, \mathrm{G}) = d$ et il possède une base hilbertienne finie ou dénombrable d'après Gram-Schmidt. Le point $g \in \mathrm{G}$ le plus proche de x existe donc, d'après ce qui précède, mais comme $\|x - g\| = d$, le point g est en même temps le point de F le plus proche de x.

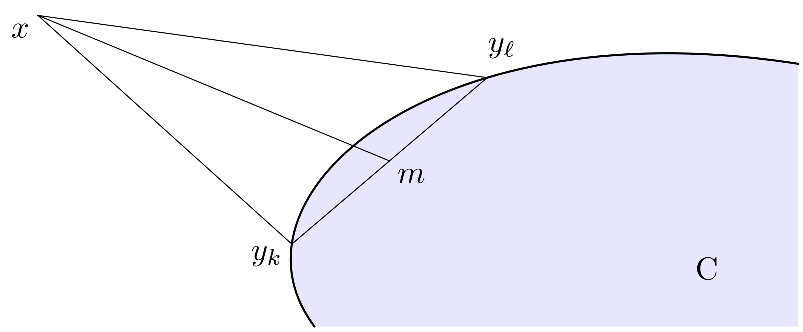
Projection en général

Considérons plus généralement un sous-ensemble convexe fermé non vide C de l'espace de Hilbert H, et un point $x \in \mathrm{H}$. On peut toujours trouver une suite $(y_n) \subset \mathrm{C}$ telle que $d(x, y_n) = \|x - y_n\|$ tende vers

$$d := d(x, \mathrm{C}) = \inf\{\|x - z\| : z \in \mathrm{C}\}.$$

On va montrer que cette suite (y_n) est de Cauchy. On pose $m = \frac{1}{2}(y_k + y_\ell)$, le milieu du segment qui joint les points y_k et y_ℓ ; ce point m est dans C, d'après la convexité de C. Posons de plus $u = x - y_k$ et $v = x - y_\ell$. On obtient par la relation du parallélogramme (2)

$$4\|x - m\|^2 + \|y_k - y_\ell\|^2 = 2\big(\|x - y_k\|^2 + \|x - y_\ell\|^2\big).$$



Comme $m \in \mathrm{C}$, on a $\|x - m\| \ge d$ et

$$\frac{1}{2}\|y_k - y_\ell\|^2 \le \|x - y_k\|^2 + \|x - y_\ell\|^2 - 2\,d^2$$

qui tend vers 0 quand $k, \ell \to +\infty$. La suite de Cauchy (y_n) converge dans l'espace complet H vers un vecteur y, qui est dans C parce que C est fermé. On a de plus

$$\|x - y\| = \lim_n \|x - y_n\| = d,$$

et on a ainsi montré l'existence d'un point $y \in \mathrm{C}$ qui réalise la plus courte distance de x à un point de C. L'unicité résulte ici de la relation du parallélogramme : si y, y' sont deux points de C tels que

$$\|x - y\| = \|x - y'\| = d = d(x, \mathrm{C}),$$

on peut appliquer les inégalités ci-dessus en prenant $y_k = y$ et $y_\ell = y'$; alors

$$\frac{1}{2}\|y - y'\|^2 = \frac{1}{2}\|y_k - y_\ell\|^2 \le \|x - y_k\|^2 + \|x - y_\ell\|^2 - 2\,d^2 = d^2 + d^2 - 2\,d^2 = 0.$$

On a donc montré le théorème qui suit.

Théorème 3. *Soient* H *un espace de Hilbert et* C *un sous-ensemble convexe, fermé et non vide de* H *; pour tout vecteur* $x \in \mathrm{H}$*, il existe un élément* y *de* C unique *qui est le point de* C *le plus proche de* x*,*

$$\|x - y\| = \min\{\|x - z\| : z \in \mathrm{C}\}.$$

Linéarité de la projection sur un sous-espace vectoriel fermé

Dans le cas où on projette sur un *sous-espace vectoriel* fermé F, la projection $\mathrm{P_F}$ est linéaire : si les points $x_1, x_2 \in \mathrm{H}$ ont pour projections $y_1 = \mathrm{P_F} x_1$ et $y_2 = \mathrm{P_F} x_2 \in \mathrm{F}$, on voit facilement que le vecteur

$$a_1 x_1 + a_2 x_2 - (a_1 y_1 + a_2 y_2) = a_1(x_1 - y_1) + a_2(x_2 - y_2)$$

est orthogonal à F, ce qui montre que $a_1 y_1 + a_2 y_2 \in \mathrm{F}$ est la projection orthogonale de $a_1 x_1 + a_2 x_2$ sur F (lemme 2). La projection $\mathrm{P_F}$ est donc linéaire,

$$\mathrm{P_F}(a_1 x_1 + a_2 x_2) = a_1 y_1 + a_2 y_2 = a_1 \, \mathrm{P_F} x_1 + a_2 \, \mathrm{P_F} x_2.$$

Notons encore que par Pythagore, $\|x\|^2 = \|x - \mathrm{P_F} x\|^2 + \|\mathrm{P_F} x\|^2 \ge \|\mathrm{P_F} x\|^2$.

Théorème 4. *Soit* F *un sous-espace vectoriel fermé de l'espace de Hilbert* H *; l'application* $\mathrm{P_F}$ *de projection orthogonale de* H *sur* F *est linéaire, et*

$$\forall x \in \mathrm{H}, \quad \|\mathrm{P_F} x\| \le \|x\|.$$

Exemples.

1. Partition finie, moyennes. Considérons un espace de probabilité $(\Omega, \mathcal{A}, \mathrm{P})$. Supposons que $\mathrm{A}_1, \ldots, \mathrm{A_N}$ soit une partition de Ω en ensembles de la tribu $\mathcal{A}$, tels que $\mathrm{P}(\mathrm{A}_j) > 0$ pour tout $j = 1, \ldots, \mathrm{N}$; le sous-espace F de dimension N de $\mathrm{L}^2(\Omega, \mathcal{A}, \mathrm{P})$ formé des fonctions qui sont constantes sur chaque ensemble de la partition admet pour base orthonormée les fonctions

$$f_j = \mathrm{P}(\mathrm{A}_j)^{-1/2} \mathbf{1}_{\mathrm{A}_j}, \quad j = 1, \ldots, \mathrm{N}.$$

La projection orthogonale de $f \in \mathrm{L}^2(\Omega, \mathcal{A}, \mathrm{P})$ sur F est donnée par

$$\mathrm{P_F} f = \sum_{j=1}^{\mathrm{N}} \langle f, f_j \rangle \, f_j = \sum_{j=1}^{\mathrm{N}} \Bigl(\frac{1}{\mathrm{P}(\mathrm{A}_j)} \int_{\mathrm{A}_j} f \, \mathrm{dP} \Bigr) \mathbf{1}_{\mathrm{A}_j} \, ;$$

la fonction $\mathrm{P_F} f$ est constante sur chaque ensemble A_j, et sa valeur sur A_j est la *moyenne de* f *sur* A_j.

2. Dans l'espace de Hilbert $\mathrm{H} = \mathrm{L}^2([0,1]^2)$ on considère le sous-espace vectoriel F formé des fonctions ne dépendant que de la variable x : la fonction g appartient à F s'il existe une fonction $\mathrm{G}(x)$ d'une seule variable, de carré intégrable sur $[0,1]$, et telle que $g(x,y) = \mathrm{G}(x)$ pour presque tout couple $(x,y) \in [0,1]^2$. Ce sous-espace F est fermé (exercice). La projection orthogonale sur F d'une fonction $f \in \mathrm{H}$ est donnée par

$$(\mathrm{P_F} f)(x,y) = \mathrm{G}(x) = \int_0^1 f(x,y) \, \mathrm{d}y.$$

Pour toute partie $\mathrm{A} \subset \mathrm{H}$, on définit l'orthogonal $\mathrm{A}^\perp$ de cette partie,

$$\mathrm{A}^\perp = \{x \in \mathrm{H} : x \perp \mathrm{A}\},$$

où la notation $x \perp \mathrm{A}$ signifie que x est orthogonal à tous les éléments de A ; si A est vide, il est naturel de poser $\mathrm{A}^\perp = \mathrm{H}$. On voit que $\mathrm{A}^\perp$ est toujours un sous-espace vectoriel fermé de H, pour toute partie $\mathrm{A} \subset \mathrm{H}$, puisque $\mathrm{A}^\perp$ est l'intersection des sous-espaces vectoriels fermés $a^\perp = \ker \ell_a$, où a varie dans l'ensemble A.

Il est clair que $\mathrm{A}^\perp \supset \mathrm{B}^\perp$ lorsque $\mathrm{A} \subset \mathrm{B}$; en particulier $\mathrm{H}^\perp = \{0\}$ est le plus petit orthogonal, correspondant à la plus grande partie possible, $\mathrm{A} = \mathrm{H}$. On a

$$\mathrm{A}^\perp = \bigl(\overline{\mathrm{Vect}\,\mathrm{A}}\bigr)^\perp. \tag{3}$$

Puisque $\mathrm{A} \subset \overline{\mathrm{Vect}\,\mathrm{A}}$, il est clair que $\mathrm{A}^\perp \supset \bigl(\overline{\mathrm{Vect}\,\mathrm{A}}\bigr)^\perp$. Supposons inversement que $y \in \mathrm{A}^\perp$, c'est-à-dire que y soit orthogonal à A ; ceci signifie que A est contenu dans le sous-espace vectoriel $\mathrm{V} = y^\perp$; si V contient A, il contient aussi le sous-espace vectoriel $\mathrm{Vect}(\mathrm{A})$ engendré par A. Mais $\mathrm{V} = y^\perp$ est aussi un ensemble fermé ; s'il contient $\mathrm{Vect}(\mathrm{A})$, il contient aussi son adhérence $\overline{\mathrm{Vect}\,\mathrm{A}}$. Mais l'inclusion $\overline{\mathrm{Vect}\,\mathrm{A}} \subset y^\perp$ signifie que $y \in \bigl(\overline{\mathrm{Vect}\,\mathrm{A}}\bigr)^\perp$, et la vérification est finie.

Lemme. *Soit* H *un espace de Hilbert ; on suppose que l'espace* H *est égal à la somme vectorielle* $\mathrm{F} + \mathrm{G}$*, où* F *et* G *sont deux sous-espaces vectoriels orthogonaux (tout vecteur* $f \in \mathrm{F}$ *est orthogonal à tout vecteur* $g \in \mathrm{G}$*) ; alors* H *admet la décomposition en somme directe*

$$\mathrm{H} = \mathrm{F} \oplus \mathrm{G},$$

on a $\mathrm{G} = \mathrm{F}^\perp$, $\mathrm{F} = \mathrm{G}^\perp$*, les sous-espaces* F *et* G *sont fermés, les deux projections* $f + g \to f \in \mathrm{F}$ *et* $f + g \to g \in \mathrm{G}$ *de la somme directe sont les projections orthogonales de* H *sur* F *et* G *respectivement.*

Preuve. — Pour commencer, si $x \in \mathrm{F} \cap \mathrm{G}$, alors x élément de F est orthogonal à lui-même, élément de G, donc $x = 0$; puisque $\mathrm{F} \cap \mathrm{G} = \{0\}$ et $\mathrm{H} = \mathrm{F} + \mathrm{G}$, l'espace H est somme directe de F et G. Par hypothèse on a $\mathrm{G} \subset \mathrm{F}^\perp$; inversement, si $v \in \mathrm{H}$ est orthogonal à F, écrivons $v = f + g$, avec $f \in \mathrm{F}$ et $g \in \mathrm{G}$; on a

$$0 = \langle v, f \rangle = \langle f + g, f \rangle = \langle f, f \rangle + \langle f, g \rangle = \langle f, f \rangle,$$

ce qui montre que $f = 0$, donc $v = g$ est dans G et on a montré que $\mathrm{G} = \mathrm{F}^\perp$. On montre de la même façon que $\mathrm{F} = \mathrm{G}^\perp$, et il en résulte que F et G sont fermés.

Si $x = f + g$, avec $f \in \mathrm{F}$ et $g \in \mathrm{G}$, le vecteur f est dans F, et $x - f = g$ est orthogonal à F, donc f est la projection orthogonale de x sur F. De même, g est la projection orthogonale de x sur G.

Remarque. Le lemme peut s'appliquer quand H est la somme de trois (ou plus) sous-espaces orthogonaux : si $\mathrm{H} = \mathrm{F}_1 + \mathrm{F}_2 + \mathrm{F}_3$ est la somme vectorielle de trois sous-espaces vectoriels $\mathrm{F}_1, \mathrm{F}_2$ et F_3 deux à deux orthogonaux, alors

$$\mathrm{H} = \mathrm{F}_1 + (\mathrm{F}_2 + \mathrm{F}_3)$$

avec F_1 orthogonal à $\mathrm{F}_2 + \mathrm{F}_3$, donc F_1 est fermé d'après le lemme, ainsi que $\mathrm{F}_2 + \mathrm{F}_3$ et on a $\mathrm{H} = \mathrm{F}_1 \oplus (\mathrm{F}_2 + \mathrm{F}_3)$; de même les autres sous-espaces sont fermés, et de plus $\mathrm{H} = \mathrm{F}_1 \oplus \mathrm{F}_2 \oplus \mathrm{F}_3$. En effet, le sous-espace fermé $\mathrm{F}_2 + \mathrm{F}_3$ peut être considéré comme un espace de Hilbert auquel appliquer le lemme, donc $\mathrm{F}_2 + \mathrm{F}_3 = \mathrm{F}_2 \oplus \mathrm{F}_3$ et finalement $\mathrm{H} = \mathrm{F}_1 \oplus (\mathrm{F}_2 \oplus \mathrm{F}_3) = \mathrm{F}_1 \oplus \mathrm{F}_2 \oplus \mathrm{F}_3$.

Théorème 5. *Soit* F *un sous-espace vectoriel fermé d'un espace de Hilbert* H *; l'espace* H *admet la décomposition en somme directe orthogonale*

$$\mathrm{H} = \mathrm{F} \oplus \mathrm{F}^{\perp}.$$

Les deux projecteurs de la somme directe sont les projections orthogonales sur les sous-espaces F *et* $\mathrm{F}^{\perp}$*. On a de plus*

$$(\mathrm{F}^{\perp})^{\perp} = \mathrm{F}.$$

Preuve. — Pour tout $x \in \mathrm{H}$ on a

$$x = \mathrm{P}_{\mathrm{F}}(x) + (x - \mathrm{P}_{\mathrm{F}}(x))$$

avec $\mathrm{P}_{\mathrm{F}} x \in \mathrm{F}$ et $x - \mathrm{P}_{\mathrm{F}} x \perp \mathrm{F}$; ceci montre que $\mathrm{H} = \mathrm{F} + \mathrm{F}^{\perp}$. D'après le lemme précédent, l'espace H est somme directe des deux sous-espaces vectoriels orthogonaux F et $\mathrm{G} = \mathrm{F}^{\perp}$, les deux projections de la somme directe sont les projections orthogonales sur les facteurs, et de plus F est l'orthogonal de $\mathrm{G} = \mathrm{F}^{\perp}$.

On notera que la projection orthogonale sur $\mathrm{F}^{\perp}$ est égale à $\mathrm{Id} - \mathrm{P}_{\mathrm{F}}$.

Proposition 6 : critère de densité. *Pour qu'une partie* A *de* H *engendre un sous-espace vectoriel* Vect(A) *dense dans* H*, il faut et il suffit que* 0 *soit le seul vecteur orthogonal à l'ensemble* A*, c'est-à-dire que*

$$\mathrm{A}^{\perp} = \{0\}.$$

Preuve. — Si $\mathrm{F} = \overline{\mathrm{Vect}(\mathrm{A})}$ est différent de H, on voit, par exemple par le théorème 5, que $\mathrm{F}^{\perp}$ n'est pas réduit à 0, donc on peut trouver un vecteur v non nul orthogonal à F, en particulier orthogonal à tous les éléments de A. Inversement, si Vect(A) est dense dans H, on a en utilisant l'équation (3)

$$\mathrm{A}^{\perp} = \bigl(\overline{\mathrm{Vect}\,\mathrm{A}}\bigr)^{\perp} = \mathrm{H}^{\perp} = \{0\}.$$

Théorème de Weierstrass

Proposition 7. *Pour tout réels* $a < b$*, les monômes* $f_n(x) = x^n$*,* $n \in \mathbb{N}$*, engendrent un sous-espace vectoriel dense dans* $\mathrm{L}^2([a, b])$*.*

Preuve. — Dans le cas contraire on pourrait, par la proposition 6, trouver une fonction $g \in \mathrm{L}^2([a, b])$ *non nulle* qui serait orthogonale à tous les monômes,

$$\int_a^b g(x)\, x^n \,\mathrm{d}x = 0$$

pour tout $n \geq 0$. Considérons la fonction $\widetilde{g}$ sur $\mathbb{R}$ qui est égale à g dans $[a, b]$ et à 0 en dehors, et la transformée de Fourier de $\widetilde{g}$

$$\mathrm{G}(y) = \int_{\mathbb{R}} \widetilde{g}(x)\, \mathrm{e}^{-\mathrm{i}xy} \,\mathrm{d}x = \int_a^b g(x)\, \mathrm{e}^{-\mathrm{i}xy} \,\mathrm{d}x.$$

On va montrer que $\mathrm{G}(y) = 0$ pour tout y : fixons y, posons pour tout $\mathrm{N} \geq 0$ et $x \in [a, b]$

$$h_{\mathrm{N}}(x) = g(x) \sum_{k=0}^{\mathrm{N}} \frac{(-\mathrm{i}xy)^k}{k!}.$$

On reconnaît les sommes partielles de la série exponentielle, donc $h_{\mathrm{N}}(x)$ converge simplement vers $g(x)\,\mathrm{e}^{-\mathrm{i}xy}$. De plus la convergence est dominée :

$$|h_{\mathrm{N}}(x)| \le \sum_{k=0}^{\mathrm{N}} \frac{|xy|^k}{k!}\,|g(x)| \le \mathrm{e}^{|xy|}\,|g(x)|$$

et le majorant $x \to \mathrm{e}^{|xy|}\,|g(x)|$ est intégrable sur l'intervalle borné $[a,b]$ (lemme A8, annexe du chapitre 1). Il en résulte par Lebesgue dominé que

$$\int_a^b g(x)\,\mathrm{e}^{-\mathrm{i}xy}\,\mathrm{d}x = \lim_{\mathrm{N}} \int_a^b h_{\mathrm{N}}(x)\,\mathrm{d}x = \lim_{\mathrm{N}} \sum_{k=0}^{\mathrm{N}} \frac{(-\mathrm{i}y)^k}{k!} \int_a^b g(x)x^k\,\mathrm{d}x = 0.$$

On en déduit par l'injectivité de Fourier que $\widetilde{g} = 0$, donc $g = 0$, contrairement à notre hypothèse initiale.

Lemme. *Pour toute fonction continue* F *sur un intervalle fermé borné* $[a,b] \subset \mathbb{R}$ *et pour tout* $\varepsilon > 0$, *il existe une fonction* $g \in \mathrm{L}^2([a,b])$ *telle que la fonction* G *définie par*

$$\forall x \in [a,b], \quad \mathrm{G}(x) = \mathrm{F}(a) + \int_a^x g(t)\,\mathrm{d}t$$

vérifie $|\mathrm{F}(x) - \mathrm{G}(x)| < \varepsilon$ *pour tout* $x \in [a,b]$, *c'est-à-dire que* $\|\mathrm{F} - \mathrm{G}\|_u < \varepsilon$.

Preuve. — Soient F une fonction continue sur $[a,b]$ et $\varepsilon > 0$; par continuité uniforme, on peut trouver $\delta > 0$ tel que $|\mathrm{F}(y) - \mathrm{F}(x)| < \varepsilon/2$, dès que $|y - x| < \delta$. Considérons une subdivision $a = x_0 < x_1 < \ldots < x_{\mathrm{N}} = b$ de l'intervalle $[a,b]$, de pas $< \delta$, c'est-à-dire que $x_j - x_{j-1} < \delta$ pour tout $j = 1, \ldots, \mathrm{N}$. Introduisons une fonction en escalier g qui prenne sur chaque intervalle $[x_j, x_{j-1}[$, $j = 1, \ldots, \mathrm{N}$ la valeur constante

$$c_j = \frac{\mathrm{F}(x_j) - \mathrm{F}(x_{j-1})}{x_j - x_{j-1}}.$$

Posons ensuite $\mathrm{G}(x) = \mathrm{F}(a) + \int_a^x g(t)\,\mathrm{d}t$. On voit que

$$\mathrm{G}(x_j) - \mathrm{G}(x_{j-1}) = \int_{x_{j-1}}^{x_j} g(t)\,\mathrm{d}t = c_j(x_j - x_{j-1}) = \mathrm{F}(x_j) - \mathrm{F}(x_{j-1}),$$

ce qui entraîne, puisque $\mathrm{G}(a) = \mathrm{F}(a)$, que

$$\mathrm{G}(x_k) = \mathrm{G}(a) + \sum_{j=1}^{k} \big(\mathrm{G}(x_j) - \mathrm{G}(x_{j-1})\big) = \mathrm{F}(a) + \sum_{j=1}^{k} \big(\mathrm{F}(x_j) - \mathrm{F}(x_{j-1})\big) = \mathrm{F}(x_k),$$

pour tout $k = 1, \ldots, \mathrm{N}$. Si x est dans l'intervalle $[x_{j-1}, x_j[$, on a

$$|\mathrm{G}(x) - \mathrm{G}(x_{j-1})| = \Big|\int_{x_{j-1}}^{x} g(t)\,\mathrm{d}t\Big| = |c_j(x - x_{j-1})| \le |c_j(x_j - x_{j-1})| = |\mathrm{F}(x_j) - \mathrm{F}(x_{j-1})|$$

qui est $< \varepsilon/2$ par le choix de δ ; on a aussi $|\mathrm{F}(x) - \mathrm{F}(x_{j-1})| < \varepsilon/2$, $\mathrm{G}(x_{j-1}) = \mathrm{F}(x_{j-1})$, donc $|\mathrm{F}(x) - \mathrm{G}(x)| < \varepsilon$. Ceci est valable pour tout x de $[a,b]$, donc $\|\mathrm{F} - \mathrm{G}\|_u < \varepsilon$.

Théorème (Weierstrass). *Pour tout intervalle fermé borné* $[a, b] \subset \mathbb{R}$, *les monômes* $f_n(x) = x^n$, $n \in \mathbb{N}$, *engendrent un sous-espace vectoriel dense dans* $\mathrm{C}([a, b])$.

Preuve. — Soient F une fonction continue sur $[a, b]$ et $\varepsilon > 0$; par le lemme précédent, il existe une fonction $g \in \mathrm{L}^2([a, b])$ telle que $\|\mathrm{F} - \mathrm{G}\|_u < \varepsilon$, où on a posé comme avant $\mathrm{G}(x) = \mathrm{F}(a) + \int_a^x g(t)\,\mathrm{d}t$. D'après la proposition 7, on peut pour tout $\varepsilon > 0$ trouver une fonction polynomiale p sur $[a, b]$ telle que $\|p - g\|_2 < \varepsilon$; alors

$$\mathrm{P}(x) = \mathrm{F}(a) + \int_a^x p(t)\,\mathrm{d}t$$

est une fonction polynomiale, et par Cauchy-Schwarz on a pour tout x de $[a, b]$

$$|\mathrm{G}(x) - \mathrm{P}(x)| = \Bigl|\int_a^x (g(t) - p(t))\,\mathrm{d}t\Bigr| \le \int_a^b |g(t) - p(t)|\,\mathrm{d}t \le \sqrt{b-a}\,\|g - p\|_2,$$

ce qui montre la possibilité d'approcher la fonction G, donc aussi F, par une fonction polynomiale P, uniformément sur $[a, b]$.

Polynômes orthogonaux

La suite des monômes est linéairement indépendante dans $\mathrm{L}^2([a, b])$, et elle engendre un espace vectoriel dense dans $\mathrm{L}^2([a, b])$. Par Gram-Schmidt, on pourra fabriquer une base hilbertienne de $\mathrm{L}^2([a, b])$ formée de fonctions polynomiales. On peut en fait écrire des formules explicites, par exemple dans le cas de l'intervalle $[-1, 1]$; la fonction polynomiale $\mathrm{P}_n(x) = \mathrm{D}^n(1 - x^2)^n$ (dérivée $n$ième de $x \to (1 - x^2)^n$; on pose $\mathrm{P}_0 = \mathbf{1}$) est de degré n, et on montre par intégration par parties que les polynômes $(\mathrm{P}_n)_{n \ge 0}$ sont deux à deux orthogonaux. Il en résulte que P_n est proportionnel au $n$ième vecteur obtenu par la méthode de Gram-Schmidt appliquée à la suite des monômes.

Dual d'un espace de Hilbert H

On a vu que pour tout vecteur $v \in \mathrm{H}$, l'application $\ell_v : x \in \mathrm{H} \to \langle x, v\rangle$ est une forme linéaire continue sur H. On va voir une réciproque.

Théorème. *Toute forme linéaire continue* ℓ *sur un espace de Hilbert* H *est de la forme* $\ell = \ell_v$ *pour un certain vecteur* $v \in \mathrm{H}$, *c'est-à-dire que*

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

ce vecteur v *est unique.*

En d'autres termes, l'application $v \to \ell_v$ est une bijection de l'espace de Hilbert H sur son dual topologique H′. *Attention !* cette application n'est pas linéaire dans le cas complexe, car l'image du vecteur $\lambda\, v$ est la forme linéaire $\overline{\lambda}\,\ell_v$ (et pas $\lambda\,\ell_v$) ; l'application $v \to \ell_v$ est une bijection *antilinéaire* de l'espace de Hilbert H sur son dual topologique H′.

Preuve. — Montrons l'unicité : si v_1 et v_2 étaient deux vecteurs de H tels que $\ell_{v_1} = \ell_{v_2}$, on aurait $\langle x, v_1\rangle = \ell(x) = \langle x, v_2\rangle$ pour tout vecteur x, donc $\langle x, v_1 - v_2\rangle = 0$; en appliquant ceci à $x = v_1 - v_2$, on déduit que $v_1 - v_2 = 0$. Pour démontrer l'existence on prouvera un lemme en apparence plus général.

Lemme. *Soient* E *un* $\mathbb{K}$*-sous-espace vectoriel de l'espace de Hilbert* H, C *un nombre réel et* $\ell : \mathrm{E} \to \mathbb{K}$ *une forme linéaire sur* E *telle que* $|\ell(x)| \le \mathrm{C}\,\|x\|$ *pour tout* $x \in \mathrm{E}$ *; il existe un vecteur* $v \in \mathrm{H}$ *tel que*

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

Preuve du lemme. — Supposons d'abord que le noyau $\mathrm{F} = \ker \ell \subset \mathrm{E}$ soit dense dans H ; pour tout $x \in \mathrm{E}$, on peut trouver une suite $(x_n) \subset \mathrm{F}$ qui tend vers x ; comme $\ell(x_n) = 0$ pour tout n, on a

$$|\ell(x)| = |\ell(x) - \ell(x_n)| = |\ell(x - x_n)| \le \mathrm{C}\, \|x - x_n\| \to 0,$$

donc $\ell(x) = 0$. Ainsi dans ce cas, la forme linéaire ℓ est nulle sur E et il suffit de prendre $v = 0$ pour finir la preuve de ce cas.

Dans le cas contraire, on peut trouver (proposition 6) un vecteur $w \in \mathrm{H}$ non nul orthogonal à F ; on peut supposer w de norme 1, en le remplaçant par un multiple scalaire convenable. Puisque $w \notin \ker \ell = \mathrm{F}$, on a $\ell(w) \ne 0$. Posons $w_1 = \ell(w)^{-1}\, w$, de sorte que $\ell(w_1) = \ell(w)^{-1}\, \ell(w) = 1$. Pour tout vecteur $x \in \mathrm{E}$, remarquons que

$$x = \bigl(x - \ell(x)\, w_1\bigr) + \ell(x)\, w_1 = y + \ell(x)\, w_1$$

et $y = x - \ell(x)\, w_1 \in \ker \ell$; en effet, $\ell(y) = \ell(x) - \ell(x)\, \ell(w_1) = 0$; puisque $y \in \ker \ell = \mathrm{F}$, le vecteur y est orthogonal à w et

$$\langle x, w\rangle = \langle y + \ell(x)\, w_1, w\rangle = \langle y, w\rangle + \ell(x)\, \langle w_1, w\rangle = \frac{\ell(x)}{\ell(w)}\, \langle w, w\rangle = \frac{\ell(x)}{\ell(w)}.$$

On vient ainsi de montrer que

$$\ell(x) = \ell(w)\, \langle x, w\rangle = \langle x, \overline{\ell(w)}\, w\rangle$$

pour tout $x \in \mathrm{E}$. On voit donc que la forme linéaire ℓ est *représentée* par le produit scalaire avec le vecteur $v = \overline{\ell(w)}\, w$.

Exercice. Pour simplifier un tout petit peu on prendra $\mathbb{K} = \mathbb{R}$ dans cet exercice. On dit qu'une fonction réelle φ définie sur un intervalle I de la droite réelle est C-*lipschitzienne* si pour tous $x, y \in \mathrm{I}$ on a

$$|\varphi(x) - \varphi(y)| \le \mathrm{C}\, |x - y|.$$

Avec une fonction 1-lipschitzienne φ sur $[0, 1]$ on fabrique une forme linéaire ℓ sur l'espace vectoriel $\mathcal{E}$ des fonctions en escalier réelles de la façon suivante : si $h = \sum_{j=0}^{n-1} c_j \mathbf{1}_{[x_j, x_{j+1}[}$, où $0 = x_0 < x_1 < \ldots < x_n = 1$, on pose

$$\ell(h) = \sum_{j=0}^{n-1} c_j \bigl(\varphi(x_{j+1}) - \varphi(x_j)\bigr).$$

On a

$$(4) \qquad |\ell(h)| \le \sum_{j=0}^{n-1} |c_j| \bigl|\varphi(x_{j+1}) - \varphi(x_j)\bigr| \le \sum_{j=0}^{n-1} |c_j|\, (x_{j+1} - x_j) = \|h\|_1 \le \|h\|_2.$$

D'après le lemme précédent appliqué à $\mathrm{E} = \mathcal{E} \subset \mathrm{L}^2([0, 1])$, la forme linéaire ℓ peut être représentée par le produit scalaire avec une fonction réelle $f \in \mathrm{L}^2([0, 1])$: pour toute fonction en escalier h,

$$\ell(h) = \langle h, f\rangle = \int_0^1 h(t) f(t)\, \mathrm{d}t.$$

En particulier, lorsque $h = \mathbf{1}_{[0, x[}$, on trouve que

$$\varphi(x) - \varphi(0) = \ell(\mathbf{1}_{[0, x]}) = \int_0^x f(t)\, \mathrm{d}t$$

pour tout $x \in [0, 1]$; on peut ensuite(1) montrer que f est dans $\mathrm{L}^\infty([0, 1])$, et donc : *toute fonction lipschitzienne est « primitive » d'une fonction mesurable bornée* (et inversement, évidemment). Le théorème de représentation du dual de L^2 a permis de faire apparaître une fonction f qui n'était pas du tout visible au départ !

Notes du chapitre 2

([a]) Le terme *semi-produit scalaire* n'est pas classique. Mais la plupart des gens exige qu'un produit scalaire sur l'espace vectoriel E vérifie la propriété suivante : si $x \neq 0_{\mathrm{E}}$, alors $\langle x, x\rangle \neq 0$. Quand nous supposons seulement $\langle x, x\rangle \geq 0$, nous disons *semi-produit scalaire* comme on dit *semi-norme* ; tout ceci n'est utilisé qu'au début du chapitre.

([b]) Le choix du côté qui est linéaire dans la fonction de deux variables $(x, y) \rightarrow \langle x, y\rangle$ (pour nous, c'est le gauche) n'est pas universel. Dans les milieux proches de la Physique, on fait souvent le choix opposé à celui qui est fait dans ces notes de cours.

([c]) En géométrie plane élémentaire, cette relation signifie que dans un parallélogramme ABCD, la somme des carrés des longueurs des deux diagonales AC et BD est égale à la somme des carrés des longueurs des quatre côtés AB, BC, CD et DA : prendre le vecteur x de la relation (2) égal au vecteur $\overrightarrow{\mathrm{AB}}$ et le vecteur y égal au vecteur $\overrightarrow{\mathrm{BC}}$.

([d]) Il s'agit maintenant d'un *vrai* produit scalaire, c'est-à-dire que la relation $\langle \widetilde{f}, \widetilde{f}\rangle = 0$ implique que $\widetilde{f} = \widetilde{0}$, le zéro de l'espace vectoriel $\mathrm{L}^2(\Omega, \mathcal{A}, \mu)$. En effet, pour n'importe quel représentant f de $\widetilde{f}$, on a

$$\langle \widetilde{f}, \widetilde{f}\rangle = \int_\Omega f(\omega)\overline{f(\omega)}\,\mathrm{d}\mu(\omega) = \int_\Omega |f(\omega)|^2\,\mathrm{d}\mu(\omega)\,;$$

quand l'intégrale d'une fonction mesurable positive telle que $|f|^2$ est nulle, cela entraîne qu'elle est nulle presque-partout, donc la fonction $\mathbf{0}$ est dans la classe de f, donc $\widetilde{f} = \widetilde{0}$.

([e]) Si (ℓ_n) est une suite de Cauchy dans E', on sait que pour tout $\varepsilon > 0$ il existe N tel que $\|\ell_m - \ell_n\| < \varepsilon$ lorsque $m, n \geq \mathrm{N}$. Pour tout vecteur $x \in \mathrm{H}$, on a

$$|\ell_m(x) - \ell_n(x)| \leq \|\ell_m - \ell_n\|\,\|x\| \leq \varepsilon\,\|x\|$$

lorsque $m, n \geq \mathrm{N}$, ce qui permet de voir que la suite scalaire $(\ell_n(x))$ est de Cauchy, donc convergente puisque le corps des scalaires est complet. Si on pose $\ell(x) = \lim_n \ell_n(x)$, pour tout $x \in \mathrm{H}$, on montre facilement que ℓ est linéaire, et on obtient en passant à la limite quand $n \rightarrow +\infty$

$$|\ell_m(x) - \ell(x)| \leq \varepsilon\,\|x\|$$

pour tout x, qui montre que $\|\ell - \ell_m\| \leq \varepsilon$ quand $m \geq \mathrm{N}$, donc ℓ est la limite dans E' de la suite (ℓ_n). On a ainsi montré que E' est complet. On montre de la même façon que l'espace des applications linéaires continues d'un espace normé E dans un espace normé *complet* F est complet.

([f]) La plupart des étudiants a déjà appris qu'un sous-espace vectoriel F de dimension finie est toujours fermé dans un espace normé E : la preuve passe par l'équivalence des normes en dimension finie, qui permet de voir que F est complet pour la norme induite par E, donc fermé dans E. La démonstration par projection orthogonale est plutôt plus courte et plus naturelle.

(g) Puisque H est de dimension infinie, il existe au moins un vecteur $w \neq 0$ dans H, et puisque (x_k) est dense, il existe un entier k tel que $\|w - x_k\| < \frac{1}{2}\|w\|$, ce qui implique que x_k n'est pas nul. On définit k_0 comme le plus petit entier k tel que $x_k \neq 0$ et on pose $v_0 = x_{k_0}$.

Si $k_0, \ldots, k_{n-1}$ ont été définis et si on a posé $v_j = x_{k_j}$ pour $0 \le j < n$, on dit que H, qui est de dimension infinie, ne peut être égal à $\mathrm{F} = \mathrm{Vect}(v_0, \ldots, v_{n-1})$; le sous-espace F est de dimension finie, donc fermé dans H (corollaire 1). Si w est un vecteur de H qui n'est pas dans le fermé F, on peut trouver une boule ouverte B qui contient w et ne rencontre pas F ; comme (x_k) est dense, il existe des vecteurs x_k qui sont dans B, donc pas dans F. On peut alors définir k_n comme le plus petit entier k tel que $x_k \notin \mathrm{F}$ et on pose $v_n = x_{k_n}$.

Pour tous les entiers $j < k_n$ on a $x_j \in \mathrm{F} \subset \mathrm{Vect}(v_0, \ldots, v_n)$, et pour $j = k_n$ on a aussi $x_{k_n} = v_n \in \mathrm{Vect}(v_0, \ldots, v_n)$. Les vecteurs (v_n) sont indépendants puisque $v_0 \neq 0$ et $v_n \notin \mathrm{Vect}(v_0, \ldots, v_{n-1})$, pour tout $n \ge 1$.

(h) Pour chaque entier $n \ge 1$ soit φ_n la fonction qui vaut 0 en $x = 0$ et en $x = 2\pi$, qui vaut 1 sur l'intervalle $[1/n, 2\pi - 1/n]$, et qui est affine sur les deux intervalles $[0, 1/n]$ et $[2\pi - 1/n, 2\pi]$. Cette suite de fonctions positives tend simplement vers l'indicatrice de l'intervalle ouvert $]0, 2\pi[$, en restant majorée par 1. Si f est une fonction continue sur $[0, 2\pi]$, la suite $|f - f\varphi_n|^2$ tend presque partout vers 0 en restant majorée par la fonction $|f|^2$, qui est intégrable sur $[0, 2\pi]$; il en résulte par convergence dominée que f est limite dans L^2 de la suite des fonctions continues $f\varphi_n$, qui sont nulles en 0 et en 2π. Comme les fonctions continues sur $[0, 2\pi]$ sont denses dans $\mathrm{L}^2([0, 2\pi])$ (cours d'intégration), on en déduit le résultat voulu.

(i) Soit f une fonction continue sur $[0, 2\pi]$ telle que $f(2\pi) = f(0)$; pour tout réel x, il existe un entier $k \in \mathbb{Z}$ unique tel que $0 \le x - 2k < 2\pi$; pour prolonger f en une fonction $\widetilde{f}$ définie sur $\mathbb{R}$, on pose simplement $\widetilde{f}(x) = f(x - 2k\pi)$, et on vérifie que $\widetilde{f}$ est 2π-périodique, et continue sur $\mathbb{R}$.

(j) Considérons la fonction f sur un intervalle de deux périodes, par exemple $[0, 4\pi]$; sur ce compact, on sait que la fonction f est uniformément continue, donc il existe $\delta > 0$ tel que tel que $|f(x) - f(y)| < \varepsilon/2$ dès que $x, y \in [0, 4\pi]$ vérifient $|x - y| < \delta$; on peut choisir $\delta < 2\pi$; si x_1, y_1 sont deux points quelconques de $\mathbb{R}$ tels que $|x_1 - y_1| < \delta < 2\pi$, on peut toujours trouver un entier k tel que $x = x_1 - 2k\pi$ et $y = y_1 - 2k\pi$ soient tous les deux dans $[0, 4\pi]$: si $x_1 \le y_1$ par exemple, on choisit k tel que $x = x_1 - 2k\pi \in [0, 2\pi[$; alors si $y = y_1 - 2k\pi$, on a $0 \le y - x = y_1 - x_1 < 2\pi$ donc

$$0 \le x \le y \le x + 2\pi < 4\pi.$$

On aura $|x - y| = |x_1 - y_1| < \delta$, et $|f(x_1) - f(y_1)| = |f(x) - f(y)| < \varepsilon/2$ par périodicité.

(k) Il s'agit de voir qu'une fonction continue h nulle presque partout sur $[0, 2\pi]$ est en fait nulle *partout* : voir le théorème A5, annexe du chapitre 1.

(l) Considérons, pour un réel $c > 1$ quelconque, l'ensemble mesurable

$$\mathrm{A}_c = \{x \in [0, 1] : |f(x)| > c\} ;$$

on peut trouver une suite (h_n) de fonctions en escalier qui tende dans L^1 et presque partout vers la fonction $\mathbf{1}_{\mathrm{A}_c} \operatorname{sign} f$ (annexe du chapitre 1, théorème A6 et théorème A3) ;

comme la fonction limite est à valeurs dans $[-1, 1]$, on peut supposer que $|h_n| \le 1$. D'après l'inégalité (4), on sait que $|\ell(h_m) - \ell(h_n)| \le \|h_m - h_n\|_1$ tend vers 0. La suite numérique $(\ell(h_n))$ est de Cauchy, donc convergente vers une limite $\xi \in \mathbb{R}$. On a d'après (4)

$$|\ell(h_n)| \le \|h_n\|_1 \to \|\mathbf{1}_{\mathrm{A}_c} \operatorname{sign} f\|_1 = \int_0^1 \mathbf{1}_{\mathrm{A}_c}(t)\,\mathrm{d}t = \lambda(\mathrm{A}_c),$$

la mesure de Lebesgue de l'ensemble A_c ; on a donc $|\xi| \le \lambda(\mathrm{A}_c)$. Par ailleurs, la suite $(h_n f)$ tend presque partout vers la limite $(\mathbf{1}_{\mathrm{A}_c} \operatorname{sign} f)\, f = \mathbf{1}_{\mathrm{A}_c}|f|$, en étant majorée par la fonction intégrable fixe $|f| \in \mathrm{L}^2([0, 1])$, donc

$$c\,\lambda(\mathrm{A}_c) = c \int_0^1 \mathbf{1}_{\mathrm{A}_c}(t)\,\mathrm{d}t \le \int_0^1 \mathbf{1}_{\mathrm{A}_c}(t)|f(t)|\,\mathrm{d}t = \lim_n \int_0^1 h_n(t) f(t)\,\mathrm{d}t = \lim_n \ell(h_n) = \xi,$$

et il en résulte que $c\,\lambda(\mathrm{A}_c) \le \lambda(\mathrm{A}_c)$, mais $c > 1$ donc $\lambda(\mathrm{A}_c) = 0$; ainsi, l'ensemble A_c est négligeable pour tout $c > 1$, et la fonction $|f|$ est donc bornée par 1 presque partout.