Analyse hilbertienne et de Fourier, séance 7

Espaces de Hilbert, suite 3

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. On peut montrer que l'espace $\mathrm{L}^2([a, b])$ est séparable. On a 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.

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 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 de A_n par un élément de B_n sera possible avec une erreur $\le \mathrm{N} \, n^{-2} \le 1/n$, et 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 B_N, suivis de tous les éléments de $\mathrm{B}_{\mathrm{N}+1}$, $\mathrm{B}_{\mathrm{N}+2}$, 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.

Produit scalaire : forme linéaire continue et conséquence. Pour tout $v \in \mathrm{H}$ 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

$$|\ell_v(x_1) - \ell_v(x_2)| = |\langle x_1 - x_2, v \rangle| \le \|v\| \, \|x_1 - x_2\|.$$

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).

Projection orthogonale

Rappel : Lemme. *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 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 : projection orthogonale sur un sous-espace vectoriel fermé, séparable et de dimension infinie. D'après le corollaire de Gram-Schmidt, on peut trouver une base hilbertienne $(f_n)_{n \ge 0}$ pour F. Considérons la série de vecteurs (deux à deux orthogonaux)

$$(*) \qquad \sum_k \langle x, f_k \rangle \, f_k \,;$$

pour tout entier N, le vecteur $y_\mathrm{N} = \sum_{k=0}^{\mathrm{N}} \langle x, f_k \rangle \, f_k \in \mathrm{F}$ est la projection orthogonale de x sur $\mathrm{Vect}(f_0, \ldots, f_\mathrm{N})$ et on sait que

$$\sum_{k=0}^{\mathrm{N}} |\langle x, f_k \rangle|^2 \le \|x\|^2.$$

Il en résulte que la série numérique $\sum |\langle x, f_k \rangle|^2$ converge, ce qui entraîne aussi la convergence de la série $(*)$ de vecteurs orthogonaux ; la suite (y_N) des sommes partielles converge vers la somme de la série, disons y ; ce vecteur limite y est dans F, puisque F est fermé. Le produit scalaire avec f_m étant continu, on obtient

$$\langle y, f_m \rangle = \lim_\mathrm{N} \langle y_\mathrm{N}, f_m \rangle = \langle x, f_m \rangle.$$

Ceci montre que $x - y$ est orthogonal à tous les vecteurs f_m ; comme l'orthogonal $(x - y)^\perp$ est un sous-espace vectoriel, il contient d'abord toutes les combinaisons linéaires des f_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 $x - y \perp \mathrm{F}$.

Projection en général, par deux méthodes

Première méthode, par bricolage : on peut trouver une suite $(y_n) \subset \mathrm{F}$ telle que $\|x - y_n\|$ tende vers

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

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.

Deuxième méthode, classique : égalité du parallélogramme. On donne deux vecteurs $u, v \in \mathrm{H}$ et on ajoute la relation

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

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

$$\|u+v\|^2 + \|u-v\|^2 = 2\,\big(\|u\|^2 + \|v\|^2\big).$$

On montre que la suite (y_n) introduite dans la première méthode est de Cauchy. On pose $m = (y_k + y_\ell)/2 \in \mathrm{F}$ (convexité de F) et $u = x - y_k$, $v = x - y_\ell$. On obtient

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

Comme $m \in \mathrm{F}$, 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 F parce que F est fermé. On a de plus

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

et on a ainsi montré l'existence d'un point $y \in \mathrm{F}$ qui réalise la plus courte distance de x à un point de F. L'unicité de la projection sur le sous-espace fermé F se montre comme on l'a déjà dit.

Ça marche aussi bien avec un convexe fermé non vide C, au lieu du sous-espace vectoriel fermé F. Le seul point un peu différent est la façon de montrer l'unicité, qui résulte ici de la relation du parallélogramme : si y, y' sont deux points de C tels que

$$\|x-y\| = \|x-y'\| = d = \min\{\|x-z\| : z \in \mathrm{C}\},$$

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

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

Théorème. *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 $a_1x_1 + a_2x_2 - (a_1y_1 + a_2y_2)$ est orthogonal à F, ce qui montre que $a_1y_1 + a_2y_2$ est la projection orthogonale de $a_1x_1 + a_2x_2$ sur F, et la projection $\mathrm{P_F}$ est donc linéaire,

$$\mathrm{P_F}(a_1x_1 + a_2x_2) = a_1y_1 + a_2y_2 = a_1\,\mathrm{P_F}x_1 + a_2\,\mathrm{P_F}x_2.$$

Théorème. *Soient* H *un espace de Hilbert et* F *un sous-espace vectoriel fermé de* 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 $\mathrm{A}_1, \ldots, \mathrm{A}_\mathrm{N}$ soit une partition de Ω en ensembles de la tribu $\mathcal{A}$, tels que $\mathrm{P}(\mathrm{A}_j) > 0$ pour tout j ; 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

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

la fonction $\mathrm{P}_\mathrm{F} f$ 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 fermé 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 tout couple $(x,y) \in [0,1]^2$. La projection orthogonale sur F d'une fonction $f \in \mathrm{H}$ est donnée par

$$(\mathrm{P}_\mathrm{F} f)(x,y) = \mathrm{G}(x) = \int_0^1 f(x,y) \, dy.$$

Décomposition en sous-espaces vectoriels orthogonaux

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 $a^\perp$, où a varie dans 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 peut montrer que

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

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

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

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

Preuve. — Tout d'abord, $F \cap F^\perp = \{0\}$, car les vecteurs v de cette intersection sont orthogonaux à eux-mêmes, donc nuls. Ensuite,

$$x = P_F(x) + (x - P_F(x))$$

avec $P_F x \in F$ et $x - P_F x \perp F$ montre que $H = F + F^\perp$.

Supposons que $H = F \oplus G$, avec F et G orthogonaux, c'est-à-dire que tout vecteur f de F est orthogonal à tout vecteur $g \in G$. Si on décompose $x \in H$ sous la forme $x = f + g$, avec $f \in F$ et $g \in G$, on a $f \in F$ et $x - f = g$ est orthogonal à F : il en résulte que f est la projection orthogonale de x sur F ; les rôles de F et G étant symétriques, g est la projection orthogonale de x sur G.

En fait dans cette situation on a $G = F^\perp$; il est clair que $G \subset F^\perp$ par l'hypothèse d'orthogonalité de F et G. Si x orthogonal à F est décomposé sous la forme $x = f + g$, alors $0 = \langle x, f\rangle = \langle f, f\rangle + \langle f, g\rangle = \langle f, f\rangle$, donc $f = 0$ et $x = g$ appartient à G, donc $F^\perp \subset G$. Si on part de $H = F^\perp \oplus F$, on voit que $F = (F^\perp)^\perp$.

Si on applique les considérations précédentes à $G = F^\perp$, on voit que la projection orthogonale sur $F^\perp$ est égale à $\mathrm{Id} - P_F$.

Proposition : 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*

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

Preuve. — Si $F = \overline{\mathrm{Vect}(A)}$ est différent de H, on voit que $F^\perp$ n'est pas réduit à 0, donc on peut trouver un vecteur v non nul orthogonal à F, donc en particulier orthogonal à tous les éléments de A.

Inversement, si Vect(A) est dense dans H, on a $A^\perp = \bigl(\overline{\mathrm{Vect}\,A}\bigr)^\perp = H^\perp = \{0\}$.

Théorème de Weierstrass

Proposition. *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* $L^2([a, b])$.

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

$$\int_a^b g(x)\, x^n \, dx = 0$$

pour tout $n \ge 0$. 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 $\widetilde{g}$

$$G(y) = \int_{\mathbb{R}} \widetilde{g}(x)\, e^{-ixy} \, dy = \int_a^b g(x)\, e^{-ixy} \, dy.$$

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

$$h_N(x) = \sum_{k=0}^{N} \frac{(-ixy)^k}{k!}\, g(x).$$

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 $[a, b]$. Il en résulte par Lebesgue dominé que

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

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

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

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

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

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

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

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

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 $[a, b]$. D'après la proposition précédente, on peut, pour tout $\varepsilon > 0$, trouver une fonction polynomiale p sur $[a, b]$ telle que $\|p - g\|_2 < \varepsilon$; alors

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

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

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

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

Polynômes orthogonaux : par Gram-Schmidt, on pourra fabriquer une base hilbertienne de $\mathrm{L}^2([a,b])$ formée de polynômes.

Dual d'un espace de Hilbert H

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

Théorème. *Pour toute forme linéaire continue ℓ sur un espace de Hilbert* H*, il existe un vecteur $v \in \mathrm{H}$ tel que $\ell = \ell_v$,*

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