Cours “Théorie spectrale”, automne 2003

1. Opérateurs compacts sur les espaces de Hilbert

Une forme du théorème de Hahn-Banach

Soit X *un espace normé, réel ou complexe ; pour tout vecteur* $x \in \mathrm{X}$*, il existe une forme linéaire continue* ξ *sur* X *telle que*

$$\xi(x) = \|x\| \, ; \quad \|\xi\| \le 1.$$

En particulier, pour tout vecteur $x \neq 0$ dans X, il existe une forme linéaire continue ξ sur X telle que $\xi(x) \neq 0$.

Opérateurs linéaires

Si X est un espace normé, on notera $\mathrm{B_X}$ sa boule unité fermée,

$$\mathrm{B_X} = \{x \in \mathrm{X} : \|x\| \le 1\}.$$

Si Y est un autre espace normé, on note $\mathcal{L}(\mathrm{X}, \mathrm{Y})$ l’espace vectoriel des applications linéaires continues de X dans Y (on dit aussi *opérateurs linéaires* bornés). Si $\mathrm{T} \in \mathcal{L}(\mathrm{X}, \mathrm{Y})$, on définit sa norme (norme d’opérateur) par

$$\|\mathrm{T}\|_{\mathcal{L}(\mathrm{X},\mathrm{Y})} = \sup\{\|\mathrm{T}(x)\| : x \in \mathrm{B_X}\} \, ;$$

pour alléger la notation, on écrira souvent $\mathrm{T}x$ pour l’image $\mathrm{T}(x)$ d’un vecteur $x \in \mathrm{X}$ par $\mathrm{T} \in \mathcal{L}(\mathrm{X}, \mathrm{Y})$. Lorsque X et Y sont de plus complets (espaces de Banach), l’espace $\mathcal{L}(\mathrm{X}, \mathrm{Y})$ est complet pour la norme précédente. On notera $\mathcal{L}(\mathrm{X})$ l’espace des endomorphismes continus de X.

Séries de vecteurs

Dans un espace normé X, une suite (x_n) de vecteurs converge vers $x \in \mathrm{X}$ si $\|x_n - x\|$ tend vers 0 ; une série de vecteurs $\sum u_n$ de X est convergente (dans X) s’il existe un vecteur $s \in \mathrm{X}$ tel que la suite $x_n = u_0 + \cdots + u_n$ converge vers s. Dans ce cas on note

$$s = \sum_{n=0}^{+\infty} u_n \in \mathrm{X}.$$

Si T est une application linéaire continue, elle transformera les sommes partielles en sommes partielles par linéarité, et la convergence par continuité. On aura

$$\mathrm{T}s = \sum_{n=0}^{+\infty} \mathrm{T}u_n.$$

Ce principe simple est utilisé souvent.

Opérateurs compacts

Définition. Soient X et Y deux espaces de Banach ; on dit que $T \in \mathcal{L}(X, Y)$ est *compact* si l'adhérence $\overline{T(B_X)}$ dans Y de l'image par T de la boule unité de X est compacte dans l'espace Y.

Si X_0 est un sous-espace vectoriel fermé de X et si $T \in \mathcal{L}(X, Y)$ est compact, sa restriction $T_{|X_0}$ est un opérateur compact de X_0 dans Y, puisque l'image de la boule unité de X_0 est contenue dans l'image de la boule de l'espace entier. Si T est un endomorphisme compact de X et si $Y \subset X$ est stable par T, la restriction $T_{|Y}$, vue comme application de Y dans Y, est un endomorphisme compact de Y.

Proposition 1.1. *Pour qu'une partie fermée* F *d'un espace métrique complet* (X, d) *soit compacte, (il faut et) il suffit que pour tout* $\varepsilon > 0$*, on puisse recouvrir* F *par un nombre fini de boules de rayon* $< \varepsilon$.

Preuve. La condition est clairement nécessaire. Pour établir la réciproque, on va montrer que toute suite $(x_n) \subset F$ possède une valeur d'adhérence dans l'ensemble F. Si F_1 est un sous-ensemble fermé de F tel que l'ensemble

$$I_{F_1} = \{n \in \mathbb{N} : x_n \in F_1\}$$

soit infini, on va trouver un fermé $F_2 \subset F_1$, tel que I_{F_2} soit encore infini et que de plus le diamètre de F_2 vérifie $\operatorname{diam}(F_2) \le \varepsilon$.

Par hypothèse on peut couvrir F par un nombre *fini* de boules de rayon $\le \varepsilon/2$, qu'on peut supposer fermées. On peut donc trouver une boule fermée B de rayon $\varepsilon/2$ telle que l'ensemble des indices $n \in I_{F_1}$ pour lesquels $x_n \in B$ soit infini, ce qui signifie que $F_2 = F_1 \cap B$ vérifie $\operatorname{diam}(F_2) \le \varepsilon$ et que I_{F_2} est encore infini.

On peut donc produire ainsi une suite décroissante de fermés de diamètres tendant vers 0, dont chacun contient une infinité de termes de la suite. Cette suite de fermés définit un point unique y dans l'espace complet X, et y est une valeur d'adhérence de la suite (x_n) donnée.

Théorème 1.2. *Les opérateurs compacts de* X *dans* Y *forment un sous-espace vectoriel fermé de* $\mathcal{L}(X, Y)$.

Preuve. Supposons que $T \in \mathcal{L}(X, Y)$ soit adhérent en norme d'opérateur à l'ensemble des opérateurs compacts. Soit $\varepsilon > 0$; on peut trouver S compact tel que $\|T - S\| < \varepsilon/2$, et un nombre fini de boules de rayon $< \varepsilon/2$ qui couvrent le compact $\overline{S(B_X)}$. Les boules de même centre, mais de rayon ε, couvrent $\overline{T(B_X)}$.

La compacité de $T_1 + T_2$ se montre par le même principe.

Remarque. Les opérateurs de rang fini sont évidemment compacts. Toute limite en norme d'opérateur d'une suite d'opérateurs de rang fini est un opérateur compact.

Proposition. *Si* T *est un opérateur compact défini sur un espace de Hilbert* H*, l'image de la boule unité fermée* B_H *est un ensemble fermé, donc compact.*

Preuve. Supposons que T soit une application linéaire continue de l'espace de Hilbert H dans un espace de Banach Y. Soit y un point de l'adhérence de $T(B_H)$ dans Y, soit $(x_n) \subset B_H$ une suite telle que $y = \lim_n Tx_n$; puisque la suite (x_n) est bornée dans l'espace de Hilbert H, elle possède des sous-suites faiblement convergentes. En passant à une sous-suite, on peut donc supposer que (x_n) tend faiblement vers $x \in H$ et que

$y = \lim_n \mathrm{T}x_n$. On en déduit que $y = \mathrm{T}x$: en effet, considérons une forme linéaire continue ξ quelconque sur Y ; la composition $\xi \circ \mathrm{T}$ est une forme linéaire continue sur H, donc

$$\xi(\mathrm{T}x) = (\xi \circ \mathrm{T})x = \lim_n (\xi \circ \mathrm{T})x_n = \lim_n \xi(\mathrm{T}x_n)$$

par convergence faible dans H, et on a aussi $\xi(y) = \lim_n \xi(\mathrm{T}x_n) = \xi(\mathrm{T}x)$; il en résulte que $y = \mathrm{T}x$ par Hahn-Banach. Enfin, on a $\|x\| \le 1$ car

$$\|x\|^2 = \langle x, x\rangle = \lim_n \langle x_n, x\rangle \le \|x\|.$$

On a bien montré que $y = \mathrm{T}x$ est dans $\mathrm{T}(\mathrm{B_H})$.

Opérateurs hermitiens

Un endomorphisme continu T d'un espace de Hilbert réel ou complexe H est dit *hermitien* si

$$\langle \mathrm{T}x, y\rangle = \langle x, \mathrm{T}y\rangle$$

pour tous $x, y \in \mathrm{H}$.

Si un sous-espace fermé $\mathrm{Y} \subset \mathrm{H}$ est stable par T hermitien, la restriction $\mathrm{T}_{|\mathrm{Y}} \in \mathcal{L}(\mathrm{Y})$ est clairement hermitienne. De plus, l'orthogonal $\mathrm{Y}^\perp$ de Y est lui aussi stable par T. En effet, si $z \perp \mathrm{Y}$ on aura pour tout $y \in \mathrm{Y}$

$$\langle \mathrm{T}z, y\rangle = \langle z, \mathrm{T}y\rangle = 0$$

puisque $\mathrm{T}y \in \mathrm{Y}$; ceci montre que $\mathrm{T}z \perp \mathrm{Y}$, donc $\mathrm{Y}^\perp$ est stable par T.

Les valeurs propres (éventuelles) d'un opérateur hermitien sont réelles.

Lemme. *Si* T *est hermitien compact sur un espace de Hilbert* $\mathrm{H} \ne \{0\}$, *il admet une valeur propre égale à* $\|\mathrm{T}\|$ *ou à* $-\|\mathrm{T}\|$.

Preuve. Pour la démonstration de ce lemme, on peut se limiter au cas des scalaires réels : si l'espace de Hilbert est complexe, on peut le considérer comme espace réel, muni du produit scalaire réel $x \cdot y = \mathrm{Re}\,\langle x, y\rangle$; en effet, ce produit scalaire réel est bien $\mathbb{R}$-bilinéaire, et on a $\|x\|^2 = x \cdot x$; l'opérateur T reste hermitien dans ce contexte réel, puisque $\mathrm{T}x \cdot y = \mathrm{Re}\,\langle \mathrm{T}x, y\rangle = \mathrm{Re}\,\langle x, \mathrm{T}y\rangle = x \cdot \mathrm{T}y$ pour tous vecteurs $x, y \in \mathrm{H}$.

Si $\mathrm{T} = 0$, tout est évident (prendre un vecteur $x \ne 0$ quelconque dans H) ; on supposera donc $\mathrm{T} \ne 0$, et on posera $\lambda = \|\mathrm{T}\| > 0$. L'image $\mathrm{K} = \mathrm{T}(\mathrm{B_H})$ est compacte, on peut donc maximiser la fonction continue $y \to \|y\|^2$ sur ce compact K ; le maximum est égal à $\|\mathrm{T}\|^2 = \lambda^2 > 0$, et il est atteint en un vecteur $y_0 \in \mathrm{K}$; on a $y_0 = \mathrm{T}x_0$ avec $\|x_0\| = 1$ (sinon on pourrait trouver plus, en poussant x_0 vers la sphère unité).

On pose $\mathrm{S} = \mathrm{T}^2$; on note que $\mathrm{T}^2 x \cdot x = \mathrm{T}x \cdot \mathrm{T}x = \|\mathrm{T}x\|^2$ pour tout x ; si $v \cdot x_0 = 0$, avec v vecteur de norme un, et si on pose $x_\theta = \cos(\theta)x_0 + \sin(\theta)v$, ce vecteur x_θ est de norme un, et

$$\mathrm{S}(x_\theta) \cdot x_\theta = \cos^2(\theta)\,\mathrm{S}(x_0) \cdot x_0 + 2\sin(\theta)\cos(\theta)\,\mathrm{S}(x_0) \cdot v + \sin^2(\theta)\,\mathrm{S}(v) \cdot v =$$
$$= \|\mathrm{T}x_\theta\|^2 \le \|\mathrm{T}x_0\|^2 = \lambda^2 = \mathrm{S}(x_0) \cdot x_0$$

est maximum pour $\theta = 0$, donc $\mathrm{S}(x_0) \cdot v = 0$; ainsi $\mathrm{S}x_0$ est orthogonal à tout vecteur v orthogonal à x_0 ; on en déduit que $\mathrm{S}x_0 \in \mathbb{R}x_0$, donc $\mathrm{S}x_0 = \lambda^2 x_0$. Pour terminer, on note que

$$0 = (\mathrm{T}^2 - \lambda^2\,\mathrm{Id})\,x_0 = (\mathrm{T} - \lambda\,\mathrm{Id})(\mathrm{T} + \lambda\,\mathrm{Id})\,x_0.$$

Si $x_1 = (\mathrm{T} + \lambda\,\mathrm{Id})\,x_0$ n'est pas nul, il est vecteur propre de T pour la valeur propre λ ; si $x_1 = 0$, alors x_0 est vecteur propre de T pour la valeur propre $-\lambda$.

Diagonalisation de T *compact hermitien*

Théorème. *Soient* H *un espace de Hilbert, réel ou complexe, et* $\mathrm{T} \in \mathcal{L}(\mathrm{H})$ *un opérateur hermitien compact ; il existe une base hilbertienne de* H *formée de vecteurs propres de l'opérateur* T *; pour tout* $\varepsilon > 0$, *il n'existe qu'un nombre fini de vecteurs de cette base pour lesquels la valeur propre est de valeur absolue* $\geq \varepsilon$.

Preuve. On supposera $\dim \mathrm{H} = +\infty$, sinon le résultat est déjà connu (en fait cette hypothèse est un peu superflue : elle sert uniquement à garantir que le procédé inductif qui suit se poursuit jusqu'à l'infini, ce qui évite d'avoir à distinguer deux cas).

On construit par récurrence une suite croissante (E_n) de sous-espaces de dimension finie et une suite décroissante (X_n) de sous-espaces de dimension infinie, stables par T. On amorce l'induction avec $\mathrm{E}_0 = \{0\}$.

Si $n \geq 0$ et si $\mathrm{E}_n = \mathrm{Vect}(e_1, \ldots, e_n)$ est donné, engendré par des vecteurs propres de T, on raisonne ainsi : on a $\mathrm{E}_n \neq \mathrm{H}$ puisque $\dim \mathrm{H} = +\infty$, donc $\mathrm{X}_{n+1} = \mathrm{E}_n^\perp$ n'est pas réduit à $\{0\}$. La restriction $\mathrm{T}_{n+1} = \mathrm{T}_{|\mathrm{X}_{n+1}} \in \mathcal{L}(\mathrm{X}_{n+1})$ est hermitienne et compacte, donc elle admet une valeur propre réelle λ_{n+1} telle que $|\lambda_{n+1}| = \|\mathrm{T}_{n+1}\|$, et il existe un vecteur $e_{n+1} \in \mathrm{X}_{n+1}$, qu'on peut choisir de norme 1, tel que $\mathrm{T}e_{n+1} = \mathrm{T}_{n+1}e_{n+1} = \lambda_{n+1}e_{n+1}$. On pose $\mathrm{E}_{n+1} = \mathrm{Vect}(e_1, \ldots, e_n, e_{n+1})$ et la boucle est bouclée.

Cette construction par récurrence produit une suite $(e_n)_{n\geq 1}$ de vecteurs propres de l'opérateur T ; de plus ces vecteurs sont de norme 1, deux à deux orthogonaux puisque e_{n+1} est orthogonal à $e_1, \ldots, e_n$ par construction.

On a $\mathrm{T}e_n = \lambda_n e_n$ avec $|\lambda_n| = \|\mathrm{T}_n\|$ pour tout $n \geq 1$. La suite $(\|\mathrm{T}_n\|)$ est décroissante (car la suite (X_n) est décroissante), donc converge vers un nombre $d \geq 0$. Par la compacité on doit avoir $d = 0$, c'est-à-dire que $\|\mathrm{T}_n\| \to 0$, sinon la suite $(\mathrm{T}e_n)$ du compact $\mathrm{T}(\mathrm{B_H})$ ne pourrait avoir aucune sous-suite convergente, puisque pour tous $m \neq n$ on a

$$\|\mathrm{T}e_m - \mathrm{T}e_n\|^2 = |\lambda_m|^2 + |\lambda_n|^2 \geq 2\,d^2.$$

L'espace fermé E_∞ engendré par cette suite de vecteurs propres (e_n) est stable par T, ainsi que son orthogonal X_∞ ; on voit que la restriction T_∞ de T à X_∞ est nulle, car $\|\mathrm{T}\|_\infty \leq \|\mathrm{T}_n\|$ pour tout n. On a diagonalisé T : il suffit de compléter la suite orthonormée (e_n) par une base hilbertienne $(f_j)_{j\in \mathrm{J}}$ de X_∞, où l'ensemble J peut être vide, dans le cas où $\mathrm{X}_\infty = \{0\}$; on aura $\mathrm{T}f_j = 0$ pour tout indice $j \in \mathrm{J}$.

Désignons par $(h_i)_{i\in \mathrm{I}}$ une base hilbertienne de H formée de vecteurs propres de T, $\mathrm{T}h_i = \mu_i h_i$, soit $\varepsilon > 0$ et $\mathrm{J}_\varepsilon = \{i \in \mathrm{I} : |\mu_i| \geq \varepsilon\}$; comme on l'a déjà vu plus haut, on a $\|\mathrm{T}h_i - \mathrm{T}h_j\|^2 \geq 2\varepsilon^2$ pour tout couple (i, j) d'éléments distincts dans J_ε, ce qui implique que l'ensemble J_ε est fini.

Exercice. Si P désigne l'opérateur de primitive sur $\mathrm{L}_2(0, 1)$, diagonaliser $\mathrm{P}^*\mathrm{P}$.

Diagonalisation de T *compact, normal complexe.*

Si $\mathrm{H}_1, \mathrm{H}_2$ sont deux espaces de Hilbert et si $\mathrm{T} \in \mathcal{L}(\mathrm{H}_1, \mathrm{H}_2)$ on définit *l'adjoint* hilbertien $\mathrm{T}^* \in \mathcal{L}(\mathrm{H}_2, \mathrm{H}_1)$ de l'opérateur T par la propriété

$$\forall x \in \mathrm{H}_1,\ \forall y \in \mathrm{H}_2, \quad \langle \mathrm{T}^* y, x\rangle = \langle y, \mathrm{T}x\rangle.$$

Définition. On dit que $\mathrm{T} \in \mathcal{L}(\mathrm{H})$ est un opérateur *normal* si T commute avec son adjoint, $\mathrm{T}^*\mathrm{T} = \mathrm{T}\,\mathrm{T}^*$.

Remarque. Si T est normal on a $\ker \mathrm{T}^* = \ker \mathrm{T}$. En effet, on a pour tout $x \in \mathrm{H}$

$$\|\mathrm{T}x\|^2 = \langle \mathrm{T}^*\mathrm{T}x, x\rangle = \langle \mathrm{T}\mathrm{T}^*x, x\rangle = \|\mathrm{T}^*x\|^2.$$

Si $\mathrm{T}x = \lambda x$, on a $\mathrm{T}^*x = \overline{\lambda}x$. En effet, en appliquant ce qui précède à l'opérateur normal $\mathrm{S} = \mathrm{T} - \lambda\,\mathrm{Id}$, et en notant que l'adjoint de $\lambda\,\mathrm{Id}$ est $\overline{\lambda}\,\mathrm{Id}$,

$$\|\mathrm{T}x - \lambda x\|^2 = \|\mathrm{T}^*x - \overline{\lambda}x\|^2.$$

Théorème. *Soient* H *un espace de Hilbert* complexe, *et* $\mathrm{T} \in \mathcal{L}(\mathrm{H})$ *un opérateur normal compact ; il existe une base hilbertienne de* H *formée de vecteurs propres de* T.

Preuve. On diagonalise $\mathrm{T}_1 = \mathrm{T}^*\mathrm{T}$, qui est hermitien compact et commute avec T ; on remarque ensuite que les espaces propres $\mathrm{E}_\lambda = \ker(\mathrm{T}_1 - \lambda\,\mathrm{Id})$ de T_1 sont invariants par T, et de dimension finie lorsque $\lambda \neq 0$, ce qui permet de les redécomposer (ici on a besoin de scalaires complexes pour résoudre cette question de dimension finie). De plus, $\ker \mathrm{T} = \ker \mathrm{T}^*$.

Compléments sur les espaces mesurés produits

On suppose que $(\mathrm{S}, \mathcal{A}, \mu)$ et $(\mathrm{T}, \mathcal{B}, \nu)$ sont deux espaces mesurés σ-finis. On définit alors la tribu $\mathcal{A} \otimes \mathcal{B}$ de parties de $\mathrm{S} \times \mathrm{T}$ engendrée par les pavés mesurables $\mathrm{A} \times \mathrm{B}$, $\mathrm{A} \in \mathcal{A}$ et $\mathrm{B} \in \mathcal{B}$. On sait définir une mesure $\mu \otimes \nu$ sur cette tribu.

Si f est une fonction sur S et g une fonction sur T, on définit une fonction $f \otimes g$ sur le produit $\mathrm{S} \times \mathrm{T}$ en posant

$$\forall (s,t) \in \mathrm{S} \times \mathrm{T}, \quad (f \otimes g)(s,t) = f(s)g(t).$$

On notera 1_A la fonction indicatrice d'un sous-ensemble A, égale à 1 sur l'ensemble A et à 0 en dehors de A.

Lemme. *Toute fonction* 1_C, $\mathrm{C} \in \mathcal{A} \otimes \mathcal{B}$, *peut être approchée dans* $\mathrm{L}_2(\mathrm{S} \times \mathrm{T}, \mu \otimes \nu)$ *par des combinaisons linéaires de fonctions* $1_\mathrm{A} \otimes 1_\mathrm{B}$, *avec des ensembles* $\mathrm{A} \in \mathcal{A}$, $\mathrm{B} \in \mathcal{B}$ *de mesure finie.*

Preuve. Comme les mesures sont σ-finies il suffit de traiter le cas où μ et ν sont finies. Désignons par $\mathcal{F}$ l'espace vectoriel des combinaisons linéaires de fonctions $1_\mathrm{A} \otimes 1_\mathrm{B}$, $\mathrm{A} \in \mathcal{A}$, $\mathrm{B} \in \mathcal{B}$. Désignons par $\mathcal{C}$ la classe des $\mathrm{C} \in \mathcal{A} \otimes \mathcal{B}$ tels que la fonction indicatrice 1_C soit limite dans $\mathrm{L}_2(\mu \otimes \nu)$ d'une suite $(\varphi_n) \subset \mathcal{F}$, telle que $0 \leq \varphi_n \leq 1$. Cette classe $\mathcal{C}$ est une tribu contenant tous les pavés $\mathrm{A} \times \mathrm{B}$, donc elle est égale à la tribu $\mathcal{A} \otimes \mathcal{B}$.

Supposons pour fixer les idées que les deux espaces $\mathrm{L}_2(\mu)$ et $\mathrm{L}_2(\nu)$ soient de dimension infinie et séparables (c'est de très loin le cas le plus important). On peut alors trouver deux bases hilbertiennes infinies dénombrables, que l'on peut indexer par $\mathbb{N}$.

Proposition 1.3. *Si* $(\varphi_n)_{n\in\mathbb{N}}$ *est une base hilbertienne de* $\mathrm{L}_2(\mathrm{S},\mu)$ *et* $(\psi_n)_{n\in\mathbb{N}}$ *une base hilbertienne de* $\mathrm{L}_2(\mathrm{T},\nu)$*, la famille* $(\varphi_m\otimes\psi_n)_{(m,n)\in\mathbb{N}^2}$ *est une base hilbertienne de* $\mathrm{L}_2(\mathrm{S}\times\mathrm{T},\mu\otimes\nu)$*.*

Preuve. Soient $f\in\mathrm{L}_2(\mathrm{S},\mu)$ et $g\in\mathrm{L}_2(\mathrm{T},\nu)$; puisqu'on a des bases hilbertiennes, on peut trouver une combinaison linéaire f_1 des vecteurs de la base (φ_m) et une combinaison linéaire g_1 des vecteurs de la base (ψ_n) telles que

$$\|f-f_1\|<\varepsilon,\quad \|g-g_1\|<\varepsilon.$$

On écrit ensuite

$$f\otimes g-f_1\otimes g_1=(f-f_1)\otimes g+f_1\otimes(g-g_1),$$

donc

$$\begin{aligned}\|f\otimes g-f_1\otimes g_1\|&\le\|(f-f_1)\otimes g\|+\|f_1\otimes(g-g_1)\|\\&=\|f-f_1\|\,\|g\|+\|f_1\|\,\|g-g_1\|\le\varepsilon\|g\|+(\|f\|+\varepsilon)\,\varepsilon,\end{aligned}$$

qui peut être rendu arbitrairement petit. La fonction $f_1\otimes g_1$ est une combinaison linéaire des vecteurs de la base $(\varphi_m\otimes\psi_n)$; ces combinaisons linéaires peuvent donc approcher autant que souhaité toute fonction de la forme $1_\mathrm{A}\otimes 1_\mathrm{B}$, donc aussi les combinaisons linéaires $\sum c_j 1_{\mathrm{A}_j}\otimes 1_{\mathrm{B}_j}$, donc toutes les fonctions 1_C, $\mathrm{C}\in\mathcal{A}\otimes\mathcal{B}$, d'après le lemme précédent. Il en résulte que la famille $(\varphi_m\otimes\psi_n)$ est totale dans $\mathrm{L}_2(\mathrm{S}\times\mathrm{T},\mu\otimes\nu)$. Il est par ailleurs clair que la famille est orthonormée, car

$$\langle f_1\otimes g_1,f_2\otimes g_2\rangle=\langle f_1,f_2\rangle\,\langle g_1,g_2\rangle$$

pour tous les couples (f_1,g_1), (f_2,g_2).

Opérateurs de Hilbert-Schmidt

On dira qu'un sous-ensemble B d'un espace de Hilbert H forme une base hilbertienne de cet espace H si les éléments de B sont de norme 1, deux à deux orthogonaux et si l'espace vectoriel engendré par B est dense dans H. Ce langage permet d'indexer les quantités liées au vecteurs de la base par les vecteurs $b\in\mathrm{B}$ eux-mêmes.

Lemme. *Soient* H_1 *et* H_2 *deux espaces de Hilbert,* B_1 *une base hilbertienne de* H_1 *et* B_2 *une base hilbertienne de* H_2 *; pour tout* $\mathrm{T}\in\mathcal{L}(\mathrm{H}_1,\mathrm{H}_2)$ *on a :*

$$\sum_{b\in\mathrm{B}_1}\|\mathrm{T}(b)\|^2=\sum_{b'\in\mathrm{B}_2}\|\mathrm{T}^*(b')\|^2$$

(valeur finie ≥ 0 *ou bien* $+\infty$*). Cette quantité ne dépend pas des bases* B_1 *et* B_2 *choisies.*

Preuve. Pour $x\in\mathrm{H}_1$ et $y\in\mathrm{H}_2$ on a

$$\|x\|^2=\sum_{b\in\mathrm{B}_1}|\langle x,b\rangle|^2,\quad \|y\|^2=\sum_{b'\in\mathrm{B}_2}|\langle b',y\rangle|^2,$$

et on en déduit que

$$\sum_{b\in\mathrm{B}_1}\|\mathrm{T}(b)\|^2=\sum_{b\in\mathrm{B}_1,b'\in\mathrm{B}_2}|\langle b',\mathrm{T}(b)\rangle|^2=\sum_{b'\in\mathrm{B}_2}\|\mathrm{T}^*(b')\|^2.$$

Il est clair que $\sum_{b\in\mathrm{B}_1}\|\mathrm{T}(b)\|^2$ ne dépend pas de B_2 et que $\sum_{b'\in\mathrm{B}_2}\|\mathrm{T}^*(b')\|^2$ ne dépend pas de B_1, d'où la deuxième assertion.

Pour $\mathrm{T} \in \mathcal{L}(\mathrm{H}_1, \mathrm{H}_2)$ on pose

$$\|\mathrm{T}\|_{\mathrm{HS}} = \Big(\sum_{b \in \mathrm{B}} \|\mathrm{T}(b)\|^2\Big)^{1/2}$$

où B est une base hilbertienne quelconque de H_1. Un opérateur $\mathrm{T} \in \mathcal{L}(\mathrm{H}_1, \mathrm{H}_2)$ tel que $\|\mathrm{T}\|_{\mathrm{HS}} < +\infty$ est dit de *Hilbert-Schmidt.*

Remarques.

1. D'après ce qui précède, on voit que T est Hilbert-Schmidt si et seulement si l'adjoint T^* est Hilbert-Schmidt, et $\|\mathrm{T}^*\|_{\mathrm{HS}} = \|\mathrm{T}\|_{\mathrm{HS}}$.

2. Si x est un vecteur de norme 1 quelconque dans H_1, on peut trouver une base hilbertienne B de H_1 qui contient x. Il en résulte que $\|\mathrm{T}x\| \leq \|\mathrm{T}\|_{\mathrm{HS}}$, donc

$$\|\mathrm{T}\| \leq \|\mathrm{T}\|_{\mathrm{HS}}.$$

3. Si S ou T est un opérateur de Hilbert-Schmidt, il en va de même pour TS lorsque la composition est définie. De plus

$$\|\mathrm{TS}\|_{\mathrm{HS}} \leq \|\mathrm{T}\| \, \|\mathrm{S}\|_{\mathrm{HS}}, \quad \|\mathrm{TS}\|_{\mathrm{HS}} \leq \|\mathrm{T}\|_{\mathrm{HS}} \|\mathrm{S}\|.$$

Soit B une base hilbertienne de H_1 ; pour tout $b \in \mathrm{B}$ on a $\|\mathrm{TS}(b)\| \leq \|\mathrm{T}\| \, \|\mathrm{S}(b)\|$ donc

$$\|\mathrm{TS}\|_{\mathrm{HS}}^2 = \sum_{b \in \mathrm{B}} \|\mathrm{TS}(b)\|^2 \leq \|\mathrm{T}\|^2 \sum_{b \in \mathrm{B}} \|\mathrm{S}(b)\|^2 = \|\mathrm{T}\|^2 \, \|\mathrm{S}\|_{\mathrm{HS}}^2.$$

La deuxième inégalité en résulte en remplaçant S et T par leurs adjoints.

Exemples.

1. Prenons d'abord $\mathrm{H}_1 = \mathrm{H}_2 = \mathbb{C}^n$. Un opérateur $\mathrm{T} \in \mathcal{L}(\mathbb{C}^n)$ est représenté par une matrice $(a_{i,j})$ dans la base canonique ; si $(e_j)_{j=1}^n$ désigne la base canonique de $\mathbb{C}^n$, on a $\|\mathrm{T}(e_j)\|^2 = \sum_{i=1}^n |a_{i,j}|^2$, donc la norme Hilbert-Schmidt de T est égale à

$$\|\mathrm{T}\|_{\mathrm{HS}} = \Big(\sum_{i,j=1}^{n} |a_{i,j}|^2\Big)^{1/2}.$$

Si $\mathrm{H}_1 = \mathrm{H}_2 = \ell_2(\mathbb{N})$, un opérateur T peut se représenter par une matrice infinie $(a_{i,j})$, et on voit de même que la norme Hilbert-Schmidt est égale à

$$\|\mathrm{T}\|_{\mathrm{HS}} = \Big(\sum_{i,j=0}^{+\infty} |a_{i,j}|^2\Big)^{1/2}.$$

2. L'exemple le plus important est celui de certains opérateurs définis par un noyau $\mathrm{K}(s,t)$ de carré intégrable sur un espace produit $\mathrm{S} \times \mathrm{T}$. Cet exemple est développé dans le paragraphe suivant.

Opérateurs de Hilbert-Schmidt définis par un noyau

Soient (S, μ) et (T, ν) deux espaces mesurés σ-finis et $\mathrm{K}(s, t)$ une fonction de carré intégrable sur $\mathrm{S} \times \mathrm{T}$. On va montrer que pour toute fonction $f \in \mathrm{L}_2(\mathrm{T}, \nu)$,

$$\int_{\mathrm{S}} \Bigl(\int_{\mathrm{T}} |\mathrm{K}(s,t)| \, |f(t)| \, d\nu(t) \Bigr)^2 \, d\mu(s) < +\infty$$

ce qui montrera que l'intégrale $\int_{\mathrm{T}} \mathrm{K}(s,t) f(t) \, d\nu(t)$ est absolument convergente pour μ-presque toute valeur de $s \in \mathrm{S}$. On définira un opérateur T_{K} en posant pour $f \in \mathrm{L}_2(\mathrm{T}, \nu)$, et pour $s \in \mathrm{S}$

$$(\mathrm{T}_{\mathrm{K}} f)(s) = \int_{\mathrm{T}} \mathrm{K}(s,t) f(t) \, d\nu(t).$$

Montrons comme annoncé que T_{K} est bien défini, et agit continûment de $\mathrm{L}_2(\mathrm{T}, \nu)$ dans $\mathrm{L}_2(\mathrm{S}, \mu)$: avec Cauchy-Schwarz, on a

$$\bigl(\int_{\mathrm{T}} |\mathrm{K}(s,t)| \, |f(t)| \, d\nu(t) \bigr)^2 \le \bigl(\int_{\mathrm{T}} |\mathrm{K}(s,t)|^2 \, d\nu(t) \bigr) \bigl(\int_{\mathrm{T}} |f(t)|^2 \, d\nu(t) \bigr)$$

ce qui donne en réintégrant

$$\int_{\mathrm{S}} \Bigl(\int_{\mathrm{T}} |\mathrm{K}(s,t)| \, |f(t)| \, d\nu(t) \Bigr)^2 \, d\mu(s) \le \bigl(\int_{\mathrm{S} \times \mathrm{T}} |\mathrm{K}(s,t)|^2 \, d\mu(s) d\nu(t) \bigr) \bigl(\int_{\mathrm{T}} |f(t)|^2 \, d\nu(t) \bigr)$$

qui est une quantité finie. Reprenant le calcul, en commençant par majorer $|(\mathrm{T}_{\mathrm{K}} f)(s)|$ par l'intégrale de la valeur absolue, on trouve

$$\int_{\mathrm{S}} |(\mathrm{T}_{\mathrm{K}} f)(s)|^2 \, d\mu(s) \le \Bigl(\int_{\mathrm{S} \times \mathrm{T}} |\mathrm{K}(s,t)|^2 \, d\mu(s) d\nu(t) \Bigr) \Bigl(\int_{\mathrm{T}} |f(t)|^2 \, d\nu(t) \Bigr) = \|\mathrm{K}\|_2^2 \, \|f\|^2.$$

On a trouvé que l'intégrale qui définit T_{K} est absolument convergente pour μ-presque tout s, et on voit de plus que $\|\mathrm{T}_{\mathrm{K}}\| \le \|\mathrm{K}\|_2$.

Proposition. *L'application* $h : \mathrm{K} \to \mathrm{T}_{\mathrm{K}}$ *est linéaire et continue de* $\mathrm{L}_2(\mathrm{S} \times \mathrm{T}, \mu \otimes \nu)$ *dans* $\mathcal{L}(\mathrm{L}_2(\mathrm{T}, \nu), \mathrm{L}_2(\mathrm{S}, \nu))$ *; on a* $\|h\| \le 1$.

Envisageons le cas particulier où $\mathrm{K} = \varphi \otimes \overline{\psi}$; dans ce cas,

$$(\mathrm{T}_{\mathrm{K}} f)(s) = \int_{\mathrm{T}} \varphi(s) \overline{\psi(t)} f(t) \, d\nu(t) = \langle f, \psi \rangle \, \varphi(s).$$

Ceci montre que $\mathrm{T}_{\varphi \otimes \overline{\psi}} f = \langle f, \psi \rangle \, \varphi$, donc l'opérateur est de rang un (son image est contenue dans la droite engendrée par φ). Si $\mathrm{K} = \sum_{j=1}^n \varphi_j \otimes \overline{\psi}_j$, on en déduit que T_{K} est de rang fini, et $\mathrm{rang}(\mathrm{T}_{\mathrm{K}}) \le n$.

Si $(f_n)_{n\geq 0}$ est une base hilbertienne de $\mathrm{L}_2(\mathrm{S},\mu)$ et $(\overline{g_n})_{n\geq 0}$ une base hilbertienne de $\mathrm{L}_2(\mathrm{T},\nu)$, on a vu que les fonctions $(s,t)\to f_m(s)\overline{g_n(t)}$ (où m,n prennent toutes les valeurs entières ≥ 0) donnent une base orthonormée de l'espace $\mathrm{L}_2(\mathrm{S}\times\mathrm{T},\mu\otimes\nu)$. Supposons donné un noyau $\mathrm{K}\in\mathrm{L}_2(\mathrm{S}\times\mathrm{T},\mu\otimes\nu)$; on peut le décomposer sur la base hilbertienne précédente,

$$(*)\qquad \mathrm{K}=\sum_{m,n=0}^{+\infty} c_{m,n}\, f_m\otimes\overline{g_n}.$$

On sait que

$$\sum_{m,n=0}^{+\infty}|c_{m,n}|^2=\|\mathrm{K}\|_2^2<+\infty,$$

et la série $(*)$ converge dans l'espace de Hilbert $\mathrm{L}_2(\mu\otimes\nu)$, quel que soit l'ordre dans lequel on énumère les termes. Par image linéaire continue, on en déduit que

$$\mathrm{T}_\mathrm{K}=\sum_{m,n=0}^{+\infty} c_{m,n}\,\mathrm{T}_{f_m\otimes\overline{g_n}}\,,$$

où la série converge en norme d'opérateur, quel que soit l'ordre dans lequel on énumère les termes. L'opérateur T_K est donc limite en norme d'opérateur des sommes partielles, qui sont des opérateurs de rang fini, donc T_K est compact. En fait, l'opérateur T_K est de Hilbert-Schmidt : il résulte de la convergence de la série ci-dessus que

$$\mathrm{T}_\mathrm{K}f=\sum_{m,n=0}^{+\infty} c_{m,n}\mathrm{T}_{f_m\otimes\overline{g_n}}\,f=\sum_{m,n=0}^{+\infty} c_{m,n}\langle f,g_n\rangle f_m.$$

Il en résulte que $\mathrm{T}_\mathrm{K}(g_p)=\sum_m c_{m,p}f_m$, donc $\|\mathrm{T}_\mathrm{K}(g_p)\|^2=\sum_m|c_{m,p}|^2$, et ensuite

$$\sum_{p=0}^{+\infty}\|\mathrm{T}_\mathrm{K}(g_p)\|^2=\sum_{m,p=0}^{+\infty}|c_{m,p}|^2=\|\mathrm{K}\|_2^2<+\infty.$$

On a donc vérifié que la norme Hilbert-Schmidt de T_K est *égale* à la norme L_2 du noyau K dans l'espace $\mathrm{L}_2(\mathrm{S}\times\mathrm{T},\mu\otimes\nu)$. En particulier, l'application $\mathrm{K}\to\mathrm{T}_\mathrm{K}$ est injective : si l'opérateur T_K est l'opérateur nul, le noyau K est nul $\mu\otimes\nu$-presque partout sur $\mathrm{S}\times\mathrm{T}$. On notera aussi la correspondance entre l'opérateur T_K et les coefficients $c_{m,n}$ du développement de K,

$$(**)\qquad c_{m,n}=\langle\mathrm{T}_\mathrm{K}g_n,f_m\rangle.$$

Exercice. Montrer que la composition de deux opérateurs $\mathrm{T}_{\mathrm{K}_1}$ et $\mathrm{T}_{\mathrm{K}_2}$ de la forme précédente est un opérateur T_K, avec

$$\mathrm{K}(s,t)=\int\mathrm{K}_1(s,u)\mathrm{K}_2(u,t)du.$$

Opérateurs hermitiens à noyau

On vérifie facilement avec Fubini que l'adjoint $(\mathrm{T_K})^*$ de $\mathrm{T_K}$ est l'opérateur de noyau $\mathrm{K}^*(t,s) = \overline{\mathrm{K}(s,t)}$. Supposons que $\mathrm{S} = \mathrm{T}$, $\mu = \nu$ et que K soit un *noyau hermitien*, c'est à dire que

$$\mathrm{K}(t,s) = \overline{\mathrm{K}(s,t)}$$

pour tous $(s,t) \in \mathrm{S}^2$; dans ce cas l'opérateur $\mathrm{T_K}$ est compact et hermitien ; il existe donc une base orthonormée (f_n) de $\mathrm{L}_2(\mathrm{S},\mu)$ formée de vecteurs propres de l'opérateur $\mathrm{T_K}$, c'est à dire telle que $\mathrm{T_K} f_n = \lambda_n f_n$ pour tout $n \ge 0$. Si on exprime le noyau K dans la base orthonormée de l'espace $\mathrm{L}_2(\mathrm{S}^2, \mu \otimes \mu)$ formée des fonctions $h_{m,n} = f_m \otimes \overline{f_n}$, on obtient un développement $\mathrm{K} = \sum_{m,n} c_{m,n} f_m \otimes \overline{f_n}$, et on a vu que

$$c_{m,n} = \langle \mathrm{T_K} f_n, f_m \rangle = \lambda_n \, \langle f_n, f_m \rangle = \lambda_n \, \delta_{m,n},$$

ce qui montre que $c_{p,p} = \lambda_p$, et que les autres coefficients $c_{m,p}$, pour $m \ne p$ sont nuls. On voit donc que tout noyau hermitien K sur S^2 se représente sous la forme

$$\mathrm{K}(s,t) = \sum_{n=0}^{+\infty} \lambda_n \, f_n \otimes \overline{f_n}$$

où les λ_n sont réels, et (f_n) une base orthonormée. La série converge **au sens de** L_2.

Remarque. Dans certains cas, le noyau K est continu, disons par exemple sur $[0,1]^2$, les fonctions propres précédentes sont continues sur $[0,1]$, et la série qui représente K est aussi normalement convergente sur $[0,1]^2$; si de plus tout ouvert non vide de $\mathrm{S} \times \mathrm{S}$ est de mesure > 0 (ce qui est le cas pour la mesure de Lebesgue), on en déduira que l'égalité ci-dessus est vraie *pour toutes* les valeurs s, t,

$$\mathrm{K}(s,t) = \sum_{n=0}^{+\infty} \lambda_n f_n(s) \overline{f_n(t)}.$$

Exemple. Traiter le cas $\mathrm{K}(s,t) = \max(s,t)$, $0 \le s,t \le 1$.