ANALYSE FONCTIONNELLE ET THÉORIE SPECTRALE

MT404

Année 2000-2001

deuxième partie

Sommaire de la deuxième partie

5. Espaces hilbertiens

On a déjà revu une bonne partie des propriétés des espaces de Hilbert dans le chapitre 2, en particulier l'existence de base orthonormée pour tout espace de Hilbert séparable. Dans ce chapitre, on va généraliser la notion de base hilbertienne au cas non séparable ; l'outil essentiel pour cette extension est la notion de *famille sommable* de vecteurs d'un espace normé. On introduira aussi dans ce chapitre les principales classes d'opérateurs bornés entre espaces de Hilbert.

5.1. Produits scalaires

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

Définition 5.1.1. Soit X un espace vectoriel complexe ; on appelle *forme sesquilinéaire* sur X une application $\mathrm{B} : \mathrm{X} \times \mathrm{X} \to \mathbb{C}$ telle que, pour tout $y \in \mathrm{X}$, l'application $x \to \mathrm{B}(x, y)$ soit linéaire et telle que pour tout $x \in \mathrm{X}$, l'application $y \to \mathrm{B}(x, y)$ soit antilinéaire (de X dans $\mathbb{C}$).

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

Proposition 5.1.1 : identité de polarisation.

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

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

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

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

Démonstration. On a $\mathrm{B}(x+y, x+y) - \mathrm{B}(x-y, x-y) = 2(\mathrm{B}(x, y) + \mathrm{B}(y, x))$. Remplaçant y par iy, on trouve : $\mathrm{B}(x+iy, x+iy) - \mathrm{B}(x-iy, x-iy) = 2(\mathrm{B}(x, iy) + \mathrm{B}(iy, x)) = 2i(-\mathrm{B}(x, y) + \mathrm{B}(y, x))$; le point (*i*) en résulte ; le point (*ii*) est laissé en exercice.

//

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

Corollaire 5.1.2. *Soient* X *un espace vectoriel complexe et* B *une forme sesquilinéaire sur* X *; les conditions suivantes sont équivalentes :*

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

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

//

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

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

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

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

Proposition 5.1.3 : inégalité de Cauchy-Schwarz. *Soit* X *un espace vectoriel muni d'un produit scalaire ; pour tous* $x, y \in \mathrm{X}$ *on a*

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

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

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

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

//

Corollaire 5.1.4. *Soit* X *un espace vectoriel muni d'un produit scalaire ; l'application* $x \to \langle x, x\rangle^{1/2}$ *est une semi-norme sur* X.

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

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

par la proposition 3.

//

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

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

5.2. Espaces hilbertiens

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

Proposition 5.2.1. *Soit* X *un espace préhilbertien ; pour tout vecteur* $x \in \mathrm{X}$ *la forme linéaire* $f_x : y \to \langle y, x \rangle$ *est continue. L'application* $x \to f_x$ *est antilinéaire et isométrique de* X *dans* X^*.

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

//

Soit (X, p) un espace normé ; on dira que la norme p est *issue d'un produit scalaire* s'il existe un produit scalaire $\langle \ , \ \rangle$ sur X tel que, pour tout $x \in \mathrm{X}$ on ait $p(x) = \langle x, x \rangle^{1/2}$. Si un tel produit scalaire existe, il est unique, par la proposition 1.1. On dira que (X, p) est un espace préhilbertien si p est issue d'un produit scalaire ; on dira que (X, p) est un espace hilbertien s'il est préhilbertien complet.

Proposition 5.2.2. *Soit* X *un espace préhilbertien ; le complété de* X *est un espace hilbertien.*

Démonstration. Notons H le complété de X et considérons pour simplifier que X est un sous-espace vectoriel de H. Si h, k sont deux éléments de H, on peut trouver deux suites (x_n) et (y_n) dans X, telles que $h = \lim_n x_n$ et $k = \lim y_n$. On remarque que la suite scalaire $(\langle x_n, y_n \rangle)$ est de Cauchy :

$$\begin{aligned} \big|\langle x_n, y_n \rangle - \langle x_m, y_m \rangle\big| &\le \big|\langle x_n - x_m, y_n \rangle\big| + \big|\langle x_m, y_n - y_m \rangle\big| \\ &\le \|x_n - x_m\| \, \|y_n\| + \|x_m\| \, \|y_n - y_m\| \end{aligned}$$

qui tend vers 0 lorsque $n, m \to +\infty$ parce que les deux suites sont de Cauchy, donc bornées. De plus, on peut voir que la limite $\lim_n \langle x_n, y_n \rangle$ ne dépend que de h et k, ce qui permet de poser

$$\langle h, k \rangle = \lim_n \langle x_n, y_n \rangle.$$

En passant à la limite, on vérifie que cette formule définit une forme hermitienne sur $\mathrm{H} \times \mathrm{H}$. Lorsque $h = k$, on aura

$$\langle h, h \rangle = \lim_n \langle x_n, x_n \rangle = \lim_n \|x_n\|^2 = \|h\|^2$$

ce qui montre que le produit scalaire défini sur H donne la norme de H.

//

Exemples 5.2.1.

1. Soit I un ensemble d'indices non dénombrable, et soit H l'espace des familles $x = (x_i)_{i \in \mathrm{I}}$ de scalaires telles que l'ensemble des $i \in \mathrm{I}$ tels que $x_i \ne 0$ soit un ensemble dénombrable $\mathrm{J}(x)$, et telles que $\sum_{i \in \mathrm{J}(x)} |x_i|^2 < +\infty$. Si x et y sont de telles familles, le produit $x_i y_i$ est nul sauf pour au plus un ensemble dénombrable d'indices J, et $|x_i y_i| \le \frac{1}{2}(x_i^2 + y_i^2)$, ce qui permet de poser

$$\langle x, y \rangle = \sum_{i \in \mathrm{J}} x_i y_i,$$

le résultat ne dépendant pas de l'ensemble dénombrable J qui contient tous les indices i tels que $x_i y_i \ne 0$. On obtient ainsi un exemple d'espace de Hilbert non séparable, qui sera traité plus en détail à la section 5.6.

2. On a vu au chapitre 2 que l'espace $\mathrm{L}_2(\Omega, \mu)$ est un espace de Hilbert pour le produit scalaire $\langle f, g \rangle = \int_\Omega f(s) \overline{g(s)} \, d\mu(s)$. Il est possible que cet espace soit non séparable, même si la mesure μ est une probabilité.

Mentionnons une conséquence du théorème de projection. On rappelle que l'orthogonal d'une partie $\mathrm{A} \subset \mathrm{H}$ est le sous-espace vectoriel $\mathrm{A}^\perp$ de H formé de tous les vecteurs x qui sont orthogonaux à tous les éléments de A.

Proposition 5.2.3. *Soit* H *un espace de Hilbert ;*

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

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

Démonstration. Montrons le point (*i*). Soit F le plus petit sous-espace vectoriel fermé de H contenant A ; tout élément de A est orthogonal à $\mathrm{A}^\perp$, donc $\mathrm{A} \subset (\mathrm{A}^\perp)^\perp$. Comme $(\mathrm{A}^\perp)^\perp$ est un sous-espace vectoriel fermé de H contenant A, on a $\mathrm{F} \subset (\mathrm{A}^\perp)^\perp$. Inversement, soit $x \in (\mathrm{A}^\perp)^\perp$; écrivons $x = y + z$ avec $y \in \mathrm{F}$ et $z \in \mathrm{F}^\perp$. Comme $\mathrm{F} \subset (\mathrm{A}^\perp)^\perp$, $z = x - y \in (\mathrm{A}^\perp)^\perp$; comme $\mathrm{A} \subset \mathrm{F}$, on a $\mathrm{F}^\perp \subset \mathrm{A}^\perp$; comme $z \in (\mathrm{A}^\perp)^\perp \cap \mathrm{A}^\perp$, il s'ensuit que $\langle z, z\rangle = 0$, donc $z = 0$. On en déduit que $x = y \in \mathrm{F}$.

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

//

5.3. Adjoint d'une application linéaire continue

Proposition 5.3.1. *Soient* E *et* F *deux espaces de Hilbert et* $\mathrm{T} \in \mathcal{L}(\mathrm{E}, \mathrm{F})$ *; il existe un unique* $\mathrm{T}^* \in \mathcal{L}(\mathrm{F}, \mathrm{E})$ *tel que pour tout* $x \in \mathrm{E}$ *et tout* $y \in \mathrm{F}$ *on ait*

$$\langle \mathrm{T}(x), y\rangle = \langle x, \mathrm{T}^*(y)\rangle.$$

On a de plus $\|\mathrm{T}^*\| = \|\mathrm{T}\|$.

Démonstration. Pour tout $y \in \mathrm{F}$, l'application $x \to \langle \mathrm{T}(x), y\rangle$ est linéaire et continue. Il existe donc un unique élément $\mathrm{T}^*(y) \in \mathrm{E}$ tel que pour tout $x \in \mathrm{E}$ on ait $\langle \mathrm{T}(x), y\rangle = \langle x, \mathrm{T}^*(y)\rangle$. Pour tout $x \in \mathrm{E}$, l'application $y \to \langle \mathrm{T}^*(y), x\rangle = \langle y, \mathrm{T}(x)\rangle$ est linéaire d'où l'on déduit que T^* est linéaire. On a, par définition de $\|\mathrm{T}\|$ et par la proposition 2.1

$$\begin{aligned}\|\mathrm{T}^*\| &= \sup\{\|\mathrm{T}^*(y)\| : y \in \mathrm{F},\ \|y\| \le 1\} \\ &= \sup\{\langle x, \mathrm{T}^*(y)\rangle : x \in \mathrm{E},\ y \in \mathrm{F},\ \|x\|, \|y\| \le 1\} \\ &= \sup\{\langle \mathrm{T}(x), y\rangle : x \in \mathrm{E},\ y \in \mathrm{F},\ \|x\|, \|y\| \le 1\} = \|\mathrm{T}\|.\end{aligned}$$

//

Rapport avec ${}^t\mathrm{T}$. Désignons par h_E l'isomorphisme antilinéaire d'un espace de Hilbert E sur son dual E^*. Si E et F sont deux espaces de Hilbert et si $\mathrm{T} \in \mathcal{L}(\mathrm{E}, \mathrm{F})$, la transposée est linéaire de F^* dans E^* et elle est reliée à l'adjoint T^* par la formule

$$\mathrm{T}^* = h_\mathrm{E}^{-1}\,({}^t\mathrm{T})\,h_\mathrm{F}.$$

Définition 5.3.1. Soient E et F deux espaces de Hilbert et $\mathrm{T} \in \mathcal{L}(\mathrm{E}, \mathrm{F})$; l'unique $\mathrm{T}^* \in \mathcal{L}(\mathrm{F}, \mathrm{E})$ tel que pour tout $x \in \mathrm{E}$ et tout $y \in \mathrm{F}$ on ait $\langle \mathrm{T}(x), y\rangle = \langle x, \mathrm{T}^*(y)\rangle$ est appelé *adjoint* de T.

Regroupons dans la proposition suivante quelques propriétés des adjoints.

Proposition 5.3.2. *Soient* E *et* F *deux espaces de Hilbert ; l'application* $\mathrm{T} \to \mathrm{T}^*$ *est antilinéaire et isométrique de* $\mathcal{L}(\mathrm{E}, \mathrm{F})$ *sur* $\mathcal{L}(\mathrm{F}, \mathrm{E})$ *; pour tout* $\mathrm{T} \in \mathcal{L}(\mathrm{E}, \mathrm{F})$ *on a* $(\mathrm{T}^*)^* = \mathrm{T}$ *et* $\|\mathrm{T}^* \circ \mathrm{T}\| = \|\mathrm{T}\|^2$*. Pour tout espace hilbertien* H, *tout* $\mathrm{S} \in \mathcal{L}(\mathrm{E}, \mathrm{F})$ *et tout* $\mathrm{T} \in \mathcal{L}(\mathrm{F}, \mathrm{H})$ *on a* $(\mathrm{T} \circ \mathrm{S})^* = \mathrm{S}^* \circ \mathrm{T}^*$.

Démonstration. Montrons que $\|\mathrm{T}^* \circ \mathrm{T}\| = \|\mathrm{T}\|^2$. On a $\|\mathrm{T}^* \circ \mathrm{T}\| \leq \|\mathrm{T}^*\| \, \|\mathrm{T}\| \leq \|\mathrm{T}\|^2$. De plus, pour tout $x \in \mathrm{E}$ tel que $\|x\| \leq 1$ on a

$$\|\mathrm{T}(x)\|^2 = \langle \mathrm{T}(x), \mathrm{T}(x) \rangle = \langle x, \mathrm{T}^* \circ \mathrm{T}(x) \rangle \leq \|\mathrm{T}^* \circ \mathrm{T}\|$$

grâce à l'inégalité de Cauchy-Schwarz. Donc

$$\|\mathrm{T}\|^2 = \sup\{\|\mathrm{T}(x)\|^2 : x \in \mathrm{H},\ \|x\| \leq 1\} \leq \|\mathrm{T}^* \circ \mathrm{T}\|.$$

Les autres propriétés sont laissées en exercice.

//

Proposition 5.3.3. *Soit* E *et* F *deux espaces de Hilbert et soit* $\mathrm{T} \in \mathcal{L}(\mathrm{E}, \mathrm{F})$ *; alors* $\ker \mathrm{T}^* = (\mathrm{T}(\mathrm{E}))^\perp$ *et l'adhérence de* $\mathrm{T}^*(\mathrm{F})$ *est* $(\ker \mathrm{T})^\perp$.

Démonstration. Si $y \in \mathrm{F}$, on voit que $y \in \ker \mathrm{T}^*$ si et seulement si pour tout $x \in \mathrm{E}$, on a $0 = \langle \mathrm{T}^*(y), x \rangle = \langle y, \mathrm{T}(x) \rangle$; clairement, ceci équivaut à dire que $y \in (\mathrm{T}(\mathrm{E}))^\perp$, d'où la première assertion. Il en résulte (par la proposition 2.3) que $\overline{\mathrm{T}(\mathrm{E})} = (\ker \mathrm{T}^*)^\perp$, d'où la deuxième assertion en remplaçant T par son adjoint.

//

Définition 5.3.2. Soient E et F deux espaces de Hilbert ; un élément $\mathrm{U} \in \mathcal{L}(\mathrm{E}, \mathrm{F})$ est appelé *unitaire* si $\mathrm{U}^* \circ \mathrm{U} = \mathrm{Id}_\mathrm{E}$ et $\mathrm{U} \circ \mathrm{U}^* = \mathrm{Id}_\mathrm{F}$. Un élément $\mathrm{T} \in \mathcal{L}(\mathrm{E})$ est appelé *normal* si $\mathrm{T}^* \circ \mathrm{T} = \mathrm{T} \circ \mathrm{T}^*$, *autoadjoint* ou *hermitien* si $\mathrm{T} = \mathrm{T}^*$ et *positif* s'il est autoadjoint et si $\langle \mathrm{T}(x), x \rangle$ est réel ≥ 0 pour tout $x \in \mathrm{E}$.

Exemples 5.3.3.

1. Soient H un espace hilbertien, $\mathrm{P} \in \mathcal{L}(\mathrm{H})$ un projecteur orthogonal ; notons F son image. Pour $x, x' \in \mathrm{F}$ et $y, y' \in \mathrm{F}^\perp$ on a $\langle \mathrm{P}(x+y), x'+y' \rangle = \langle x, x' \rangle = \langle x+y, \mathrm{P}(x'+y') \rangle$; donc $\mathrm{P} = \mathrm{P}^*$. De plus, $\langle \mathrm{P}(x+y), x+y \rangle = \langle x, x \rangle \in \mathbb{R}_+$, donc P est positif.

2. On dit que $\mathrm{U} \in \mathcal{L}(\mathrm{E}, \mathrm{F})$ est *isométrique* si $\|\mathrm{U}(x)\| = \|x\|$ pour tout $x \in \mathrm{E}$. On vérifie facilement que U est isométrique si et seulement si $\mathrm{U}^* \mathrm{U} = \mathrm{Id}_\mathrm{E}$. En effet, si $\mathrm{U}^* \circ \mathrm{U} = \mathrm{Id}_\mathrm{E}$, alors pour tout $x \in \mathrm{E}$ on a $\|\mathrm{U}(x)\|^2 = \langle \mathrm{U}(x), \mathrm{U}(x) \rangle = \langle x, \mathrm{U}^*(\mathrm{U}(x)) \rangle = \|x\|^2$. La réciproque utilise la formule de polarisation (voir la démonstration ci-dessous).

Proposition 5.3.4. *Soient* E *et* F *deux espaces de Hilbert et* $\mathrm{T} \in \mathcal{L}(\mathrm{E}, \mathrm{F})$ *; les conditions suivantes sont équivalentes :*

(*i*) *l'opérateur* T *est unitaire ;*
(*ii*) *l'opérateur* T *est surjectif et* $\mathrm{T}^* \circ \mathrm{T} = \mathrm{Id}_\mathrm{E}$ *;*
(*iii*) T *est une isométrie de* E *sur* F.

Démonstration. Si T est unitaire, comme $\mathrm{T} \circ \mathrm{T}^* = \mathrm{Id}_\mathrm{F}$, l'opérateur T est surjectif, donc (*i*) $\Rightarrow$ (*ii*). Si $\mathrm{T}^* \circ \mathrm{T} = \mathrm{Id}_\mathrm{E}$, alors pour tout $x \in \mathrm{E}$ on a $\|\mathrm{T}(x)\|^2 = \|x\|^2$, donc

$(ii) \Rightarrow (iii)$. Enfin, supposons que T soit une isométrie de E sur F, c'est à dire que pour tout $x \in \mathrm{E}$ on ait $\langle x, x\rangle = \langle \mathrm{T}(x), \mathrm{T}(x)\rangle$; comme

$$(x, y) \to \langle x, \mathrm{T}^*(\mathrm{T}(y))\rangle = \langle \mathrm{T}x, \mathrm{T}y\rangle$$

est un produit scalaire sur E, il résulte de la proposition 1.1 que, pour tous $x, y \in \mathrm{E}$, on a $\langle x, \mathrm{T}^*(\mathrm{T}(y))\rangle = \langle x, y\rangle$, ce qui implique $\mathrm{T}^*(\mathrm{T}(y)) = y$, c'est à dire que $\mathrm{T}^* \circ \mathrm{T} = \mathrm{Id}_{\mathrm{E}}$. Comme par l'hypothèse (iii) l'application T est bijective, $\mathrm{T}^* = \mathrm{T}^{-1}$, d'où (i). //

5.4. Familles sommables dans un espace de Banach

Définition 5.4.1. Soient X un espace normé, I un ensemble et $(x_i)_{i\in\mathrm{I}}$ une famille d'éléments de X ; on dit que la famille $(x_i)_{i\in\mathrm{I}}$ est *sommable* de somme $\mathrm{S} \in \mathrm{X}$ et on écrit $\mathrm{S} = \sum_{i\in\mathrm{I}} x_i$ si, pour tout $\varepsilon > 0$, il existe une partie finie J de I telle que, pour toute partie finie K de I contenant J on ait $\left\|\mathrm{S} - \sum_{i\in\mathrm{K}} x_i\right\| < \varepsilon$.

La somme S est unique : remarquons que si S et S′ sont deux éléments de E vérifiant ces conditions, alors pour tout $\varepsilon > 0$, il existe des parties finies J et J′ de I telles que, pour toute partie finie K de I contenant J (resp. J′) on ait $\left\|\mathrm{S} - \sum_{i\in\mathrm{K}} x_i\right\| < \varepsilon$ (resp. $\left\|\mathrm{S}' - \sum_{i\in\mathrm{K}} x_i\right\| < \varepsilon$). Prenant $\mathrm{K} = \mathrm{J} \cup \mathrm{J}'$ on trouve $\|\mathrm{S} - \mathrm{S}'\| < 2\varepsilon$. Cela étant vrai pour tout $\varepsilon > 0$, on en déduit $\mathrm{S} = \mathrm{S}'$.

Il est facile de vérifier que si $(x_i)_{i\in\mathrm{I}}$ et $(y_i)_{i\in\mathrm{I}}$ sont deux familles sommables, la famille $(x_i + y_i)_{i\in\mathrm{I}}$ est elle aussi sommable, avec une somme égale à la somme des deux sommes. Considérons pour commencer le cas des familles sommables de nombres réels, et d'abord de réels ≥ 0.

Proposition 5.4.1. *Une famille de nombres réels à termes positifs est sommable si et seulement si les sommes finies sont majorées ; sa somme est alors la borne supérieure de l'ensemble des sommes finies. Une famille $(x_i)_{i\in\mathrm{I}}$ à termes réels est sommable si et seulement si elle est absolument sommable, c'est à dire si la famille $(|x_i|)_{i\in\mathrm{I}}$ est sommable.*

Démonstration. Si $(x_i)_{i\in\mathrm{I}}$ est sommable, et $x_i \geq 0$ pour tout $i \in \mathrm{I}$, on aura, en prenant $\varepsilon = 1$ dans la définition, l'existence d'un ensemble fini $\mathrm{J}_0 \subset \mathrm{I}$ tel que $|\mathrm{S} - \sum_{i\in\mathrm{K}} x_i| \leq 1$ pour tout ensemble fini K contenant J_0. Pour tout $\mathrm{L} \subset \mathrm{I}$ fini on aura alors

$$\sum_{i\in\mathrm{L}} x_i \leq \sum_{i\in\mathrm{L}\cup\mathrm{J}_0} x_i \leq \mathrm{S} + 1,$$

ce qui montre que l'ensemble Σ des sommes finies est majoré. Réciproquement, lorsque l'ensemble Σ est majoré, il est assez facile de montrer que la somme de la famille est égale à la borne supérieure de l'ensemble Σ.

Supposons maintenant que $(x_i)_{i\in\mathrm{I}}$ est sommable à termes réels quelconques, et choisissons J_0 comme ci-dessus. Soit I^+ l'ensemble des indices $i \in \mathrm{I}$ tels que $x_i > 0$, et soit $\mathrm{L} \subset \mathrm{I}^+$ un ensemble fini, disjoint de J_0. On aura

$$\Big|\mathrm{S} - \sum_{i\in\mathrm{L}\cup\mathrm{J}_0} x_i\Big| = \Big|\mathrm{S} - \sum_{i\in\mathrm{L}} x_i - \sum_{i\in\mathrm{J}_0} x_i\Big| \leq \varepsilon = 1,$$

ce qui implique que les sommes finies dans $\mathrm{I}^+ \setminus \mathrm{J}_0$ sont majorées par la quantité $|\mathrm{S}| + |\sum_{i\in \mathrm{J}_0} x_i| + 1$. Il en résulte par la première partie que la famille $(x_i)_{i\in \mathrm{I}^+}$ est sommable, et le même raisonnement s'applique aux termes < 0 de la famille. Il en résulte que la famille $(|x_i|)_{i\in \mathrm{I}}$ est sommable.

//

Plus généralement, si E est de dimension finie, la famille $(x_n)_{n\geq 0}$ est sommable si et seulement si la série $\sum x_n$ est normalement convergente (exercice). En revanche, dans un espace de Banach de dimension infinie, il existe toujours des familles sommables de vecteurs qui ne sont pas sommables en norme (exercice infaisable).

La proposition suivante est laissée en exercice :

Proposition 5.4.2. *Soient* X *et* Y *deux espaces normés,* $f \in \mathcal{L}(\mathrm{X}, \mathrm{Y})$ *et* $(x_i)_{i\in \mathrm{I}}$ *une famille sommable d'éléments de* X *; alors la famille* $(f(x_i))_{i\in \mathrm{I}}$ *est sommable dans* Y *et on a* $\sum_{i\in \mathrm{I}} f(x_i) = f(\sum_{i\in \mathrm{I}} x_i)$.

Soient X un espace normé et $(x_i)_{i\in \mathrm{I}}$ une famille d'éléments de X ; on dit que la famille $(x_i)_{i\in \mathrm{I}}$ vérifie le *critère de sommabilité de Cauchy* si, pour tout $\varepsilon > 0$, il existe une partie finie J de I telle que, pour toute partie finie L de I disjointe de J, on ait $\left\|\sum_{i\in \mathrm{L}} x_i\right\| < \varepsilon$.

Proposition 5.4.3.
(i) *Toute famille sommable d'un espace normé vérifie le critère de sommabilité de Cauchy.*

(ii) *Dans un espace de Banach, toute famille vérifiant le critère de sommabilité de Cauchy est sommable.*

Démonstration. Soit $(x_i)_{i\in \mathrm{I}}$ une famille sommable d'éléments d'un espace normé X, et notons S sa somme ; soient $\varepsilon > 0$ et J une partie finie de I telle que, pour toute partie finie K de I contenant J on ait $\left\|\mathrm{S} - \sum_{i\in \mathrm{K}} x_i\right\| < \varepsilon/2$; si L est une partie finie de I disjointe de J, on a

$$\left\|\sum_{i\in \mathrm{L}} x_i\right\| = \left\|\Big(\mathrm{S} - \sum_{i\in \mathrm{J}} x_i\Big) - \Big(\mathrm{S} - \sum_{i\in \mathrm{L}\cup \mathrm{J}} x_i\Big)\right\| < \varepsilon,$$

ce qui montre que le critère de Cauchy est vérifié. Soient E un espace de Banach et $(x_i)_{i\in \mathrm{I}}$ une famille d'éléments de E vérifiant le critère de sommabilité de Cauchy ; il existe une suite (J_n) de parties finies de I telles que, pour tout entier $n \geq 0$ et toute partie finie J de I disjointe de J_n, on ait $\left\|\sum_{i\in \mathrm{J}} x_i\right\| < 2^{-n}$. Posons $\mathrm{K}_n = \bigcup_{k=0}^n \mathrm{J}_k$ et $\mathrm{S}_n = \sum_{i\in \mathrm{K}_n} x_i$; pour n, m entiers, avec $n \leq m$, on a

$$\|\mathrm{S}_m - \mathrm{S}_n\| = \left\|\sum_{i\in \mathrm{K}_m\setminus \mathrm{K}_n} x_i\right\| \leq 2^{-n},$$

vu que $\mathrm{K}_m \setminus \mathrm{K}_n$ est disjoint de J_n. La suite (S_n) étant de Cauchy, elle converge. Soit S sa limite ; pour $m \geq n$ on a $\|\mathrm{S}_m - \mathrm{S}_n\| \leq 2^{-n}$. Faisant tendre m vers l'infini, on en déduit que $\|\mathrm{S} - \mathrm{S}_n\| \leq 2^{-n}$. Si J est une partie finie contenant K_n, on a alors

$$\left\|\mathrm{S} - \sum_{i\in \mathrm{J}} x_i\right\| = \left\|(\mathrm{S} - \mathrm{S}_n) + \Big(\mathrm{S}_n - \sum_{i\in \mathrm{J}} x_i\Big)\right\| \leq 2^{-n} + \left\|\sum_{i\in \mathrm{J}\setminus \mathrm{K}_n} x_i\right\| \leq 2^{-n+1}$$

vu que $J \setminus K_n$ est disjoint de J_n. On a montré que la famille $(x_i)_{i\in I}$ est sommable de somme S.

//

Remarque 5.4.2. Soient X un espace normé, $(x_i)_{i\in I}$ une famille d'éléments de X, $\varepsilon > 0$ et J une partie finie de I telle que, pour toute partie finie L de I disjointe de J, on ait $\left\|\sum_{i\in L} x_i\right\| < \varepsilon$. En considérant des parties L à un seul élément, on voit que, pour tout $i \in I\setminus J$, $\|x_i\| < \varepsilon$. Donc, si $(x_i)_{i\in I}$ vérifie le critère de sommabilité de Cauchy, pour tout entier $n \geq 1$, il n'y a qu'un nombre fini d'indices i tels que x_i soit de norme $\geq 1/n$; donc il y a un nombre dénombrable de x_i non nuls. En d'autres termes, on peut toujours se ramener au cas $I = \mathbb{N}$.

Soit $(x_n)_{n\geq 0}$ une famille sommable, et notons S sa somme ; soient $\varepsilon > 0$ et $J \subset \mathbb{N}$ une partie finie telle que, pour toute partie finie K de $\mathbb{N}$ contenant J on ait $\left\|S - \sum_{i\in K} x_i\right\| < \varepsilon$; soit $m \in \mathbb{N}$ un majorant de J ; pour tout $n \geq m$, comme $J \subset \{0, 1, \dots, n\}$ on a $\left\|S - \sum_{i=0}^{n} x_i\right\| < \varepsilon$. Donc la série de terme général (x_n) est convergente et sa somme est égale à S. La notation $\sum_{n\in\mathbb{N}} x_n$ n'est donc pas en conflit avec la notation des séries.

Corollaire 5.4.4. *Soient* E *un espace de Banach et* $(x_i)_{i\in I}$ *une famille d'éléments de* E *;*

(i) soit $(t_i)_{i\in I}$ une famille de nombres réels ; si pour tout $i \in I$, on a $\|x_i\| \leq t_i$ et si la famille $(t_i)_{i\in I}$ est sommable, alors la famille $(x_i)_{i\in I}$ est sommable. En particuler, si la famille $(\|x_i\|)_{i\in I}$ est sommable, alors la famille $(x_i)_{i\in I}$ est sommable ;

(ii) si la famille $(x_i)_{i\in I}$ est sommable, pour toute partie $J \subset I$, la famille $(x_i)_{i\in J}$ est sommable.

Démonstration. Si la famille de réels $(t_i)_{i\in I}$ vérifie le critère de sommabilité de Cauchy, il en va clairement de même pour la famille de vecteurs $(x_i)_{i\in I}$ (inégalité triangulaire), et ceci montre le point (i). Si $(x_i)_{i\in I}$ vérifie le critère de sommabilité de Cauchy, il en va clairement de même pour la sous-famille $(x_i)_{i\in J}$.

//

Remarque 5.4.3. On peut plus généralement définir les familles sommables dans un espace vectoriel topologique séparé X : une famille $(x_i)_{i\in I}$ est dite sommable de somme $S \in X$, si, pour tout voisinage V de S dans X, il existe une partie finie J de I telle que, pour toute partie finie K de I contenant J on ait $\sum_{i\in K} x_i \in V$. Dans ce contexte vectoriel topologique, il peut arriver qu'une famille sommable ait une infinité non dénombrable d'éléments non nuls.

Dans un espace de Hilbert, on dispose d'un outil très simple pour tester la sommabilité d'une famille de vecteurs deux à deux orthogonaux, appelée aussi *système orthogonal.*

Lemme 5.4.5. *Soit $(x_i)_{i\in I}$ un système orthogonal dans un espace de Hilbert ; la famille (x_i) est sommable si et seulement si la famille $(\|x_i\|^2)$ est sommable ; dans ce cas, on a*

$$\Big\|\sum_{i\in I} x_i\Big\|^2 = \sum_{i\in I} \|x_i\|^2.$$

Démonstration. Pour toute partie finie J de I, on a $\|\sum_{i\in J} x_i\|^2 = \sum_{i\in J} \|x_i\|^2$. On en déduit que la famille (x_i) vérifie le critère de Cauchy de sommabilité si et seulement si la famille $(\|x_i\|^2)$ vérifie le critère de Cauchy de sommabilité. Dans ce cas, il existe une suite croissante J_n de parties finies de I telles que $\mathrm{S} = \sum_{i\in \mathrm{I}} x_i$ soit la limite de $\mathrm{S}_n = \sum_{i\in \mathrm{J}_n} x_i$ et $\sum_{i\in \mathrm{I}} \|x_i\|^2$ soit la limite de $\sum_{i\in \mathrm{J}_n} \|x_i\|^2$. Mais alors

$$\|\mathrm{S}\|^2 = \lim \|\mathrm{S}_n\|^2 = \sum_{i\in \mathrm{I}} \|x_i\|^2.$$

//

5.5. Bases hilbertiennes

Définition 5.5.1. Soient E un espace de Hilbert et $(x_i)_{i\in \mathrm{I}}$ un système de vecteurs de E ; on dit que le système $(x_i)_{i\in \mathrm{I}}$ est *orthogonal* si les x_i sont deux à deux orthogonaux ; on dit que c'est un *système orthonormal* si de plus, pour tout $i \in \mathrm{I}$, on a $\|x_i\| = 1$; on appelle *base hilbertienne* de E un système orthonormal total dans E.

Un sous-ensemble B de E définit un système $(b)_{b\in \mathrm{B}}$. On dira que le sous-ensemble B est orthogonal, orthonormal, ou que c'est une base hilbertienne si le système $(b)_{b\in \mathrm{B}}$ est orthogonal, orthonormal, ou est une base hilbertienne. Ce procédé *d'auto-indexation* simplifie l'écriture de la démonstration qui suit.

Théorème 5.5.1. *Tout espace de Hilbert admet une base hilbertienne.*

Démonstration. Soit H un espace de Hilbert ; notons $\mathrm{U} \subset \mathcal{P}(\mathrm{H})$ l'ensemble des parties orthonormales. Montrons que, muni de l'inclusion, U est inductif. Soit $\{\mathrm{B}_i : i \in \mathrm{I}\}$ une partie totalement ordonnée de U ; si $x, y \in \bigcup_{i\in \mathrm{I}} \mathrm{B}_i$, il existe $i \in \mathrm{I}$ tel que $x, y \in \mathrm{B}_i$ donc $\langle x, x\rangle = 1$ et si $x \neq y$ alors $\langle x, y\rangle = 0$; il s'ensuit que $\bigcup_{i\in \mathrm{I}} \mathrm{B}_i$ est un élément de U majorant $\{\mathrm{B}_i : i \in \mathrm{I}\}$.

Soit B un élément maximal de U ; on veut montrer que B est total, et pour cela, on montre que $\mathrm{B}^\perp = \{0\}$; sinon, il existerait un vecteur x non nul et orthogonal à B (en particulier $x \notin \mathrm{B}$), et quitte à multiplier x par un scalaire convenable on peut supposer $\|x\| = 1$; alors $\mathrm{B} \cup \{x\} \in \mathrm{U}$, ce qui contredirait la maximalité de B. Donc $\mathrm{B}^\perp = \{0\}$, ce qui entraîne que $(\mathrm{B}^\perp)^\perp = \mathrm{H}$. Par la proposition 2.3, B est total. C'est donc une base hilbertienne.

//

Remarque 5.5.2. Dans le cas séparable, on a donné une démonstration plus concrète de ce théorème dans le chapitre 2.

Théorème 5.5.2 : inégalité de Bessel. *Soient* E *un espace de Hilbert et* $(e_i)_{i\in \mathrm{I}}$ *un système orthonormal dans* E *; pour tout* $x \in \mathrm{E}$ *la famille* $(|\langle x, e_i\rangle|^2)_{i\in \mathrm{I}}$ *est sommable et*

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

Démonstration. Par la proposition 4.1, il suffit de montrer que, pour toute partie finie J de I, on a $\sum_{i\in \mathrm{J}} |\langle x, e_i\rangle|^2 \leq \langle x, x\rangle$. Ce résultat a été vu au lemme 2.2.3.

//

Proposition 5.5.3. *Le cardinal de tout système orthonormal d'un espace de Hilbert est inférieur ou égal à celui de tout système total.*

Démonstration. Soient $(e_i)_{i\in \mathrm{I}}$ un système orthonormal et $(x_j)_{j\in \mathrm{J}}$ un système total dans E hilbertien ; si J est fini, l'espace E est de dimension finie puisque $(x_j)_{j\in \mathrm{J}}$ est un système générateur fini pour E. Comme le système $(e_i)_{i\in \mathrm{I}}$ est libre, on trouve que le cardinal de I est inférieur ou égal à celui de J.

Supposons J infini. Posons $\mathrm{X} = \{(i,j) \in \mathrm{I}\times \mathrm{J} : \langle e_i, x_j\rangle \neq 0\}$. Pour $i \in \mathrm{I}$, comme $(x_j)_{j\in \mathrm{J}}$ est total, e_i n'est pas orthogonal à tous les x_j. Donc l'application $(i,j) \to i$ est surjective de X sur I. Donc le cardinal de I est majoré par le cardinal de X. Pour $j \in \mathrm{J}$, la famille $(|\langle e_i, x_j\rangle|^2)_{i\in \mathrm{I}}$ est sommable (par le théorème 2), donc $\{i : (i,j) \in \mathrm{X}\}$ est dénombrable. On en déduit qu'il existe une injection de X dans $\mathrm{J}\times \mathbb{N}$. Donc le cardinal de X est inférieur ou égal à celui de J.

//

Corollaire 5.5.4. *Deux bases hilbertiennes d'un espace de Hilbert* E *ont le même cardinal.*

Définition 5.5.3. On appelle *dimension hilbertienne* d'un espace de Hilbert E le cardinal d'une base hilbertienne quelconque de E.

Théorème 5.5.5 : identité de Parseval. *Soient* E *un espace de Hilbert,* $(e_i)_{i\in \mathrm{I}}$ *une base hilbertienne de* E *et* $x \in \mathrm{E}$ *; la famille de nombres réels* $(|\langle x, e_i\rangle|^2)_{i\in \mathrm{I}}$ *est sommable, la famille de vecteurs* $(\langle x, e_i\rangle\, e_i)_{i\in \mathrm{I}}$ *est sommable dans* E *et*

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

Démonstration. Comme $(e_i)_{i\in \mathrm{I}}$ est un système orthonormal, il résulte du théorème 2 que la famille de réels $(|\langle x, e_i\rangle|^2)_{i\in \mathrm{I}}$ est sommable. Par le lemme 4.5, la famille $(\langle x, e_i\rangle\, e_i)_{i\in \mathrm{I}}$ est sommable dans E et, si on note y sa somme, on a $\sum_{i\in \mathrm{I}} |\langle x, e_i\rangle|^2 = \langle y, y\rangle$. Pour tout $j \in \mathrm{I}$, appliquant la proposition 4.2 à la forme linéaire $z \to \langle z, e_j\rangle$, on trouve $\langle y, e_j\rangle = \sum_{i\in \mathrm{I}} \langle x, e_i\rangle\langle e_i, e_j\rangle = \langle x, e_j\rangle$. Donc $x - y$ est orthogonal aux e_i, donc à l'espace vectoriel engendré par les (e_i) ; comme le système e_i est total, $x = y$.

//

Corollaire 5.5.6. *Soient* H *un espace de Hilbert,* F *un sous-espace vectoriel fermé de* H*, et* $(e_i)_{i\in \mathrm{I}}$ *une base hibertienne du sous-espace* F *; pour tout vecteur* $x \in \mathrm{H}$*, la projection orthogonale de* x *sur* F *est donnée par*

$$\mathrm{P}_{\mathrm{F}}(x) = \sum_{i\in \mathrm{I}} \langle x, e_i\rangle\, e_i.$$

Démonstration. Posons $y = \mathrm{P}_{\mathrm{F}}(x)$; puisque $x-y$ est orthogonal à F, on a $\langle x-y, e_i\rangle = 0$ pour tout $i \in \mathrm{I}$, donc $\langle x, e_i\rangle = \langle y, e_i\rangle$. D'après le théorème précédent, appliqué à F et à y, on a

$$y = \sum_{i\in \mathrm{I}} \langle y, e_i\rangle\, e_i = \sum_{i\in \mathrm{I}} \langle x, e_i\rangle\, e_i.$$

//

Remarque 5.5.4. Soient E un espace de Hilbert et $(e_i)_{i\in \mathrm{I}}$ un système orthonormal de vecteurs de E ; on a l'égalité $\|x\|^2 = \sum_{i\in\mathrm{I}} |\langle x, e_i\rangle|^2$ si et seulement si

$$x = \sum_{i\in\mathrm{I}} \langle x, e_i\rangle\, e_i.$$

Soit F le plus petit sous-espace vectoriel fermé de E contenant les (e_i). Soit $x \in \mathrm{E}$; écrivons $x = y+z$ où $y = \mathrm{P_F}(x) \in \mathrm{F}$ et $z \in \mathrm{F}^\perp$. Comme (e_i) est une base hilbertienne de F, on a $y = \sum_{i\in\mathrm{I}} \langle x, e_i\rangle\, e_i$ (corollaire précédent), donc $\|y\|^2 = \sum_{i\in\mathrm{I}} |\langle x, e_i\rangle|^2$. De plus $\|x\|^2 = \|y\|^2 + \|z\|^2$; donc

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

En conclusion, $\|x\|^2 = \sum_{i\in\mathrm{I}} |\langle x, e_i\rangle|^2$ si et seulement si $z = 0$, c'est à dire si et seulement si $x = \sum_{i\in\mathrm{I}} \langle x, e_i\rangle\, e_i$.

5.6. L'espace hilbertien $\ell^2(\mathrm{I})$

Soit I un ensemble ; notons $\ell^2(\mathrm{I})$ l'ensemble des familles de scalaires $(x_i)_{i\in\mathrm{I}}$ telles que la famille de nombres réels positifs $|x_i|^2$ soit sommable. Si $\eta = (y_i)_{i\in\mathrm{I}}$ est un autre élément de $\ell^2(\mathrm{I})$, la relation $|x_i + y_i|^2 \le 2\big(|x_i|^2 + |y_i|^2\big)$ montre que $\xi + \eta$ est encore dans $\ell_2(\mathrm{I})$, et on en déduit facilement que $\ell_2(\mathrm{I})$ est un espace vectoriel. Pour tout $\xi = (x_i)_{i\in\mathrm{I}} \in \ell^2(\mathrm{I})$ on pose

$$\|\xi\|_2 = \Big(\sum_{i\in\mathrm{I}} |x_i|^2\Big)^{1/2}.$$

On voit que cette quantité définit une norme sur l'espace vectoriel $\ell_2(\mathrm{I})$; en fait la relation $2|x_i\overline{y_i}| \le |x_i|^2 + |y_i|^2$ montre que la famille $(x_i\overline{y_i})_{i\in\mathrm{I}}$ est sommable, et si on pose

$$\langle \xi, \eta\rangle = \sum_{i\in\mathrm{I}} x_i\overline{y}_i$$

on définit sur $\ell_2(\mathrm{I})$ un produit scalaire pour lequel $\langle \xi, \xi\rangle = \|\xi\|^2$.

Pour $j \in \mathrm{I}$, notons $\epsilon_j \in \ell^2(\mathrm{I})$ la famille $(x_i)_{i\in\mathrm{I}}$ telle que $x_j = 1$ et $x_i = 0$ si $i \in \mathrm{I}\setminus\{j\}$.

Proposition 5.6.1. *Muni du produit scalaire précédent, l'espace vectoriel $\ell^2(\mathrm{I})$ est un espace de Hilbert. La famille $(\epsilon_i)_{i\in\mathrm{I}}$ est une base hilbertienne de $\ell^2(\mathrm{I})$.*

Démonstration. Il est clair que $(\epsilon_i)_{i\in\mathrm{I}}$ est un système orthonormal. Pour $\xi = (x_i)_{i\in\mathrm{I}}$ et $i \in \mathrm{I}$ on a $\langle \xi, \epsilon_i\rangle = x_i$, donc $\|\xi\|_2^2 = \sum_i |\langle \xi, \epsilon_i\rangle|^2$. Par la remarque 5.4, $(\epsilon_i)_{i\in\mathrm{I}}$ est une base hilbertienne de $\ell^2(\mathrm{I})$.

Montrons enfin que $\ell^2(\mathrm{I})$ est complet. Notons $u : \ell^2(\mathrm{I}) \to \mathrm{H}$ l'application isométrique de $\ell^2(\mathrm{I})$ dans son complété. Il suffit de montrer que u est surjective. Soit $x \in \mathrm{H}$; posons $\xi = (\langle x, u(\epsilon_i)\rangle)_{i\in\mathrm{I}}$; par le théorème 5.2, on sait que $\xi \in \ell^2(\mathrm{I})$. Pour tout $i \in \mathrm{I}$, on a $\langle \xi, \epsilon_i\rangle = \langle x, u(\epsilon_i)\rangle$, donc $x - u(\xi)$ est orthogonal à $u(\epsilon_i)$. Or $(\epsilon_i)_{i\in\mathrm{I}}$ est total dans $\ell^2(\mathrm{I})$; comme l'image de u est dense dans H, $(u(\epsilon_i))_{i\in\mathrm{I}}$ est total dans H ; on en déduit que $x = u(\xi)$. Il s'ensuit que u est isométrique et bijective. Alors $\ell^2(\mathrm{I})$ est isométrique à H, donc est complet.

//

Théorème 5.6.2. *Soient* H *un espace de Hilbert et* $\mathrm{B} = (e_i)_{i\in\mathrm{I}}$ *une base hilbertienne de* H *; l'application* $\mathrm{U} : x \to (\langle x, e_i\rangle)$ *est une bijection linéaire isométrique de* H *sur* $\ell^2(\mathrm{I})$.

Démonstration. Il est clair que U est une application linéaire de H dans $\mathbb{K}^{\mathrm{I}}$. Par le théorème 5.5, U est isométrique de H dans $\ell^2(\mathrm{I})$. Soit $(\lambda_i)_{i\in\mathrm{I}} \in \ell^2(\mathrm{I})$; par le lemme 4.5, la famille $(\lambda_i e_i)_{i\in\mathrm{I}}$ est sommable dans H ; posons $x = \sum_i \lambda_i e_i$. Pour tout $j \in \mathrm{I}$, appliquant la proposition 4.2 à la forme linéaire $z \to \langle z, e_j\rangle$, on trouve $\langle x, e_j\rangle = \sum_{i\in\mathrm{I}} \lambda_i\langle e_i, e_j\rangle = \lambda_j$. Donc $\mathrm{U}(x) = (\lambda_i)_{i\in\mathrm{I}}$, donc U est surjective.

//

Il est clair que, pour tout $i \in \mathrm{I}$, on a $\mathrm{U}(e_i) = \epsilon_i$, où $(\epsilon_i)_{i\in\mathrm{I}}$ désigne la base hilbertienne canonique de $\ell^2(\mathrm{I})$.

Corollaire 5.6.3. *Soient* E *et* F *deux espaces de Hilbert,* $(e_i)_{i\in\mathrm{I}}$ *une base hilbertienne de* E, $(f_j)_{j\in\mathrm{J}}$ *une base hilbertienne de* F *et* σ *une bijection de* I *sur* J *; il existe une bijection linéaire isométrique* U *de* E *sur* F *telle que, pour tout* $i \in \mathrm{I}$, *on ait* $\mathrm{U}(e_i) = f_{\sigma(i)}$.

Démonstration. Remarquons que $(f_{\sigma(i)})_{i\in\mathrm{I}}$ est une base hilbertienne de F. Notons $u : \mathrm{E} \to \ell^2(\mathrm{I})$ l'application $x \to (\langle x, e_i\rangle)_{i\in\mathrm{I}}$ et $v : \mathrm{F} \to \ell^2(\mathrm{I})$ l'application $y \to (\langle y, f_{\sigma(i)}\rangle)_{i\in\mathrm{I}}$. Ce sont des bijections isométriques par le théorème 2. La bijection isométrique $\mathrm{U} = v^{-1} \circ u$ convient.

//

En d'autres termes, deux espaces hilbertiens ayant même dimension hilbertienne sont isomorphes.

6. Théorie spectrale

Un certain nombre de résultats de ce chapitre et des suivants n'ont de sens que pour les espaces de Banach complexes, mais quelques énoncés seront valables aussi dans le cas réel. Quand nous dirons simplement “espace de Banach” ou “algèbre de Banach” cela signifiera que le résultat est valable dans le cas réel ou complexe.

6.1. Algèbres de Banach, spectre et résolvante

Une *algèbre de Banach unitaire* est un espace de Banach A muni d'un produit $(a, b) \in \mathrm{A} \times \mathrm{A} \to ab \in \mathrm{A}$, bilinéaire et associatif, tel qu'il existe un élément neutre 1_{A} pour la multiplication ($1_{\mathrm{A}} a = a 1_{\mathrm{A}} = a$ pour tout $a \in \mathrm{A}$) et que de plus

$$\|1_{\mathrm{A}}\| = 1\,; \quad \|ab\| \le \|a\|\,\|b\|$$

pour tous $a, b \in \mathrm{A}$. L'application $(a, b) \to ab$ est donc continue de $\mathrm{A} \times \mathrm{A}$ dans A, et il en résulte que les applications $b \to ab$ et $b \to ba$ sont continues de A dans A.

On remarquera que notre définition exclut $\mathrm{A} = \{0\}$, puisqu'on ne pourrait pas y trouver un élément 1_{A} de norme 1 !

Pour $a \in \mathrm{A}$ et n entier ≥ 0, on définit a^n par récurrence en posant $a^0 = 1_{\mathrm{A}}$ et $a^{n+1} = aa^n = a^n a$ pour tout entier $n \ge 0$.

Exemples 6.1.1.

1. L'exemple de loin le plus important sera $\mathrm{A} = \mathcal{L}(\mathrm{E})$, où E est un espace de Banach ; si $\mathrm{E} \ne \{0\}$, il s'agit bien d'une algèbre de Banach unitaire. Le produit est la composition des applications linéaires, la norme de A est la norme d'application linéaire et $1_{\mathrm{A}} = \mathrm{Id}_{\mathrm{E}}$ est l'élément neutre du produit ; il est de norme 1 quand $\mathrm{E} \ne \{0\}$.

Pour $\mathrm{T} \in \mathcal{L}(\mathrm{E}, \mathrm{F})$ on notera im(T) le sous-espace de F *image de l'application* T, noté aussi T(E),

$$\mathrm{im(T)} = \mathrm{T(E)} = \{y \in \mathrm{F} : \exists x \in \mathrm{E},\ y = \mathrm{T}(x)\}.$$

Si E, F, G sont des espaces normés, $\mathrm{S} \in \mathcal{L}(\mathrm{E}, \mathrm{F})$ et $\mathrm{T} \in \mathcal{L}(\mathrm{F}, \mathrm{G})$, nous noterons TS la composée $\mathrm{T} \circ \mathrm{S}$ de ces applications.

2. Soit K un espace compact métrique non vide ; considérons l'espace de Banach $\mathrm{A} = \mathrm{C(K)}$ des fonctions continues sur K à valeurs complexes, muni du produit usuel et de la norme de convergence uniforme (exemples 1.1.5) ; c'est une algèbre de Banach unitaire. L'élément 1_{A} est la fonction constante égale à 1. Cet exemple donne une algèbre commutative.

3. L'espace $\mathrm{L}_\infty([0, 1], dx)$ donne un autre exemple d'algèbre de Banach commutative. Si $(\Omega, \mathcal{A})$ est un espace mesurable, l'espace $\mathcal{L}_\infty(\Omega, \mathcal{A})$ des fonctions $\mathcal{A}$-mesurables bornées est une algèbre de Banach pour la norme du sup.

Définition 6.1.2. Soient A une algèbre de Banach unitaire, et $a \in \mathrm{A}$; on dit que a est *inversible dans* A s'il existe $b \in \mathrm{A}$ tel que $ab = ba = 1_{\mathrm{A}}$.

Exemples 6.1.3.

1. Soit E un espace de Banach et considérons $A = \mathcal{L}(E)$; une application linéaire continue $T \in A$ est inversible dans A s'il existe $S \in \mathcal{L}(E)$ telle que $ST = \mathrm{Id}_E = 1_A$ et $TS = \mathrm{Id}_E$. Cela signifie que l'application T est bijective et que T^{-1} est continue, et correspond bien à la définition usuelle de l'inversibilité d'une application linéaire continue.

Remarquons que par le théorème des isomorphismes, si T est bijective et continue, T^{-1} est automatiquement continue. En d'autres termes, T est inversible si et seulement si elle est bijective et continue.

2. Soit $f \in A = C(K)$; si f est inversible il existe une fonction continue g telle que $f(s)g(s) = 1$ pour tout $s \in K$ donc $f(s) \neq 0$ pour tout $s \in K$. Inversement, si f ne s'annule pas sur K, la fonction $s \to 1/f(s)$ est définie et continue sur K, et elle est l'inverse de f dans $A = C(K)$. On voit donc que f est inversible dans C(K) si et seulement si elle ne s'annule pas sur K.

Exercice. Exprimer l'inversibilité dans L_∞ de $f \in L_\infty$.

Lemme 6.1.1. *Soient* A *une algèbre de Banach unitaire et* $a \in A$ *tel que* $\|a\| < 1$ *; alors, la série* $\sum_k a^k$ *est convergente dans* A *et sa somme est l'inverse de* $1_A - a$*,*

$$(1_A - a)^{-1} = \sum_{k=0}^{+\infty} a^k.$$

On a de plus l'estimation

$$\|(1_A - a)^{-1}\| \le \frac{1}{1 - \|a\|}.$$

Démonstration. Comme $\|a^k\| \le \|a\|^k$ pour tout entier $k \ge 0$ et que $\|a\| < 1$, la série $\sum_{k=0}^{+\infty} a^k$ est normalement convergente, donc convergente dans l'espace complet A. Notons S sa somme. On vérifie facilement que

$$S\,a = a\,S = \sum_{k=0}^{+\infty} a^{k+1} = S - 1_A$$

ce qui implique que $S(1_A - a) = (1_A - a)S = 1_A$. En majorant la norme de la série par la série des normes, on obtient $\|(1_A - a)^{-1}\| \le \sum_{k=0}^{+\infty} \|a\|^k = (1 - \|a\|)^{-1}$.

//

Remarque 6.1.4. Soient encore A une algèbre de Banach et $a \in A$ tel que $\|a\| < 1$; en enlevant le premier terme de la série géométrique $\sum a^k$, on a obtenu ci-dessus l'égalité $(1_A - a)^{-1} - 1_A = a\,(1_A - a)^{-1}$; de même, en enlevant les deux premiers termes on obtient la relation $(1_A - a)^{-1} - 1_A - a = a^2\,(1_A - a)^{-1}$. On en déduit les inégalités :

$$\|(1_A - a)^{-1} - 1_A\| \le \|a\|\,(1 - \|a\|)^{-1}\,; \quad \|(1_A - a)^{-1} - 1_A - a\| \le \|a\|^2(1 - \|a\|)^{-1}.$$

Proposition 6.1.2. *Soit* A *une algèbre de Banach ; l'ensemble des éléments inversibles dans* A *est un ouvert* U *(non vide) de* A*. L'application* $\varphi : a \to a^{-1}$ *est continue et différentiable de* U *dans* A *; sa différentielle en* $a \in \mathrm{U}$ *est* $(d\varphi)_a : b \to -a^{-1} b\, a^{-1}$.

Démonstration. Si $a \in \mathrm{A}$ est inversible et si $b \in \mathrm{A}$ est tel que $\|b\| < \|a^{-1}\|^{-1}$, on aura $a + b = a(1_\mathrm{A} + a^{-1}b)$, et si on pose $u_b = -a^{-1}b$ on aura $\|u_b\| = \|a^{-1}b\| \le \|a^{-1}\|\,\|b\| < 1$, ce qui implique que $1_\mathrm{A} - u_b = 1_\mathrm{A} + a^{-1}b$ est inversible dans A, donc $a + b$ aussi, et $(a+b)^{-1} = (1_\mathrm{A} + a^{-1}b)^{-1}a^{-1}$.

En utilisant le développement en série obtenu au lemme 1, on obtient que lorsque $\|b\| < \|a^{-1}\|^{-1}$, on a $(a+b)^{-1} = \bigl(\sum_{k=0}^{+\infty} u_b^k\bigr)\, a^{-1}$, ce qui peut s'écrire

$$(a+b)^{-1} = a^{-1} - a^{-1}b\, a^{-1} + a^{-1}ba^{-1}b\, a^{-1} - \cdots$$

Considérons que a est fixé, b variable et petit, et gardons en évidence les deux premiers termes du développement, sous la forme

$$(*) \qquad (a+b)^{-1} = a^{-1} - a^{-1}b\, a^{-1} + \mathrm{V}(b)$$

où $\mathrm{V}(b) = \bigl(\sum_{k=2}^{+\infty} u_b^k\bigr)\, a^{-1}$. Comme dans la remarque 4, on obtient la majoration de norme $\|\mathrm{V}(b)\| \le \|u_b\|^2\,(1 - \|u_b\|)^{-1}\,\|a^{-1}\|$, qui montre que $\|\mathrm{V}(b)\| = \mathrm{O}(\|b\|^2)$ lorsque $b \to 0_\mathrm{A}$. Puisque $\psi : b \to -a^{-1}b\, a^{-1}$ est une application linéaire continue de A dans elle-même, la relation $(*)$ montre que l'application $u \in \mathrm{U} \to u^{-1}$ est différentiable au point a (donc continue au point a) et que sa différentielle au point a est ψ.

//

Remarque 6.1.5. Dans le cas complexe, la différentiabilité d'une fonction f au point a signifie que df_a est $\mathbb{C}$-linéaire ; cette différence anodine a en fait des conséquences considérables (penser aux fonctions holomorphes, qui ne sont rien d'autre que des fonctions $\mathbb{C}$-différentiables).

Dans le cas d'applications linéaires, on peut étendre légèrement le résultat au cas d'applications entre deux espaces de Banach distincts. La démonstration est identique.

Corollaire 6.1.3. *Soient* E, F *deux espaces de Banach ; l'ensemble* $\mathrm{U} \subset \mathcal{L}(\mathrm{E},\mathrm{F})$ *des applications linéaires continues inversibles est ouvert dans l'espace* $\mathcal{L}(\mathrm{E},\mathrm{F})$*. L'application* $\varphi : \mathrm{A} \to \mathrm{A}^{-1}$ *est continue et différentiable de* U *dans* $\mathcal{L}(\mathrm{F},\mathrm{E})$ *; sa différentielle en* $\mathrm{T} \in \mathrm{U}$ *est* $(d\varphi)_\mathrm{T} : \mathrm{S} \to -\mathrm{T}^{-1}\mathrm{S}\mathrm{T}^{-1}$.

Ce qui a été dit jusqu'ici est valable aussi bien dans le cas réel que complexe. En revanche, la théorie du spectre n'est vraiment satisfaisante que dans le cas $\mathbb{K} = \mathbb{C}$. Nous prendrons donc des algèbres de Banach sur $\mathbb{C}$.

Définition 6.1.6. Soient A une algèbre de Banach unitaire complexe et $a \in \mathrm{A}$; on appelle *spectre* de a et l'on note $\mathrm{Sp}(a)$ l'ensemble des $\lambda \in \mathbb{C}$ tels que $a - \lambda 1_\mathrm{A}$ ne soit pas inversible. On appelle *résolvante* de a l'application qui à $\lambda \in \mathbb{C} \setminus \mathrm{Sp}(a)$ associe l'inverse $(a - \lambda 1_\mathrm{A})^{-1}$. On notera si $\lambda \notin \mathrm{Sp}(\mathrm{T})$

$$\mathrm{R}_\lambda(a) = (a - \lambda 1_\mathrm{A})^{-1}.$$

Si $|\lambda| > \|a\|$, on peut écrire $a - \lambda 1_{\mathrm{A}} = -\lambda(1_{\mathrm{A}} - a/\lambda)$, et $\|a/\lambda\| < 1$, ce qui montre que $a - \lambda 1_{\mathrm{A}}$ est inversible dans ce cas. On voit donc que $\mathrm{Sp}(a)$ est contenu dans le disque fermé du plan complexe centré en 0 et de rayon $\|a\|$. De plus, d'après le lemme 1

$$\text{(R)} \qquad \text{si } \ |\lambda| > \|a\|, \quad \|\mathrm{R}_\lambda(a)\| \le \frac{1}{|\lambda| - \|a\|}.$$

Exemples 6.1.7.

a. Munissons $\mathbb{C}^n$ d'une norme (complexe) quelconque et considérons $\mathrm{M}_n(\mathbb{C})$ comme l'algèbre de Banach $\mathrm{A} = \mathcal{L}(\mathbb{C}^n)$; le spectre d'une matrice $\mathrm{M} \in \mathrm{A}$ est l'ensemble des valeurs propres de la matrice.

b. Soit K un espace compact métrique non vide ; considérons l'espace de Banach $\mathrm{A} = \mathrm{C}(\mathrm{K})$; on a vu que $f - \lambda$ est inversible si et seulement si $f - \lambda$ ne s'annule pas, donc si et seulement si $\lambda \notin f(\mathrm{K})$; on a donc $\mathrm{Sp}(f) = f(\mathrm{K})$.

On dit qu'une application f d'un ouvert U de $\mathbb{C}$ dans un espace de Banach complexe F est $\mathbb{C}$-*dérivable* au point λ (ou bien *dérivable au sens complexe*), de dérivée $f'(\lambda)$, si

$$f'(\lambda) = \lim_{z \in \mathbb{C}, z \to 0} \frac{1}{z}\big(f(\lambda + z) - f(\lambda)\big).$$

On dit qu'une application f d'un ouvert U de $\mathbb{C}$ dans un espace de Banach complexe F, dérivable au sens complexe en tout point de l'ouvert U, est une *fonction holomorphe* de U dans F.

Théorème 6.1.4. *Soient* A *une algèbre de Banach unitaire complexe et* $a \in \mathrm{A}$ *; le spectre de* a *est une partie compacte non vide de* $\mathbb{C}$*, l'application* $\mathrm{R}(a) : \lambda \to (a - \lambda 1_{\mathrm{A}})^{-1} = \mathrm{R}_\lambda(a)$ *est holomorphe sur* $\mathbb{C} \setminus \mathrm{Sp}(a)$*, avec*

$$\mathrm{R}(a)'(\lambda) = (\mathrm{R}_\lambda(a))^2.$$

(et la dérivée est donc continue).

Démonstration. Si λ n'est pas dans $\mathrm{Sp}(a)$, l'élément $a - \lambda 1_{\mathrm{A}}$ est inversible ; d'après la proposition 3, $a - \lambda' 1_{\mathrm{A}}$ sera encore inversible pour tout λ' dans un voisinage de λ, ce qui montre que le complémentaire du spectre est ouvert dans $\mathbb{C}$, donc $\mathrm{Sp}(a)$ est fermé dans $\mathbb{C}$; on a vu ci-dessus que $\mathrm{Sp}(a)$ est contenu dans le disque de rayon $\|a\|$, donc le spectre est borné.

Soit $\lambda \in \mathbb{C} \setminus \mathrm{Sp}(a)$; posons $b = a - \lambda 1_{\mathrm{A}}$; alors b est inversible, $b^{-1} = \mathrm{R}_\lambda(a)$ et on sait que pour z assez petit, $a - (\lambda + z)1_{\mathrm{A}} = b - z 1_{\mathrm{A}}$ est inversible et

$$\mathrm{R}_{\lambda+z}(a) = b^{-1} + z b^{-2} + z^2 b^{-3} + z^3 b^{-4} + \cdots$$

En écrivant comme précédemment $\mathrm{R}_{\lambda+z}(a) = \mathrm{R}_\lambda(a) + z(\mathrm{R}_\lambda(a))^2 + \mathrm{V}(z)$ on montre que $\|\mathrm{V}(z)\| = \mathrm{O}(|z|^2)$ lorsque $z \to 0$, ce qui entraîne que l'application résolvante est dérivable (complexe) au point λ, avec $(\mathrm{R}_\lambda(a))^2$ pour dérivée en ce point.

Il reste à montrer que $\mathrm{Sp}(a) \ne \emptyset$. Choisissons λ_0 hors du spectre ; alors $\mathrm{R}_{\lambda_0}(a)$ est non nul puisqu'inversible et on peut trouver une forme linéaire a^* continue sur A telle que $a^*(\mathrm{R}_{\lambda_0}(a)) \ne 0$ (corollaire 3.2.5) ; l'application $g : \lambda \to a^*(\mathrm{R}_\lambda(a))$ est une fonction holomorphe scalaire définie sur $\mathbb{C} \setminus \mathrm{Sp}(a)$, telle que $g(\lambda_0) \ne 0$. Si $\mathrm{Sp}(a)$ était vide, cette fonction serait entière (holomorphe sur $\mathbb{C}$ tout entier) ; or d'après la relation (R) on voit que $g(\lambda) = a^*(\mathrm{R}_\lambda(a))$ tend vers 0 quand $\lambda \to \infty$. Par le théorème de Liouville on aurait $g(\lambda) = 0$ pour tout $\lambda \in \mathbb{C}$, ce qui n'est pas vrai, donc $\mathrm{Sp}(a) \ne \emptyset$.

//

Exemples 6.1.8.

1. Soit K un espace compact métrique ; considérons l'espace de Banach $E = C(K)$ des fonctions continues sur K à valeurs complexes, muni de la norme de convergence uniforme (exemples 1.1.5). Soit $f \in E$; l'application $M_f : g \to fg$ est linéaire de E dans E et continue puisque pour tout $g \in E$, on a $\|fg\|_\infty \le \|f\|_\infty \|g\|_\infty$. De plus la relation $\|M_f(f)\|_\infty = \|f\|_\infty^2$ implique $\|M_f\| \ge \|f\|_\infty$, donc $\|M_f\| = \|f\|_\infty$. Soit $\lambda \in \mathbb{C}$;
– si pour tout $s \in K$, on a $f(s) \neq \lambda$, alors la fonction $h : s \to (f(s) - \lambda)^{-1}$ est continue de K dans $\mathbb{C}$. On voit alors que $M_f - \lambda \,\mathrm{Id}_E$ est inversible et que son inverse $R_\lambda(M_f)$ est l'application $M_h : g \to hg$;
– s'il existe $s \in K$ tel que $f(s) = \lambda$, alors pour tout $g \in E$, la fonction $fg - \lambda g$ s'annule au point s, donc $\mathrm{im}(M_f - \lambda \,\mathrm{Id}_E) \subset \{g \in E : g(s) = 0\}$, qui est un sous-espace fermé de E, distinct de E. On en déduit que l'image de $M_f - \lambda \,\mathrm{Id}_E$ n'est pas dense, donc $M_f - \lambda \,\mathrm{Id}_E$ n'est pas inversible puisqu'il n'est pas surjectif. En résumé, le spectre de M_f est l'ensemble $\mathrm{Sp}(M_f) = f(K) = \{f(s) : s \in K\}$ des valeurs de f. C'est aussi le spectre de f dans l'algèbre C(K). On verra plus loin que ça n'est pas un hasard !

2. Soit p un nombre réel tel que $1 \le p < +\infty$, et soit $S \in \mathcal{L}(\ell_p)$ l'application qui à une suite $(x_n)_{n\ge0}$ associe la suite $(y_n)_{n\ge0}$ définie par $y_0 = 0$ et $y_n = x_{n-1}$ pour $n \ge 1$ (on décale d'un cran vers la droite, en introduisant un 0 à la place 0 ; en bon français, cet opérateur s'appelle *opérateur de décalage (à droite)*, ou *opérateur de shift* en langage mathématique usuel) ; l'application S est clairement isométrique. Comme $\|S\| = 1$, on a $\mathrm{Sp}(S) \subset \{\lambda \in \mathbb{C} : |\lambda| \le 1\}$.

Si y est un élément de ℓ_q (exposant conjugué de p) et si $x \in \ell_p$ on notera l'action de dualité de ℓ_q sur ℓ_p par

$$(y, x) = j_q(y)(x) = \sum_{n=0}^{+\infty} y_n x_n.$$

Avec cette notation on va chercher à exprimer la transposée de S, considérée comme endomorphisme de ℓ_q. Soit T l'opération de décalage à gauche, définie par $T((y_n)_{n\ge0}) = (y_{n+1})_{n\ge0}$. On constate sans peine que $(T(y), x) = (y, S(x))$ pour tous $x \in \ell_p$, $y \in \ell_q$. L'application T "est" donc la transposée de S. Quand un opérateur V sur ℓ_p est inversible, il est clair que sa transposée est inversible dans $\mathcal{L}(\ell_q)$, ce qui entraîne que $\mathrm{Sp}({}^tS) \subset \mathrm{Sp}(S)$. Soit $\lambda \in \mathbb{C}$; si $|\lambda| < 1$, posons $y = (\lambda^n)_{n\ge0}$; c'est un élément non nul de ℓ_q et ${}^tS(y) = \lambda y$. Il en résulte que ${}^tS - \lambda \,\mathrm{Id}_{\ell_q}$ n'est pas inversible, donc le spectre de ${}^tS$ contient le disque unité ouvert, et il est contenu dans Sp(S) qui est contenu dans le disque unité fermé ; puisque le spectre est fermé,

$$\mathrm{Sp}(S) = \mathrm{Sp}({}^tS) = \{\lambda \in \mathbb{C} : |\lambda| \le 1\}.$$

6.2. Rayon spectral

Soit A une algèbre de Banach unitaire complexe. La quantité

$$\rho(a) = \max\{|\lambda| : \lambda \in \mathrm{Sp}(a)\}$$

s'appelle le *rayon spectral* de $a \in A$. On a déjà remarqué que le spectre de a est contenu dans le disque de $\mathbb{C}$ centré en 0 et de rayon $\|a\|$, donc

$$\rho(a) \le \|a\|.$$

On va obtenir au théorème 1 une formule importante qui précise cette remarque simple et qui permet d'estimer, sinon de calculer, ce rayon spectral.

Théorème 6.2.1. *Soient* A *une algèbre de Banach unitaire complexe et* $a \in \mathrm{A}$ *; la suite* $(\|a^n\|^{1/n})$ *est convergente et on a*

$$\lim_{n\to\infty} \|a^n\|^{1/n} = \rho(a) = \sup\{|\lambda| : \lambda \in \mathrm{Sp}(a)\}.$$

Démonstration. On démontre d'abord que $\rho(a) \le \limsup_n \|a^n\|^{1/n}$; remarquons tout de suite que $\|a^n\|^{1/n} \le \|a\|$ pour tout $n \ge 1$, donc ce que nous devons démontrer est un raffinement de l'estimation $\rho(a) \le \|a\|$ que nous avons déjà vue ; on obtiendra ce raffinement en reprenant les arguments déjà employés ; si $b \in \mathrm{A}$ est tel que $\beta = \limsup_n \|b^n\|^{1/n} < 1$, choisissons t réel tel que $\beta < t < 1$; on aura alors $\|b^n\|^{1/n} < t$ pour n grand, donc $\|b^n\| \le t^n$, donc la série $\sum_k b^k$ sera normalement convergente, donc convergente dans le Banach A, et la démonstration déjà vue pour le lemme 1.1 nous dira que $1_\mathrm{A} - b$ est inversible ; si on écrit comme avant $a - \lambda 1_\mathrm{A} = -\lambda(1_\mathrm{A} - a/\lambda)$, cet élément sera inversible dès que $b = a/\lambda$ vérifiera $\limsup_n \|b^n\|^{1/n} < 1$, ce qui se produit quand $\limsup_n \|a^n\|^{1/n} < |\lambda|$. Ceci signifie que $\rho(a) \le \limsup_n \|a^n\|^{1/n}$.

La deuxième inégalité demande de se rappeler le cours de fonctions holomorphes ; si $g(z) : \mathrm{B}(0,\mathrm{R}) \to \mathbb{C}$ est holomorphe ($\mathrm{R} = +\infty$ admis), alors elle est développable en série entière $\sum_{k=0}^{+\infty} b_k z^k$ dans ce disque ouvert $\mathrm{B}(0,\mathrm{R})$; pour tout r tel que $0 < r < \mathrm{R}$ la formule de Cauchy appliquée au cercle γ_r de rayon r donne pour tout $n \ge 0$

$$r^n b_n = r^n \,\frac{1}{2i\pi} \int_{\gamma_r} \frac{g(z)}{z^{n+1}}\,dz = \int_0^{2\pi} g(r\,\mathrm{e}^{i\theta})\,\mathrm{e}^{-in\theta}\,\frac{d\theta}{2\pi},$$

ce qui donne les inégalités de Cauchy

$$|b_n| r^n \le \mathrm{M}(r) = \max\{|g(z)| : |z| = r\}.$$

Considérons la fonction $f(z) = (1_\mathrm{A} - za)^{-1}$; elle est définie pour tout z tel que $1/z$ ne soit pas dans le spectre de a, ce qui est le cas lorsque $|z| < \mathrm{R} = \rho(a)^{-1}$; de plus $z \to f(z)$ est holomorphe de $\mathrm{B}(0,\mathrm{R})$ dans A. Soit r tel que $0 < r < \mathrm{R}$; la fonction $z \to \|f(z)\|$ est continue sur le cercle de rayon r qui est compact, donc cette fonction est bornée par un certain $\mathrm{M}_0(r)$; soit $a^* \in \mathrm{A}^*$ telle que $\|a^*\| \le 1$, et posons $g(z) = a^*(f(z))$. Alors g est holomorphe scalaire dans $\mathrm{B}(0,\mathrm{R})$ donc développable en série entière $\sum_{k=0}^{+\infty} b_k z^k$; mais par ailleurs, pour z assez petit on sait $f(z) = \sum_{k=0}^{+\infty} z^k a^k$ (lemme 1.1), donc $g(z) = \sum_k z^k a^*(a^k)$; par l'unicité des coefficients de Taylor il résulte que $b_n = a^*(a^n)$ pour tout n, et les inégalités de Cauchy donnent $|a^*(a^n)| \le \mathrm{M}_0(r)/r^n$ (remarquer que $|g(z)| \le \|f(z)\|$) ; en appliquant Hahn-Banach on aura $\|a^n\| \le \mathrm{M}_0(r)/r^n$, ce qui implique $\limsup_n \|a^n\|^{1/n} \le 1/r$, d'où $\limsup_n \|a^n\|^{1/n} \le \rho(a)$ en faisant tendre r vers $\mathrm{R} = 1/\rho(a)$.

La convergence de la suite $(\|a^n\|^{1/n})$ résulte immédiatement du lemme qui suit et du fait que pour tous $p, q \ge 1$, on a $\|a^{p+q}\| = \|a^p a^q\| \le \|a^p\|\,\|a^q\|$.

//

Lemme 6.2.2. *Soit* (u_n) *une suite de nombres réels positifs ou nuls telle que, pour tous entiers* $p, q \ge 1$ *on ait* $(u_{p+q})^{p+q} \le u_p^p\, u_q^q$ *; alors la suite* (u_n) *converge vers* $\inf_{n \ge 1} u_n$.

Démonstration. Montrons d'abord que pour tous entiers $p, k \ge 1$, on a $u_{pk} \le u_k$, par récurrence sur $p \ge 1$; c'est clair pour $p = 1$; si on connaît cette inégalité pour un certain $p \ge 1$, alors

$$(u_{(p+1)k})^{(p+1)k} \le u_{kp}^{kp}\, u_k^k \le u_k^{kp}\, u_k^k = u_k^{(p+1)k}.$$

Notons $m = \inf_{n\geq 1} u_n$; s'il existe un entier $k \geq 1$ tel que $u_k = 0$, alors $m = 0$ et, pour tout $p \geq 1$, on a $u_{k+p} = 0$, donc (u_n) converge vers m. Supposons désormais que l'on ait $u_k \neq 0$ pour tout $k > 0$. Soit $\varepsilon > 0$; par définition de m, il existe un entier $k \geq 1$ tel que $u_k < m + \varepsilon$. Soit $n \geq 1$ et écrivons $n = kp + r$ avec p, r entiers ≥ 0, $r < k$; alors $u_n^n \leq u_{kp}^{kp}\, u_r^r \leq u_k^{kp}\, u_1^r$ d'après notre première étape, donc

$$u_n \leq u_k^{kp/n} u_1^{r/n} = u_k \left(\frac{u_1}{u_k}\right)^{r/n} \leq u_k \left(\frac{u_1}{u_k}\right)^{(k-1)/n}.$$

Comme la suite $n \to u_k(u_1/u_k)^{(k-1)/n}$ converge vers $u_k < m+\varepsilon$, on aura $u_n < m+\varepsilon$ pour n assez grand, mais aussi $m \leq u_n$, d'où la convergence vers m de la suite (u_n).

//

Remarque 6.2.1. Ce n'est pas vraiment le lieu ici de développer les théories de l'intégration et des fonctions holomorphes pour les fonctions à valeurs dans un espace de Banach. Disons cependant que la théorie de Cauchy se généralise sans peine aux fonctions holomorphes à valeurs dans un espace de Banach (complexe, bien sûr) : soient $\mathrm{D} \subset \mathbb{C}$ un disque ouvert, F un espace de Banach complexe et $f : \overline{\mathrm{D}} \to \mathrm{F}$ une application continue, holomorphe sur D ; alors, pour tout $\lambda \in \mathrm{D}$ on a

$$f(\lambda) = \frac{1}{2i\pi} \int_{\partial \mathrm{D}} (z-\lambda)^{-1} f(z)\, dz,$$

l'intégrale étant prise sur le bord ∂D de D. De plus, on a en fait

$$a_n = \frac{1}{2\pi r^n} \int_0^{2\pi} \mathrm{e}^{-int}\, f(re^{it})\, dt.$$

Cependant, on n'a pas expliqué le sens de ces intégrales...

Ces intégrales sont des intégrales de fonctions continues définies sur un intervalle $[a, b]$, à valeurs dans un espace de Banach F. Il ne serait pas bien difficile de définir l'intégrale de Riemann dans ce cadre. Si f est continue de $[a, b]$ dans F, elle est uniformément continue puisque $[a, b]$ est compact, et on en déduit facilement que les sommes de Riemann (vectorielles) $\sum_{i=1}^{m} (x_{i+1} - x_i) f(\xi_i)$ convergent dans F lorsque le pas de la subdivision $\pi = (a = x_0 < x_1 < \ldots < x_m = b)$ de $[a, b]$ tend vers 0 (on commence par montrer que si (π_n) est une suite de subdivisions dont le pas tend vers 0, les sommes de Riemann correspondantes forment une suite de Cauchy, donc convergente puisque F est complet ; on montre ensuite que la limite ne dépend pas de la suite (π_n) choisie). Il est tout à fait raisonnable d'appeler $\int_a^b f(t)\, dt$ cette limite.

Proposition 6.2.3. *Soit* H *un espace hilbertien complexe ; le rayon spectral de tout élément normal de* $\mathcal{L}(\mathrm{H})$ *est égal à sa norme.*

Démonstration. Soit d'abord A un élément autoadjoint ; on a $\|\mathrm{A}^2\| = \|\mathrm{A}^*\mathrm{A}\| = \|\mathrm{A}\|^2$ (proposition 5.3.2) ; on en déduit par récurrence que $\|\mathrm{A}^{2^n}\| = \|\mathrm{A}\|^{2^n}$ pour tout $n \geq 0$, donc $\rho(\mathrm{A}) = \|\mathrm{A}\|$. Soit maintenant T un élément normal de $\mathcal{L}(\mathrm{H})$; par récurrence sur n, on a $(\mathrm{T}^*\mathrm{T})^n = (\mathrm{T}^*)^n\mathrm{T}^n$ donc $\|(\mathrm{T}^*\mathrm{T})^n\| = \|\mathrm{T}^n\|^2$ et $\rho(\mathrm{T}^*\mathrm{T}) = \rho(\mathrm{T})^2$. Or $\mathrm{A} = \mathrm{T}^*\mathrm{T}$ est autoadjoint, donc $\rho(\mathrm{T})^2 = \rho(\mathrm{T}^*\mathrm{T}) = \|\mathrm{T}^*\mathrm{T}\| = \|\mathrm{T}\|^2$.

//

Exemple 6.2.2.

Posons $\mathrm{H} = \mathrm{L}_2([0,1])$; pour $f \in \mathrm{H}$ et $s \in [0,1]$, on pose $\mathrm{V}(f)(s) = \int_0^s f(t)\,dt$. Puisque $\mathrm{L}_2([0,1]) \subset \mathrm{L}_1([0,1])$, la fonction f est intégrable et on en déduit que $\mathrm{V}(f)$ est continue (appliquer par exemple le théorème de Lebesgue à une suite de la forme $(\mathbf{1}_{[0,s_n]} f)$ pour une suite (s_n) tendant vers s) ; en appliquant Cauchy-Schwarz au produit $\mathbf{1}_{[0,s]} f$ on voit que $|\mathrm{V}(f)(s)| \le \sqrt{s}\,\|f\|_2$, ce qui implique que

$$\|\mathrm{V}(f)\|_2^2 \le \|f\|_2^2 \int_0^1 s\,ds = \frac{1}{2}\,\|f\|_2^2,$$

donc V définit une application linéaire continue (encore notée V) de $\mathrm{L}_2([0,1])$ dans lui-même.

Soit $f \in \mathrm{H}$ telle que $\|f\|_2 \le 1$; on a montré que $|\mathrm{V}(f)(s)| \le \sqrt{s}\,\|f\|_2 \le 1$ pour tout réel $s \in [0,1]$; on en déduit que $|\mathrm{V}(\mathrm{V}(f))(s)| = |\int_0^s \mathrm{V}(f)(t)\,dt| \le s$, puis, par récurrence sur n, que $|\mathrm{V}^{n+1}(f)(s)| \le s^n/n!$ donc

$$\|\mathrm{V}^{n+1}(f)\|_2^2 \le \frac{1}{(n!)^2} \int_0^1 s^{2n}\,ds \le \frac{1}{(n!)^2},$$

ce qui donne $\|\mathrm{V}^{n+1}\| \le (n!)^{-1}$. Comme $\lim_n (n!)^{-1/n} = 0$, il s'ensuit que le rayon spectral de V est nul, donc $\mathrm{Sp}(\mathrm{V}) = \{0\}$.

Exercice. Retrouver le spectre de V en trouvant explicitement la résolvante $\mathrm{R}_\lambda(\mathrm{V})$ pour tout $\lambda \ne 0$ (exercice d'équations différentielles !).

Changement d'algèbre, homomorphismes

Définition 6.2.3. Un *homomorphisme d'algèbres de Banach unitaires* est une application linéaire continue $\varphi : \mathrm{A} \to \mathrm{B}$ entre deux algèbres de Banach unitaires A et B, telle que $\varphi(ab) = \varphi(a)\varphi(b)$ pour tous $a, b \in \mathrm{A}$ et que $\varphi(1_\mathrm{A}) = 1_\mathrm{B}$.

Si a est inversible dans A, son image est inversible dans B et l'inverse de l'image est l'image de l'inverse. De plus $\varphi(a - \lambda 1_\mathrm{A}) = \varphi(a) - \lambda 1_\mathrm{B}$. Il en résulte que

$$\mathrm{Sp}(\varphi(a)) \subset \mathrm{Sp}(a).$$

On dira qu'on a un plongement isométrique de A dans B si φ est de plus isométrique. On écrira parfois $\mathrm{A} \subset \mathrm{B}$ dans ce cas. Même dans ce cas de plongement, il est possible qu'un élément non inversible dans A devienne inversible dans B.

Exemple-exercice. Soit A la sous-algèbre de $\mathrm{B} = \mathrm{C}(\mathbb{T})$ engendrée par les fonctions $(z^n)_{n \ge 0}$; montrer que la fonction z n'est pas inversible dans A (alors qu'elle le devient dans B ; indication : utiliser la norme L_2 et la base de Fourier).

Il y a cependant un cas où un élément a non inversible dans A ne peut jamais avoir une image inversible $\varphi(a)$, par un plongement isométrique φ de A dans B : disons que a est un *diviseur de zéro approché* (à gauche) dans A s'il existe une suite (u_n) dans A telle que $\|u_n\| = 1$ pour tout n, mais $a\,u_n \to 0$ (on peut aussi considérer la propriété analogue à droite). Il est clair qu'un d.z.a. ne peut pas être inversible, et que l'image d'un d.z.a. par une isométrie est encore un d.z.a.

Dans l'algèbre C(K), il est facile de voir que les non-inversibles sont exactement les d.z.a. Il en résulte que pour tout homomorphisme isométrique φ de C(K) dans une algèbre de Banach B, on a $\mathrm{Sp}(\varphi(f)) = \mathrm{Sp}(f)$.

6.3. Décomposition du spectre d'un opérateur borné

Proposition 6.3.1. *Soient* E, F *deux espaces de Banach et soit* $\mathrm{T} \in \mathcal{L}(\mathrm{E}, \mathrm{F})$ *; alors les conditions suivantes sont équivalentes :*

- (i) *l'application* T *est injective d'image fermée ;*
- (ii) *il existe un nombre* $c > 0$ *tel que pour tout* $x \in \mathrm{E}$ *on ait* $\|\mathrm{T}(x)\| \geq c\,\|x\|$ *;*
- (iii) *il n'existe pas de suite* (x_n) *dans* E *telle que* $\|x_n\| = 1$ *et* $\lim_n \|\mathrm{T}(x_n)\| = 0$.

Démonstration. Si (i) est satisfaite, T détermine une application continue bijective T_1 de E sur l'espace de Banach $\mathrm{im}(\mathrm{T})$. Par le théorème des isomorphismes (théorème 3.1.3), T_1 est un isomorphisme : on obtient (ii) avec $c = \|\mathrm{T}_1^{-1}\|^{-1}$. Il est évident que (ii) implique (iii) ; montrons que $(iii) \Rightarrow (ii)$: si (ii) n'est pas satisfaite, il existe pour tout entier $n \geq 1$ un vecteur $y_n \in \mathrm{E}$ tel que $n^{-1}\,\|y_n\| > \|\mathrm{T}(y_n)\|$; si on pose $x_n = \|y_n\|^{-1} y_n$, on a $\|x_n\| = 1$ et $\|\mathrm{T}(x_n)\| < 1/n$, donc (iii) n'est pas satisfaite.

Si (ii) est satisfaite, il est clair que T est injective ; si (y_n) est une suite dans $\mathrm{im}(\mathrm{T})$ qui converge vers $y \in \mathrm{F}$, écrivons $y_n = \mathrm{T}(x_n)$ avec $x_n \in \mathrm{E}$; on a $\|x_n - x_m\| \leq c^{-1}\,\|y_n - y_m\|$, donc la suite (x_n) est de Cauchy, donc convergente vers $x \in \mathrm{E}$ puisque E est complet ; alors la suite $y_n = \mathrm{T}(x_n)$ converge vers $\mathrm{T}(x)$, donc $y = \mathrm{T}(x)$ est dans $\mathrm{im}(\mathrm{T})$, qui est donc fermée dans l'espace F.

//

Spectre et transposition dans $\mathcal{L}(\mathrm{E})$

On va maintenant s'intéresser au rapport entre le spectre d'un opérateur $\mathrm{T} \in \mathcal{L}(\mathrm{E})$ et celui de son transposé ${}^t\mathrm{T} \in \mathcal{L}(\mathrm{E}^*)$. Ce rapport sera très simple : ces deux spectres sont égaux.

Lemme 6.3.2. *Soient* E, F *deux espaces de Banach et* $\mathrm{T} \in \mathcal{L}(\mathrm{E}, \mathrm{F})$ *; l'application* ${}^t\mathrm{T}$ *est injective si et seulement si* $\mathrm{im}(\mathrm{T})$ *est dense dans* F*. De plus, si* $\mathrm{im}({}^t\mathrm{T})$ *est dense dans* E^**, l'application* T *est injective.*

Démonstration. Si $\mathrm{im}(\mathrm{T})$ n'est pas dense, son adhérence G est un sous-espace fermé de F, distinct de F. D'après le théorème de Hahn-Banach, il existe une forme linéaire y^* non nulle sur F, mais dont la restriction à G est nulle ; en particulier, $y^*(\mathrm{T}(x)) = 0$ pour tout $x \in \mathrm{E}$ puisque G contient l'image de T. On a donc $({}^t\mathrm{T})(y^*)(x) = 0$ pour tout $x \in \mathrm{E}$, ce qui signifie que ${}^t\mathrm{T}(y^*) = 0$, donc ${}^t\mathrm{T}$ n'est pas injective.

Si ${}^t\mathrm{T}$ n'est pas injective, il existe $y^* \in \mathrm{F}^*$ non nulle telle que ${}^t\mathrm{T}(y^*) = 0$, ce qui signifie que $y^*(\mathrm{T}(x)) = 0$ pour tout $x \in \mathrm{E}$. On voit alors que l'image de T est contenue dans le noyau de y^*, qui est un sous-espace fermé de F, distinct de F. Il en résulte que $\mathrm{im}(\mathrm{T})$ n'est pas dense dans F.

Si $\mathrm{T}(x) = 0$, on a ${}^t\mathrm{T}(y^*)(x) = y^*(\mathrm{T}(x)) = 0$ pour tout $y^* \in \mathrm{F}^*$, ce qui montre que $x^*(x) = 0$ pour tout $x^* = {}^t\mathrm{T}(y^*) \in \mathrm{im}({}^t\mathrm{T})$; si $\mathrm{im}({}^t\mathrm{T})$ est dense dans E^*, on en déduit par continuité que $x^*(x) = 0$ pour tout $x^* \in \mathrm{E}^*$, donc $x = 0$ par Hahn-Banach, donc T est injective.

//

Proposition 6.3.3. *Soient* E, F *deux espaces de Banach et soit* $\mathrm{T} \in \mathcal{L}(\mathrm{E}, \mathrm{F})$ *; l'opérateur transposé* ${}^t\mathrm{T} \in \mathcal{L}(\mathrm{F}^*, \mathrm{E}^*)$ *est inversible si et seulement* T *est inversible.*

Démonstration. Si T est inversible, comme $\mathrm{T}^{-1}\mathrm{T} = \mathrm{Id}_\mathrm{E}$ et $\mathrm{T}\mathrm{T}^{-1} = \mathrm{Id}_\mathrm{F}$, on trouve ${}^t\mathrm{T}\,{}^t(\mathrm{T}^{-1}) = \mathrm{Id}_{\mathrm{E}^*}$ et ${}^t(\mathrm{T}^{-1})\,{}^t\mathrm{T} = \mathrm{Id}_{\mathrm{F}^*}$. Donc ${}^t\mathrm{T}$ est inversible et $({}^t\mathrm{T})^{-1} = {}^t(\mathrm{T}^{-1})$.

L'implication inverse est évidente si E est réflexif, parce que T est alors "en gros" la transposée de ${}^t\mathrm{T}$ dans ce cas. Sinon il faut travailler un peu : supposons donc inversement que ${}^t\mathrm{T}$ soit inversible. On montre d'abord qu'il existe une constante $c > 0$ telle que pour tout $x \in \mathrm{E}$ on ait $\|\mathrm{T}(x)\| \geq c\,\|x\|$. Soit $x \in \mathrm{E}$; il existe $x^* \in \mathrm{E}^*$ tel que $\|x^*\| \leq 1$ et $x^*(x) = \|x\|$ (Hahn-Banach) ; puisque ${}^t\mathrm{T}$ est inversible, il existe $y^* \in \mathrm{F}^*$ tel que $x^* = {}^t\mathrm{T}(y^*)$ et $\|y^*\| \leq \|({}^t\mathrm{T})^{-1}\| = \mathrm{M}$. Alors $\mathrm{M}\,\|\mathrm{T}(x)\| \geq y^*(\mathrm{T}(x)) = x^*(x) = \|x\|$, d'où le résultat voulu, avec $c = \mathrm{M}^{-1}$. Il en résulte que T est un isomorphisme de E sur im(T), qui est donc complet, donc fermé dans F. Par ailleurs, im(T) est dense puisque ${}^t\mathrm{T}$ est injective, donc im(T) = F et T est un isomorphisme de E sur F.

//

On en déduit immédiatement :

Corollaire 6.3.4. *Soient* E *un espace de Banach complexe et* $\mathrm{T} \in \mathcal{L}(\mathrm{E})$ *; on a*

$$\mathrm{Sp}({}^t\mathrm{T}) = \mathrm{Sp}(\mathrm{T}).$$

Démonstration. Il suffit de remarquer que ${}^t(\mathrm{T} - \lambda\,\mathrm{Id}_\mathrm{E}) = {}^t\mathrm{T} - \lambda\,\mathrm{Id}_{\mathrm{E}^*}$ pour tout $\lambda \in \mathbb{C}$.

//

Dans le cas hilbertien, on préfère souvent exprimer le résultat précédent en utilisant l'adjoint $\mathrm{T}^* \in \mathcal{L}(\mathrm{H})$ plutôt que la transposée ${}^t\mathrm{T} \in \mathcal{L}(\mathrm{H}^*)$. Le seul petit piège à éviter est que $(\mathrm{T} - \lambda\,\mathrm{Id}_\mathrm{H})^* = \mathrm{T}^* - \overline{\lambda}\,\mathrm{Id}_\mathrm{H}$ (il y a une barre de conjugaison !).

Corollaire 6.3.5. *Soient* H *un espace de Hilbert complexe et* $\mathrm{T} \in \mathcal{L}(\mathrm{H})$ *; on a*

$$\mathrm{Sp}(\mathrm{T}^*) = \overline{\mathrm{Sp}(\mathrm{T})} = \{\overline{\lambda} : \lambda \in \mathrm{Sp}(\mathrm{T})\}.$$

On va maintenant distinguer plusieurs sous-ensembles intéressants du spectre d'un opérateur borné, correspondant à plusieurs façons pour $\mathrm{T} - \lambda\,\mathrm{Id}_\mathrm{E}$ de ne pas être inversible. Soient E un espace de Banach complexe, $\mathrm{T} \in \mathcal{L}(\mathrm{E})$ et $\lambda \in \mathrm{Sp}(\mathrm{T})$; nous distinguerons trois cas :

1. Le scalaire λ est une valeur propre de T, autrement dit $\mathrm{T} - \lambda\,\mathrm{Id}_\mathrm{E}$ n'est pas injectif.

2. Le scalaire λ est une valeur propre de ${}^t\mathrm{T}$, mais n'est pas une valeur propre de T ; autrement dit $\mathrm{T} - \lambda\,\mathrm{Id}_\mathrm{E}$ est injectif et il existe une forme linéaire non nulle $x^* \in \mathrm{E}^*$ telle que $({}^t\mathrm{T} - \lambda\,\mathrm{Id}_{\mathrm{E}^*})(x^*) = 0$, c'est à dire que ${}^t(\mathrm{T} - \lambda\,\mathrm{Id}_\mathrm{E})$ n'est pas injectif ; d'après le lemme 2, cela se produit si et seulement si $\mathrm{T} - \lambda\,\mathrm{Id}_\mathrm{E}$ est injectif mais n'a pas une image dense dans E.

3. Le scalaire λ n'est une valeur propre ni de T, ni de ${}^t\mathrm{T}$, mais λ est quand même dans le spectre de T. Alors, $\mathrm{T} - \lambda\,\mathrm{Id}_\mathrm{E}$ est injectif, son image est dense mais n'est pas fermée.

Définition 6.3.1. Soient E un espace de Banach complexe et $T \in \mathcal{L}(E)$; on appelle *spectre ponctuel* de T l'ensemble $\mathrm{Sp}_p(T)$ des $\lambda \in \mathbb{C}$ tels que $T - \lambda \,\mathrm{Id}_E$ ne soit pas injectif (c'est l'ensemble des valeurs propres de T). On appelle *spectre résiduel* de T l'ensemble $\mathrm{Sp}_r(T)$ des $\lambda \in \mathbb{C}$ tels que $T - \lambda \,\mathrm{Id}_E$ soit injectif, mais son image ne soit pas dense. On appelle *spectre continu* de T l'ensemble $\mathrm{Sp}_c(T)$ des $\lambda \in \mathbb{C}$ tels que $T - \lambda \,\mathrm{Id}_E$ soit injectif, à image dense mais pas fermée.

On voit que l'on a $\lambda \in \mathrm{Sp}_c(T)$ si et seulement si : $\lambda \in \mathrm{Sp}(T)$ **et** $T - \lambda \,\mathrm{Id}_E$ est injectif à image dense ; en effet, l'image de $T - \lambda \,\mathrm{Id}_E$ n'est alors pas fermée : si elle était fermée, elle serait égale à E, l'opérateur $T - \lambda \,\mathrm{Id}_E$ serait un isomorphisme et λ ne serait pas dans le spectre de T.

Proposition 6.3.6. *Soient* E *un espace de Banach complexe et* $T \in \mathcal{L}(E)$ *; on a*

$$\mathrm{Sp}_r(T) = \mathrm{Sp}_p({}^tT) \setminus \mathrm{Sp}_p(T) \quad \textit{et} \quad \mathrm{Sp}_c({}^tT) \subset \mathrm{Sp}_c(T).$$

Si E *est réflexif, on a l'égalité* $\mathrm{Sp}_c({}^tT) = \mathrm{Sp}_c(T)$.

Démonstration. On a vu que λ est dans le spectre résiduel de T si et seulement si λ est une valeur propre de ${}^tT$, mais n'est pas une valeur propre de T, d'où la première assertion. Si $\lambda \in \mathrm{Sp}_c({}^tT)$, on sait que ${}^tT - \lambda \,\mathrm{Id}_{E^*}$ est injectif à image dense, donc $T - \lambda \,\mathrm{Id}_E$ est injectif à image dense par le lemme 2, et puisque $\mathrm{Sp}_c({}^tT) \subset \mathrm{Sp}({}^tT) = \mathrm{Sp}(T)$, on a $\lambda \in \mathrm{Sp}(T)$, par conséquent $\lambda \in \mathrm{Sp}_c(T)$.

Dans le cas où E est réflexif, T "s'identifie" à la transposée de ${}^tT$, et il en résulte que $\mathrm{Sp}_c(T) \subset \mathrm{Sp}_c({}^tT)$. Plus précisément, on vérifie que $T = I_E^{-1} \circ {}^t({}^tT) \circ I_E$, où I_E désigne l'isomorphisme de E sur E^{**} (on devra remarquer que si U est un isomorphisme de E sur F et si $T \in \mathcal{L}(F)$, toutes les notions de spectre introduites sont les mêmes pour les deux opérateurs T et $U^{-1}TU \in \mathcal{L}(E)$).

//

Dans le cas hilbertien, on a :

Proposition 6.3.7. *Soient* H *un espace hilbertien complexe et* $T \in \mathcal{L}(H)$ *un opérateur borné ; on a* $\mathrm{Sp}(T^*) = \{\overline{\lambda} : \lambda \in \mathrm{Sp}(T)\}$ *et* $\mathrm{Sp}_r(T) = \{\lambda \in \mathbb{C} \setminus \mathrm{Sp}_p(T) : \overline{\lambda} \in \mathrm{Sp}_p(T^*)\}$.

Démonstration. La première assertion est un rappel. Par la proposition 5.3.3, l'image de $T - \lambda \,\mathrm{Id}_H$ est dense si et seulement si $T^* - \overline{\lambda} \,\mathrm{Id}_H$ est injectif. La deuxième assertion en résulte.

//

Proposition 6.3.8. *Le spectre résiduel d'un opérateur normal est vide.*

Démonstration. Soit $T \in \mathcal{L}(H)$ un opérateur normal ; pour tout vecteur $x \in H$, on a $\|T^*(x)\|^2 = \langle T^*(x), T^*(x)\rangle = \langle x, T(T^*(x))\rangle = \langle x, T^*(T(x))\rangle = \langle T(x), T(x)\rangle = \|T(x)\|^2$, donc $\ker(T^*) = \ker(T)$. Pour tout $\lambda \in \mathbb{C}$, l'opérateur $T - \lambda \,\mathrm{Id}_H$ est encore normal, donc $\ker(T^* - \overline{\lambda} \,\mathrm{Id}_H) = \ker(T - \lambda \,\mathrm{Id}_H)$. Il en résulte que $\overline{\lambda} \in \mathrm{Sp}_p(T^*)$ si et seulement si $\lambda \in \mathrm{Sp}_p(T)$, d'où le résultat.

//

Exemples 6.3.2.

a. Soient K un espace compact métrique, $\mathrm{E} = \mathrm{C}(\mathrm{K})$ et soit $f \in \mathrm{E}$; on a vu dans l'exemple 2.2 que l'application $\mathrm{T} = \mathrm{M}_f$ de multiplication par f vérifie $\mathrm{Sp}(\mathrm{T}) = f(\mathrm{K})$; on a vu aussi que s'il existe $s \in \mathrm{K}$ tel que $f(s) = \lambda$, l'image de $\mathrm{T} - \lambda\,\mathrm{Id}_\mathrm{E}$ n'est pas dense, donc $\lambda \in \mathrm{Sp}_p(\mathrm{T}) \cup \mathrm{Sp}_r(\mathrm{T})$. Remarquons que λ est une valeur propre de T si et seulement s'il existe $g \in \mathrm{E}$ non nulle telle que $\mathrm{T}(g) = \lambda g$, c'est à dire $(f - \lambda)g = 0$. L'ensemble des $s \in \mathrm{K}$ tels que $g(s) \neq 0$ est alors un ouvert non vide U de K et f est égale à λ sur U. Supposons inversement qu'il existe un ouvert non vide U de K tel que f soit égale à λ sur U ; notons $g \in \mathrm{E}$ la fonction qui à $s \in \mathrm{K}$ associe sa distance au complémentaire de U. On a $(\mathrm{T} - \lambda\,\mathrm{Id}_\mathrm{E})(g) = 0$.

En résumé, le spectre de T est l'ensemble $\mathrm{Sp}(\mathrm{T}) = \{f(s) : s \in \mathrm{K}\}$, le spectre ponctuel de T est l'ensemble des $\lambda \in \mathbb{C}$ tels que l'intérieur de $f^{-1}(\{\lambda\})$ soit non vide, le spectre résiduel de T est $\mathrm{Sp}(\mathrm{T}) \setminus \mathrm{Sp}_p(\mathrm{T})$ et le spectre continu de T est vide.

b. Soit $\mathrm{S} \in \mathcal{L}(\ell_2)$ l'application de décalage à droite. On a vu que Sp(S) est le disque unité fermé $\{\lambda \in \mathbb{C} : |\lambda| \leq 1\}$. Soient $\xi = (x_n)_{n\geq 0} \in \ell_2$ et $\lambda \in \mathbb{C}$ tels que $\mathrm{S}(\xi) = \lambda\xi$; on trouve alors $\lambda x_0 = 0$ et, pour tout $n \geq 1$, $\lambda x_n = x_{n-1}$; si $\lambda \neq 0$, on trouve alors par récurrence sur n que $x_n = 0$ pour tout $n \geq 0$; si $\lambda = 0$, on trouve, pour tout $n \geq 1$, $x_{n-1} = 0$. Dans les deux cas, $\xi = 0$. Donc $\mathrm{Sp}_p(\mathrm{S}) = \emptyset$.

On a vu que tout λ tel que $|\lambda| < 1$ est valeur propre de ${}^t\mathrm{S}$. Supposons que $|\lambda| = 1$ et soit $\eta = (x_n)_{n\geq 0} \in \ell_2$ tel que ${}^t\mathrm{S}(\eta) = \lambda\eta$; alors, pour tout $n \geq 0$, on a $x_{n+1} = \lambda x_n$; il s'ensuit alors que $x_n = \lambda^n x_0$; comme la suite $(\lambda^n)_{n\geq 0}$ n'est pas dans ℓ_2 (vu que $|\lambda| = 1$), on a nécessairement $x_0 = 0$, et enfin, $\eta = 0$; donc $\mathrm{Sp}_p({}^t\mathrm{S}) = \{\lambda \in \mathbb{C} : |\lambda| < 1\}$. Il résulte alors de la proposition 7 que $\mathrm{Sp}_r(\mathrm{S}) = \{\lambda \in \mathbb{C} : |\lambda| < 1\}$; on a alors pour terminer $\mathrm{Sp}_c(\mathrm{S}) = \{\lambda \in \mathbb{C} : |\lambda| = 1\}$.

c. Posons $\mathrm{H} = \mathrm{L}_2([0,1])$ et reprenons l'opérateur V de l'exemple 2.2, défini par $\mathrm{V}(f)(s) = \int_0^s f(t)\,dt$ pour $f \in \mathrm{H}$ et $s \in [0,1]$. On a montré que le rayon spectral de V est nul, donc $\mathrm{Sp}(\mathrm{V}) = \{0\}$. Remarquons que l'application qui à une fonction continue associe sa classe dans $\mathrm{L}_2([0,1])$ est injective ; donc si $\mathrm{V}(f) = 0$, alors $\mathrm{V}(f)(s) = 0$ pour tout $s \in [0,1]$, ce qui signifie que f est orthogonale à toutes les fonctions $\mathbf{1}_{[0,s]}$, donc à toutes les fonctions en escalier. Comme celles-ci forment un sous-espace dense dans $\mathrm{L}_2([0,1])$ il s'ensuit que V est injective. Il est clair que l'image de V contient l'ensemble des fonctions de classe C^1 nulles en 0. Or celles-ci forment un sous-espace dense de $\mathrm{L}_2([0,1])$. On a montré que $\mathrm{Sp}_p(\mathrm{V}) = \mathrm{Sp}_r(\mathrm{V}) = \emptyset$ et $\mathrm{Sp}_c(\mathrm{V}) = \mathrm{Sp}(\mathrm{V}) = \{0\}$.

Valeurs propres approchées

Lemme 6.3.9. *Soient* E *un espace de Banach complexe,* $\mathrm{T} \in \mathcal{L}(\mathrm{E})$ *et soit* $\lambda \in \partial\,\mathrm{Sp}(\mathrm{T})$ *(la frontière du spectre de* T*). Il existe une suite* $(x_n) \subset \mathrm{E}$ *de vecteurs de norme* 1 *telle que* $(\mathrm{T} - \lambda\,\mathrm{Id}_\mathrm{E})(x_n)$ *tende vers* 0.

Démonstration. En posant $\mathrm{S} = \mathrm{T} - \lambda\,\mathrm{Id}_\mathrm{E}$, on se ramène à montrer que si $0 \in \partial\,\mathrm{Sp}(\mathrm{S})$, il existe une suite $(x_n) \subset \mathrm{E}$ de vecteurs de norme 1 telle que $\mathrm{S}(x_n)$ tende vers 0. Puisque Sp(S) est fermé, sa frontière est contenue dans Sp(S), donc $0 \in \mathrm{Sp}(\mathrm{S})$ et S n'est pas inversible. Si on ne pouvait pas trouver la suite (x_n), on aurait $\|\mathrm{S}(x)\| \geq c\,\|x\|$ pour un $c > 0$ et tout $x \in \mathrm{E}$ (proposition 1), et $\mathrm{S}(\mathrm{E}) \neq \mathrm{E}$ puisque S n'est pas inversible. On pourrait alors trouver $y \notin \mathrm{S}(\mathrm{E})$; puisque 0 est à la frontière du spectre de S, il existe une suite (μ_n) hors du spectre et qui tend vers 0 ; alors $\mathrm{S} - \mu_n\,\mathrm{Id}_\mathrm{E}$ est inversible pour tout n. Il existe donc un vecteur $z_n \in \mathrm{E}$ tel que $(\mathrm{S} - \mu_n \operatorname{Id}_{\mathrm{E}})(z_n) = y$. Si (z_n) était bornée, on aurait $\mu_n z_n \to 0$ et y serait limite de la suite $(\mathrm{S}(z_n)) \subset \mathrm{S}(\mathrm{E})$, ce qui est impossible puisque $\mathrm{S}(\mathrm{E})$ est supposé fermé et $y \notin \mathrm{S}(\mathrm{E})$. Il existe donc une sous-suite (z'_n) telle que $\|z'_n\|$ tende vers $+\infty$; en posant $x_n = \|z'_n\|^{-1} z'_n$, on voit que $\mathrm{S}(x_n) - \mu_n x_n = \|z'_n\|^{-1} y$ tend vers 0, donc $\mathrm{S}(x_n) \to 0$.

//

Exemple 6.3.3. Exemple de valeurs propres approchées : soit S le shift à droite sur ℓ_2 ; on sait que le spectre de S est égal au disque unité fermé, sa frontière est donc le cercle unité $\mathbb{T}$. Soit λ de module un un point de $\partial \operatorname{Sp}(\mathrm{S})$; on considère pour tout $n \geq 1$ le vecteur de norme 1 de ℓ_2

$$x_n = n^{-1/2}(1, \lambda^{-1}, \lambda^{-2}, \ldots, \lambda^{-n+1}, 0, \ldots)$$

et on note que $\|\mathrm{S}(x_n) - \lambda x_n\| \leq 2n^{-1/2} \to 0$, ce qui donne des presque vecteurs propres pour la valeur $\lambda \in \mathbb{T}$.

Proposition 6.3.10. *Soit* H *un espace hilbertien complexe ; le spectre de tout élément hermitien de* $\mathcal{L}(\mathrm{H})$ *est réel ; si de plus* T *est un opérateur positif, son spectre est contenu dans* $[0, +\infty[$. *Le spectre de tout élément unitaire de* $\mathcal{L}(\mathrm{H})$ *est contenu dans le cercle unité.*

Démonstration. Soit β le max de $|b|$ pour $a + ib = \lambda$ dans le spectre de T hermitien, et soit $\lambda = a + ib \in \operatorname{Sp}(\mathrm{T})$ tel que $|b| = \beta$. Alors λ est point frontière du spectre, donc il existe une suite (x_n) de vecteurs de norme un telle que $\mathrm{T}(x_n) - \lambda x_n$ tende vers 0. On voit donc que

$$\langle \mathrm{T}(x_n), x_n \rangle - \lambda \langle x_n, x_n \rangle = \langle \mathrm{T}(x_n), x_n \rangle - \lambda$$

tend vers 0, et $\langle \mathrm{T}(x_n), x_n \rangle = \langle x_n, \mathrm{T}(x_n) \rangle$ est réel, donc λ est réel, $b = \beta = 0$ et tout le spectre de T est réel. Si T est positif, son spectre est contenu dans $\mathbb{R}$, donc tous les points λ de $\operatorname{Sp}(\mathrm{T})$ sont points frontière et sont donc limite de suites de la forme $(\langle \mathrm{T}(x_n), x_n \rangle)$ comme on l'a vu ci-dessus. Mais quand T est positif, tous ces nombres sont ≥ 0, donc $\lambda \geq 0$.

Si $\mathrm{U} \in \mathcal{L}(\mathrm{H})$ est unitaire, considérons de même τ, le min de $|r|$ pour $\lambda = r\,\mathrm{e}^{i\theta}$ dans le spectre de U. Par compacité, on peut trouver un point du spectre de la forme $\lambda = \tau\,\mathrm{e}^{i\theta}$; puisque U est inversible, $0 \notin \operatorname{Sp}(\mathrm{U})$ donc $\tau > 0$. Comme précédemment, λ est un point frontière et on peut trouver une suite (x_n) de vecteurs de norme un telle que $\mathrm{U}x_n - \lambda x_n \to 0$. Comme U est isométrique, il en résulte que $|\lambda| = 1$. Le spectre de U ne contient donc aucun point du disque unité ouvert. Par ailleurs, $\operatorname{Sp}(\mathrm{U})$ est contenu dans le disque unité puisque $\|\mathrm{U}\| = 1$. Le résultat annoncé en découle.

//

Le résultat sur les isométries bijectives est valable pour tout espace de Banach. Par ailleurs, on peut aussi le démontrer très simplement en remarquant que le spectre de U^{-1} est formé des inverses des éléments de $\operatorname{Sp}(\mathrm{U})$, et qu'il est lui aussi contenu dans le disque unité puisque $\|\mathrm{U}^{-1}\| = 1$.

Remarque 6.3.4. Voici une méthode plus orthodoxe pour traiter le cas hilbertien. Soient E un espace de Hilbert et S un opérateur normal sur E ; si S n'est pas inversible et s'il n'existe pas de suite $(x_n) \subset \mathrm{E}$ de vecteurs de norme un telle que $\mathrm{S}(x_n)$ tende vers 0, alors S est un isomorphisme de E sur son image fermée $\mathrm{S(E)} \neq \mathrm{E}$ (proposition 1), donc le noyau de S^* n'est pas nul. Mais quand S est normal, on a pour tout $x \in \mathrm{E}$

$$\|\mathrm{S}(x)\|^2 = \langle \mathrm{S}^*\mathrm{S}(x), x\rangle = \langle \mathrm{S}\mathrm{S}^*(x), x\rangle = \|\mathrm{S}^*(x)\|^2$$

ce qui montre que $\ker(\mathrm{S}) = \ker(\mathrm{S}^*)$ et aboutit à une contradiction. Si T est normal, $\mathrm{S} = \mathrm{T} - \lambda\,\mathrm{Id_E}$ est normal pour tout λ, et ce qui précède montre que toutes les valeurs du spectre de T sont valeurs propres approchées. De plus, comme $\mathrm{S}^* = \mathrm{T}^* - \overline{\lambda}\,\mathrm{Id_E}$, la relation précédente signifie que

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

et montre que quand x est presque vecteur propre de T pour la valeur λ, le même vecteur x est presque vecteur propre de T^* pour la valeur conjuguée $\overline{\lambda}$.

7. Quelques classes d'opérateurs

7.1. Applications linéaires compactes

Définition 7.1.1. Soient E et F deux espaces de Banach ; une application linéaire continue $T \in \mathcal{L}(E, F)$ est dite *compacte* si l'image $T(B_E)$ par l'application T de la boule unité fermée B_E de l'espace E est relativement compacte (en norme) dans F. On note $\mathcal{K}(E, F)$ l'ensemble des applications linéaires compactes de E dans F. On pose $\mathcal{K}(E) = \mathcal{K}(E, E)$.

Rappelons qu'une partie A d'un espace topologique séparé X est dite *relativement compacte* dans X s'il existe une partie compacte B de X contenant A. Dans ce cas B est fermée dans X donc contient $\overline{A}$ et $\overline{A}$ est alors fermé dans B donc est compacte. Autrement dit, A est relativement compacte si et seulement si $\overline{A}$ est compacte. Rappelons qu'un espace métrique X est dit *précompact* si, pour tout $\varepsilon > 0$ il existe un recouvrement fini de X par des parties de diamètre $\leq \varepsilon$; un espace métrique est compact si et seulement s'il est précompact est complet. En particulier, dans un espace métrique complet, les parties relativement compactes sont les parties précompactes.

Dans le cas d'un sous-ensemble A d'un espace de Banach E, il est agréable de retenir un critère qui utilise le caractère vectoriel de l'espace ambiant : pour que l'adhérence de A soit compacte dans l'espace de Banach E, il faut et il suffit que A vérifie les deux conditions suivantes :

– l'ensemble A est borné ;

– pour tout $\varepsilon > 0$, il existe un sous-espace vectoriel $L_\varepsilon \subset E$ *de dimension finie* tel que tout point de A soit à une distance $< \varepsilon$ de L_ε :

$$\forall x \in A, \quad \operatorname{dist}(x, L_\varepsilon) < \varepsilon.$$

Si l'adhérence de A est compacte il est facile de vérifier que le critère est satisfait ; on va esquisser la démonstration de l'autre direction. Supposons les deux conditions du critère vérifiées et soit (x_n) une suite dans $\overline{A}$. On va montrer d'abord que pour tout $\varepsilon > 0$ on peut extraire une sous-suite telle que $\|x_{n_k} - x_{n_\ell}\| < \varepsilon$ pour tous k, ℓ. On commence par trouver x'_n dans A tel que $\|x_n - x'_n\| < \varepsilon/8$ pour tout n ; ensuite d'après le critère appliqué avec $\varepsilon/8$ il existe un espace de dimension finie L tel que tout point de A soit à distance $< \varepsilon/8$ d'un point de L ; on peut donc trouver $y_n \in L$ tel que $\|y_n - x'_n\| < \varepsilon/8$, ce qui donne $\|x_n - y_n\| < \varepsilon/4$. Par ailleurs la suite (x_n) est bornée, donc il existe M tel que $\|x_n\| \leq M$ pour tout n, donc $\|y_n\| \leq M + \varepsilon/4$ est bornée aussi, et située dans l'espace de dimension finie L ; d'après Bolzano-Weierstrass, on peut trouver une sous-suite (y_{n_k}) convergente, qui vérifiera donc pour k, ℓ assez grands $\|y_{n_k} - y_{n_\ell}\| < \varepsilon/2$. En revenant aux (x_n) on obtient $\|x_{n_k} - x_{n_\ell}\| < \varepsilon$ avec l'inégalité triangulaire.

On appliquera ce premier pas successivement avec $\varepsilon = 1/2, 1/4$, etc... en prenant à chaque fois une sous-suite de la sous-suite précédente, puis on prendra une sous-suite diagonale qui sera de Cauchy, donc convergente dans l'espace complet E.

Proposition 7.1.1. *Soient* E *et* F *deux espaces de Banach ; l'ensemble* $\mathcal{K}(\mathrm{E},\mathrm{F})$ *est un sous-espace vectoriel fermé de* $\mathcal{L}(\mathrm{E},\mathrm{F})$.

Soient E, F *et* G *des espaces de Banach,* $\mathrm{S} \in \mathcal{L}(\mathrm{E},\mathrm{F})$ *et* $\mathrm{T} \in \mathcal{L}(\mathrm{F},\mathrm{G})$ *; si* S *ou* T *est compacte alors* TS *est compacte. En particulier,* $\mathcal{K}(\mathrm{E})$ *est un idéal bilatère de* $\mathcal{L}(\mathrm{E})$.

Démonstration. Il est clair que si $\mathrm{T} \in \mathcal{K}(\mathrm{E},\mathrm{F})$ et $\lambda \in \mathbb{K}$, alors $\lambda\mathrm{T} \in \mathcal{K}(\mathrm{E},\mathrm{F})$. Soient maintenant T_1 et T_2 deux applications linéaires compactes de E dans F, et considérons les ensembles $\mathrm{A}_1 = \mathrm{T}_1(\mathrm{B}_\mathrm{E})$, $\mathrm{A}_2 = \mathrm{T}_2(\mathrm{B}_\mathrm{E})$ et $\mathrm{A} = (\mathrm{T}_1+\mathrm{T}_2)(\mathrm{B}_\mathrm{E})$; appliquons le critère précédent à l'ensemble A ; tout d'abord, $\mathrm{T}_1+\mathrm{T}_2$ est continue, donc A est borné ; ensuite, pour tout $\varepsilon > 0$ il existe deux sous-espaces vectoriels L_1 et L_2 de dimension finie de F tels que, pour $j = 1, 2$, tout point de A_j soit à une distance $< \varepsilon/2$ de l'espace L_j. Le sous-espace $\mathrm{L} = \mathrm{L}_1 + \mathrm{L}_2$ est de dimension finie et les points de A_j sont *a fortiori* à une distance $< \varepsilon/2$ de L. Soit y un point quelconque de A ; on peut écrire $y = \mathrm{T}_1(x) + \mathrm{T}_2(x)$, avec $x \in \mathrm{B}_\mathrm{E}$, donc $\mathrm{T}_j(x) \in \mathrm{A}_j$. Il existe $z_1, z_2 \in \mathrm{L}$ tels que $\|\mathrm{T}_j(x) - z_j\| < \varepsilon/2$, d'où résulte que $\|x - (z_1+z_2)\| < \varepsilon$ et $\mathrm{dist}(y,\mathrm{L}) < \varepsilon$.

Supposons que $\mathrm{T} \in \mathcal{L}(\mathrm{E},\mathrm{F})$ soit adhérent à $\mathcal{K}(\mathrm{E},\mathrm{F})$. L'ensemble $\mathrm{T}(\mathrm{B}_\mathrm{E})$ est borné et vérifie la deuxième condition du critère précédent : pour tout $\varepsilon > 0$ donné, on peut trouver S compacte telle que $\|\mathrm{T} - \mathrm{S}\| < \varepsilon/2$, puis un sous-espace L de dimension finie qui approche $\mathrm{S}(\mathrm{B}_\mathrm{E})$ à $\varepsilon/2$. Il en résulte facilement que L approche $\mathrm{T}(\mathrm{B}_\mathrm{E})$ à moins de ε.

Montrons pour finir les propriétés de composition. Supposons $\mathrm{S} \in \mathcal{L}(\mathrm{E},\mathrm{F})$ compacte ; si $\mathrm{K} \subset \mathrm{F}$ est compact et contient l'image $\mathrm{S}(\mathrm{B}_\mathrm{E})$, alors $\mathrm{T}(\mathrm{K})$ est compact et contient l'image $\mathrm{TS}(\mathrm{B}_\mathrm{E})$, donc TS est compacte. Pour l'autre cas, remarquons que l'image $\mathrm{S}(\mathrm{B}_\mathrm{E})$ est contenue dans la boule de F de centre 0 et de rayon $r = \|\mathrm{S}\|$; si $\mathrm{K} \subset \mathrm{G}$ est compact et contient l'image par T de la boule unité de F, alors $r\,\mathrm{K}$ est compact et contient l'image par TS de B_E.

//

Exemples 7.1.2.

1. Il est clair que tout opérateur T de rang fini est compact : en effet, l'ensemble $\mathrm{T}(\mathrm{B}_\mathrm{E})$ est alors un ensemble borné d'un espace vectoriel de dimension finie. D'après le résultat précédent, toute limite T en norme d'opérateur d'une suite (T_n) d'opérateurs de rang fini est compacte. C'est une méthode assez efficace pour vérifier que certains opérateurs sont compacts ; on montre par exemple que si $c_n \to 0$, l'opérateur Δ_c de ℓ_p dans ℓ_p défini par $\Delta_c((x_n)) = (c_n x_n)$ est compact :

on commence par remarquer que la norme de Δ_c dans $\mathcal{L}(\ell_p)$ est majorée par $\|c\|_\infty$. Ensuite, pour tout entier N on considère la suite $c^{(\mathrm{N})}$ telle que $c_n^{(\mathrm{N})} = c_n$ si $n \le \mathrm{N}$ et $c_n^{(\mathrm{N})} = 0$ sinon ; l'opérateur $\mathrm{T}_\mathrm{N} = \Delta_{c^{(\mathrm{N})}}$ est de rang fini, et $\|\Delta_c - \mathrm{T}_\mathrm{N}\| = \|\Delta_{c-c^{(\mathrm{N})}}\|$ est majoré par $\|c - c^{(\mathrm{N})}\|_\infty = \sup_{n>\mathrm{N}} |c_n|$ qui tend vers 0 parce que la suite (c_n) tend vers 0.

2. Pour toute fonction f intégrable sur $[0, 1]$ définissons la fonction continue $\mathrm{V}(f)$ comme dans l'exemple 6.2.2,

$$(\mathrm{V}f)(t) = \int_0^t f(s)\,ds\,;$$

pour tout p tel que $1 \leq p \leq +\infty$, désignons par V_p l'opérateur de $\mathrm{L}_p = \mathrm{L}_p(0,1)$ dans $\mathrm{C}([0,1])$ qui associe à $f \in \mathrm{L}_p$ la fonction continue $\mathrm{V}(f)$; alors V_p est compact lorsque $p > 1$:

> on voit en effet en appliquant Hölder que $|\mathrm{V}(f)(s) - \mathrm{V}(f)(t)| \leq |s-t|^{1/q}$ pour toute $f \in \mathrm{B}_{\mathrm{L}_p}$ (où $1/p + 1/q = 1$), donc $\mathrm{A} = \mathrm{V}_p(\mathrm{B}_{\mathrm{L}_p})$ est borné dans $\mathrm{C}([0,1])$ et équicontinu (ici $q < +\infty$, donc $1/q > 0$ et la fonction $\delta(t) = t^{1/q}$ tend vers 0 avec t), donc A est relativement compact dans $\mathrm{C}([0,1])$ par Ascoli.

On peut voir que V_1 n'est pas compact de L_1 dans $\mathrm{C}([0,1])$.

Proposition 7.1.2. *Soient* E *et* F *deux espaces de Banach ; si* $\mathrm{T} \in \mathcal{L}(\mathrm{E}, \mathrm{F})$ *est compacte, sa transposée* ${}^t\mathrm{T}$ *est compacte de* F^* *dans* E^*.

Démonstration. Soit $\mathrm{K} \subset \mathrm{F}$ un compact qui contienne $\mathrm{T}(\mathrm{B}_\mathrm{E})$; on munit K de la distance induite par F, c'est à dire $d(y_1, y_2) = \|y_1 - y_2\|_\mathrm{F}$. Considérons l'application linéaire $\mathrm{V} : \mathrm{F}^* \to \mathrm{C}(\mathrm{K})$ qui associe à chaque $y^* \in \mathrm{F}^*$ la fonction $\mathrm{V}(y^*) : y \in \mathrm{K} \to y^*(y)$. Si M est le maximum de $\|y\|$ lorsque y varie dans K, on voit que $\|\mathrm{V}(y^*)\|_{\mathrm{C}(\mathrm{K})} \leq \mathrm{M}\, \|y^*\|$, donc V est bornée. Par ailleurs si $x \in \mathrm{B}_\mathrm{E}$, on a $\mathrm{T}(x) \in \mathrm{K}$, donc

$$|({}^t\mathrm{T}(y^*))(x)| = |y^*(\mathrm{T}(x))| = |\mathrm{V}(y^*)(\mathrm{T}(x))| \leq \|\mathrm{V}(y^*)\|_{\mathrm{C}(\mathrm{K})}$$

ce qui montre en prenant le sup sur $x \in \mathrm{B}_\mathrm{E}$ que $\|{}^t\mathrm{T}(y^*)\| \leq \|\mathrm{V}(y^*)\|_{\mathrm{C}(\mathrm{K})}$. Soit $\mathcal{G} \subset \mathrm{C}(\mathrm{K})$ l'ensemble $\mathrm{V}(\mathrm{B}_{\mathrm{F}^*})$, formé de toutes les fonctions sur K de la forme $\mathrm{V}(y^*)$, où y^* varie dans la boule unité de F^*. Cet ensemble $\mathcal{G}$ est uniformément borné et formé de fonctions uniformément lipschitziennes sur (K, d) : on a en effet pour toute fonction $f = \mathrm{V}(y^*) \in \mathcal{G}$, et $y_1, y_2 \in \mathrm{K}$

$$|f(y_1) - f(y_2)| = |y^*(y_1) - y^*(y_2)| = |y^*(y_1 - y_2)| \leq d(y_1, y_2).$$

Il résulte du théorème d'Ascoli que $\mathcal{G}$ est relativement compact dans C(K). Soit maintenant (y_n^*) une suite dans $\mathrm{B}_{\mathrm{F}^*}$, et montrons que la suite $({}^t\mathrm{T}(y_n^*)) \subset \mathrm{E}^*$ admet une sous-suite de Cauchy (en norme) dans E^* ; d'après ce qui précède, il existe une sous-suite $(\mathrm{V}(y_{n_k}^*))$ qui converge uniformément dans C(K), donc qui est de Cauchy dans C(K). Mais $\|{}^t\mathrm{T}(y^*)\| \leq \|\mathrm{V}(y^*)\|_{\mathrm{C}(\mathrm{K})}$, ce qui implique que $({}^t\mathrm{T}(y_{n_k}))$ est de Cauchy dans E^*.

//

Remarque 7.1.3. Si ${}^t\mathrm{T}$ est compacte, il en résulte que T est compacte, donc le résultat précédent est en fait une équivalence ; ceci provient du fait que ${}^t({}^t\mathrm{T})$ est compacte de E^{**} dans F^{**}, et des rapports entre la bitransposée et les injections canoniques dans les biduaux.

Rappelons que la topologie *faible* sur un espace de Banach E est la topologie $\sigma(\mathrm{E}, \mathrm{E}^*)$.

Lemme 7.1.3.

(*i*) *Dans un espace normé toute suite faiblement convergente est bornée.*

(*ii*) *Soit* $(e_n)_{n \geq 0}$ *un système orthonormal dans un espace hilbertien ; alors* e_n *converge faiblement vers* 0.

Démonstration. Soit (x_n) une suite faiblement convergente ; en plongeant isométriquement E dans E^{**} on peut considérer (x_n) comme une suite d'applications linéaires de l'espace de Banach E^* dans $\mathbb{K}$ qui converge en tout point $x^* \in \mathrm{E}^*$; il résulte alors du théorème de Banach-Steinhaus (théorème 3.1.8) que $\{\|x_n\| : n \geq 0\}$ est borné. Si E est un Hilbert et (e_n) un système orthonormal, on a $\sum_{n \geq 0} |\langle e_n, x\rangle|^2 \leq \|x\|^2$ pour tout $x \in \mathrm{E}$ (inégalité de Bessel - théorème 5.5.2) ; la suite $(\langle e_n, x\rangle)$ est de carré sommable donc tend vers 0.

//

Proposition 7.1.4. *Soient* E *et* F *deux espaces de Banach et* $\mathrm{T} \in \mathcal{L}(\mathrm{E}, \mathrm{F})$ *; notons* $\mathrm{B_E}$ *la boule unité fermée de* E.

(i) *Supposons* T *compact ; alors* T *est continu de* $\mathrm{B_E}$*, munie de la topologie faible, dans* F *muni de la topologie de la norme ; en conséquence, pour toute suite* (x_n) *de points de* E *convergeant faiblement vers* 0 *la suite* $(\mathrm{T}(x_n))$ *converge en norme vers* 0.

(ii) *Supposons* E *réflexif ; alors* T *est compact si et seulement si : pour toute suite* (x_n) *de points de* E *convergeant faiblement vers* 0*, la suite* $(\mathrm{T}(x_n))$ *converge en norme vers* 0 *; de plus, l'ensemble* $\mathrm{T}(\mathrm{B_E})$ *est compact (en norme) dans* F *lorsque* T *est compact.*

Démonstration. Supposons T compact, et soit K un compact de F contenant $\mathrm{T}(\mathrm{B_E})$; l'identité de K, muni de la topologie de la norme, dans K muni de la topologie faible est continue ; comme K est compact, c'est un homéomorphisme. Comme T est continu de $\mathrm{B_E}$ muni de la topologie faible dans K muni de la topologie faible, il en résulte que T est continu de $\mathrm{B_E}$ faible dans F muni de la norme. Si (x_n) est une suite qui converge faiblement vers 0 dans E, elle est bornée dans E, donc $(\mathrm{T}(x_n))$ tend vers 0 en norme par ce qui précède.

Lorsque E est réflexif, la boule $\mathrm{B_E}$ est faiblement compacte, donc son image $\mathrm{T}(\mathrm{B_E})$ est faiblement compacte dans F, donc faiblement fermée, donc fermée ; puisque $\mathrm{T}(\mathrm{B_E})$ est relativement compacte, elle est en fait compacte. Supposons encore E réflexif et que $(\mathrm{T}(x_n))$ converge vers 0 en norme dans F pour toute suite (x_n) qui tend faiblement vers 0_E ; soit (x_n) une suite dans $\mathrm{B_E}$; d'après le théorème 4.3.3, il existe une sous-suite (x_{n_k}) qui converge faiblement vers un point $x \in \mathrm{B_E}$; alors $(x_{n_k} - x)$ converge faiblement vers 0, donc $\mathrm{T}(x_{n_k}) - \mathrm{T}(x)$ converge en norme vers 0 d'après l'hypothèse ; on a ainsi montré que pour toute suite $(x_n) \subset \mathrm{B_E}$, il existe une sous-suite $(\mathrm{T}(x_{n_k}))$ qui converge en norme, donc T est compact.

//

Pour tout espace de Banach E et tout $\mathrm{T} \in \mathcal{L}(\mathrm{E}, \mathrm{F})$, le fait que T soit continu de $\mathrm{B_E}$ munie de la topologie faible dans F muni de la norme implique que T est compact ; en revanche la propriété des suites n'est pas suffisante en général (l'application identique de ℓ_1 la vérifie). Supposons T continu de $\mathrm{B_E}$ faible dans F normé, et utilisons seulement la continuité en 0 : pour tout $\varepsilon > 0$, il existe un voisinage faible W de 0 dans $\mathrm{B_E}$ tel que $\mathrm{T}(\mathrm{W}) \subset \mathrm{B}(0_\mathrm{F}, \varepsilon)$; on peut choisir W de la forme

$$\mathrm{W} = \{x \in \mathrm{B_E} : \forall j = 1, \ldots, n, \quad |x_j^*(x)| < \delta\}$$

pour un certain $\delta > 0$ et $x_1^*, \ldots, x_n^* \in \mathrm{E}^*$. Le lecteur utilisera la précompacité des bornés de $\mathbb{K}^n$ pour montrer que cette condition permet de recouvrir $\mathrm{T}(\mathrm{B_E})$ par un nombre fini de boules de rayon ε.

Dans le cas où l'espace de départ est hilbertien, on peut donner des caractérisations plus précises de la compacité.

Théorème 7.1.5. *Soient* E *un espace de Hilbert,* F *un espace de Banach et* $\mathrm{T} \in \mathcal{L}(\mathrm{E}, \mathrm{F})$ *; notons* $\mathrm{B_E}$ *la boule unité fermée de* E*. Les propriétés suivantes sont équivalentes :*

(i) *l'opérateur* T *est compact de* E *dans* F *;*

(ii) *l'ensemble* $\mathrm{T}(\mathrm{B_E})$ *est compact (en norme) dans* F *;*

(iii) *l'opérateur* T *est continu de* $\mathrm{B_E}$*, munie de la topologie faible, dans* F *muni de la topologie de la norme ;*

(*iv*) *pour toute suite* (x_n) *de points de* E *convergeant faiblement vers* 0 *la suite* $(\mathrm{T}(x_n))$ *converge en norme vers* 0 *;*

(*v*) *l'opérateur* T *est adhérent (en norme d'opérateur) à l'espace des applications linéaires continues de rang fini ;*

(*vi*) *pour tout système orthonormal* $(e_n)_{n\geq 0}$ *dans* E *on a* $\lim_n \|\mathrm{T}(e_n)\| = 0$.

Démonstration. Puisque E est réflexif, on sait que (*i*), (*ii*), (*iii*) et (*iv*) sont équivalents. De plus, (*v*) $\Rightarrow$ (*i*) en général.

Supposons que (*v*) ne soit pas vérifiée. Il existe alors $\varepsilon > 0$ tel que pour toute application linéaire continue de rang fini R on ait $\|\mathrm{T} - \mathrm{R}\| > \varepsilon$. Construisons alors par récurrence sur n un système orthonormal $(e_n)_{n\geq 0}$ tel que $\|\mathrm{T}(e_n)\| > \varepsilon$ pour tout $n \geq 0$: comme $\|\mathrm{T}\| > \varepsilon$, il existe $e_0 \in \mathrm{E}$ tel que $\|e_0\| = 1$ et $\|\mathrm{T}(e_0)\| > \varepsilon$; supposons e_k construit pour $k < n$ et soit P le projecteur orthogonal sur le sous-espace de E engendré par $\{e_k : k < n\}$; alors TP est de rang fini donc $\|\mathrm{T} - \mathrm{TP}\| > \varepsilon$; il existe donc $y_n \in \mathrm{E}$ tel que $\|\mathrm{T}(\mathrm{Id}_\mathrm{E} - \mathrm{P})(y_n)\| > \varepsilon \|y_n\| \geq \varepsilon \|(\mathrm{Id}_\mathrm{E} - \mathrm{P})(y_n)\|$; on pose alors $z_n = (\mathrm{Id}_\mathrm{E} - \mathrm{P})(y_n)$, puis $e_n = \|z_n\|^{-1} z_n$. On a alors $\|\mathrm{T}(e_n)\| = \|z_n\|^{-1}\|\mathrm{T}(z_n)\| > \varepsilon$; donc (*vi*) n'est pas vérifiée. On a montré que (*vi*) $\Rightarrow$ (*v*). Enfin, (*iv*) $\Rightarrow$ (*vi*) résulte du lemme 3.

//

Remarque 7.1.4. Il existe des espaces de Banach tels que l'adhérence des opérateurs de rang fini soit strictement plus petite que l'espace des opérateurs compacts (P. Enflo, 1972).

Proposition 7.1.6. *Soient* E, F *deux espaces hilbertiens et* $\mathrm{T} \in \mathcal{L}(\mathrm{E}, \mathrm{F})$ *; on a équivalence entre*

(*i*) *l'application* T *est compacte ;*

(*ii*) *l'application* T^* *est compacte ;*

(*iii*) *l'application* $\mathrm{T}^*\mathrm{T}$ *est compacte ;*

(*iv*) *l'application* $|\mathrm{T}|$ *est compacte.*

Démonstration. (*ii*) $\Rightarrow$ (*iii*) résulte de la proposition 1 ; pour tout système orthonormal $(e_n)_{n\geq 0}$ dans l'espace de Hilbert E on a $\|\mathrm{T}(e_n)\|^2 = \langle \mathrm{T}^*\mathrm{T}(e_n), e_n\rangle \leq \|\mathrm{T}^*\mathrm{T}(e_n)\|$ donc si $\lim \|\mathrm{T}^*\mathrm{T}(e_n)\| = 0$, alors $\lim_n \|\mathrm{T}(e_n)\| = 0$; donc (*iii*) $\Rightarrow$ (*i*) (propriété équivalente (*vi*) du théorème 5). On a montré que (*ii*) $\Rightarrow$ (*i*). Appliquant cela à T^*, on en déduit que (*i*) $\Rightarrow$ (*ii*). Appliquant (*i*) $\iff$ (*iii*) à $|\mathrm{T}|$, on trouve (*iv*) $\iff$ (*iii*).

//

7.2. Théorie spectrale des opérateurs compacts

Cette théorie est pour l'essentiel la création du mathématicien hongrois F. Riesz, aux alentours de 1910. Le théorème 3.4.2 (avec son corollaire) est l'un des points-clés de cette théorie.

Lemme 7.2.1. *Soit* E *un espace de Banach ; pour tout sous-espace vectoriel* L *de dimension finie de* E, *il existe un projecteur continu* P *de* E *sur* L, *c'est à dire qu'il existe un sous-espace fermé* F *tel que* $\mathrm{E} = \mathrm{L} \oplus \mathrm{F}$.

Démonstration. Soit $(e_1, \ldots, e_n)$ une base de L et soit $(e_1^*, \ldots, e_n^*)$ la base duale pour le dual L^* ; par le théorème de Hahn-Banach, on peut prolonger chaque forme linéaire e_j^* en une forme linéaire continue $x_j^* \in \mathrm{E}^*$. Il suffit alors de poser

$$\forall x \in \mathrm{E}, \quad \mathrm{P}(x) = \sum_{j=1}^{n} x_j^*(x)\, e_j,$$

et de poser pour finir $\mathrm{F} = \ker(\mathrm{P})$.

//

Lemme 7.2.2. *Soit* $\mathrm{S} \in \mathcal{L}(\mathrm{E})$ *un opérateur compact, et posons* $\mathrm{T} = \mathrm{Id_E} - \mathrm{S}$ *; si* F *est un sous-espace fermé de* E *tel que* T *soit injectif de* F *dans* E, *il existe une constante* $c > 0$ *telle que* $\|\mathrm{T}(x)\| \ge c\,\|x\|$ *pour tout* $x \in \mathrm{F}$ *; il en résulte que l'image* T(F) *est fermée.*

Démonstration. En cas contraire, on pourrait trouver une suite $(x_n) \subset \mathrm{F}$ de vecteurs de norme 1 telle que $\mathrm{T}(x_n) \to 0$. Puisque S est compact, on peut trouver une sous-suite (x_{n_k}) telle que $\mathrm{S}(x_{n_k})$ converge ; mais $\mathrm{T}(x_{n_k}) = x_{n_k} - \mathrm{S}(x_{n_k})$ tend vers 0, donc x_{n_k} converge vers un vecteur $x \in \mathrm{F}$ (puisque F est fermé) tel que $\|x\| = 1$, et à la limite $\mathrm{T}(x) = 0$, ce qui contredit l'hypothèse T injectif sur F.

Désignons par T_1 la restriction de T à F ; on a vu dans la proposition 6.3.1 que la minoration $\|\mathrm{T}_1(x)\| \ge c\|x\|$ (pour tout $x \in \mathrm{F}$, et avec $c > 0$) implique que $\mathrm{im}(\mathrm{T}_1) = \mathrm{T}(\mathrm{F})$ est fermée.

//

Proposition 7.2.3. *Soit* $\mathrm{S} \in \mathcal{L}(\mathrm{E})$ *un opérateur compact, et posons* $\mathrm{T} = \mathrm{Id_E} - \mathrm{S}$ *; le noyau de* T *est de dimension finie et l'image* T(E) *est fermée.*

On remarquera que $\mathrm{T}^n = (\mathrm{Id_E} - \mathrm{S})^n$ est de la forme $\mathrm{Id_E} - \mathrm{S}_n$, avec S_n compact, en utilisant la formule du binôme et la propriété d'idéal de $\mathcal{K}(\mathrm{E})$, donc les images de T^n sont fermées pour tout $n \ge 0$.

Démonstration. Le noyau de T est le sous-espace propre de l'opérateur compact S pour la valeur propre 1, il est donc de dimension finie d'après le théorème de Riesz (corollaire du théorème 3.4.2). Soit F un sous-espace fermé de E tel que $\mathrm{E} = \ker(\mathrm{T}) \oplus \mathrm{F}$; alors T est injectif sur F, donc $\mathrm{T}(\mathrm{E}) = \mathrm{T}(\mathrm{F})$ est fermé.

//

Lemme 7.2.4. *Soit* $\mathrm{S} \in \mathcal{L}(\mathrm{E})$ *un opérateur compact, et posons* $\mathrm{T} = \mathrm{Id}_\mathrm{E} - \mathrm{S}$ *; il n'existe pas de chaîne infinie* $(\mathrm{F}_n)_{n \geq 0}$ *(resp :* $(\mathrm{F}_n)_{n \leq 0}$*) de sous-espaces vectoriels fermés de* E *telle que*

$$\mathrm{F}_n \subset \mathrm{F}_{n+1},\ \mathrm{F}_n \neq \mathrm{F}_{n+1} \ \ \textit{et} \ \ \mathrm{T}(\mathrm{F}_{n+1}) \subset \mathrm{F}_n$$

pour tout $n \geq 0$ *(resp :* $n < 0$*).*

Démonstration. Traitons le cas $n \geq 0$, le cas $n < 0$ est identique. Supposons au contraire que $\mathrm{F}_n \neq \mathrm{F}_{n+1}$ pour tout $n \geq 0$; on peut trouver pour tout $n \geq 0$ un vecteur $x_{n+1} \in \mathrm{F}_{n+1}$ tel que $\|x_{n+1}\| = 1$ et $\mathrm{dist}(x_{n+1}, \mathrm{F}_n) > 1 - \varepsilon$. Puisque $\mathrm{T}(\mathrm{F}_{n+1}) \subset \mathrm{T}(\mathrm{F}_n)$ et $\mathrm{S} = \mathrm{Id}_\mathrm{E} - \mathrm{T}$, on a $\mathrm{S}(\mathrm{F}_{n+1}) \subset \mathrm{F}_{n+1}$. Soient alors k, ℓ deux entiers tels que $0 < k < \ell$; le vecteur $\mathrm{T}(x_\ell)$ est dans $\mathrm{F}_{\ell-1}$ et $\mathrm{S}(x_k) \in \mathrm{F}_k \subset \mathrm{F}_{\ell-1}$, donc $\mathrm{T}(x_\ell) + \mathrm{S}(x_k) \in \mathrm{F}_{\ell-1}$, donc $\|x_\ell - (\mathrm{T}(x_\ell) + \mathrm{S}(x_k))\| > 1 - \varepsilon$. Mais cette quantité est égale à $\|\mathrm{S}(x_\ell) - \mathrm{S}(x_k)\|$. L'image $\mathrm{S}(\mathrm{B}_\mathrm{E})$ contiendrait donc une suite infinie de points dont les distances mutuelles seraient $\geq 1 - \varepsilon$, ce qui contredirait la compacité de S.

//

Corollaire 7.2.5. *Soit* $\mathrm{S} \in \mathcal{L}(\mathrm{E})$ *un opérateur compact, et posons* $\mathrm{T} = \mathrm{Id}_\mathrm{E} - \mathrm{S}$ *; la suite croissante des noyaux* $(\ker(\mathrm{T}^n))_{n \geq 0}$ *est stationnaire. La suite décroissante des images* $(\mathrm{im}(\mathrm{T}^n))_{n \geq 0}$ *est stationnaire.*

Démonstration. Posons $\mathrm{F}_n = \ker(\mathrm{T}^n)$. On a bien F_n fermé, $\mathrm{F}_n \subset \mathrm{F}_{n+1}$ et de plus $\mathrm{T}(\mathrm{F}_{n+1}) \subset \mathrm{F}_n$ pour tout $n \geq 0$; si la suite n'était pas stationnaire, elle contredirait le lemme précédent. Pour le cas des images on posera $\mathrm{F}_{-n} = \mathrm{im}(\mathrm{T}^n)$; on a vu à la proposition 3 que toutes ces images sont fermées.

//

Corollaire 7.2.6. *Soit* $\mathrm{S} \in \mathcal{L}(\mathrm{E})$ *un opérateur compact, et posons* $\mathrm{T} = \mathrm{Id}_\mathrm{E} - \mathrm{S}$ *; si* T *est surjectif, alors* $\ker(\mathrm{T}) = \{0\}$ *; si* T *est injectif, alors* $\mathrm{im}(\mathrm{T}) = \mathrm{E}$.

Démonstration. Si l'opérateur T est surjectif et si $\ker(\mathrm{T}) \neq \{0\}$, on montre par récurrence que $\ker(\mathrm{T}^n) \neq \ker(\mathrm{T}^{n+1})$ pour tout $n \geq 1$: si $x \in \ker(\mathrm{T}^{n+1}) \setminus \ker(\mathrm{T}^n)$, on a $\mathrm{T}^{n+1}(x) = 0$ et $\mathrm{T}^n(x) \neq 0$. Puisque T est surjectif, il existe y tel que $\mathrm{T}(y) = x$. Il en résulte que $\mathrm{T}^{n+2}(y) = \mathrm{T}^{n+1}(x) = 0$ mais $\mathrm{T}^{n+1}(y) = \mathrm{T}^n(x) \neq 0$. Ceci est impossible quand $\mathrm{T} = \mathrm{Id}_\mathrm{E} - \mathrm{S}$, avec S compact, par le corollaire précédent.

Si T est injectif et $\mathrm{T}(\mathrm{E}) \neq \mathrm{E}$, on vérifie que $\mathrm{im}(\mathrm{T}^{n+1}) \neq \mathrm{im}(\mathrm{T}^n)$ pour tout $n \geq 0$, ce qui est à nouveau impossible quand $\mathrm{T} = \mathrm{Id}_\mathrm{E} - \mathrm{S}$, avec S compact.

//

Si F est un sous-espace vectoriel fermé de E, on appelle *codimension* de F la dimension du quotient E/F (finie ou $+\infty$). Si F est de codimension finie n, on peut trouver un sous-espace vectoriel G de dimension n tel que $\mathrm{E} = \mathrm{F} \oplus \mathrm{G}$, et pour tout sous-espace G' tel que $\dim(\mathrm{G}') > n$, on a $\mathrm{F} \cap \mathrm{G}' \neq \{0\}$.

Théorème 7.2.7 : Alternative de Fredholm. *Soient* E *un espace de Banach complexe et* $\mathrm{S} \in \mathcal{K}(\mathrm{E})$ *;*

(*i*) *pour tout* $\lambda \in \mathbb{C} \setminus \{0\}$*, l'image de* $\lambda \operatorname{Id}_{\mathrm{E}} - \mathrm{S}$ *est fermée et de codimension finie et l'on a*

$$\operatorname{codim} \operatorname{im}(\lambda \operatorname{Id}_{\mathrm{E}} - \mathrm{S}) = \dim \ker(\lambda \operatorname{Id}_{\mathrm{E}} - \mathrm{S})\,;$$

(*ii*) *le spectre* $\mathrm{Sp(S)}$ *est fini ou formé d'une suite tendant vers* 0.

Pour un opérateur T à image fermée et à noyau de dimension finie, la différence $\dim \ker(\mathrm{T}) - \operatorname{codim} \operatorname{im}(\mathrm{T})$ s'appelle *l'indice* de l'opérateur T et se note $\operatorname{ind}(\mathrm{T})$. La première partie du théorème dit que $\mathrm{S} - \lambda \operatorname{Id}_{\mathrm{E}}$ est d'indice nul pour tout $\lambda \neq 0$.

Démonstration. En remplaçant S par S/λ on se ramène à $\lambda = 1$, et à étudier l'opérateur $\mathrm{T} = \lambda \operatorname{Id}_{\mathrm{E}} - \mathrm{S}$, égal à $\operatorname{Id}_{\mathrm{E}} - \mathrm{S}$ dans ce cas : on a vu que $\ker(\mathrm{T})$ est de dimension finie et $\operatorname{im}(\mathrm{T})$ fermée. On doit montrer de plus que $\dim \ker(\mathrm{T}) = \operatorname{codim} \mathrm{T(E)}$, c'est à dire que l'indice de T est nul. On va procéder par récurrence sur la dimension de $\ker(\mathrm{T})$. Si $\dim \ker(\mathrm{T}) = 0$, on sait que T est surjectif d'après le corollaire 6, donc l'indice est nul dans ce cas ; on suppose donc que n est un entier > 0 et que $\operatorname{ind}(\mathrm{T}') = 0$ pour tout opérateur $\mathrm{T}' = \operatorname{Id}_{\mathrm{E}} - \mathrm{S}'$, où S' est compact et $\dim \ker(\mathrm{T}') < n$. Soit $\mathrm{T} = \operatorname{Id}_{\mathrm{E}} - \mathrm{S}$ avec S compact et $\dim \ker(\mathrm{T}) = n > 0$; d'après le corollaire 6, on a $\operatorname{im}(\mathrm{T}) \neq \mathrm{E}$; soit donc $y_0 \notin \operatorname{im}(\mathrm{T})$, et soit $x_0 \in \ker(\mathrm{T})$, x_0 non nul ; on va construire T' de la forme voulue tel que $\operatorname{ind}(\mathrm{T}') = \operatorname{ind}(\mathrm{T})$ et $\dim \ker(\mathrm{T}') < \dim \ker(\mathrm{T})$; d'après l'hypothèse de récurrence, on aura $0 = \operatorname{ind}(\mathrm{T}') = \operatorname{ind}(\mathrm{T})$, d'où le résultat.

Soit $x_0^* \in \mathrm{E}^*$ telle que $x_0^*(x_0) = 1$ et posons $\mathrm{T}'(x) = \mathrm{T}(x) + x_0^*(x)\, y_0$ pour tout $x \in \mathrm{E}$. L'opérateur T' est obtenu en ajoutant à T l'opérateur de rang un $\mathrm{R} : x \to x_0^*(x)\, y_0$, donc $\mathrm{T}' = \operatorname{Id}_{\mathrm{E}} - \mathrm{S}'$ avec $\mathrm{S}' = \mathrm{S} - \mathrm{R}$ compact. Déterminons le noyau de T' ; la relation $\mathrm{T}'(x) = 0$ entraîne $\mathrm{T}(x) = 0$ et $x_0^*(x) = 0$ (parce que $y_0 \notin \operatorname{im}(\mathrm{T})$) ; le noyau de T' est donc le sous-espace de $\ker(\mathrm{T})$ défini par l'équation $x_0^*(x) = 0$, qui élimine x_0 du noyau de T' et montre que $\dim \ker(\mathrm{T}') = \dim \ker(\mathrm{T}) - 1$. On a donc $\operatorname{ind}(\mathrm{T}') = 0$ d'après l'hypothèse de récurrence, ce qui montre déjà que $\operatorname{codim} \operatorname{im}(\mathrm{T}')$ est finie. Il est clair que $\operatorname{im}(\mathrm{T}') \subset \operatorname{im}(\mathrm{T}) \oplus \mathbb{K} y_0$; de plus $\mathrm{T}'(x_0) = y_0$ (puisque $\mathrm{T}(x_0) = 0$), et pour tout $x \in \mathrm{E}$ on observe que $x' = x - x_0^*(x) x_0$ annule x_0^*, donc $\mathrm{T}'(x') = \mathrm{T}(x') = \mathrm{T}(x)$, ce qui montre que $\operatorname{im}(\mathrm{T}')$ contient $\operatorname{im}(\mathrm{T})$. Finalement $\operatorname{im}(\mathrm{T}') = \operatorname{im}(\mathrm{T}) \oplus \mathbb{K} y_0$, donc $\operatorname{codim} \operatorname{im}(\mathrm{T}) = \operatorname{codim} \operatorname{im}(\mathrm{T}') + 1$, et $\operatorname{ind}(\mathrm{T}) = \operatorname{ind}(\mathrm{T}') = 0$.

On va maintenant montrer que si 1 est dans le spectre de S, alors 1 est valeur propre et 1 est isolé dans le spectre de S. Si 1 n'est pas valeur propre de S, l'opérateur T est injectif, donc surjectif d'après le corollaire 6, donc $\operatorname{Id}_{\mathrm{E}} - \mathrm{S}$ est inversible et 1 n'est pas dans le spectre de S. Posons $\mathrm{T} = \operatorname{Id}_{\mathrm{E}} - \mathrm{S}$; on a vu que $\dim \ker(\mathrm{T}) = \operatorname{codim} \operatorname{im}(\mathrm{T})$. Remarquons que $\mathrm{T}^k = (\operatorname{Id}_{\mathrm{E}} - \mathrm{S})^k = \operatorname{Id}_{\mathrm{E}} - \mathrm{S}_k$ avec S_k compact (utiliser la formule du binôme), donc $\dim \ker(\mathrm{T}^n) = \operatorname{codim} \operatorname{im}(\mathrm{T}^n)$ pour tout $n \geq 0$ (pour $n = 0$, c'est une évidence). On a vu qu'il existe un entier k tel que $\ker(\mathrm{T}^k) = \ker(\mathrm{T}^{k+1})$, et on peut prendre pour k le plus petit entier vérifiant cette propriété ; on a $k \geq 1$ puisque 1 est valeur propre de S ; alors $\ker(\mathrm{T}^k) \cap \operatorname{im}(\mathrm{T}) = \{0\}$, sinon $\ker(\mathrm{T}^k) \neq \ker(\mathrm{T}^{k+1})$; on a *a fortiori* $\ker(\mathrm{T}^k) \cap \operatorname{im}(\mathrm{T}^k) = \{0\}$, et d'après l'égalité dimension-codimension il en résulte que

$$\mathrm{E} = \ker(\mathrm{T}^k) \oplus \operatorname{im}(\mathrm{T}^k).$$

L'espace E se trouve décomposé en deux sous-espaces fermés T-invariants. La restriction T_2 de T à $\operatorname{im}(\mathrm{T}^k)$ est injective, donc c'est un isomorphisme de $\operatorname{im}(\mathrm{T}^k)$ sur $\operatorname{im}(\mathrm{T}^k)$ d'après le point (*i*). La restriction T_1 de T à $\ker(\mathrm{T}^k)$ est un endomorphisme en dimension finie, dont la seule valeur propre est 0 ; pour tout $\lambda \neq 0$, $\mathrm{T}_1 - \lambda$ est donc bijective de $\ker(\mathrm{T}^k)$ sur $\ker(\mathrm{T}^k)$, et pour λ assez petit, $\mathrm{T}_2 - \lambda$ est encore un isomorphisme ; il en résulte que $\mathrm{T} - \lambda$ est un isomorphisme pour $\lambda \neq 0$ et assez petit, ce qui signifie que 0 est isolé dans le spectre de T, ou encore que 1 est isolé dans le spectre de S. On en déduit que pour tout $\varepsilon > 0$ il y a un nombre fini de valeurs spectrales telles que $|\lambda| \geq \varepsilon$, autrement dit que toute suite (λ_n) de valeurs spectrales distinctes tend vers 0.

//

On peut donner une démonstration courte mais un peu artificielle du point (ii) du théorème précédent. Si $(\lambda_n)_{n\geq 0}$ était une suite de valeurs propres de S distinctes de λ et qui converge vers $\lambda \neq 0$, on aurait pour tout $n \geq 0$ un vecteur x_n de norme un tel que $\mathrm{S}(x_n) = \lambda_n\, x_n$. Soit F le sous-espace fermé engendré par la suite $(x_n)_{n\geq 0}$; il est clair que F est S-invariant, ce qui permet de considérer la restriction S' de S à F. Alors $x_n = (\lambda - \lambda_n)^{-1}(\lambda\,\mathrm{Id}_\mathrm{F} - \mathrm{S}')(x_n)$ pour tout $n \geq 0$, ce qui montre que $\mathrm{T}' = \lambda\,\mathrm{Id}_\mathrm{F} - \mathrm{S}'$ a une image dense dans F (l'image contient tous les vecteurs $(x_n)_{n\geq 0}$), donc égale à F. Il en résulte que T' est un isomorphisme, donc $\lambda \notin \mathrm{Sp}(\mathrm{S}')$, ce qui est impossible puisque $\mathrm{Sp}(\mathrm{S}')$ est fermé et contient les (λ_n).

Formulation classique de l'alternative de Fredholm. A l'époque de l'article de Fredholm (1903), il n'y avait pas plus d'espaces de Banach que de théorie de Riesz des opérateurs compacts. Cependant, quelques années après, sous l'influence de F. Riesz, on est arrivé à peu de chose près à la formulation “classique” suivante : soit S un opérateur compact de E. On rappelle que ${}^t\mathrm{S}$ est compacte de E^* dans E^*. On a *l'alternative* suivante :

– ou bien les deux équations $x - \mathrm{S}(x) = y$, $x^* - {}^t\mathrm{S}(x^*) = y^*$ admettent pour tous seconds membres $y \in \mathrm{E}$, $y^* \in \mathrm{E}^*$ une solution unique $x \in \mathrm{E}$, $x^* \in \mathrm{E}^*$.

– ou bien les équations homogènes $x - \mathrm{S}(x) = 0$, $x^* - {}^t\mathrm{S}(x^*) = 0$ admettent un même nombre fini $k > 0$ de solutions indépendantes, $x_1, \ldots, x_k$ et $x_1^*, \ldots, x_k^*$. Dans ce cas, pour que l'équation $x - \mathrm{S}(x) = y$ admette une solution $x \in \mathrm{E}$, il faut et il suffit que $x_1^*(y) = x_2^*(y) = \cdots = x_k^*(y) = 0$, et pour que l'équation $x^* - {}^t\mathrm{S}(x^*) = y^*$ admette une solution $x^* \in \mathrm{E}^*$, il faut et il suffit que $y^*(x_1) = y^*(x_2) = \cdots = y^*(x_k) = 0$.

Pour ce point de vue classique, on pourra consulter le livre de F. Riesz (Leçons d'Analyse Fonctionnelle).

Théorème 7.2.8. *Toute application linéaire compacte normale d'un espace de Hilbert complexe* H *admet une base hilbertienne formée de vecteurs propres.*

Démonstration. Soient H un espace de Hilbert complexe et $\mathrm{T} \in \mathcal{L}(\mathrm{H})$ une application linéaire compacte normale ; soit K son spectre. Pour $\lambda \in \mathrm{K}$ notons E_λ l'espace propre de T associé. On doit démontrer que :

– les E_λ sont deux à deux orthogonaux ;
– le sous-espace engendré par les E_λ est dense.

Alors si B_λ est une base hilbertienne de E_λ, la famille $\bigcup_{\lambda\in\mathrm{S}} \mathrm{B}_\lambda$ sera la base voulue. Si $x \in \mathrm{E}_\lambda$, comme $\mathrm{T}\mathrm{T}^*(x) = \mathrm{T}^*\mathrm{T}(x) = \lambda\mathrm{T}^*(x)$ on trouve $\mathrm{T}^*(x) \in \mathrm{E}_\lambda$. Pour tout $y \in \mathrm{E}_\lambda$ on a $\langle \mathrm{T}^*(x), y\rangle = \langle x, \mathrm{T}(y)\rangle = \langle \overline{\lambda} x, y\rangle$; donc $\mathrm{T}^*(x) - \overline{\lambda} x \in \mathrm{E}_\lambda \cap \mathrm{E}_\lambda^\perp$ donc $\mathrm{T}^*(x) = \overline{\lambda} x$. Si $x \in \mathrm{E}_\lambda$ et $y \in \mathrm{E}_\mu$ alors $\langle \mathrm{T}(x), y\rangle = \mu\langle x, y\rangle = \langle x, \mathrm{T}^* y\rangle = \lambda\langle x, y\rangle$ ce qui montre que les sous-espaces propres sont deux à deux orthogonaux.

Notons F le sous-espace fermé de H engendré par les E_λ, pour $\lambda \in \mathrm{K} \setminus \{0\}$. Montrons que H est la somme du noyau E_0 de T et de F. On a $\mathrm{T}(\mathrm{F}) \subset \mathrm{F}$ et

$T^*(F) \subset F$. Il s'ensuit que $T(F^\perp) \subset F^\perp$ et $T^*(F^\perp) \subset F^\perp$. Notons $T_1 \in \mathcal{L}(F^\perp)$ la restriction de T. Alors T_1^* est la restriction de T^*, donc T_1 est normal. Remarquons que T_1 est compacte, qu'elle n'a pas de valeur propre non nulle ; par l'alternative de Fredholm, T_1 n'a pas de valeur spectrale non nulle ; par la proposition 6.2.3, $T_1 = 0$, donc $F^\perp = E_0$, d'où le résultat. Pour obtenir une base orthonormée de H, on rassemble des bases orthonormées de chaque espace E_λ, qui sont des bases finies, et éventuellement, une base orthonormée du noyau E_0.

//

Remarque 7.2.1. Il s'agit ici d'un théorème qui demande que le corps de base soit $\mathbb{C}$; déjà en dimension réelle deux, une matrice normale 2×2 n'est pas forcément diagonalisable sur $\mathbb{R}$ (prendre tout simplement une rotation d'angle différent de $k\pi$) ; on peut cependant décomposer l'espace réel en sous-espaces invariants de dimension ≤ 2. Si H est un espace de Hilbert réel et $T \in \mathcal{L}(H)$ un opérateur autoadjoint compact, il existe une base orthonormée de H formée de vecteurs propres de T.

7.3. Opérateurs de Hilbert-Schmidt

Lemme 7.3.1. *Soient* E *et* F *deux espaces hilbertiens,* B *une base hilbertienne de* E *et* B' *une base hilbertienne de* F *; pour tout* $T \in \mathcal{L}(E, F)$ *on a :*

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

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

Démonstration. Pour $x \in E$ et $y \in F$ on a

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

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

//

Pour $T \in \mathcal{L}(E, F)$ on pose $\|T\|_2 = \left(\sum_{b \in B} \|T(b)\|^2\right)^{1/2}$ où B est une base hilbertienne de E. Posons $\mathcal{L}^2(E, F) = \{T \in \mathcal{L}(E, F) : \|T\|_2 < +\infty\}$.

Théorème 7.3.2. *Soient* E *et* F *deux espaces de Hilbert ;*

(i) *l'ensemble* $\mathcal{L}^2(E, F)$ *est un sous-espace vectoriel de* $\mathcal{L}(E, F)$ *;*

(ii) *pour tous opérateurs* $S, T \in \mathcal{L}^2(E, F)$ *et toute base hilbertienne* B *de* E*, la famille* $(\langle S(b), T(b) \rangle)_{b \in B}$ *est sommable ; l'application* $(S, T) \to \sum_{b \in B} \langle S(b), T(b) \rangle$ *est un produit scalaire sur* $\mathcal{L}^2(E, F)$*, indépendant de la base* B.

On note $(S, T) \to (S, T)_2$ *ce produit scalaire ;*

(iii) *muni de ce produit scalaire,* $\mathcal{L}^2(E, F)$ *est un espace de Hilbert ;*

(iv) *on a* $\mathcal{L}^2(E, F) \subset \mathcal{K}(E, F)$ *;*

(v) *soit* $T \in \mathcal{K}(E, F)$ *; notons* $(\lambda_n)_{n \geq 0}$ *les valeurs propres de* $|T|$ *(qui est compact par la proposition 1.6) comptées avec leur multiplicité. Alors* $\|T\|_2^2 = \sum_{n=0}^{+\infty} \lambda_n^2$.

Démonstration. Le point (i) est évident. Soient $\mathrm{S}, \mathrm{T} \in \mathcal{L}(\mathrm{E}, \mathrm{F})$ et B une base hilbertienne de E ; pour $b \in \mathrm{B}$ on a

$$|\langle \mathrm{S}(b), \mathrm{T}(b)\rangle| \leq \|\mathrm{S}(b)\| \, \|\mathrm{T}(b)\| \leq \frac{1}{2} \left(\|\mathrm{S}(b)\|^2 + \|\mathrm{T}(b)\|^2\right).$$

On en déduit que la famille $(\langle \mathrm{S}(b), \mathrm{T}(b)\rangle)_{b \in \mathrm{B}}$ est sommable. Il est clair que l'application $(\mathrm{S}, \mathrm{T}) \to \sum_{b \in \mathrm{B}} \langle \mathrm{S}(b), \mathrm{T}(b)\rangle$ est un produit scalaire, indépendant de la base d'après le lemme précédent et l'identité de polarisation de la proposition 5.1.1.

Remarquons que, pour tout $\mathrm{T} \in \mathcal{L}^2(\mathrm{E}, \mathrm{F})$ et tout $x \in \mathrm{E}$ de norme 1, prenant une base hilbertienne contenant x, on a $\|\mathrm{T}\|_2 \geq \|\mathrm{T}(x)\|$; ceci ayant lieu pour tout x il en résulte que $\|\mathrm{T}\|_2 \geq \|\mathrm{T}\|$, donc $\| \; \|_2$ est une norme sur $\mathcal{L}^2(\mathrm{E}, \mathrm{F})$. Du point (ii) il résulte alors que $\mathcal{L}^2(\mathrm{E}, \mathrm{F})$ est un espace préhilbertien ; on doit montrer qu'il est de plus complet. Soit (T_n) une suite de Cauchy dans $\mathcal{L}^2(\mathrm{E}, \mathrm{F})$; comme $\| \; \| \leq \| \; \|_2$, la suite (T_n) est de Cauchy dans $\mathcal{L}(\mathrm{E}, \mathrm{F})$ qui est complet, donc la suite (T_n) converge en norme vers un opérateur T. Pour tout ensemble fini $\mathrm{I} \subset \mathrm{B}$ on aura

$$\sum_{b \in \mathrm{I}} \|\mathrm{T}(b)\|^2 = \lim_n \sum_{b \in \mathrm{I}} \|\mathrm{T}_n(b)\|^2 \leq \sup_n \|\mathrm{T}_n\|_2^2 = \mathrm{M},$$

donc $\sum_{b \in \mathrm{B}} \|\mathrm{T}(b)\|^2 \leq \mathrm{M} < +\infty$ et $\mathrm{T} \in \mathcal{L}^2(\mathrm{E}, \mathrm{F})$. Comme (T_n) est de Cauchy dans $\mathcal{L}^2(\mathrm{E}, \mathrm{F})$, il existe un entier N tel que pour tout $n \geq \mathrm{N}$ on ait $\|\mathrm{T}_n - \mathrm{T}_\mathrm{N}\|_2 \leq \varepsilon$. L'argument précédent appliqué à la suite $(\mathrm{T}_n - \mathrm{T}_\mathrm{N})_{n \geq \mathrm{N}}$, qui converge vers $\mathrm{T} - \mathrm{T}_\mathrm{N}$ montre que $\|\mathrm{T} - \mathrm{T}_\mathrm{N}\|_2 \leq \sup_{n \geq \mathrm{N}} \|\mathrm{T}_n - \mathrm{T}_\mathrm{N}\|_2 \leq \varepsilon$, d'où le résultat.

Soient $\mathrm{T} \in \mathcal{L}^2(\mathrm{E}, \mathrm{F})$ et (e_n) un système orthonormal ; soit B une base contenant les vecteurs e_n ; la famille $(\|\mathrm{T}(b)\|^2)$ est sommable, donc la suite $\|\mathrm{T}(e_n)\|$ tend vers 0. Donc T est compact par la caractérisation (vi) du théorème 1.5. Le point (iv) est clair si on choisit une base B formée de vecteurs propres pour $|\mathrm{T}|$.

//

Définition 7.3.1. Soient E et F deux espaces hilbertiens ; un opérateur $\mathrm{T} \in \mathcal{L}^2(\mathrm{E}, \mathrm{F})$ est dit de *Hilbert-Schmidt.*

Exemples 7.3.2.

1. Prenons d'abord $\mathrm{E} = \mathrm{F} = \mathbb{C}^n$. Un opérateur T est représenté par une matrice $(a_{i,j})$ dans la base canonique ; si (e_j) désigne la base canonique, 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

$$\|\mathrm{T}\|_2 = \Big(\sum_{i,j=1}^n |a_{i,j}|^2 \Big)^{1/2}.$$

Si $\mathrm{E} = \mathrm{F} = \ell_2$, 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

$$\|\mathrm{T}\|_2 = \Big(\sum_{i,j=0}^{+\infty} |a_{i,j}|^2 \Big)^{1/2}.$$

2. Soient (X, μ) et (Y, ν) deux espaces mesurés et $\mathrm{K}(s,t)$ une fonction de carré intégrable sur $\mathrm{X} \times \mathrm{Y}$. On définit un opérateur T_K par

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

On montre que $\mathrm{T_K}$ est bien défini, et agit continument de $\mathrm{L_2(Y,\nu)}$ dans $\mathrm{L_2(X,\mu)}$:

avec Cauchy-Schwarz, on a

$$|\mathrm{T_K} f(s)|^2 \le \Big(\int_{\mathrm{Y}} |\mathrm{K}(s,t)|\,|f(t)|\,d\nu(t)\Big)^2 \le \Big(\int_{\mathrm{Y}} |\mathrm{K}(s,t)|^2\,d\nu(t)\Big)\Big(\int_{\mathrm{Y}} |f(t)|^2\,d\nu(t)\Big)$$

ce qui donne en réintégrant

$$\int_{\mathrm{X}} |\mathrm{T_K} f(s)|^2\,d\mu(s) \le \Big(\int_{\mathrm{X\times Y}} |\mathrm{K}(s,t)|^2\,d\mu(s)\,d\nu(t)\Big)\Big(\int_{\mathrm{Y}} |f(t)|^2\,d\nu(t)\Big).$$

On trouve *a posteriori* que l'intégrale qui définit $\mathrm{T_K}$ est absolument convergente pour presque tout s, et on voit que $\|\mathrm{T_K}\| \le \|\mathrm{K}\|_2$.

Si $(f_n)_{n\ge 0}$ est une base hilbertienne de $\mathrm{L_2(X,\mu)}$ et $(g_n)_{n\ge 0}$ une base hilbertienne de $\mathrm{L_2(Y,\nu)}$, il en résulte que les fonctions $(s,t)\to f_m(s)g_n(t)$ (où m,n prennent toutes les valeurs entières ≥ 0) donnent une base orthonormée de l'espace $\mathrm{L_2(X\times Y,\mu\otimes\nu)}$. Si $\mathrm{K_N}$ est de la forme $\sum_{m,n=0}^{\mathrm{N}} a_{m,n} f_m(s)g_n(t)$, on voit facilement que $\mathrm{T_{K_N}}$ est de rang fini (l'image est contenue dans $\mathrm{Vect}(f_0,\dots,f_{\mathrm{N}})$). Si K_n tend vers K en norme $\mathrm{L_2}$, il en résulte que les opérateurs de rang fini $\mathrm{T_{K_{\mathit{n}}}}$ convergent en norme d'opérateur vers $\mathrm{T_K}$, qui est donc compact. En fait, l'opérateur $\mathrm{T_K}$ est de Hilbert-Schmidt :

si on écrit

$$\mathrm{K}(s,t) = \sum_{m,n=0}^{+\infty} c_{m,n} f_m(s)\overline{g_n(t)},$$

on constate que $\mathrm{T_K}(g_p) = \sum_m c_{m,p} f_m$, donc $\|\mathrm{T_K}(g_p)\|^2 = \sum_m |c_{m,p}|^2$, et ensuite

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

On peut vérifier que l'adjoint de $\mathrm{T_K}$ est l'opérateur de noyau $\mathrm{K}^*(t,s) = \overline{\mathrm{K}(s,t)}$. Supposons que X = Y et que K soit un *noyau hermitien*, c'est à dire que $\mathrm{K}(t,s) = \overline{\mathrm{K}(s,t)}$; il existe alors une base orthonormée (f_n) telle que $\mathrm{T_K}(f_n) = \lambda_n f_n$ pour tout $n\ge 0$. Si on exprime $\mathrm{K}(s,t) = \sum_{m,n} c_{m,n} f_m(s)\overline{f_n(t)}$, on voit que

$$\mathrm{T_K}(f_p)(s) = \sum_{m,n} c_{m,n}\int f_m(s)\overline{f_n(t)} f_p(t)\,dt = \sum_m c_{m,p} f_m(s) = \lambda_p f_p(s),$$

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

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

où les λ_n sont réels, et (f_n) une base orthonormée.

Exercice 7.3.3. Réciproquement, soit T un opérateur de Hilbert-Schmidt de $\mathrm{L_2(Y)}$ dans $\mathrm{L_2(X)}$, et soit (g_n) une base hilbertienne de $\mathrm{L_2(Y)}$. Soit $\mathrm{F}_n = \mathrm{T}(g_n)$. Montrer que

$$\mathrm{K}(s,t) = \sum_{n=0}^{+\infty} \mathrm{F}_n(s)\overline{g_n(t)}$$

définit une fonction de $\mathrm{L_2(X\times Y)}$, et $\mathrm{T} = \mathrm{T_K}$.

Proposition 7.3.3. *Soient* E*,* F *et* H *des espaces hilbertiens ; pour tout* $\mathrm{S} \in \mathcal{L}(\mathrm{E}, \mathrm{F})$ *et* $\mathrm{T} \in \mathcal{L}(\mathrm{F}, \mathrm{H})$ *on a :*

(*i*) $\|\mathrm{S}\|_2 = \|\mathrm{S}^*\|_2$ *;*

(*ii*) $\|\mathrm{TS}\|_2 \leq \|\mathrm{T}\| \, \|\mathrm{S}\|_2$ *et* $\|\mathrm{TS}\|_2 \leq \|\mathrm{T}\|_2 \|\mathrm{S}\|$ *;*

(*iii*) *si* S *ou* T *est un opérateur de Hilbert-Schmidt alors il en va de même pour* TS*.*

Démonstration. Le point (*i*) résulte de la définition de $\|\mathrm{S}\|_2$ (lemme 1). Pour le point (*ii*), soit B une base hilbertienne de E ; pour tout $b \in \mathrm{B}$ on a $\|\mathrm{TS}(b)\| \leq \|\mathrm{T}\| \, \|\mathrm{S}(b)\|$ donc $\|\mathrm{TS}\|_2^2 = \sum_{b \in \mathrm{B}} \|\mathrm{TS}(b)\|^2 \leq \|\mathrm{T}\|^2 \sum_{b \in \mathrm{B}} \|\mathrm{S}(b)\|^2 = \|\mathrm{T}\| \, \|\mathrm{S}\|_2$. La deuxième assertion en résulte en remplaçant S et T par leurs adjoints. Le point (*iii*) résulte aussitôt de (*ii*).

//

En particulier l'espace $\mathcal{L}^2(\mathrm{E}) = \mathcal{L}^2(\mathrm{E}, \mathrm{E})$ est un idéal bilatère de $\mathcal{L}(\mathrm{E})$.

7.4. Opérateurs nucléaires

Proposition 7.4.1. *Soit* E *un espace de Hilbert ;*

(*i*) *soit* $\mathrm{T} \in \mathcal{L}(\mathrm{E})_+$ *; la quantité (finie ou égale à* $+\infty$*)* $\mathrm{Tr}(\mathrm{T}) = \sum_{b \in \mathrm{B}} \langle \mathrm{T}(b), b \rangle$ *ne dépend pas de la base hilbertienne* B *de l'espace* E *;*

(*ii*) *pour tous* $\mathrm{S}, \mathrm{T} \in \mathcal{L}(\mathrm{E})_+$ *et tout* $\lambda \geq 0$ *on a* $\mathrm{Tr}(\mathrm{S} + \mathrm{T}) = \mathrm{Tr}(\mathrm{S}) + \mathrm{Tr}(\mathrm{T})$ *et* $\mathrm{Tr}(\lambda \mathrm{S}) = \lambda \, \mathrm{Tr}(\mathrm{S})$ *;*

(*iii*) *soient* F *un espace hilbertien,* $\mathrm{U} \in \mathcal{L}(\mathrm{E}, \mathrm{F})$ *un opérateur unitaire et* $\mathrm{T} \in \mathcal{L}(\mathrm{E})_+$ *; alors* $\mathrm{Tr}(\mathrm{U}\,\mathrm{T}\,\mathrm{U}^*) = \mathrm{Tr}(\mathrm{T})$ *;*

(*iv*) *si* $\mathrm{T} \in \mathcal{L}(\mathrm{E})_+$ *est compact,* $\mathrm{Tr}(\mathrm{T})$ *est la somme des valeurs propres de* T *comptées avec leur multiplicité.*

Démonstration. Ecrivons $\mathrm{T} = \mathrm{S}^*\mathrm{S}$ alors $\mathrm{Tr}(\mathrm{T}) = \|\mathrm{S}\|_2^2$ ne dépend pas de la base, d'où (*i*). Le point (*ii*) est clair. Soit B une base hilbertienne de E ; alors U(B) est une base hilbertienne de F. On a $\sum_{b \in \mathrm{U(B)}} \langle \mathrm{UTU}^*(b), b \rangle = \sum_{b \in \mathrm{B}} \langle \mathrm{T}(b), b \rangle$, d'où (*iii*). Enfin, (*iv*) est clair si on choisit une base B formée de vecteurs propres pour T.

//

Soient E, F deux espaces hilbertiens ; pour $\mathrm{T} \in \mathcal{L}(\mathrm{E}, \mathrm{F})$ posons $\|\mathrm{T}\|_1 = \mathrm{Tr}(|\mathrm{T}|)$ et $\mathcal{L}^1(\mathrm{E}, \mathrm{F}) = \{\mathrm{T} \in \mathcal{L}(\mathrm{E}, \mathrm{F}) : \|\mathrm{T}\|_1 < +\infty\}$.

Lemme 7.4.2. *Soit* $\mathrm{T} \in \mathcal{L}(\mathrm{E}, \mathrm{F})$ *;*

(*i*) *soient* H *un espace de Hilbert et* $\mathrm{S} \in \mathcal{L}(\mathrm{E}, \mathrm{H})$ *tels que* $|\mathrm{T}| = \mathrm{S}^*\mathrm{S}$ *; on a alors l'égalité* $\|\mathrm{T}\|_1 = \|\mathrm{S}\|_2^2$ *;*

(*ii*) *on a* $\|\mathrm{T}\|_1 = \sup\{|(\mathrm{TR}, \mathrm{S})_2| : \mathrm{R} \in \mathcal{L}^2(\mathrm{E}), \, \mathrm{S} \in \mathcal{L}^2(\mathrm{E}, \mathrm{F}), \, \|\mathrm{R}\| \, \|\mathrm{S}\| \leq 1\}$.

Démonstration. Le point (*i*) est clair. Si $\mathrm{T} \in \mathcal{L}^1(\mathrm{E}, \mathrm{F})$, alors $|\mathrm{T}|^{1/2} \in \mathcal{L}^2(\mathrm{E})$; soient $\mathrm{R} \in \mathcal{L}^2(\mathrm{E})$ et $\mathrm{S} \in \mathcal{L}^2(\mathrm{E}, \mathrm{F})$ tels que $\|\mathrm{S}\| \, \|\mathrm{R}\| \leq 1$; soit B une base hilbertienne de E ; on a $(\mathrm{TR}, \mathrm{S})_2 = \sum_{\mathrm{B}} \langle \mathrm{TR}(b), \mathrm{S}(b) \rangle = (|\mathrm{T}|^{1/2}\mathrm{R}, |\mathrm{T}|^{1/2} u^* \mathrm{S})_2$ où u est la phase de T. On en déduit que

$$|(\mathrm{TR}, \mathrm{S})_2| \leq \||\mathrm{T}|^{1/2}\mathrm{R}\|_2 \||\mathrm{T}|^{1/2} u^*\mathrm{S}\|_2 \leq \|u^*\mathrm{S}\| \, \|\mathrm{R}\| \, \||\mathrm{T}|^{1/2}\|_2^2 \leq \||\mathrm{T}|^{1/2}\|_2^2 = \|\mathrm{T}\|_1.$$

Donc $\{|(\mathrm{TR}, \mathrm{S})_2| : \mathrm{R} \in \mathcal{L}^2(\mathrm{E}), \ \mathrm{S} \in \mathcal{L}^2(\mathrm{E}, \mathrm{F}), \ \|\mathrm{R}\| \, \|\mathrm{S}\| \leq 1\}$ est majoré par $\|\mathrm{T}\|_1$. Soient B une base hilbertienne de E et I une partie finie de B ; notons P le projecteur orthogonal sur le sous-espace de E engendré par I ; alors

$$\sum_{b \in \mathrm{I}} \langle |\mathrm{T}|(b), b \rangle = \sum_{b \in \mathrm{I}} \langle u^* \mathrm{T}(b), b \rangle = (\mathrm{TP}, u\mathrm{P})_2 \leq$$

$$\leq \sup\{|(\mathrm{TR},\mathrm{S})_2| : \mathrm{R} \in \mathcal{L}^2(\mathrm{E}),\ \mathrm{S} \in \mathcal{L}^2(\mathrm{E},\mathrm{F}),\ \|\mathrm{R}\|\,\|\mathrm{S}\| \leq 1\}.$$

Prenant le « sup » sur les parties finies de B on trouve

$$\|\mathrm{T}\|_1 \leq \sup\{|(\mathrm{TR},\mathrm{S})_2| : \mathrm{R} \in \mathcal{L}^2(\mathrm{E}),\ \mathrm{S} \in \mathcal{L}^2(\mathrm{E},\mathrm{F}),\ \|\mathrm{R}\|\,\|\mathrm{S}\| \leq 1\}.$$

//

Théorème 7.4.3. *Soient* E, F *et* H *deux espaces de Hilbert ;*

(*i*) *l'ensemble* $\mathcal{L}^1(\mathrm{E},\mathrm{F})$ *est un sous-espace vectoriel de* $\mathcal{L}(\mathrm{E},\mathrm{F})$ *;*

(*ii*) *l'application* $\mathrm{T} \to \|\mathrm{T}\|_1$ *est une norme sur* $\mathcal{L}^1(\mathrm{E},\mathrm{F})$ *pour laquelle* $\mathcal{L}^1(\mathrm{E},\mathrm{F})$ *est complet :*

(*iii*) *pour tout* $\mathrm{T} \in \mathcal{L}(\mathrm{E},\mathrm{F})$ *on a* $\|\mathrm{T}\|_1 = \|\mathrm{T}^*\|_1$ *;*

(*iv*) *pour tout* $\mathrm{S} \in \mathcal{L}(\mathrm{E},\mathrm{F})$ *et tout* $\mathrm{T} \in \mathcal{L}(\mathrm{F},\mathrm{H})$ *on a* $\|\mathrm{TS}\|_1 \leq \|\mathrm{S}\|_1\|\mathrm{T}\|$ *et aussi* $\|\mathrm{TS}\|_1 \leq \|\mathrm{S}\|\,\|\mathrm{T}\|_1$.

Démonstration. Soient $\mathrm{S},\mathrm{T} \in \mathcal{L}(\mathrm{E},\mathrm{F})$; pour tous $\mathrm{R}_1 \in \mathcal{L}^2(\mathrm{E})$ et $\mathrm{R}_2 \in \mathcal{L}^2(\mathrm{E},\mathrm{F})$ tels que $\|\mathrm{R}_1\|\,\|\mathrm{R}_2\| \leq 1$, on a $|((\mathrm{S}+\mathrm{T})\mathrm{R}_1,\mathrm{R}_2)_2| \leq |(\mathrm{SR}_1,\mathrm{R}_2)_2| + |(\mathrm{TR}_1,\mathrm{R}_2)_2|$. Par le lemme 2, $\|\mathrm{S}+\mathrm{T}\|_1 \leq \|\mathrm{S}\|_1 + \|\mathrm{T}\|_1$; on en déduit immédiatement que $\mathcal{L}^1(\mathrm{E},\mathrm{F})$ est un sous-espace vectoriel de $\mathcal{L}(\mathrm{E},\mathrm{F})$ et que $\|\ \|_1$ est une semi-norme. Par le lemme 2, $\|\mathrm{T}\|_1 = \||\mathrm{T}|^{1/2}\|_2^2 \geq \||\mathrm{T}|^{1/2}\|^2 = \|\mathrm{T}\|$ (proposition 5.3.2). En particulier, $\|\ \|_1$ est une norme.

Par le lemme 2 l'application $\mathrm{T} \to \mathrm{Tr}(|\mathrm{T}|)$ est semi-continue inférieurement et on en déduit comme dans le théorème 3.2 que $\mathcal{L}^1(\mathrm{E})$, muni de cette norme, est complet. On a $|\mathrm{T}^*| = u\,|\mathrm{T}|\,u^* = (|\mathrm{T}|^{1/2}u^*)^*\,(|\mathrm{T}|^{1/2}u^*)$. Donc

$$\|\mathrm{T}^*\|_1 = \||\mathrm{T}|^{1/2}u^*\|_2^2 \leq \||\mathrm{T}|^{1/2}\|_2^2 = \|\mathrm{T}\|_1\,;$$

remplaçant T par T^*, on en déduit (*iii*). Soient $\mathrm{R}_1 \in \mathcal{L}^2(\mathrm{E})$ et $\mathrm{R}_2 \in \mathcal{L}^2(\mathrm{E},\mathrm{H})$; on a $(\mathrm{TSR}_1,\mathrm{R}_2)_2 = (\mathrm{SR}_1,\mathrm{T}^*\mathrm{R}_2)_2$. Il résulte alors du lemme 2 que $\|\mathrm{TS}\|_1 \leq \|\mathrm{S}\|_1\,\|\mathrm{T}\|$; remplaçant S et T par leurs adjoints il résulte de (*iii*) que $\|\mathrm{TS}\|_1 \leq \|\mathrm{S}\|\,\|\mathrm{T}\|_1$.

//

Théorème 7.4.4. *Soient* E, F *deux espaces de Hilbert ;*

(*i*) *pour tout opérateur* $\mathrm{T} \in \mathcal{L}^1(\mathrm{E})$ *et pour toute base hilbertienne* B *de* E *la famille* $(\langle \mathrm{T}(b), b\rangle)_{b\in\mathrm{B}}$ *est sommable et la quantité* $\mathrm{Tr}(\mathrm{T}) = \sum_{b\in\mathrm{B}}\langle \mathrm{T}(b), b\rangle$ *ne dépend pas de la base hilbertienne* B *;*

(*ii*) *on a* $|\mathrm{Tr}(\mathrm{T})| \leq \mathrm{Tr}(|\mathrm{T}|)$ *;*

(*iii*) *pour tous* $\mathrm{S} \in \mathcal{L}(\mathrm{E},\mathrm{F})$, $\mathrm{T} \in \mathcal{L}(\mathrm{F},\mathrm{E})$ *si* $\mathrm{S} \in \mathcal{L}^1(\mathrm{E},\mathrm{F})$ *ou si* S *et* T *sont de Hilbert-Schmidt, alors* $\mathrm{Tr}(\mathrm{TS}) = \mathrm{Tr}(\mathrm{ST})$.

Démonstration. On a $\langle \mathrm{T}(b), b\rangle = \langle |\mathrm{T}|^{1/2}(b), |\mathrm{T}|^{1/2}u^*(b)\rangle$. Comme les opérateurs $|\mathrm{T}|^{1/2}$ et $|\mathrm{T}|^{1/2}u^*$ sont dans $\mathcal{L}^2(\mathrm{E})$, (*i*) résulte du théorème 3.2. De plus

$$|\mathrm{Tr}(\mathrm{T})| = |(|\mathrm{T}|^{1/2}, |\mathrm{T}|^{1/2}u^*)_2| \leq \||\mathrm{T}|^{1/2}\|_2\,\||\mathrm{T}|^{1/2}u^*\|_2 \leq \|\mathrm{T}\|_1,$$

d'où (*ii*). L'application $\mathrm{S} \to \mathrm{S}^*$ est une isométrie antilinéaire de $\mathcal{L}^2(\mathrm{E},\mathrm{F})$ dans $\mathcal{L}^2(\mathrm{F},\mathrm{E})$; pour $\mathrm{S} \in \mathcal{L}^2(\mathrm{E},\mathrm{F})$, on a alors $\mathrm{Tr}(\mathrm{S}^*\mathrm{S}) = \mathrm{Tr}(\mathrm{SS}^*)$. Par l'identité de polarisation on en déduit que, pour $\mathrm{S},\mathrm{S}_1 \in \mathcal{L}^2(\mathrm{E},\mathrm{F})$, on a $\mathrm{Tr}(\mathrm{S}_1^*\mathrm{S}) = \mathrm{Tr}(\mathrm{SS}_1^*)$. Soit $\mathrm{T} \in \mathcal{L}^2(\mathrm{F},\mathrm{E})$; posant $\mathrm{S}_1 = \mathrm{T}^*$, on en déduit $\mathrm{Tr}(\mathrm{TS}) = \mathrm{Tr}(\mathrm{ST})$.

Enfin, soient $\mathrm{T} \in \mathcal{L}(\mathrm{F},\mathrm{E})$ et $\mathrm{S} \in \mathcal{L}^1(\mathrm{E},\mathrm{F})$; donnons nous $\mathrm{S}_1 \in \mathcal{L}^2(\mathrm{E},\mathrm{F})$, $\mathrm{S}_2 \in \mathcal{L}^2(\mathrm{E})$ tels que $\mathrm{S} = \mathrm{S}_1\mathrm{S}_2$; on a $\mathrm{Tr}(\mathrm{ST}) = \mathrm{Tr}(\mathrm{S}_1\mathrm{S}_2\mathrm{T}) = \mathrm{Tr}(\mathrm{S}_2\mathrm{TS}_1) = \mathrm{Tr}(\mathrm{TS})$.

//

Définition 7.4.1. Un opérateur $\mathrm{T} \in \mathcal{L}^1(\mathrm{E})$ est appelé *nucléaire* ou *à trace*.

Il est clair que $\mathcal{L}^1(\mathrm{E},\mathrm{F}) \subset \mathcal{L}^2(\mathrm{E},\mathrm{F})$. Tout opérateur de rang fini est nucléaire.

8. Calcul fonctionnel continu

L'un des objectifs du chapitre est de construire un homomorphisme isométrique φ_{T} de $\mathrm{C}(\mathrm{Sp}(\mathrm{T}))$ dans $\mathcal{L}(\mathrm{H})$ lorsque T est un opérateur autoadjoint sur un espace de Hilbert H. Cet homomorphisme sera difficile à visualiser dans le cas le plus général, mais il est peut-être utile de montrer le cas le plus évident, celui d'un espace de Hilbert de dimension finie. Si $\mathrm{H} = \mathbb{C}^3$, considérons un endomorphisme hermitien T, c'est à dire qui peut se décrire par une matrice $\mathrm{M} = \mathrm{U}^*\Delta\mathrm{U}$, où U est une matrice unitaire, Δ une matrice diagonale dont les coefficients diagonaux $\lambda_1, \lambda_2, \lambda_3$ sont réels, disons distincts pour fixer les idées. Alors $\mathrm{K} = \mathrm{Sp}(\mathrm{T}) = \{\lambda_1, \lambda_2, \lambda_3\}$. Pour toute fonction f (continue !) sur K on considérera la matrice

$$\varphi_{\mathrm{M}}(f) = \mathrm{U}^* \begin{pmatrix} f(\lambda_1) & 0 & 0 \\ 0 & f(\lambda_2) & 0 \\ 0 & 0 & f(\lambda_3) \end{pmatrix} \mathrm{U}.$$

Il est clair que φ_{M} est un homomorphisme d'algèbres unitaires de $\mathrm{C}(\mathrm{K})$ dans $\mathrm{M}_3(\mathbb{C})$.

8.1. Calcul fonctionnel polynomial

Cette section est de nature purement algébrique. On considère d'abord une algèbre unitaire A sur $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$ (quand on en viendra au spectre, on imposera $\mathbb{K} = \mathbb{C}$ comme d'habitude). Soient

$$\mathrm{P} = c_0 + c_1\mathrm{X} + \cdots + c_n\mathrm{X}^n$$

un polynôme de $\mathbb{K}[\mathrm{X}]$ et $a \in \mathrm{A}$; on pose

$$\varphi_a(\mathrm{P}) = \mathrm{P}(a) = c_0 1_{\mathrm{A}} + c_1 a + \cdots + c_n a^n \in \mathrm{A}.$$

Il est évident que $(\mathrm{P}+\mathrm{Q})(a) = \mathrm{P}(a)+\mathrm{Q}(a)$ et $(\lambda\mathrm{P})(a) = \lambda\mathrm{P}(a)$; l'application φ_a est donc linéaire ; si $\mathrm{Q} = \mathrm{X}^k$ on vérifie que $(\mathrm{PQ})(a) = \mathrm{P}(a)\mathrm{Q}(a)$ et on en déduit le cas général en décomposant Q en combinaison linéaire de monômes. On a obtenu :

Proposition 8.1.1. *Soient* A *une algèbre de Banach unitaire et* $a \in \mathrm{A}$ *; il existe un unique homomorphisme d'algèbres unitaires* φ_a *de* $\mathbb{K}[\mathrm{X}]$ *dans* A *tel que* $\varphi_a(\mathrm{X}) = a$ *; cet homomorphisme est donné par* $\varphi_a(\mathrm{P}) = \mathrm{P}(a)$.

Remarque. Si P et Q sont deux polynômes, on a $\mathrm{P}(a)\mathrm{Q}(a) = (\mathrm{PQ})(a) = (\mathrm{QP})(a) = \mathrm{Q}(a)\mathrm{P}(a)$: tous les éléments de la forme $\mathrm{P}(a)$ commutent (pour a fixé). Si $ab = ba$, on en déduit que $\mathrm{P}(a)b = b\mathrm{P}(a)$.

Lemme 8.1.2. *Si* $a_1\, a_2 = a_2\, a_1$ *est inversible dans* A, *alors* a_1 *est inversible dans* A.

Démonstration. Il existe un élément c tel que $c(a_1a_2) = 1_{\mathrm{A}} = (a_1a_2)c$; on voit que a_1 est inversible à gauche et à droite : $1_{\mathrm{A}} = a_1(a_2c)$ et $1_{\mathrm{A}} = c(a_1a_2) = (ca_2)a_1$; il en résulte que $a_2c = ca_2$ est l'inverse de a_1 :

$$a_2c = (a_2c)(a_1a_2c) = a_2(ca_1a_2)c = a_2c.$$

//

Corollaire 8.1.3. *Si $c \in \mathbb{K}$, $c \neq 0$ et si $c(a - \mu_1 1_\mathrm{A}) \ldots (a - \mu_k 1_\mathrm{A})$ est inversible, alors chaque $a - \mu_j 1_\mathrm{A}$ est inversible.*

Démonstration. Montrons le pour $a - \mu_1 1_\mathrm{A}$ par exemple ; considérons le produit $a_2 = c(a - \mu_2 1_\mathrm{A}) \ldots (a - \mu_k 1_\mathrm{A})$; alors $a_1 = a - \mu_1 1_\mathrm{A}$ et a_2 commutent, et $a_1 a_2$ est inversible, donc a_1 est inversible.

//

Théorème 8.1.4 : Petit théorème spectral. *Soit* A *une algèbre de Banach* complexe *; pour tout $a \in \mathrm{A}$, on a*

$$\mathrm{Sp}(\mathrm{P}(a)) = \mathrm{P}(\mathrm{Sp}(a)).$$

Démonstration. Posons $\mathrm{K} = \mathrm{Sp}(a)$, et supposons P non constant (ce cas particulier est évident). Puisqu'on est sur $\mathbb{C}$, on peut factoriser le polynôme $\mathrm{P} - \lambda$ sous la forme

$$\mathrm{P} - \lambda = c \prod_{i=1}^{k} (\mathrm{X} - \mu_i)$$

avec $c \neq 0$ et $k = \deg \mathrm{P} \geq 1$. Supposons d'abord que $\lambda \notin \mathrm{P}(\mathrm{K})$. On voit que $\mathrm{P}(\mu_i) = \lambda$ pour chaque racine μ_i ; puisque $\lambda \notin \mathrm{P}(\mathrm{K})$, chacun des μ_i est en dehors de K, donc chaque $a - \mu_i 1_\mathrm{A}$ est inversible, donc $\mathrm{P}(a) - \lambda 1_\mathrm{A} = c \prod_i (a - \mu_i 1_\mathrm{A})$ est inversible et $\lambda \notin \mathrm{Sp}(\mathrm{P}(a))$.

Supposons ensuite que $\lambda \notin \mathrm{Sp}(\mathrm{P}(a))$; alors $\mathrm{P}(a) - \lambda 1_\mathrm{A}$ est inversible, et égal au produit $c(a - \mu_1 1_\mathrm{A}) \ldots (a - \mu_k 1_\mathrm{A})$; on a vu que cela implique que chaque $a - \mu_j 1_\mathrm{A}$ est inversible ; si $z \in \mathbb{C}$ est tel que $\lambda = \mathrm{P}(z)$, le polynôme $\mathrm{P} - \lambda$ s'annule au point z, donc z est l'une des racines μ_i, donc $\lambda \notin \mathrm{P}(\mathrm{K})$.

//

Remarque 8.1.1. On peut aussi définir un calcul fonctionnel pour les fractions rationnelles. C'est très facile à partir du cas polynomial. Si $\mathrm{Q} \in \mathbb{C}[\mathrm{X}]$ ne s'annule pas sur $\mathrm{Sp}(a)$, l'élément $\mathrm{Q}(a)$ est inversible. Si $\mathrm{F} = \mathrm{P}/\mathrm{Q}$ est une fraction rationnelle, il est raisonnable de définir $\mathrm{F}(a) = \mathrm{P}(a)\mathrm{Q}(a)^{-1} = \mathrm{Q}(a)^{-1}\mathrm{P}(a)$.

8.2. Calcul fonctionnel continu pour les opérateurs hermitiens

Polynômes et adjoints

Soient H un espace de Hilbert complexe et $\mathrm{T} \in \mathcal{L}(\mathrm{H})$; il résulte des propriétés des adjoints que $(\mathrm{T}^k)^* = (\mathrm{T}^*)^k$ pour tout entier $k \geq 0$. Si $\mathrm{P} = \sum_{j=0}^n c_j \mathrm{X}^j$ est un polynôme à coefficients complexes, on peut considérer le polynôme dont les coefficients sont les complexes conjugués des coefficients de P. On notera $\widetilde{\mathrm{P}} = \sum_{j=0}^n \overline{c}_j \, \mathrm{X}^j$ ce polynôme ; alors

$$(\mathrm{P}(\mathrm{T}))^* = \Big(\sum c_k \mathrm{T}^k\Big)^* = \sum \overline{c}_k (\mathrm{T}^*)^k = \widetilde{\mathrm{P}}(\mathrm{T}^*)$$

ce qui montre que l'adjoint de $\mathrm{P}(\mathrm{T})$ est $\widetilde{\mathrm{P}}(\mathrm{T}^*)$. On notera que la fonction polynomiale $z \in \mathbb{C} \to \widetilde{\mathrm{P}}(z)$ **n'est pas** la fonction complexe conjuguée de la fonction $z \to \mathrm{P}(z)$ (on a en fait $\widetilde{\mathrm{P}}(\overline{z}) = \overline{\mathrm{P}(z)}$.

Si T est normal, P(T) est normal : en effet, T^* commute avec P(T) puisque T^* commute avec T, puis $\widetilde{P}(T^*)$ commute avec P(T) pour la même raison. Si T est hermitien et si P est un polynôme à coefficients réels, alors P(T) est hermitien. Le résultat essentiel pour la suite est le suivant :

Lemme 8.2.1. *Si* H *est un espace de Hilbert complexe et si* $T \in \mathcal{L}(H)$ *est normal, on a*

$$\|P(T)\| = \|P\|_{C(Sp(T))} = \max\{|P(\lambda)| : \lambda \in Sp(T)\}.$$

pour tout polynôme $P \in \mathbb{C}[X]$.

Démonstration. Soit $K = Sp(T)$. On a vu dans le théorème 1.4 que le spectre de P(T) est P(K). Par ailleurs P(T) est normal, donc

$$\|P(T)\| = \rho(P(T)) = \max\{|z| : z \in Sp(P(T))\} = \max\{|P(\lambda)| : \lambda \in K\}$$

d'après la proposition 6.2.3.

//

Lemme 8.2.2. *Soit* φ *un homomorphisme isométrique de* C(K) *dans une algèbre de Banach unitaire* B. *Alors* $f \in C(K)$ *est inversible dans* C(K) *si et seulement si* $\varphi(f)$ *est inversible dans* B.

Démonstration. On a déjà dit que si f est inversible, alors $\varphi(f)$ est inversible. Supposons maintenant f non inversible dans C(K) ; on a vu qu'il existe $s_0 \in K$ tel que $f(s_0) = 0$; posons $U_n = \{s \in K : |f(s)| < 2^{-n}\}$; c'est un ouvert qui contient s_0 ; soit h_n la fonction continue $\mathrm{dist}(s, U_n^c)$; cette fonction est non nulle, mais nulle en dehors de U_n ; si $g_n = \|h_n\|^{-1} h_n$, on a une fonction de norme 1 nulle en dehors de U_n. Alors $\|f g_n\| \le 2^{-n}$; posons $b = \varphi(f)$ et $x_n = \varphi(g_n)$; on a $\|x_n\| = 1$ et $\|b\, x_n\| \le 2^{-n}$; il est impossible que b soit inversible : si b^{-1} existait dans B, la multiplication par b^{-1} serait continue, donc $b^{-1}(b\, x_n) = x_n$ tendrait vers 0, ce qui n'est pas le cas puisque $\|x_n\|_B = \|g_n\|_{C(K)} = 1$.

//

Corollaire 8.2.3. *Pour tout homomorphisme isométrique* φ *de* C(K) *(complexe) dans une algèbre de Banach unitaire complexe* B, *on a*

$$Sp(\varphi(f)) = Sp(f) = f(K)$$

pour toute $f \in C(K)$.

Démonstration. On sait déjà que $Sp(\varphi(f)) \subset Sp(f)$. Inversement, si $\lambda \in Sp(f) = f(K)$ la fonction $f - \lambda$ est non inversible dans C(K), donc son image $\varphi(f) - \lambda 1_B$ est non inversible dans B, donc $\lambda \in Sp(\varphi(f))$.

//

Dans notre exemple matriciel à l'introduction du chapitre, on avait bien

$$Sp(\varphi_M(f)) = f(K) = \{f(\lambda_1), f(\lambda_2), f(\lambda_3)\}.$$

Passons au théorème sur le calcul fonctionnel continu. Si K est un compact de $\mathbb{C}$, on notera i_{K} la fonction $z \in \mathrm{K} \to z \in \mathbb{C}$.

Théorème 8.2.4. *Soient* H *un espace de Hilbert complexe et* $\mathrm{T} \in \mathcal{L}(\mathrm{H})$ *hermitien ; posons* $\mathrm{K} = \mathrm{Sp}(\mathrm{T})$. *Il existe un et un seul homomorphisme d'algèbres de Banach unitaires complexes* $\varphi_{\mathrm{T}} : \mathrm{C}(\mathrm{K}) \to \mathcal{L}(\mathrm{H})$ *tel que* $\varphi_{\mathrm{T}}(i_{\mathrm{K}}) = \mathrm{T}$.

L'homomorphisme φ_{T} *est isométrique. Si on note* $f(\mathrm{T}) = \varphi_{\mathrm{T}}(f)$, *on a* $f(\mathrm{T})^* = \overline{f}(\mathrm{T})$ *et* $f(\mathrm{T})$ *commute avec tout opérateur* S *qui commute avec* T *(donc* $f(\mathrm{T})$ *est normal), pour toute fonction* f *continue sur* K. *On a de plus*

$$\mathrm{Sp}(f(\mathrm{T})) = \mathrm{Sp}(f) = f(\mathrm{Sp}(\mathrm{T})).$$

Si f *est réelle continue sur* K *et* g *continue sur* $f(\mathrm{K})$, *on a* $(g \circ f)(\mathrm{T}) = g(f(\mathrm{T}))$.

Démonstration. On a vu que $\mathrm{K} = \mathrm{Sp}(\mathrm{T})$ est contenu dans $\mathbb{R}$. Désignons par A l'ensemble des fonctions continues f sur K de la forme $f : s \to \mathrm{P}(s)$ pour un $\mathrm{P} \in \mathbb{C}[\mathrm{X}]$. Pour toute $f \in \mathrm{A}$ l'élément $f(\mathrm{T})$ peut être défini de façon unique puisque si $f = \mathrm{P}_1 = \mathrm{P}_2$ sur K,

$$\|\mathrm{P}_1(\mathrm{T}) - \mathrm{P}_2(\mathrm{T})\| = \|\mathrm{P}_1 - \mathrm{P}_2\|_{\mathrm{C}(\mathrm{K})} = 0$$

d'après le lemme 1 (si le spectre K est un ensemble infini, la vérification précédente est inutile, puisque le polynôme formel $\mathrm{P}_1 - \mathrm{P}_2$ sera nul s'il a une infinité de racines ; mais le spectre de T pourrait être fini). On posera donc $f(\mathrm{T}) = \mathrm{P}(\mathrm{T})$, où P est n'importe quel polynôme qui représente la fonction f sur K.

L'ensemble A est dense dans C(K) d'après le théorème de Weierstrass –les polynômes sont uniformément denses dans $\mathrm{C}([-a, a])$– et puisque $\mathrm{K} \subset [-a, a]$ avec par exemple $a = \|\mathrm{T}\|$. On a un homomorphisme isométrique ψ de A dans $\mathcal{L}(\mathrm{H})$, que l'on va prolonger par densité en homomorphisme isométrique de C(K) dans $\mathcal{L}(\mathrm{H})$: d'après le lemme 1.4.1, il existe un prolongement unique φ_{T} de ψ en application linéaire continue de C(K) dans $\mathcal{L}(\mathrm{H})$. Posons $f(\mathrm{T}) = \varphi_{\mathrm{T}}(f)$ pour toute $f \in \mathrm{C}(\mathrm{K})$. Pour toute suite (P_n) de polynômes qui converge uniformément sur K vers la fonction f, la suite $(\mathrm{P}_n(\mathrm{T}))$ tend en norme dans $\mathcal{L}(\mathrm{H})$ vers $f(\mathrm{T})$, puisque φ_{T} est continu. Il en résulte par continuité de la norme que

$$\|f(\mathrm{T})\| = \lim_n \|\mathrm{P}_n(\mathrm{T})\| = \lim_n \|\mathrm{P}_n\|_{\mathrm{C}(\mathrm{K})} = \|f\|_{\mathrm{C}(\mathrm{K})},$$

ce qui montre que l'application $\varphi_{\mathrm{T}} : f \in \mathrm{C}(\mathrm{K}) \to f(\mathrm{T}) \in \mathcal{L}(\mathrm{H})$ est isométrique.

Par construction on a $\varphi_{\mathrm{T}}(i_{\mathrm{K}}) = \mathrm{T}$ puisque la fonction i_{K} correspond au monôme X dont l'image est T en calcul polynomial. Il reste à voir que φ_{T} est un homomorphisme. Si (P_n) converge uniformément vers f sur K et (Q_n) converge uniformément vers g sur K, alors $f(\mathrm{T})g(\mathrm{T}) = \lim(\mathrm{P}_n\mathrm{Q}_n)(\mathrm{T}) = (fg)(\mathrm{T})$ (utiliser la continuité du produit par rapport au couple de variables), donc φ_{T} est un homomorphisme d'algèbres de Banach unitaires complexes, isométrique. Il en résulte que $\mathrm{Sp}\, f(\mathrm{T}) = f(\mathrm{K})$, d'après un principe général sur C(K) (corollaire 3).

Si $\mathrm{ST} = \mathrm{TS}$, on en déduit que $\mathrm{SP}_n(\mathrm{T}) = \mathrm{P}_n(\mathrm{T})\mathrm{S}$ pour tout n, donc $\mathrm{S}f(\mathrm{T}) = f(\mathrm{T})\mathrm{S}$ par continuité du produit par S, à droite et à gauche. Ainsi $f(\mathrm{T})$ commute avec tout opérateur borné S qui commute avec T. L'adjoint $f(\mathrm{T})^*$ est la limite de $\mathrm{P}_n(\mathrm{T})^* = \widetilde{\mathrm{P}}_n(\mathrm{T})$; mais puisque $\mathrm{K} \subset \mathbb{R}$, la fonction $s \in \mathbb{K} \to \sum_j \overline{c}_j s^j$ est bien la complexe conjuguée de la fonction polynomiale P_n, donc $\widetilde{\mathrm{P}}_n$ tend uniformément sur K vers la fonction $\overline{f}$, donc $\overline{f}(\mathrm{T}) = f(\mathrm{T})^*$; il en résulte que $f(\mathrm{T})^* f(\mathrm{T}) = (\overline{f}f)(\mathrm{T}) = (f\overline{f})(\mathrm{T}) = f(\mathrm{T})f(\mathrm{T})^*$ donc $f(\mathrm{T})$ est normal.

Montrons l'unicité de φ_{T}. Soit φ un homomorphisme d'algèbres de Banach unitaires de C(K) dans $\mathcal{L}(\mathrm{H})$ tel que $\varphi(i_{\mathrm{K}}) = \mathrm{T}$; par définition, on a $\varphi(i_{\mathrm{K}}^0) = \varphi(1) = \mathrm{Id}_{\mathrm{H}} = \mathrm{T}^0$, et $\varphi(i_{\mathrm{K}}^k) = \mathrm{T}^k$ pour tout $k \geq 1$ (la fonction i_{K}^k est la fonction monôme $s \to s^k$) ; il en résulte puisque φ est de plus linéaire que pour toute fonction polynomiale $f : s \to \mathrm{P}(s)$, l'image $\varphi(f)$ est $\mathrm{P}(\mathrm{T}) = \varphi_{\mathrm{T}}(f)$. De plus, φ est continu par définition des homomorphismes d'algèbres de Banach ; si (P_n) tend uniformément vers f sur K, on aura $\varphi(f) = \lim_n \varphi(\mathrm{P}_n) = \lim_n \mathrm{P}_n(\mathrm{T}) = \varphi_{\mathrm{T}}(f)$. On a donc $\varphi = \varphi_{\mathrm{T}}$.

Supposons que f soit une fonction réelle continue sur $\mathrm{K} = \mathrm{Sp}(\mathrm{T})$. L'ensemble $\mathrm{L} = f(\mathrm{K}) \subset \mathbb{R}$ est compact, et c'est le spectre de $f(\mathrm{T})$. L'application $g \in \mathrm{C}(\mathrm{L}) \to g \circ f \in \mathrm{C}(\mathrm{K})$ est un homomorphisme χ d'algèbres de C(L) dans C(K), qui transforme i_{L} en $f(\mathrm{T})$; d'après l'unicité, la composition $\varphi_{\mathrm{T}} \circ \chi$ est égale à l'homomorphisme $\varphi_{f(\mathrm{T})}$ associé à l'opérateur hermitien $f(\mathrm{T})$. On a donc $(g \circ f)(\mathrm{T}) = g(f(\mathrm{T}))$ pour toute fonction continue g sur L.

//

Exemples 8.2.1.

1. Supposons que T soit diagonal dans une base orthonormée, avec coefficients diagonaux (λ_n) (réels) ; l'opérateur $f(\mathrm{T})$ est l'opérateur diagonal de coefficients $(f(\lambda_n))$; démonstration : passer à la limite à partir du cas polynomial.

2. Supposons que T soit l'opérateur $\mathrm{M}_\varphi : \mathrm{L}_2(0,1) \to \mathrm{L}_2(0,1)$ de multiplication par une fonction φ réelle continue. On voit que pour tout polynôme P l'opérateur $\mathrm{P}(\mathrm{M}_\varphi)$ est l'opérateur de multiplication par la fonction $s \in [0,1] \to \mathrm{P}(\varphi(t))$, donc à la limite $f(\mathrm{M}_\varphi)$ est l'opérateur de multiplication par $s \to f(\varphi(s))$, c'est à dire que $f(\mathrm{M}_\varphi) = \mathrm{M}_{f\circ\varphi}$.

3. Si f est une fonction réelle positive sur $\mathrm{K} = \mathrm{Sp}(\mathrm{T})$, alors $f(\mathrm{T})$ est hermitien positif : il suffit de considérer $g(s) = \sqrt{f(s)}$ qui est une fonction continue sur K. Alors $g(\mathrm{T})$ est hermitien et $f(\mathrm{T}) = (g(\mathrm{T}))^2$ est hermitien positif.

Le cas hermitien sur un espace réel. Complexification

Soit H un espace de Hilbert réel, dont le produit scalaire sera noté $x \,.\, y$ pour éviter les confusions avec le produit scalaire dans le complexifié ; le complexifié de H est l'espace $\mathrm{H}_{\mathbb{C}} = \mathrm{H} + i\mathrm{H}$ de tous les vecteurs $z = x + iy$ où $x, y \in \mathrm{H}$; cette écriture est simplement une écriture symbolique commode pour un couple $(x, y) \in \mathrm{H} \times \mathrm{H}$. Si $\lambda = a + ib \in \mathbb{C}$, on pose $\lambda z = (ax - by) + i(by + ax)$. On vérifiera les axiomes d'espace vectoriel complexe. On définit le produit scalaire (complexe) sur $\mathrm{H}_{\mathbb{C}}$ en posant

$$\langle x + iy, x' + iy' \rangle = (x + iy) \,.\, (x' - iy') = (x \,.\, x' + y \,.\, y') + i(y \,.\, x' - x \,.\, y').$$

Lorsque $z = x + iy$, on voit que $\langle z, z \rangle = x \,.\, x + y \,.\, y = \|x\|^2 + \|y\|^2$, ce qui donne un produit scalaire sur $\mathrm{H} \times \mathrm{H}$ dont la norme associée est $\|z\| = \bigl(\|x\|^2 + \|y\|^2\bigr)^{1/2}$. Si $z = x + i0$, on a $\|z\| = \|x\|$. A tout opérateur $\mathrm{T} \in \mathcal{L}(\mathrm{H})$ on associe l'application $\mathrm{T}_{\mathbb{C}}$ de $\mathrm{H}_{\mathbb{C}}$ dans lui-même définie par $\mathrm{T}_{\mathbb{C}}(x + iy) = \mathrm{T}(x) + i\mathrm{T}(y)$; on vérifie facilement que $\mathrm{T}_{\mathbb{C}}$ est $\mathbb{C}$-linéaire. On voit que :

l'application $\mathrm{T} \to \mathrm{T}_{\mathbb{C}}$ *est un homomorphisme isométrique de* $\mathbb{R}$*-algèbres de Banach unitaires. De plus,* $(\mathrm{T}^*)_{\mathbb{C}} = (\mathrm{T}_{\mathbb{C}})^*$ *et* $\mathrm{T}_{\mathbb{C}}$ *est inversible si et seulement si* T *est inversible.*

Il est clair que l'application est $\mathbb{R}$-linéaire, que $(\mathrm{ST})_{\mathbb{C}} = \mathrm{S}_{\mathbb{C}}\mathrm{T}_{\mathbb{C}}$, $(\mathrm{Id}_{\mathrm{H}})_{\mathbb{C}} = \mathrm{Id}_{\mathrm{H}_{\mathbb{C}}}$; si $z = x + iy$,

$$\|\mathrm{T}_{\mathbb{C}}(z)\|^2 = \|\mathrm{T}(x)\|^2 + \|\mathrm{T}(y)\|^2 \leq \|\mathrm{T}\|^2\bigl(\|x\|^2 + \|y\|^2\bigr) = \|\mathrm{T}\|^2\,\|z\|^2,$$

donc $\|T_{\mathbb{C}}\| \le \|T\|$; l'inégalité inverse est claire en regardant les vecteurs $z = x + i0$. Pour l'adjoint,

$$\langle T(x) + iT(y), x' + iy' \rangle = (T(x) + iT(y)) \,.\, (x' - iy') \,;$$

En développant, chaque produit scalaire $T(u) \,.\, v'$, avec $u = x, y$ et $v' = x', y'$ sera transformé en $u \,.\, T^*(v')$ et il n'y a plus qu'à remonter les morceaux.

Vérifions que le complexifié $T_{\mathbb{C}}$ est inversible dans $\mathcal{L}(H_{\mathbb{C}})$ si et seulement si T est inversible dans $\mathcal{L}(H)$. S'il existe $S \in \mathcal{L}(H)$ tel que $ST = TS = \mathrm{Id}_H$ il en résulte que $S_{\mathbb{C}} T_{\mathbb{C}} = T_{\mathbb{C}} S_{\mathbb{C}} = \mathrm{Id}_{H_{\mathbb{C}}}$, donc $T_{\mathbb{C}}$ est inversible. Inversement, supposons $T_{\mathbb{C}}$ inversible ; alors T est injectif : si $T(x) = 0_H$, alors $T_{\mathbb{C}}(x + i0_H) = 0_{H_{\mathbb{C}}}$, donc $x + i0 = 0$, donc $x = 0_H$; de plus T est surjectif : pour tout $x \in H$ il existe $z = x' + iy'$ tel que $T_{\mathbb{C}}(z') = x + i0$, ce qui donne $T(x') = x$; on en déduit que l'inverse est continu, soit par le théorème des isomorphismes, soit en utilisant la norme de $(T_{\mathbb{C}})^{-1}$.

Si $P \in \mathbb{R}[X]$, il résulte de la propriété d'homomorphisme unitaire que $(P(T))_{\mathbb{C}} = P(T_{\mathbb{C}})$. Si T est hermitien, son image $T_{\mathbb{C}}$ est un opérateur hermitien de spectre $K = \mathrm{Sp}(T_{\mathbb{C}})$ réel. Il en résulte que pour tout $\lambda \in \mathbb{R}$, l'opérateur $T - \lambda\,\mathrm{Id}_H$ est inversible si et seulement si $T_{\mathbb{C}} - \lambda\,\mathrm{Id}_{H_{\mathbb{C}}}$ est inversible. Si on introduit le spectre réel de T en posant

$$\mathrm{Sp}_{\mathbb{R}}(T) = \{\lambda \in \mathbb{R} : T - \lambda\,\mathrm{Id}_H \quad \text{non inversible } \}$$

on voit que $\mathrm{Sp}_{\mathbb{R}}(T) = \mathbb{R} \cap \mathrm{Sp}(T_{\mathbb{C}})$. Cette notion de spectre réel n'est pas très intéressante en général, car il est possible que $\mathrm{Sp}_{\mathbb{R}}(T)$ soit vide et ne donne aucune information. Mais dans le cas où T est hermitien, on sait que $T_{\mathbb{C}}$ est hermitien aussi, donc son spectre est réel et $\mathrm{Sp}_{\mathbb{R}}(T) = \mathrm{Sp}(T_{\mathbb{C}})$ dans ce cas.

Passons au calcul fonctionnel continu pour les hermitiens réels. Si P est un polynôme réel et si $T \in \mathcal{L}(H)$ est hermitien, on a $T_{\mathbb{C}}$ hermitien, $P(T_{\mathbb{C}})$ hermitien, donc

$$\|P(T)\| = \|(P(T))_{\mathbb{C}}\| = \|P(T_{\mathbb{C}})\| = \|P\|_{C(\mathrm{Sp}(T_{\mathbb{C}}))}.$$

Si on pose $K = \mathrm{Sp}(T_{\mathbb{C}}) = \mathrm{Sp}_{\mathbb{R}}(T)$ et si f est une fonction réelle continue sur K, on a dit qu'il existe un polynôme $P \in \mathbb{C}[X]$ tel que $|f(s) - P(s)| < \varepsilon$ pour tout $s \in K$. Comme s est réel, il est clair que si $Q \in \mathbb{R}[X]$ est le polynôme obtenu à partir de P en prenant comme coefficients les parties réelles des coefficients de P, alors $Q(s) = \mathrm{Re}\,P(s)$, donc $|f(s) - Q(s)| = |\,\mathrm{Re}(f(s) - P(s))| \le |f(s) - P(s)| < \varepsilon$. On voit donc que l'algèbre $A_{\mathbb{R}}$ des fonctions polynomiales à coefficients réels est dense dans $C_{\mathbb{R}}(K)$. On continue la démonstration comme avant. On obtient donc

Corollaire 8.2.5. *Soient* H *un espace de Hilbert réel et* $T \in \mathcal{L}(H)$ *hermitien ; désignons par* K *le spectre de* T. *Il existe un et un seul homomorphisme d'algèbres de Banach unitaires réelles* $\varphi_T : C_{\mathbb{R}}(K) \to \mathcal{L}(H)$ *tel que* $\varphi_T(i_K) = T$.

L'homomorphisme φ_T *est isométrique. Si on note* $f(T) = \varphi_T(f)$, *on a que* $f(T)^* = f(T)$ *est hermitien pour toute* f *(forcément réelle dans ce contexte) continue sur* K *et* $f(T)$ *commute avec tout opérateur* S *qui commute avec* T. *On a*

$$\mathrm{Sp}_{\mathbb{R}}(f(T)) = f(\mathrm{Sp}_{\mathbb{R}}(T)).$$

Si f *est continue sur* K *et* g *continue sur* $f(K)$, *on a* $(g \circ f)(T) = g(f(T))$.

Lemme 8.2.6. *Soit* T *un opérateur hermitien sur un espace de Hilbert* H *; pour tout* $\lambda \in \mathrm{Sp}(\mathrm{T})$ *et tout* $\varepsilon > 0$, *il existe un sous-espace fermé* $\mathrm{F} \neq \{0\}$ *de* H, *stable par tout opérateur* S *qui commute avec* T, *et tel que*

$$\forall x \in \mathrm{F}, \quad \|\mathrm{T}(x) - \lambda x\| \leq \varepsilon \, \|x\|.$$

Démonstration. Soit g une fonction réelle continue sur $\mathbb{R}$ telle que $g(\lambda) = 1$ et $g = 0$ en dehors de l'intervalle $[\lambda - \varepsilon, \lambda + \varepsilon]$. On a $0 \neq g(\lambda) \in g(\mathrm{S}p(\mathrm{T})) = \mathrm{Sp}(g(\mathrm{T}))$, ce qui montre que $g(\mathrm{T}) \neq 0$. On peut donc trouver un vecteur $x_0 = g(\mathrm{T})(y)$ non nul.

Soit ensuite f une fonction réelle continue sur $\mathbb{R}$, qui soit égale à λ sur l'intervalle $[\lambda - \varepsilon, \lambda + \varepsilon]$ et telle que $|f(t) - t| \leq \varepsilon$ pour tout $t \in \mathbb{R}$ (par exemple, $f(t) = t - \varepsilon$ quand $t > \lambda + \varepsilon$ et $f(t) = t + \varepsilon$ quand $t < \lambda - \varepsilon$). On a $(f - \lambda)\, g = 0$, donc $fg = \lambda g$ et en utilisant l'homomorphisme φ_T on voit que $f(\mathrm{T})(x_0) = f(\mathrm{T})(g(\mathrm{T})(y)) = \lambda g(\mathrm{T})(y) = \lambda x_0$, donc x_0 est vecteur propre de $f(\mathrm{T})$ pour la valeur propre λ. Posons $\mathrm{F} = \ker(f(\mathrm{T}) - \lambda \,\mathrm{Id_H})$. L'espace F est $\neq \{0\}$, et puisque tout opérateur S qui commute avec T commute avec $f(\mathrm{T})$, le sous-espace propre F de $f(\mathrm{T})$ est stable par S. Finalement, on a $\|f(\mathrm{T}) - \mathrm{T}\| \leq \varepsilon$ puisque $\|f - i_{\mathrm{Sp(T)}}\|_\infty \leq \varepsilon$. Pour tout $x \in \mathrm{F}$, on aura $f(\mathrm{T})(x) = \lambda x$ et $\|f(\mathrm{T})(x) - \mathrm{T}(x)\| \leq \varepsilon \, \|x\|$, et c'est fini.

//

8.3. Application aux hermitiens positifs. La racine carrée

Lemme 8.3.1. *Soient* H *un espace de Hilbert* **complexe** *et* $\mathrm{T} \in \mathcal{L}(\mathrm{H})$ *; si pour tout* $x \in \mathrm{H}$, *le scalaire* $\langle \mathrm{T}(x), x \rangle$ *est réel, il en résulte que* T *est hermitien.*

Démonstration. Supposons que $\langle \mathrm{T}(x), x \rangle$ soit réel pour tout $x \in \mathrm{H}$. L'application $(x, y) \to \langle \mathrm{T}(x), y \rangle$ est sesquilinéaire. Par le corollaire 5.1.2, on a $\langle \mathrm{T}(y), x \rangle = \overline{\langle \mathrm{T}(x), y \rangle}$, pour tous $x, y \in \mathrm{H}$, donc T est hermitien.

//

Si H est un espace de Hilbert réel, la condition $\langle \mathrm{T}(x), x \rangle$ réel n'implique évidemment pas que T soit hermitien ; même si on a $\langle \mathrm{T}(x), x \rangle \geq 0$ pour tout $x \in \mathrm{H}$, cela n'entraîne pas que T soit hermitien.

Théorème 8.3.2. *Soient* H *un espace de Hilbert (réel ou complexe) et* $\mathrm{T} \in \mathcal{L}(\mathrm{H})$ *; les conditions suivantes sont équivalentes :*

(*i*) *l'opérateur* T *est hermitien et* $\langle \mathrm{T}(x), x \rangle$ *est réel* ≥ 0 *pour tout* $x \in \mathrm{H}$ *;*
(*ii*) *il existe* $\mathrm{S} \in \mathcal{L}(\mathrm{H})$ *tel que* $\mathrm{T} = \mathrm{S}^*\mathrm{S}$ *;*
(*iii*) *il existe* $\mathrm{S} \in \mathcal{L}(\mathrm{H})$ *tel que* $\mathrm{S} = \mathrm{S}^*$ *et* $\mathrm{T} = \mathrm{S}^2$ *;*
(*iv*) *on a* $\mathrm{T} = \mathrm{T}^*$ *et* $\mathrm{Sp}(\mathrm{T}) \subset [0, +\infty[$.

Démonstration. Supposons (*ii*) vérifiée ; alors l'opérateur $\mathrm{T} = \mathrm{S}^*\mathrm{S}$ est hermitien et on a $\langle \mathrm{S}^*\mathrm{S}(x), x \rangle = \langle \mathrm{S}(x), \mathrm{S}(x) \rangle \geq 0$ pour tout $x \in \mathrm{H}$, donc (*ii*) $\Rightarrow$ (*i*). L'implication (*iii*) $\Rightarrow$ (*ii*) est évidente. Supposons que $\mathrm{T} = \mathrm{T}^*$ et $\mathrm{K} = \mathrm{Sp}(\mathrm{T}) \subset [0, +\infty[$; notons $f \in \mathrm{C}(\mathrm{K})$ l'application $t \to \sqrt{t}$; par le théorème 2.4 ou le corollaire 2.5, on a $f(\mathrm{T}) = f(\mathrm{T})^*$; de plus $f^2 = i_\mathrm{K}$, donc $f(\mathrm{T})^2 = \mathrm{T}$, donc (*iv*) $\Rightarrow$ (*iii*). Enfin, par la proposition 6.3.10, on sait que (*i*) implique que $\mathrm{Sp}(\mathrm{T}) \subset [0, +\infty[$.

//

Un élément hermitien de $\mathcal{L}(\mathrm{H})$ satisfaisant aux conditions équivalentes du théorème 2 est appelé *positif* (définition 5.3.2). On note $\mathcal{L}(\mathrm{H})_+$ l'ensemble des éléments positifs de $\mathcal{L}(\mathrm{H})$. Pour $\mathrm{T} \in \mathcal{L}(\mathrm{H})_+$ et $\alpha > 0$, on pose $\mathrm{T}^\alpha = f(\mathrm{T})$, où $f \in \mathrm{C}_\mathbb{R}(\mathrm{Sp}(\mathrm{T}))$ est l'application $t \to t^\alpha$. Pour $\alpha, \beta > 0$ on a $\mathrm{T}^{\alpha+\beta} = \mathrm{T}^\alpha \mathrm{T}^\beta$ et, par la dernière partie du théorème 2.4 (ou du corollaire 2.5), on a $(\mathrm{T}^\alpha)^\beta = \mathrm{T}^{\alpha\beta}$.

Proposition 8.3.3. *Pour* $\mathrm{T} \in \mathcal{L}(\mathrm{H})_+$*, il existe un et seul* $\mathrm{S} \in \mathcal{L}(\mathrm{H})_+$ *tel que* $\mathrm{S}^2 = \mathrm{T}$.

Démonstration. Soit S un opérateur hermitien positif tel que $\mathrm{S}^2 = \mathrm{T}$; considérons le spectre $\mathrm{K} = \mathrm{Sp}(\mathrm{S}) \subset [0, +\infty[$, et considérons sur K la fonction $f : s \to s^2$, puis sur $\mathrm{L} = f(\mathrm{K}) \subset [0, +\infty[$ la fonction $g(t) = \sqrt{t}$. Du fait que $\mathrm{K} \subset [0, +\infty[$, on vérifie que $g(f(s)) = \sqrt{s^2} = s$ pour tout $s \in \mathrm{K}$, donc le résultat de composition nous donne, puisque $g \circ f = i_\mathrm{K}$

$$\mathrm{S} = (g \circ f)(\mathrm{S}) = g(f(\mathrm{S})) = g(\mathrm{S}^2) = g(\mathrm{T}) = \sqrt{\mathrm{T}}.$$

//

Bien entendu il n'y a pas unicité si on ne demande pas que la racine soit positive : il suffit de considérer $-\sqrt{\mathrm{T}}$ pour avoir une autre racine hermitienne.

Décomposition polaire

Soient E et F deux espaces de Hilbert et $\mathrm{T} \in \mathcal{L}(\mathrm{E}, \mathrm{F})$; on appelle *module* de T et on note $|\mathrm{T}|$ l'unique $\mathrm{S} \in \mathcal{L}(\mathrm{E})_+$ tel que $\mathrm{S}^2 = \mathrm{T}^*\mathrm{T}$, c'est à dire que $|\mathrm{T}| = \sqrt{\mathrm{T}^*\mathrm{T}}$.

Proposition 8.3.4. *Soient* E *et* F *deux espaces de Hilbert et* $\mathrm{T} \in \mathcal{L}(\mathrm{E}, \mathrm{F})$ *; il existe un et un seul* $u \in \mathcal{L}(\mathrm{E}, \mathrm{F})$*, nul sur* $\ker(\mathrm{T})$ *tel que* $\mathrm{T} = u\,|\mathrm{T}|$.

Démonstration. Pour $x \in \mathrm{E}$ on a $\|\mathrm{T}(x)\|^2 = \langle \mathrm{T}^*\mathrm{T}(x), x\rangle = \big\| |\mathrm{T}|\,(x)\big\|^2$; en particulier, $\ker(\mathrm{T}) = \ker(|\mathrm{T}|)$; par la proposition 5.3.3, l'adhérence de l'image de $|\mathrm{T}|$ est donc l'orthogonal de $\ker(\mathrm{T})$; pour $x \in \mathrm{E}$ et $y \in \ker(\mathrm{T})$, on a alors

$$\|\mathrm{T}(x)\|^2 = \big\||\mathrm{T}|(x)\big\|^2 \le \big\||\mathrm{T}|(x)\big\|^2 + \|y\|^2 = \big\||\mathrm{T}|(x) + y\big\|^2.$$

Comme $\mathrm{im}(|\mathrm{T}|) + \ker(\mathrm{T})$ est dense dans E et que F est complet, il existe un unique élément $u \in \mathcal{L}(\mathrm{E}, \mathrm{F})$ tel que, pour tout $x \in \mathrm{E}$ et tout $y \in \ker(\mathrm{T})$ on ait $u(|\mathrm{T}|(x)+y) = \mathrm{T}(x)$ (utiliser le lemme 1.4.1).

//

Soient E et F deux espaces de Hilbert et $\mathrm{T} \in \mathcal{L}(\mathrm{E}, \mathrm{F})$; on appelle *phase* de T l'unique $u \in \mathcal{L}(\mathrm{E}, \mathrm{F})$ nul sur $\ker(\mathrm{T})$ tel que $\mathrm{T} = u\,|\mathrm{T}|$. La décomposition $\mathrm{T} = u\,|\mathrm{T}|$ s'appelle *décomposition polaire* de T.

Proposition 8.3.5. *Soient* E *et* F *deux espaces de Hilbert et* $\mathrm{T} \in \mathcal{L}(\mathrm{E}, \mathrm{F})$ *; notons* $\mathrm{T} = u\,|\mathrm{T}|$ *sa décomposition polaire.*

(*i*) *On a* $u^*\,\mathrm{T} = |\mathrm{T}|$ *; de plus,* u^*u *et* uu^* *sont les projecteurs orthogonaux sur l'orthogonal du noyau de* T *et sur l'adhérence de l'image de* T *respectivement.*

(*ii*) *Le module de* T^* *est* $u\,\mathrm{T}^* = \mathrm{T}\,u^*$ *; sa phase est* u^*.

Démonstration. Soient $x, y \in \mathrm{E}$, $z_1 \in \ker(\mathrm{T})$ et $z_2 \in \ker(\mathrm{T}^*)$; on a

$$\begin{aligned}
\langle u^*(\mathrm{T}(x)+z_2), (|\mathrm{T}|(y)+z_1)\rangle &= \langle(\mathrm{T}(x)+z_2), u(|\mathrm{T}|(y)+z_1)\rangle \\
&= \langle(\mathrm{T}(x)+z_2), \mathrm{T}(y)\rangle \\
&= \langle \mathrm{T}(x), \mathrm{T}(y)\rangle & \text{car } z_2 \in \operatorname{im}(\mathrm{T})^\perp \\
&= \langle |\mathrm{T}|(x), |\mathrm{T}|(y)\rangle \\
&= \langle |\mathrm{T}|(x), (|\mathrm{T}|(y)+z_1)\rangle & \text{car } z_1 \in \operatorname{im}(|\mathrm{T}|)^\perp.
\end{aligned}$$

Comme l'ensemble $\{|\mathrm{T}|(y)+z_1 : y \in \mathrm{E},\, z_1 \in \ker(\mathrm{T})\}$ est dense dans E, on en déduit que, pour tout $x \in \mathrm{E}$ et tout $z_2 \in \ker(\mathrm{T}^*)$ on a $u^*(\mathrm{T}(x)+z_2) = |\mathrm{T}|(x)$. En particulier, $u^*\mathrm{T} = |\mathrm{T}|$.

Notons p le projecteur orthogonal sur l'adhérence de l'image de $|\mathrm{T}|$ (c'est à dire l'orthogonal de $\ker(\mathrm{T})$). On a les relations $u^*u|\mathrm{T}| = u^*\mathrm{T} = |\mathrm{T}|$; comme u^*u s'annule sur $\ker(\mathrm{T})$, il coïncide avec p sur $\operatorname{im}(|\mathrm{T}|) + \ker(\mathrm{T})$; comme ce sous-espace est dense dans E, on a $u^*u = p$. Notons q le projecteur orthogonal sur l'adhérence de l'image de T. On a $uu^*\mathrm{T} = u|\mathrm{T}| = \mathrm{T}$; comme uu^* s'annule sur $\ker(\mathrm{T}^*) = \operatorname{im}(\mathrm{T})^\perp$, il coïncide avec q sur $\operatorname{im}(\mathrm{T}) + \ker(\mathrm{T}^*)$; comme ce sous-espace est dense dans F, on a $uu^* = q$.

Pour démontrer le point (*ii*) écrivons $|\mathrm{T}| = \mathrm{S}^*\mathrm{S}$ (théorème 2) ; on peut alors écrire $\mathrm{T}u^* = u\,|\mathrm{T}|\,u^* = (\mathrm{S}u^*)^*(\mathrm{S}u^*)$. On en déduit immédiatement que $\mathrm{T}u^*$ est positif ; alors $\mathrm{T}u^* = (\mathrm{T}u^*)^* = u\mathrm{T}^*$; de plus $(\mathrm{T}u^*)^2 = u\,\mathrm{T}^*\mathrm{T}\,u^* = u|\mathrm{T}|^2u^* = \mathrm{T}\mathrm{T}^*$; par définition du module, $|\mathrm{T}^*| = \mathrm{T}u^* = u\mathrm{T}^*$. De plus, u^* s'annule sur $\ker(\mathrm{T}^*)$ et $u^*u(y) = y$ pour tout $y \in \operatorname{im}(\mathrm{T}^*) \subset \ker(\mathrm{T})^\perp$, donc $u^*(u\mathrm{T}^*) = \mathrm{T}^*$; par définition, la phase de T^* est u^*.

//

Remarque 8.3.1. On déduit aisément des calculs ci-dessus les identités suivantes :

$$\mathrm{T} = u|\mathrm{T}| = |\mathrm{T}^*|u = u\mathrm{T}^*u, \quad |\mathrm{T}| = u^*\mathrm{T} = \mathrm{T}^*u = u^*|\mathrm{T}^*|u,$$

$$\mathrm{T}^* = u^*|\mathrm{T}^*| = |\mathrm{T}|u^* = u^*\mathrm{T}u^*, \quad |\mathrm{T}^*| = u|\mathrm{T}|u^* = \mathrm{T}u^* = u\mathrm{T}^*.$$

En particulier T^* et $|\mathrm{T}|$ ont même image et u est un isomorphisme de l'image de T^* sur celle de T.

8.4. Le cas général : opérateurs normaux

Il n'y a pas de raison de s'arrêter aux opérateurs hermitiens pour le calcul fonctionnel continu. Ce n'est qu'un cas particulier des opérateurs normaux, et la théorie du calcul fonctionnel continu se généralise dans son bon cadre à ces opérateurs. Il y a cependant des difficultés supplémentaires.

Dans deux cas particuliers de calcul fonctionnel, l'opérateur adjoint T^* est directement une fonction de T : trivialement dans le cas hermitien, puisqu'alors $\mathrm{T}^* = \mathrm{T}$, mais aussi dans le cas unitaire où $\mathrm{T}^* = \mathrm{T}^{-1}$. Notons comme d'habitude $\mathrm{K} = \mathrm{Sp}(\mathrm{T})$; on voit clairement de quelle fonction de $\mathrm{C}(\mathrm{K})$ l'opérateur T^* doit être l'image par φ_T dans ces deux cas : c'est $i_\mathrm{K} = \overline{i_\mathrm{K}}$ dans le cas hermitien et $i_\mathrm{K}^{-1} = \overline{i_\mathrm{K}}$ dans le cas unitaire. Cela ne sera plus si clair dans le cas général d'un opérateur normal, et on demandera explicitement que l'homomorphisme φ_T envoie la fonction $\overline{i_\mathrm{K}}$ sur T^*.

Il faut aussi généraliser nos polynômes : si la fonction i_K est envoyée sur T et la fonction $\overline{i_K}$ sur T^*, alors l'image de $i_K\overline{i_K}$ doit être TT^* ; dans le cas hermitien ou unitaire, la fonction $i_K\overline{i_K}$ s'exprime à partir d'un polynôme en i_K (i_K^2 dans le cas hermitien et 1 dans le cas unitaire) ; ceci n'est plus vrai maintenant, et la fonction $i_K\overline{i_K}$ est une nouvelle fonction qui doit être gardée dans notre algèbre de "polynômes" ; bien sûr le problème ne s'arrête pas là, et nous devons considérer $i_K^p\,\overline{i_K}^q$ pour tous entiers $p, q \geq 0$. Nous allons donc considérer l'algèbre $\mathbb{C}[X, Y]$ des polynômes en deux variables, puis prendre l'ensemble des fonctions sur $K = \mathrm{Sp}(T)$ obtenues en remplaçant X par i_K et Y par $\overline{i_K}$. Notre algèbre de base A qui remplacera l'algèbre des polynômes sera l'algèbre de toutes les fonctions f sur K de la forme

$$\forall z \in K, \quad f(z) = \sum_{p,q=0}^{N} c_{p,q} z^p\, \overline{z}^q.$$

avec $c_{p,q} \in \mathbb{C}$. On a envie de poser ensuite

$$f(T) = \sum_{p,q=0}^{N} c_{p,q} T^p\, (T^*)^q,$$

mais on n'est pas encore sûr que l'opérateur ainsi écrit ne dépend que de la fonction f sur K. La stratégie de démonstration sera toujours la même : l'algèbre A considérée est dense dans C(K) par Stone-Weierstrass (facile), et l'application que nous avons en tête sera isométrique.

Théorème 8.4.1. *Soient* H *un espace de Hilbert complexe et* $T \in \mathcal{L}(H)$ *normal ; posons* $K = \mathrm{Sp}(T)$ *; il existe un et un seul homomorphisme d'algèbres de Banach unitaires complexes* $\varphi_T : C(K) \to \mathcal{L}(H)$ *tel que* $\varphi_T(i_K) = T$ *et* $\varphi_T(\overline{i_K}) = T^*$.

L'homomorphisme φ_T *est isométrique. Si on note* $f(T) = \varphi_T(f)$, *on a* $f(T)^* = \overline{f}(T)$ *(donc* $f(T)$ *est normal) et* $f(T)$ *commute avec tout opérateur* S *qui commute avec* T *et avec* T^*. *On a*

$$\mathrm{Sp}(f(T)) = \mathrm{Sp}(f) = f(\mathrm{Sp}(T)).$$

Pour toute fonction continue f *sur* K *et toute fonction continue* g *sur* $f(K)$, *on a* $g(f(T)) = (g \circ f)(T)$.

Comme dans le cas hermitien, le point clé est de montrer que

$$\|P(T, T^*)\|_{\mathcal{L}(H)} = \max\{|P(\lambda, \overline{\lambda}) : \lambda \in K\}$$

pour tout polynôme de deux variables P. Ce point délicat sera présenté plus loin. Ce point établi, la démonstration est très semblable à celle du cas hermitien. On applique tout d'abord le théorème de Stone-Weierstrass pour dire que l'algèbre A est dense dans C(K), ce qui permet d'étendre la correspondance de A à C(K), et on vérifie comme avant la propriété d'homomorphisme. Expliquons la formule de composition. Posons $L = f(K) \subset \mathbb{C}$; on a un homomorphisme d'algèbres de Banach ψ de C(L) dans C(K) défini par $\psi(g) = g \circ f$ pour toute $g \in C(L)$. Considérons l'homomorphisme $\varphi = \psi \circ \varphi_T$, et cherchons les images de i_L et $\bar{i}_L$. On voit que $\psi(i_L)(t) = i_L(f(t)) = f(t)$ pour tout $t \in K$, donc $\psi(i_L) = f$, et $\psi(\bar{i}_L)(t) = \overline{f(t)}$, donc $\psi(\bar{i}_L) = \overline{f}$. On a donc $\varphi(i_L) = f(T)$ et $\varphi(\bar{i}_L) = \overline{f}(T) = (f(T))^*$; d'après l'unicité, on en déduit que $\varphi = \varphi_{f(T)}$. On obtient donc $g(f(T)) = (g \circ f)(T)$.

Considérons maintenant un opérateur normal T et

$$f(z) = \sum_{p,q=0}^{\mathrm{N}} c_{p,q} z^p \, \overline{z}^q, \quad \mathrm{V} = \sum_{p,q=0}^{\mathrm{N}} c_{p,q} \mathrm{T}^p \, (\mathrm{T}^*)^q.$$

Posons $\mathrm{K} = \mathrm{Sp}(\mathrm{T}) \subset \mathbb{C}$; nous voulons montrer que la norme de V est égale à la norme de f dans C(K). On sait que $\mathrm{V}^*\mathrm{V}$ est hermitien positif et $r = \|\mathrm{V}\|^2 = \|\mathrm{V}^*\mathrm{V}\| \in \mathrm{Sp}(\mathrm{V}^*\mathrm{V})$; d'après le lemme 2.6, il existe un sous-espace fermé F non nul, stable par tout opérateur qui commute avec $\mathrm{V}^*\mathrm{V}$, sur lequel $\mathrm{V}^*\mathrm{V} \sim r\, \mathrm{Id}_\mathrm{H}$; puisque T est normal, T et T^* commutent avec $\mathrm{V}^*\mathrm{V}$, par conséquent F est stable par T et T^* ; ceci permet de considérer la restriction T_1 de T à F, et de voir que son adjoint T_1^* est la restriction de T^*. Il en résulte que T_1 est un opérateur normal sur F. Soit λ une valeur spectrale de T_1 ; on sait d'après la remarque 6.3.4 qu'on peut trouver un vecteur $x \in \mathrm{F}$ de norme un tel que $\mathrm{T}_1(x) = \mathrm{T}(x) \sim \lambda\, x$ (ceci nous dit en passant que $\lambda \in \mathrm{Sp}(\mathrm{T}) = \mathrm{K}$), et $\mathrm{T}^*(x) \sim \overline{\lambda}\, x$. Il en résulte que $\mathrm{V}(x) \sim (\sum_{p,q} c_{p,q} \lambda^p \overline{\lambda}^q)\, x = f(\lambda)\, x$, donc $\|\mathrm{V}(x)\| \sim |f(\lambda)|$, mais par ailleurs $\|\mathrm{V}(x)\|^2 = \langle \mathrm{V}^*\mathrm{V}x, x\rangle \sim r\, \langle x, x\rangle = \|\mathrm{V}\|^2$. On en déduit que $\|\mathrm{V}\| \le \|f\|_{\mathrm{C(K)}}$.

A partir de là nous avons déjà une justification du fait que l'opérateur V ne dépend que de la fonction f sur K, ce qui permet de poser $\mathrm{V} = f(\mathrm{T})$, et de plus nous pouvons étendre par continuité la définition donnée sur l'algèbre A. Mais en fait il y a encore isométrie de A (munie de la norme de C(K)) dans $\mathcal{L}(\mathrm{H})$: soit en effet $\lambda \in \mathrm{K}$; d'après la remarque 6.3.4 il existe x de norme un tel que $\mathrm{T}(x) \sim \lambda x$ et $\mathrm{T}^*(x) \sim \overline{\lambda} x$, donc $f(\mathrm{T})(x) \sim f(\lambda) x$ par le même argument que précédemment, donc $\|f(\mathrm{T})\| \ge |f(\lambda)|$, pour tout $\lambda \in \mathrm{K}$, ce qui donne $\|f(\mathrm{T})\| \ge \|f\|_{\mathrm{C(K)}}$ et termine notre programme.

Théorème 8.4.2. *Soient* H *un espace de Hilbert complexe,* $\mathrm{T} \in \mathcal{L}(\mathrm{H})$ *un opérateur normal et* f *une fonction continue sur* $\mathrm{Sp}(\mathrm{T})$ *; alors, si* $f(\mathrm{Sp}(\mathrm{T})) \subset \mathbb{R}$*, on a* $f(\mathrm{T}) = f(\mathrm{T})^*$ *; si* $f(\mathrm{Sp}(\mathrm{T})) \subset \mathbb{T}$*, alors* $f(\mathrm{T})$ *est unitaire.*

Démonstration. Si $f = \overline{f}$ (sur Sp(T)) alors $f(\mathrm{T}) = \overline{f}(\mathrm{T}) = f(\mathrm{T})^*$; si $f(\mathrm{Sp}(\mathrm{T})) \subset \mathbb{T}$, alors $f\overline{f} = 1$, donc $f(\mathrm{T})^* f(\mathrm{T}) = f(\mathrm{T}) f(\mathrm{T})^* = (f\overline{f})(\mathrm{T}) = \mathrm{Id}_\mathrm{H}$, donc l'opérateur $f(\mathrm{T})$ est unitaire.

//

Exemples 8.4.1.

1. Supposons que U soit un opérateur unitaire sur H tel que $-1 \notin \mathrm{Sp}(\mathrm{U})$. La fonction $f(z) = i(z-1)/(z+1)$ est alors définie et continue sur Sp(U), et à valeurs réelles. Il en résulte que $f(\mathrm{U})$ est hermitien. L'opérateur $f(\mathrm{U})$ est égal à $i(\mathrm{U} - \mathrm{Id}_\mathrm{H})(\mathrm{U} + \mathrm{Id}_\mathrm{H})^{-1}$.

Inversement, si T est hermitien, on peut considérer la fonction $g(t) = (i+t)/(i-t)$ qui envoie $\mathbb{R}$ dans le cercle unité de $\mathbb{C}$. L'opérateur $g(\mathrm{T}) = (i + \mathrm{T})(i - \mathrm{T})^{-1}$ est unitaire.

2. Pour tout $s \in \mathbb{R}$ considérons la fonction f_s définie sur $\mathbb{R}$ par $f_s(t) = \mathrm{e}^{ist}$. Si T est hermitien, on peut considérer pour tout s l'opérateur $\mathrm{U}_s = f_s(\mathrm{T}) = \mathrm{e}^{is\mathrm{T}}$. C'est un opérateur unitaire puisque f_s est à valeurs dans $\mathbb{T}$. De plus $f_{s_1} f_{s_2} = f_{s_1+s_2}$ pour tous s_1, s_2, donc $\mathrm{U}_{s_1}\mathrm{U}_{s_2} = \mathrm{U}_{s_1+s_2}$. On dit qu'on a un *groupe* d'opérateurs unitaires. On verra plus loin dans le cours (dernière page du poly, théorème de Stone) une réciproque de ce fait, mais elle demandera de considérer des opérateurs autoadjoints *non bornés*.

3. Soient T un opérateur normal sur H et F un sous-espace fermé de H, stable par T et par T^* ; alors la projection orthogonale P_F commute avec T et T^*, donc avec tout opérateur $f(\mathrm{T})$. Si S désigne la restriction de T à F, alors S est un opérateur normal sur

F, et $\mathrm{Sp}(\mathrm{S}) \subset \mathrm{Sp}(\mathrm{T})$; pour toute fonction continue f sur le spectre de T l'opérateur $f(\mathrm{S})$ est la restriction de $f(\mathrm{T})$ à F.

4. Soit T un opérateur normal sur H ; supposons que le spectre de T puisse être découpé en deux compacts de $\mathbb{C}$ disjoints, disons $\mathrm{Sp}(\mathrm{T}) = \mathrm{K}_1 \cup \mathrm{K}_2$. Dans ce cas, la fonction f_1 qui est égale à 1 sur K_1 et à 0 sur K_2 est une fonction réelle continue sur Sp(T). Il en résulte que $\mathrm{P}_1 = f_1(\mathrm{T})$ est hermitien, et $\mathrm{P}_1^2 = \mathrm{P}_1$ puisque $f_1^2 = f_1$, donc P_1 est un projecteur orthogonal, qui commute avec T et avec T^*. On peut décomposer H en somme directe orthogonale $\mathrm{H}_1 \oplus \mathrm{H}_2$, où $\mathrm{H}_1 = \mathrm{P}_1(\mathrm{H})$; la restriction de T_1 à H_1 est un opérateur normal T_1 sur H_1 dont le spectre est égal à K_1 :

> si $\lambda \notin \mathrm{K}_1$, la fonction g définie par $g(z) = z - \lambda$ en tout point $z \in \mathrm{K}_1$ et $g(z) = 1$ pour tout $z \in \mathrm{K}_2$ est une fonction de $\mathrm{C}(\mathrm{Sp}(\mathrm{T}))$, inversible dans cette algèbre, donc son image est un opérateur inversible dans $\mathcal{L}(\mathrm{H})$. L'image est l'opérateur égal à $\mathrm{T}_1 - \lambda\,\mathrm{Id}_{\mathrm{H}_1}$ sur H_1 et à $\mathrm{Id}_{\mathrm{H}_2}$ sur H_2. Il en résulte que $\mathrm{T}_1 - \lambda\,\mathrm{Id}_{\mathrm{H}_1}$ est inversible pour tout $\lambda \notin \mathrm{K}_1$. On a donc $\mathrm{Sp}(\mathrm{T}_1) \subset \mathrm{K}_1$ et de la même façon $\mathrm{Sp}(\mathrm{T}_2) \subset \mathrm{K}_2$. Inversement si $\lambda \in \mathrm{K}_1$, on sait que $\mathrm{T}_2 - \lambda\,\mathrm{Id}_{\mathrm{H}_2}$ est inversible ; alors $\mathrm{T}_1 - \lambda\,\mathrm{Id}_{\mathrm{H}_1}$ ne peut pas être inversible, sinon $\mathrm{T} - \lambda\,\mathrm{Id}_{\mathrm{H}}$ le serait aussi.

9. Décomposition spectrale des opérateurs normaux

Dans ce chapitre, tous les espaces considérés sont complexes.

9.1. Opérateurs unitairement équivalents

Définition 9.1.1. Soient H_1 et H_2 deux espaces de Hilbert, et $\mathrm{T}_1 \in \mathcal{L}(\mathrm{H}_1)$, $\mathrm{T}_2 \in \mathcal{L}(\mathrm{H}_2)$; on dit que T_1 et T_2 sont *unitairement équivalents* s'il existe un opérateur unitaire $\mathrm{U} : \mathrm{H}_1 \to \mathrm{H}_2$ tel que $\mathrm{T}_1 = \mathrm{U}^* \circ \mathrm{T}_2 \circ \mathrm{U}$.

Soient H_1 et H_2 deux espaces de Hilbert, et $\mathrm{U} : \mathrm{H}_1 \to \mathrm{H}_2$ un opérateur unitaire. A tout opérateur $\mathrm{S}_2 \in \mathcal{L}(\mathrm{H}_2)$ associons l'opérateur $\mathrm{S}_1 = \mathrm{U}^*\mathrm{S}_2\mathrm{U} \in \mathcal{L}(\mathrm{H}_1)$. On vérifie que $\mathrm{S}_2 \to \mathrm{S}_1$ est un homomorphisme ψ d'algèbres de Banach unitaires de $\mathcal{L}(\mathrm{H}_2)$ dans $\mathcal{L}(\mathrm{H}_1)$, et de plus $\mathrm{Sp}(\mathrm{S}_2) = \mathrm{Sp}(\mathrm{S}_1)$ (parce que $\mathrm{S}_1 - \lambda\,\mathrm{Id}_{\mathrm{H}_1} = \mathrm{U}^*(\mathrm{S}_2 - \lambda\,\mathrm{Id}_{\mathrm{H}_2})\mathrm{U}$ pour tout $\lambda \in \mathbb{C}$). Si $\mathrm{T}_2 \in \mathcal{L}(\mathrm{H}_2)$ est normal, si on pose $\mathrm{K} = \mathrm{Sp}(\mathrm{T}_2) = \mathrm{Sp}(\mathrm{T}_1)$, et si on considère $\varphi = \psi \circ \varphi_{\mathrm{T}_2}$, on obtient un homomorphisme φ d'algèbres de Banach unitaires complexes de $\mathrm{C}(\mathrm{K})$ dans $\mathcal{L}(\mathrm{H}_1)$ tel que $\varphi(i_{\mathrm{K}}) = \mathrm{T}_1$, donc $\varphi = \varphi_{\mathrm{T}_1}$. Il en résulte d'après le théorème 8.4.1 que pour toute fonction continue f sur K,

$$f(\mathrm{T}_1) = \psi(f(\mathrm{T}_2)) = \mathrm{U}^* f(\mathrm{T}_2)\mathrm{U}.$$

Exemple 9.1.2. Le shift bilatéral sur $\ell_2(\mathbb{Z})$. A toute suite $x = (x_n)_{n\in\mathbb{Z}}$ dans $\ell_2(\mathbb{Z})$ on associe $\mathrm{S}(x) \in \ell_2(\mathbb{Z})$ définie par $\mathrm{S}(x)_n = x_{n-1}$ (décalage d'un cran vers la droite). Il est clair que S est isométrique, et bijective (l'inverse est le décalage à gauche), donc S est unitaire. On va voir que son spectre est le cercle unité entier. Pour tout $\lambda \in \mathbb{T}$, considérons un vecteur x_n ayant n coordonnées successives non nulles égales à $n^{-1/2}(1, \lambda^{-1}, \ldots, \lambda^{n-1})$. On voit facilement que $\|x_n\| = 1$ et $\|\mathrm{S}(x_n) - \lambda x_n\| = 2/\sqrt{n} \to 0$, donc $\mathrm{S} - \lambda\,\mathrm{Id}_{\mathrm{H}}$ n'est pas inversible. On va montrer un autre *modèle* pour l'opérateur S, qui rend son calcul fonctionnel facile.

Considérons l'espace $\mathrm{H}_2 = \mathrm{L}_2([0, 2\pi])$ muni de la mesure $dt/2\pi$ et de la base $(h_n)_{n\in\mathbb{Z}}$ où h_n est la fonction définie par $h_n(t) = \mathrm{e}^{int}$ (base de Fourier). Considérons sur H_2 l'opérateur T_2 de multiplication par la fonction g définie par $g(t) = \mathrm{e}^{it}$. Par ailleurs on considère l'espace $\mathrm{H}_1 = \ell_2(\mathbb{Z})$, avec sa base naturelle $(e_n)_{n\in\mathbb{Z}}$. Considérons l'isométrie surjective U de H_1 sur H_2 définie par $\mathrm{U}(e_n) = h_n$ pour tout $n \in \mathbb{Z}$, puis l'opérateur $\mathrm{U}^*\mathrm{T}_2\mathrm{U}$. On voit que cet opérateur envoie e_n sur e_{n+1} pour tout n : c'est le shift à droite S. On voit donc que le shift S sur $\ell_2(\mathbb{Z})$ est unitairement équivalent à l'opérateur de multiplication par la fonction $t \to \mathrm{e}^{it}$ sur H_2.

On a dit que le calcul fonctionnel des opérateurs de multiplication est simple : si f est une fonction continue sur $\mathbb{T}$, l'opérateur $f(\mathrm{T}_2)$ est l'opérateur de multiplication par la fonction $t \to f(\mathrm{e}^{it})$.

Les considérations de ce paragraphe sont un prélude au contenu de ce chapitre : on y verra que les longs développements du chapitre 8, un peu prolongés, permettent de dire que tout opérateur normal est unitairement équivalent à un modèle canonique : la multiplication par une fonction mesurable bornée sur un espace $\mathrm{L}_2(\Omega, \mu)$, pour lequel on a dit que le calcul fonctionnel était simple ! Autrement dit : on verra que tout était simple, mais c'était difficile de s'en apercevoir.

9.2. Opérateurs de multiplication et spectre

Soient $(\mathrm{X}, \mathcal{A}, \mu)$ un espace mesuré et $f \in \mathrm{L}_\infty(\mathrm{X}, \mathcal{A}, \mu)$ (pour abréger nous noterons simplement (X, μ) dans la suite, lorsque la mention de la tribu $\mathcal{A}$ ne sera pas utile) ; on rappelle que la norme $\|f\|_\infty$ de f dans l'espace de Banach L_∞ est donnée par le sup essentiel de la classe de fonctions f, c'est à dire le plus petit nombre réel $\mathrm{M} \geq 0$ tel que l'on ait $|f(s)| \leq \mathrm{M}$ pour μ-presque tout $s \in \mathrm{X}$. Si ξ est une fonction de $\mathrm{L}_2(\mathrm{X}, \mu)$, il est clair que le produit $f\xi$ est bien défini en tant que classe de fonctions et que $f\xi$ est de carré intégrable, avec $\|f\xi\|_2 \leq \|f\|_\infty \|\xi\|_2$.

Notons $\mathrm{M}_f : \mathrm{L}_2(\mathrm{X}, \mu) \to \mathrm{L}_2(\mathrm{X}, \mu)$ l'application qui à $\xi \in \mathrm{L}_2$ associe la fonction $f\xi$. Par ce qui précède, on sait que $\mathrm{M}_f \in \mathcal{L}(\mathrm{L}_2(\mathrm{X}, \mu))$ et $\|\mathrm{M}_f\| \leq \|f\|_\infty$. Si f, g sont deux fonctions mesurables bornées, il est clair que $\mathrm{M}_g \, \mathrm{M}_f = \mathrm{M}_{fg}$. Pour $\xi, \eta \in \mathrm{L}_2(\mathrm{X}, \mu)$, on a

$$\langle f\xi, \eta \rangle = \int_{\mathrm{X}} f(s)\xi(s)\overline{\eta(s)} \, d\mu(s) = \langle \xi, \overline{f}\eta \rangle,$$

donc $(\mathrm{M}_f)^* = \mathrm{M}_{\overline{f}}$. On en déduit immédiatement que M_f est normal. Remarquons que si f est réelle, alors M_f est autoadjoint, et que si $|f(s)| = 1$ pour μ-presque tout $s \in \mathrm{X}$, alors $(\mathrm{M}_f)^* \, \mathrm{M}_f = \mathrm{M}_{|f|^2} = \mathrm{Id}_{\mathrm{L}_2(\mathrm{X},\mu)}$, c'est à dire que M_f est unitaire.

Proposition 9.2.1. *Le spectre de* M_f *est l'ensemble des* $\lambda \in \mathbb{C}$ *tels que, pour tout* $\varepsilon > 0$ *l'ensemble* $\{s \in \mathrm{X} : |f(s) - \lambda| < \varepsilon\}$ *ne soit pas* μ*-négligeable.*

Démonstration. Pour lire plus facilement la démonstration on pourra penser que f est une *vraie* fonction mesurable bornée (c'est à dire qu'on pourrait choisir un représentant de la classe, etc. . .). Soit $\lambda \in \mathbb{C}$; s'il existe $\varepsilon > 0$ tel que l'ensemble $\{s \in \mathrm{X} : |f(s) - \lambda| < \varepsilon\}$ soit μ-négligeable, notons h la fonction définie par $h(s) = (f(s) - \lambda)^{-1}$ si $f(s) \neq \lambda$ et $h(s) = 0$ sinon. Alors pour μ-presque tout $s \in \mathrm{X}$ on a $|h(s)| \leq \varepsilon^{-1}$ et $h(s)(f(s) - \lambda) = 1$. On en déduit que $h \in \mathrm{L}_\infty(\mathrm{X}, \mu)$ et

$$\mathrm{M}_h \, (\mathrm{M}_f - \lambda \, \mathrm{Id}_{\mathrm{L}_2(\mathrm{X},\mu)}) = (\mathrm{M}_f - \lambda \, \mathrm{Id}_{\mathrm{L}_2(\mathrm{X},\mu)}) \, \mathrm{M}_h = \mathrm{Id}_{\mathrm{L}_2(\mathrm{X},\mu)} \, .$$

Inversement supposons que pour tout $\varepsilon > 0$ l'ensemble $\mathrm{A}_\varepsilon = \{s \in \mathrm{X} : |f(s) - \lambda| < \varepsilon\}$ soit tel que $\mu(\mathrm{A}_\varepsilon) > 0$; soit $\xi \in \mathrm{L}_2(\mathrm{X}, \mu)$ nulle hors de A_ε et telle que $\|\xi\|_2 = 1$ (par exemple un multiple convenable de la fonction indicatrice de l'ensemble A_ε) ; on voit alors que $|(\mathrm{M}_f - \lambda \, \mathrm{Id}_{\mathrm{L}_2(\mathrm{X},\mu)})(\xi)| \leq \varepsilon \, |\xi|$, donc $\|(\mathrm{M}_f - \lambda \, \mathrm{Id}_{\mathrm{L}_2(\mathrm{X},\mu)})(\xi)\| \leq \varepsilon$; il est clair que $\mathrm{M}_f - \lambda \, \mathrm{Id}_{\mathrm{L}_2(\mathrm{X},\mu)}$ n'est pas inversible (son inverse devrait avoir une norme $\geq \varepsilon^{-1}$, pour tout $\varepsilon > 0$).

//

Proposition 9.2.2. *Pour toute fonction* $g \in \mathrm{C}(\mathrm{Sp}(\mathrm{M}_f))$ *on a* $g(\mathrm{M}_f) = \mathrm{M}_{g \circ f}$.

Démonstration. L'application $g \to \mathrm{M}_{g \circ f}$ est un homomorphisme d'algèbres de Banach unitaires qui envoie i_{K} sur M_f, d'où le résultat par l'unicité dans le théorème 8.4.1.

//

9.3. Décomposition spectrale

Soient H un espace de Hilbert complexe et T un opérateur normal sur H ; étant donné un vecteur non nul x, on va s'intéresser au plus petit sous-espace vectoriel fermé F_x de H contenant x et qui soit stable par T et par T^*. Il est clair que ce sous-espace F_x doit contenir tous les vecteurs de la forme $\mathrm{T}^k(\mathrm{T}^*)^\ell(x)$, avec $k, \ell \geq 0$. Inversement, le sous-espace fermé engendré par tous ces vecteurs est stable par T et T^* (grâce à la commutation $\mathrm{TT}^* = \mathrm{T}^*\mathrm{T}$). L'espace F_x est donc égal à l'adhérence de $\mathrm{Vect}\{\mathrm{T}^k(\mathrm{T}^*)^\ell(x) : k, \ell \geq 0\}$. Il est stable par tout opérateur de la forme $f(\mathrm{T})$ obtenu par calcul fonctionnel continu.

Convenons de dire (entre nous) que T est *monogène* s'il existe un vecteur $x_0 \in \mathrm{H}$ tel que $\mathrm{H} = \mathrm{F}_{x_0}$. Dans ce cas particulier, le théorème de représentation prend une forme bien sympathique.

Proposition 9.3.1. *Soient* H *un espace de Hilbert,* $\mathrm{T} \in \mathcal{L}(\mathrm{H})$ *un opérateur normal* monogène *et* $\mathrm{K} = \mathrm{Sp}(\mathrm{T})$ *; il existe une probabilité* μ *sur* K *telle que* T *soit unitairement équivalent à l'opérateur* $\mathrm{M}_{i_\mathrm{K}} \in \mathcal{L}(\mathrm{L}_2(\mathrm{K}, \mu))$ *de multiplication par la fonction* i_K.

Démonstration. Soit x_0 un vecteur tel que $\mathrm{F}_{x_0} = \mathrm{H}$; on peut choisir $\|x_0\| = 1$ si on veut. Considérons la forme linéaire $\varphi : f \to \langle f(\mathrm{T})(x_0), x_0\rangle$ sur $\mathrm{C}(\mathrm{K})$; si $f \in \mathrm{C}(\mathrm{K})$ est réelle positive, on sait par le calcul fonctionnel (théorème 8.4.1) que $f(\mathrm{T})$ est hermitien positif, donc $\langle f(\mathrm{T})(x_0), x_0\rangle \geq 0$. On voit ainsi que φ est une forme linéaire positive sur $\mathrm{C}(\mathrm{K})$; donc il existe une unique mesure positive μ sur K telle que, pour toute fonction $f \in \mathrm{C}(\mathrm{K})$ on ait $\varphi(f) = \int_\mathrm{K} f(t)\, d\mu(t)$; on a $\mu(\mathrm{K}) = \int 1\, d\mu = \langle 1(\mathrm{T})(x_0), x_0\rangle = \|x_0\|^2 = 1$, donc μ est une probabilité. Si y est un vecteur de F_{x_0} de la forme $y = f(\mathrm{T})(x_0)$, avec $f \in \mathrm{C}(\mathrm{K})$, on a

$$\|y\|^2 = \langle f(\mathrm{T})(x_0), f(\mathrm{T})(x_0)\rangle = \langle \overline{f}(\mathrm{T})f(\mathrm{T})(x_0), x_0\rangle = \int_\mathrm{K} |f(t)|^2\, d\mu(t)\,;$$

la relation précédente montre que y ne dépend que de la classe $\widehat{f}$ de f dans $\mathrm{L}_2(\mathrm{K}, \mu)$, et que l'application $u_0 : \widehat{f} \to f(\mathrm{T})(x_0)$, définie sur l'image Y de $\mathrm{C}(\mathrm{K})$ dans $\mathrm{L}_2(\mathrm{K}, \mu)$, est isométrique, de Y muni de la norme de $\mathrm{L}_2(\mathrm{K}, \mu)$ vers la norme de H ; par ailleurs, le sous-espace Y est dense dans $\mathrm{L}_2(\mathrm{K}, \mu)$ par un résultat général d'intégration. On peut donc prolonger u_0 en une isométrie u de $\mathrm{L}_2(\mathrm{K}, \mu)$ dans H. Pour tous $k, \ell \geq 0$, le vecteur $y = \mathrm{T}^k(\mathrm{T}^*)^\ell(x_0)$ est dans l'image de u, puisqu'il provient de la fonction continue $i_\mathrm{K}^k(\overline{i_\mathrm{K}})^\ell$; l'image de u est donc dense puisque T est monogène, et u est unitaire de $\mathrm{L}_2(\mathrm{K}, \mu)$ dans H. Si $y = u(f)$ avec f continue, on a $y = f(\mathrm{T})(x_0)$, donc

$$\mathrm{T}(u(f)) = \mathrm{T}f(\mathrm{T})(x_0) = (i_\mathrm{K} f)(\mathrm{T})(x_0) = u(\mathrm{M}_{i_\mathrm{K}}(f)).$$

Cette relation $\mathrm{T} \circ u = u \circ \mathrm{M}_{i_\mathrm{K}}$, vraie sur le sous-espace dense Y de $\mathrm{L}_2(\mathrm{K}, \mu)$, se prolonge à $\mathrm{L}_2(\mathrm{K}, \mu)$ tout entier et dit que T et la multiplication par i_K sont conjugués par l'opérateur unitaire u, donc unitairement équivalents.

//

Remarque 9.3.1. Puisque $\mathrm{M}_{i_\mathrm{K}}$ est unitairement équivalente à T, son spectre est K tout entier. D'après la proposition 2.1, on voit que $\mu(\mathrm{U}) > 0$ pour tout ouvert non vide de K. On dit que le *support* de la mesure μ est égal à K. Il en résulte que l'application naturelle de $\mathrm{C}(\mathrm{K})$ dans $\mathrm{L}_2(\mathrm{K}, \mu)$ est injective.

Si on donne un opérateur normal $\mathrm{T} \in \mathcal{L}(\mathrm{H})$ et un vecteur $x \in \mathrm{H}$ non nul, on peut considérer la restriction S de T au sous-espace F_x ; puisque F_x est aussi stable par T^*, on voit facilement que $\mathrm{S}^* \in \mathcal{L}(\mathrm{F}_x)$ est la restriction de T^* à F_x, et que S est normal. De plus, S est évidemment monogène, en tant qu'opérateur de l'espace de Hilbert F_x. On voit ainsi qu'on a des tas de bouts de l'opérateur T qui peuvent être représentés au moyen de la proposition précédente. Le résultat général (théorème 3) consistera simplement à décomposer H en somme directe orthogonale de tels morceaux, et à exprimer le recollement des morceaux.

Donnons une illustration de la proposition précédente dans un cas simple. Supposons que T soit un opérateur autoadjoint compact sur H ; il existe alors une suite $(\lambda_n)_{n\ge 0}$ de valeurs propres de T, tendant vers 0 ; supposons ces valeurs propres deux à deux distinctes et non nulles ; le spectre K de T est l'adhérence de l'ensemble des points de la suite (λ_n), c'est à dire que

$$\mathrm{K} = \{0\} \cup \{\lambda_n : n \ge 0\}.$$

Par ailleurs, on peut trouver une base orthonormée $(e_n)_{n\ge 0}$ de H formée de vecteurs propres de T, qui vérifient donc $\mathrm{T}(e_n) = \lambda_n e_n$ pour tout $n \ge 0$. Choisissons un vecteur $x_0 = \sum_{n=0}^{+\infty} c_n e_n$ tel que $c_n \ne 0$ pour tout $n \ge 0$ et considérons l'espace F_{x_0} ; on va voir que $\mathrm{F}_{x_0} = \mathrm{H}$. Pour toute fonction f continue sur K, l'opérateur $f(\mathrm{T})$ est l'opérateur qui vérifie $f(\mathrm{T})(e_n) = f(\lambda_n)e_n$ pour tout $n \ge 0$. Pour tout n fixé, la fonction f_n sur K qui vaut 1 au point λ_n et 0 ailleurs est une fonction continue sur K, et $f_n(\mathrm{T})(x_0) = e_n$ d'après la formule précédente. Puisque F_{x_0} est stable par $f_n(\mathrm{T})$, on voit que $e_n \in \mathrm{F}_{x_0}$ pour tout $n \ge 0$, donc $\mathrm{H} = \mathrm{F}_{x_0}$ dans le cas présent. On est donc dans le cas monogène et

$$\langle f(\mathrm{T})(x_0), x_0\rangle = \sum_{n=0}^{+\infty} f(\lambda_n)|c_n|^2 = \int_{\mathrm{K}} f(\lambda)\, d\mu(\lambda),$$

ce qui montre que la mesure μ est la mesure sur K donnée par $\mu = \sum_{n=0}^{+\infty} |c_n|^2 \delta_{\lambda_n}$. Décrivons l'isomorphisme u : chaque fonction $g \in \mathrm{L}_2(\mathrm{K}, \mu)$ se réduit à la donnée de ses valeurs aux points λ_n, disons $g(\lambda_n) = d_n$, qui vérifient $\sum_{n=0}^{+\infty} |d_n|^2|c_n|^2 = \|g\|_2^2 < \infty$. L'opérateur u agit par $u(g) = \sum_{n=0}^{+\infty} d_n c_n e_n$. On voit que u est surjectif sous les hypothèses que nous avons faites.

Mais ce n'est pas toujours le cas : nous aurions pu prendre bêtement un vecteur $x = \sum_{n=0}^{+\infty} c_n e_n$ tel que $c_{n_0} = 0$. Dans ce cas, le vecteur de base e_{n_0} ne serait pas dans F_x, donc pas dans l'image de l'application u construite dans la proposition précédente ; une difficulté plus subtile vient de la possible multiplicité des valeurs propres ; dans ce cas l'ensemble K des valeurs propres ne traduit pas cette multiplicité ; soit donc (λ_n) la suite des valeurs propres, répétées selon leur multiplicité éventuelle ; supposons que $\lambda \ne 0$ est une valeur propre correspondant à un sous-espace propre E_λ de dimension deux exactement, disons $\lambda_{n_1} = \lambda_{n_2} = \lambda$ avec $n_1 \ne n_2$, et posons $x = \sum_{n=0}^{+\infty} c_n e_n$ comme avant (c'est à dire $c_n \ne 0$ pour tout n). On peut vérifier que le vecteur $c_{n_2}e_{n_1} - c_{n_1}e_{n_2} \in \mathrm{E}_\lambda$ est orthogonal à F_x, et la construction de la proposition précédente ne donne rien pour ce vecteur. Cette longue explication est censée éclairer ce qui se passe réellement dans le lemme 2 ci-dessous.

Lemme 9.3.2. *Soient* H *un espace de Hilbert et* T *un opérateur normal sur* H *; il existe une partie* D *de la sphère unité de* H *telle que les* F_x *pour* $x \in \mathrm{D}$ *soient deux à deux orthogonaux et que la somme des* F_x *pour* $x \in \mathrm{D}$ *soit dense dans* H. *Tout vecteur* $z \in \mathrm{H}$ *admet donc une représentation unique* $z = \sum_{x\in\mathrm{D}} z_x$, *où chaque* $z_x \in \mathrm{F}_x$ *est la projection orthogonale de* z *sur* F_x.

Démonstration. Si $y \in \mathrm{F}_x^\perp$, alors y est orthogonal à tout vecteur de la forme $\mathrm{T}^k(\mathrm{T}^*)^\ell(x)$, ce qui donne pour tous $m, n, p, q \ge 0$

$$\langle \mathrm{T}^m(\mathrm{T}^*)^n y, \mathrm{T}^p(\mathrm{T}^*)^q(x)\rangle = \langle y, \mathrm{T}^{n+p}(\mathrm{T}^*)^{m+q}(x)\rangle = 0$$

et montre que tous les vecteurs de F_y sont orthogonaux à F_x. Notons $\mathcal{D}$ l'ensemble des parties D de la sphère unité de H telles que, pour tout $x, y \in D$, les espaces F_x et F_y soient orthogonaux. Muni de l'ordre de l'inclusion, $\mathcal{D}$ est inductif : soit $\{D_i : i \in I\}$ une partie totalement ordonnée de $\mathcal{D}$; si $x, y \in \bigcup_{i \in I} D_i$, tels que $x \neq y$, il existe $i \in I$ tel que $x, y \in D_i$ donc les espaces F_x et F_y sont orthogonaux ; il s'ensuit que $\bigcup_{i \in I} D_i$ est un élément de $\mathcal{D}$ majorant l'ensemble $\{D_i : i \in I\}$.

Soit D un élément maximal de $\mathcal{D}$ et soit F l'adhérence de la somme des $(F_x)_{x \in D}$; si on avait $F \neq H$, on pourrait trouver $y \in H$ de norme un orthogonal aux $(F_x)_{x \in D}$; clairement, on a $y \notin D$. Par le premier point, F_y est orthogonal aux espaces F_x, pour tout $x \in D$. Alors $D \cup \{y\} \in \mathcal{D}$, ce qui contredit la maximalité de D. On a ainsi montré que $F = H$.

Soit $z \in H$ et désignons par z_x la projection orthogonale de z sur le sous-espace fermé F_x, pour tout $x \in D$. Pour tout sous-ensemble fini $J \subset D$, on vérifie que $\sum_{x \in J} z_x$ est la projection orthogonale de z sur le sous-espace fermé $\sum_{x \in J} F_x$, donc $\sum_{x \in J} \|z_x\|^2 \leq \|z\|^2$. Il en résulte que $\sum_{x \in D} \|z_x\|^2 < +\infty$, donc la famille $\sum_{x \in D} z_x$ est sommable. Sa somme y est telle que $z - y$ est orthogonal à chaque espace F_x, donc $z - y = 0$ puisque la somme des F_x est dense dans H.

//

Théorème 9.3.3. *Soient* H *un espace de Hilbert séparable et* $T \in \mathcal{L}(H)$ *un opérateur normal ; il existe un espace mesuré* (X, μ)*, une fonction* $f \in L_\infty(X, \mu)$ *et un isomorphisme unitaire* $u : L_2(X, \mu) \to H$ *tels que* $T = u\, M_f\, u^*$.

Démonstration. Soit D comme dans le lemme 2 ; il résulte de ce lemme que le système D est orthogonal ; comme H est séparable, D est (fini ou) dénombrable. Pour chaque $y \in D$, la restriction T_y de T à F_y est un opérateur normal dont le spectre est clairement contenu dans celui de T ; notons μ_y la mesure sur $K_y = \mathrm{Sp}(T_y) \subset K = \mathrm{Sp}(T)$ et $u_y : L_2(K, \mu_y) \to F_y$ l'isométrie décrite dans la proposition 1, qui conjugue T_y et la multiplication par i_{K_y} dans $L_2(K_y, \mu_y)$. La stratégie de la démonstration est essentiellement triviale : notons $(y_n)_{n \geq 0}$ la famille des éléments de D et posons $F_n = F_{y_n}$, $T_n = T_{y_n} \in \mathcal{L}(F_n)$ pour tout $n \geq 0$. Chaque T_n est unitairement équivalent à une multiplication M_n sur un compact K_n muni d'une probabilité μ_n. On va fabriquer un espace X qui sera la réunion de copies des K_n, mais disjointes, et une mesure μ sur X telle que $L_2(X, \mu)$ soit la somme directe orthogonale d'espaces $E_n \simeq L_2(K_n, \mu_n)$. Puisque de l'autre côté H est somme directe orthogonale des F_n, et que pour chaque n on a une isométrie surjective u_n de E_n sur F_n, on construira facilement une isométrie u de $L_2(K, \mu)$ sur H, qui conjuguera l'opérateur M sur $L_2(X, \mu)$, égal sur chaque E_n à la multiplication M_n, avec l'opérateur T. Mais M est une multiplication sur X, puisque sur chacun des morceaux disjoints qui compose X l'opérateur M_n était une multiplication.

Pour simplifier la description on peut considérer que les mesures μ_n sont des mesures sur K (qui contient tous les K_n). Posons $X = K \times \mathbb{N}$; l'ensemble X est la réunion dénombrable des ensembles disjoints $X_n = K \times \{n\}$, qui sont des “copies” du compact K, indexées par $n \geq 0$. Considérons la tribu $\mathcal{A}$ de parties de X formée des ensembles $A \subset X$ dont l'intersection A_n avec chaque sous-ensemble X_n, $n \geq 0$, est un borélien du compact X_n ; l'ensemble A_n est de la forme $A_n = B_n \times \{n\}$, où B_n est un borélien de K. On définit alors une mesure μ (σ-finie) sur $(X, \mathcal{A})$ en posant

$$\mu(A) = \sum_{n \geq 0} \mu_n(B_n)$$

pour tout $A \in \mathcal{A}$. Si g est une fonction sur X, pour tout $n \geq 0$ on note g_n la fonction $s \to g(s,n)$ sur K. On voit alors que si $g \in L_2(X,\mu)$, chaque fonction g_n, $n \geq 0$ est dans $L_2(K,\mu_n)$ et

$$\int_X |g|^2\, d\mu = \sum_{n\geq 0} \int_K |g_n(s)|^2\, d\mu_n(s).$$

On désigne par E_n le sous-espace de $L_2(X,\mu)$ formé des fonctions h nulles hors de X_n, c'est à dire telles que $h(s,m) = 0$ pour tout $m \neq n$; on a une isométrie j_n de $L_2(K,\mu_n)$ sur E_n qui associe à $f \in L_2(K,\mu_n)$ la fonction $j_n(f)(s,m) = f(s)$ si $m = n$ et $j_n(f)(s,m) = 0$ si $m \neq n$. On voit que $L_2(X,\mu)$ est l'adhérence de la somme des sous-espaces deux à deux orthogonaux $(E_n)_{n\geq 0}$; si $g \in E_n$, on a $g = j_n(g_n)$.

Pour chaque $n \geq 0$ soit v_n l'application qui associe à $g \in E_n$ le vecteur $u_n(g_n) \in F_n$. Il est facile de vérifier que v_n est une isométrie de E_n sur F_n. Si on définit u de $L_2(X,\mu)$ dans H en posant

$$u(g) = \sum_{n\geq 0} u_n(g_n)$$

pour toute fonction g dans $L_2(X,\mu)$, on aura

$$\|u(g)\|^2 = \sum_{n\geq 0} \|u_n(g_n)\|^2 = \sum_{n\geq 0} \int_K |g_n(s)|^2\, d\mu_n(s) = \|g\|_2^2,$$

ce qui montre que u est isométrique de $L_2(X,\mu)$ dans H. Montrons que u est surjective. Si z est un vecteur de H, on peut écrire $z = \sum_{n\geq 0} z_n$ avec $z_n \in F_n$; pour chaque $n \geq 0$ on peut trouver une fonction $g_n \in L_2(K,\mu_n)$ telle que $z_n = u_n(g_n)$; définissons une fonction mesurable g sur X en posant $g(s,n) = g_n(s)$ pour tout $(s,n) \in K \times \mathbb{N}$. On vérifie que $g \in L_2(X,\mu)$ et $u(g) = \sum_{n\geq 0} u_n(g_n) = z$, donc u est surjective.

Notons $f : X \to \mathbb{C}$ l'application $(s,n) \to s \in K \subset \mathbb{C}$. Si $g \in L_2(X,\mu)$, on aura $(fg)_n = i_K g_n$ pour tout $n \geq 0$, donc également $(fg)_n(T) = T g_n(T)$. On en déduit :

$$u(M_f(g)) = u(fg) = \sum_{n\geq 0} (fg)_n(T)(y) = T\, u(g)\,;$$

Cela donne $uM_f = Tu$, donc $T = T\,u\,u^* = u\,M_f\,u^*$.

//

Soient H un espace de Hilbert séparable et $T \in \mathcal{L}(H)$ un opérateur normal ; écrivons $T = u\,M_f\,u^*$. L'application $g \to u\,M_{g\circ f}\,u^*$ est un morphisme continu d'algèbres de $C(\mathrm{Sp}(T))$ dans $\mathcal{L}(H)$; par l'unicité dans le théorème 8.4.1, pour toute fonction continue $g \in C(\mathrm{Sp}(T))$, on a $g(T) = u\,M_{g\circ f}\,u^*$.

Théorème 9.3.4. *Soient* H *un espace de Hilbert séparable et* $T \in \mathcal{L}(H)$ *un opérateur normal ; notons* $K = \mathrm{Sp}(T)$ *et* $\mathcal{L}_\infty(K,\mathcal{B})$ *l'espace vectoriel des fonctions boréliennes bornées sur* K. *Il existe un unique homomorphisme d'algèbres unitaires* $\Phi : \mathcal{L}_\infty(K,\mathcal{B}) \to \mathcal{L}(H)$ *satisfaisant* $\Phi(i_K) = T$, $\Phi(\overline{i_K}) = T^*$, *et tel que pour toute suite bornée* $g_n \in \mathcal{L}_\infty(K,\mathcal{B})$ *convergeant simplement vers* $g \in \mathcal{L}_\infty(K,\mathcal{B})$ *la suite* $\Phi(g_n)$ *converge fortement vers* $\Phi(g)$. *Si* $g \in C(K)$ *on a* $\Phi(g) = g(T)$.

Démonstration. Montrons d'abord l'existence ; soient (X,μ) un espace mesuré, f une fonction de $L_\infty(X,\mu)$ et soit $u : L_2(X,\mu) \to H$ un isomorphisme tels que $T = u\,M_f\,u^*$ (théorème 3) ; notons $\Phi : \mathcal{L}_\infty(K,\mathcal{B}) \to \mathcal{L}(H)$ l'application $g \to u\,M_{g\circ f}\,u^*$;

l'application Φ est clairement un homomorphisme d'algèbres unitaires satisfaisant $\Phi(i_{\mathrm{K}}) = \mathrm{T}$ et $\Phi(\overline{i_{\mathrm{K}}}) = \mathrm{T}^*$. Soit $g_n \in \mathcal{L}_\infty(\mathrm{K}, \mathcal{B})$ une suite bornée par un réel $\mathrm{M} \geq 0$ et convergeant simplement vers $g \in \mathcal{L}_\infty(\mathrm{K}, \mathcal{B})$; alors, pour tout $x \in \mathrm{H}$, la suite $(g_n \circ f)u^*(x)$ est une suite dans $\mathrm{L}_2(\mathrm{X}, \mu)$ dominée par $\mathrm{M}u^*(x)$ et convergeant partout vers $(g \circ f)u^*(x)$. Par le théorème de convergence dominée, la suite $(g_n \circ f)u^*(x)$ converge vers $(g \circ f)u^*(x)$ dans L_2, donc $\Phi(g_n)(x) = u(g_n \circ f)u^*(x)$ converge (en norme) vers $u(g \circ f)u^*(x) = \Phi(g)(x)$. Remarquons que si $g \in \mathrm{C}(\mathrm{K})$ on a $\Phi(g) = g(\mathrm{T})$.

Montrons l'unicité. Soient Φ_1 et Φ_2 deux applications vérifiant les conditions ci-dessus ; posons $\mathrm{A} = \{g \in \mathcal{L}_\infty(\mathrm{K}, \mathcal{B}) : \Phi_1(g) = \Phi_2(g)\}$. On doit montrer que l'ensemble A est égal à $\mathcal{L}_\infty(\mathrm{K}, \mathcal{B})$. Par hypothèse, A est un sous-espace vectoriel de $\mathcal{L}_\infty(\mathrm{K}, \mathcal{B})$ et pour toute suite bornée $g_n \in \mathrm{A}$ convergeant simplement vers $g \in \mathcal{L}_\infty(\mathrm{K}, \mathcal{B})$, on a $g \in \mathrm{A}$; de plus, par l'unicité dans le théorème 8.4.1, A contient toutes les fonctions continues sur K. Toute fonction borélienne bornée est limite uniforme de fonctions boréliennes prenant un nombre fini de valeurs ; celles-ci sont combinaisons linéaires de fonctions caractéristiques de sous-ensembles boréliens. Il suffit donc de prouver que A contient toutes les fonctions caractéristiques. Notons $\mathcal{C}$ l'ensemble des boréliens de K dont la fonction caractéristique est dans A. Comme A est un sous-anneau, $\mathcal{C}$ est stable par intersection finie ; comme A est stable par limite de suites bornées, $\mathcal{C}$ est stable par intersection dénombrable. De plus, si $g \in \mathrm{A}$, on a aussi $1 - g \in \mathrm{A}$, donc $\mathcal{C}$ est stable par passage au complémentaire. Donc $\mathcal{C}$ est une tribu. Pour savoir que $\mathcal{C}$ contient tous les boréliens, il suffit de montrer que $\mathcal{C}$ contient tous les fermés. Soit F un fermé non vide de K ; notons $h_0 : \mathrm{K} \to \mathbb{R}$ la fonction qui à $x \in \mathrm{K}$ associe sa distance à F. Posons $g = \sup(1 - h_0, 0)$ et $g_n = g^n$. Pour tout n on a $g_n \in \mathrm{C}(\mathrm{K}) \subset \mathrm{A}$; de plus la suite g_n converge vers la fonction caractéristique de F. On a montré que $\mathrm{K} \in \mathcal{C}$.

//

Soient H un espace de Hilbert, $\mathrm{T} \in \mathcal{L}(\mathrm{H})$ normal et f une fonction borélienne sur $\mathrm{Sp}(\mathrm{T})$; l'élément $\Phi(f)$ défini dans le théorème 4 se note encore $f(\mathrm{T})$.

Remarque 9.3.2. On a démontré une espèce de théorème de représentation de Riesz : quand on a une application linéaire continue de $\mathrm{C}(\mathrm{K})$ dans $\mathbb{C}$, on en déduit une mesure μ sur K qui donne une application linéaire continue de $\mathcal{L}_\infty(\mathrm{K}, \mathcal{B})$, l'espace des fonctions boréliennes bornées sur K, à valeurs dans $\mathbb{C}$. On a réalisé le même programme en partant de l'application linéaire continue $f \to f(\mathrm{T})$ de $\mathrm{C}(\mathrm{K})$ dans $\mathcal{L}(\mathrm{H})$: l'extension opère de $\mathcal{L}_\infty(\mathrm{K}, \mathcal{B})$ dans $\mathcal{L}(\mathrm{H})$.

Remarque 9.3.3. Dans le cas d'un espace de Hilbert réel H, le théorème 4 s'applique encore aux opérateurs hermitiens, mais il ne s'applique plus aux normaux réels. On peut déduire le résultat réel hermitien en utilisant une complexification de l'espace réel H, comme on l'a déjà fait pour le calcul fonctionnel, ou bien en recopiant la démonstration de la proposition 1, mais dans un contexte réel.

Exemple 9.3.4. Soit T un opérateur normal sur un espace de Hilbert complexe (resp : un opérateur hermitien sur un Hilbert réel ou complexe) ; si $f = \mathbf{1}_{\mathrm{A}}$ est l'indicatrice d'un borélien de $\mathbb{C}$ (resp : de $\mathbb{R}$) contenu dans le spectre de T, et si $\mathrm{P} = f(\mathrm{T})$, on aura $\mathrm{P}^* = \mathrm{P}$ parce que f est réelle, et $\mathrm{P}^2 = \mathrm{P}$ parce que $f^2 = f$. L'opérateur P est donc un projecteur orthogonal. On dit que P est un *projecteur spectral*. Puisque P est obtenu par calcul fonctionnel, il commute avec T et T^*, donc l'image $\mathrm{P}(\mathrm{H})$ est stable par T et T^*, ce qui permet de considérer la restriction S de T à $\mathrm{F} = \mathrm{P}(\mathrm{H})$, qui est un opérateur normal (resp : hermitien) sur le sous-espace F. On peut voir que si F n'est pas nul, le spectre de S est contenu dans l'adhérence de l'ensemble A.

Si $A = \{\lambda\}$ est un singleton, on peut voir qu'il est possible que le projecteur $\mathbf{1}_A(T)$ soit nul, alors que la fonction $\mathbf{1}_A$ n'est pas nulle. L'homomorphisme du théorème 4 n'est pas isométrique de $\mathcal{L}_\infty(K, \mathcal{B})$ dans $\mathcal{L}(H)$; pour obtenir une isométrie, il faudrait travailler avec une norme $L_\infty(\nu)$ et une mesure ν provenant du théorème de représentation 3. Si T est monogène et si μ est la probabilité sur Sp(T) obtenue à la proposition 1, la norme de $f(T)$, quand $f \in \mathcal{L}_\infty(\mathrm{Sp}(T), \mathcal{B})$, est exactement la norme de f dans $L_\infty(\mathrm{Sp}(T), \mu)$.

Exercice 9.3.5. Montrer que $\mathbf{1}_{\{\lambda\}}(T)$ est le projecteur orthogonal sur $\ker(T - \lambda\,\mathrm{Id}_H)$. Il est donc nul chaque fois que λ est dans le spectre de T sans être valeur propre de T.

10. Décomposition spectrale d'un opérateur autoadjoint (non borné)

10.1. Opérateurs non bornés

Préliminaires algébriques

Dans cette sous-section il ne sera question que d'algèbre linéaire : pas un poil de topologie. Soient E et F deux espaces vectoriels ; une *application linéaire partiellement définie* (un peu plus loin, on dira un *opérateur*) T de E dans F est donnée par un sous-espace vectoriel dom(T) de E appelé *domaine* de T et par une application linéaire (usuelle) $\mathrm{L_T}$ de dom(T) dans F.

Autrement dit, la donnée T est celle de $(\mathrm{E}, \mathrm{F}, \mathrm{dom(T)}, \mathrm{L_T})$. Le *graphe* de T est le sous-espace vectoriel du produit $\mathrm{E} \times \mathrm{F}$ égal à $\mathrm{Gr(T)} = \{(x, \mathrm{L_T}(x)) : x \in \mathrm{dom(T)}\}$. On verra que T est complètement déterminé par Gr(T), qui est un sous-espace vectoriel G de $\mathrm{E} \times \mathrm{F}$, avec la propriété $\big((0_\mathrm{E}, y) \in \mathrm{G}\big) \Rightarrow (y = 0_\mathrm{F})$.

Dans la suite, pour tout $x \in \mathrm{dom(T)}$ on posera $\mathrm{T}(x) = \mathrm{L_T}(x)$ et on ne fera plus la distinction entre $\mathrm{L_T}(x)$ et $\mathrm{T}(x)$. On laissera donc tomber complètement $\mathrm{L_T}$. Si T est une application linéaire partiellement définie, le *graphe* de T est donc le sous-espace vectoriel du produit $\mathrm{E} \times \mathrm{F}$ égal à $\mathrm{Gr(T)} = \{(x, \mathrm{T}(x)) : x \in \mathrm{dom(T)}\}$. La restriction à Gr(T) de la première projection est injective. Réciproquement, appelons *graphe partiel* tout sous-espace vectoriel G de $\mathrm{E} \times \mathrm{F}$ tel que la restriction de la première projection à G soit injective. Autrement dit, si $(x, y) \in \mathrm{G}$ et $(x, y') \in \mathrm{G}$, alors $y = y'$; ou encore : si $(0, y) \in \mathrm{G}$, alors $y = 0$. On voit que tout graphe partiel est le graphe d'une unique application linéaire partiellement définie T. La correspondance qui à T associe son graphe est une correspondance bijective entre applications linéaires partiellement définies et graphes partiels :

> soit $\mathrm{G} \subset \mathrm{E} \times \mathrm{F}$ un graphe partiel. Notons $p_1 : \mathrm{G} \to \mathrm{E}$ et $p_2 : \mathrm{G} \to \mathrm{F}$ les projections et définissons un opérateur T en posant $\mathrm{dom(T)} = p_1(\mathrm{G})$ et $\mathrm{T}(p_1(x)) = p_2(x)$ pour tout $x \in \mathrm{G}$. Il est clair que $\mathrm{Gr(T)} = \mathrm{G}$. Comme le noyau de la première projection de $\mathrm{E} \times \mathrm{F}$ dans E est le sous-espace $\{0\} \times \mathrm{F}$ de $\mathrm{E} \times \mathrm{F}$, la correspondance entre opérateur et graphe partiel est bijective.

Désormais on dira *opérateur* au lieu d'application linéaire partiellement définie. On appelle respectivement *noyau* et *image* de T le sous-espace $\ker(\mathrm{T}) = \{x \in \mathrm{dom(T)} : \mathrm{T}(x) = 0\}$ de E et le sous-espace $\mathrm{im(T)} = \mathrm{T(dom(T))}$ de F. On appelle *extension* d'un opérateur T tout opérateur S tel que $\mathrm{Gr(T)} \subset \mathrm{Gr(S)}$. On écrit alors $\mathrm{T} \subset \mathrm{S}$. Soient T un opérateur et D un sous-espace vectoriel de dom(T) ; on note $\mathrm{T}_{|\mathrm{D}}$ l'opérateur tel que $\mathrm{T}_{|\mathrm{D}} \subset \mathrm{T}$ et $\mathrm{dom}(\mathrm{T}_{|\mathrm{D}}) = \mathrm{D}$.

Soient S et T deux opérateurs de E dans F ; on définit l'opérateur $\mathrm{S} + \mathrm{T}$ en posant $\mathrm{dom(S + T)} = \mathrm{dom(S)} \cap \mathrm{dom(T)}$ et en posant $(\mathrm{S} + \mathrm{T})(x) = \mathrm{S}(x) + \mathrm{T}(x)$ pour tout vecteur $x \in \mathrm{dom(S + T)}$. Si R, S et T sont des opérateurs de E dans F, on a clairement $\mathrm{R} + \mathrm{S} = \mathrm{S} + \mathrm{R}$ et $(\mathrm{R} + \mathrm{S}) + \mathrm{T} = \mathrm{R} + (\mathrm{S} + \mathrm{T})$.

Soient E, F et G des espaces vectoriels, T un opérateur de E dans F et S un opérateur de F dans G ; on définit la composition ST de ces deux opérateurs en posant d'abord $\mathrm{dom(ST)} = \{x \in \mathrm{dom(T)} : \mathrm{T}(x) \in \mathrm{dom(S)}\}$ et en posant $(\mathrm{ST})(x) = \mathrm{S}(\mathrm{T}(x))$ pour tout $x \in \mathrm{dom(ST)}$. Si R est un opérateur de G dans un quatrième espace vectoriel H, on a $(RS)T = R(ST)$. De plus, si T est un opérateur de E dans F et si R et S sont des opérateurs de F dans G, on a $(R + S)T = RT + ST$; cependant, si R et S sont des opérateurs de E dans F et T est un opérateur de F dans G, on a $TR + TS \subset T(R + S)$ sans avoir en général l'égalité.

Un opérateur T de E dans F est dit *injectif* si l'application $T : \mathrm{dom}(T) \to F$ est injective. Soit T un opérateur injectif de E dans F ; le sous-ensemble de $F \times E$ égal à $\{(y, x) \in F \times E : (x, y) \in \mathrm{Gr}(T)\}$ est le graphe d'un opérateur T^{-1} (de domaine $\mathrm{im}(T)$) appelé *inverse* de T. Clairement T^{-1} est injectif et $(T^{-1})^{-1} = T$. Si $T : E \to F$ et $S : F \to G$ sont injectifs, alors ST est injectif et $(ST)^{-1} = T^{-1}S^{-1}$.

Exemples 10.1.1.

A. On prend $E = F = L_2(\mathbb{R})$, $\mathrm{dom}(T)$ est l'espace des fonctions C^1 à support compact et on pose $T(f) = f'$ pour $f \in \mathrm{dom}(T)$.

B. Cet exemple se décline en trois variantes.

– B1 : on prend $E = F = L_2([0, 1])$, $\mathrm{dom}(T_1)$ est l'espace des fonctions C^1 sur $[0, 1]$ et $T_1(f) = f'$ pour $f \in \mathrm{dom}(T_1)$.

– B2 : on prend $E = F = L_2([0, 1])$, $\mathrm{dom}(T_2)$ est l'espace des fonctions f qui sont C^1 sur $[0, 1]$ et telles que $f(0) = f(1) = 0$, et $T_2(f) = f'$ pour $f \in \mathrm{dom}(T_2)$.

– B3 : on prend $E = F = L_2([0, 1])$, $\mathrm{dom}(T_3)$ est l'espace des fonctions f qui sont C^1 sur $[0, 1]$ et telles que $f(0) = f(1)$, et $T_3(f) = f'$ pour $f \in \mathrm{dom}(T_3)$.

Ça a l'air de pinaillages ridicules, mais on verra plus loin à propos des adjoints qu'il y a des différences importantes dans les propriétés de T_1, T_2 et T_3.

La topologie revient

Soient E et F deux espaces de Banach ; un opérateur de E dans F est dit *densément défini* si son domaine est dense dans E.

Définition 10.1.2. Soient E et F deux espaces de Banach ; un opérateur de E dans F est dit *fermé* si son graphe est un sous-espace fermé de $E \times F$. Un opérateur de E dans F est dit *fermable* s'il admet une extension fermée.

Soit S une extension fermée de l'opérateur T ; alors $\mathrm{Gr}(S)$ contient $\mathrm{Gr}(T)$, donc son adhérence $\overline{\mathrm{Gr}(T)}$. Il s'ensuit qu'un opérateur T est fermable si et seulement si $\overline{\mathrm{Gr}(T)}$ est le graphe d'un opérateur. On appellera *fermeture* de l'opérateur T l'opérateur $\overline{T}$ tel que $\mathrm{Gr}(\overline{T}) = \overline{\mathrm{Gr}(T)}$. En particulier, pour que l'opérateur T soit fermable il faut et il suffit que l'on ait $\overline{\mathrm{Gr}(T)} \cap (\{0\} \times F) = \{(0, 0)\}$. On en déduit immédiatement :

Proposition 10.1.1. *Soient* E *et* F *deux espaces de Banach et* T *un opérateur de* E *dans* F *; pour que* T *soit fermable il faut et il suffit que pour toute suite* (x_n) *de* $\mathrm{dom}(T)$ *qui converge vers* 0 *dans* E *et telle que* $T(x_n)$ *converge dans* F *vers un vecteur* y*, on ait* $y = 0$.

Exemple 10.1.3. Fermetures des opérateurs T, T_1, T_2 et T_3 de l'exemple 1. Commençons par l'opérateur T de l'exemple **A**, défini sur $L_2(\mathbb{R})$. On va montrer que T est fermable et isoler un candidat pour la fermeture.

Supposons que (f, g) soit dans l'adhérence de $\mathrm{Gr}(T)$; il existe une suite $(f_n) \subset C^1_{\mathrm{comp}}$ telle que $f_n \to f$ dans L_2 et $f'_n \to g$ dans L_2. Quitte à passer à une sous-suite on peut supposer qu'il existe $\mathrm{E} \subset \mathbb{R}$ tel que $\mathbb{R} \setminus \mathrm{E}$ soit négligeable et tel que $f_n(t)$ converge vers $f(t)$ pour tout $t \in \mathrm{E}$. En particulier, E est non vide, et il est même dense dans $\mathbb{R}$. Fixons $a \in \mathrm{E}$, et soit $x \in \mathrm{E}$; pour tout n on a

$$f_n(x) = f_n(a) + \int_a^x f_n'(t)\, dt,$$

et la convergence dans L_2 implique la convergence des intégrales sur les segments bornés, donc compte tenu de tout

$$f(x) = f(a) + \int_a^x g(t)\, dt.$$

Mais la fonction $\mathrm{G}(x) = \int_a^x g(t)\, dt$ est continue pour tout x, et on peut redéfinir f sur l'ensemble négligeable $\mathbb{R} \setminus \mathrm{E}$ par la formule $f(x) = f(a) + \int_a^x g(t)\, dt$. Il en résulte que f est continue, et qu'il existe une fonction $g \in \mathrm{L}_2$ telle que

$$\forall x < y, \quad f(y) = f(x) + \int_x^y g(t)\, dt.$$

On introduit l'ensemble

$$\mathrm{G}_{\mathbf{A}} = \{(f, g) \in \mathrm{L}_2(\mathbb{R}) \times \mathrm{L}_2(\mathbb{R}) : \forall x < y, \quad f(y) = f(x) + \int_x^y g(t)\, dt\}.$$

On vient de montrer que l'adhérence de $\mathrm{Gr(T)}$ est contenue dans $\mathrm{G}_{\mathbf{A}}$; pour savoir que T est fermable, il suffit de voir que $\mathrm{G}_{\mathbf{A}}$ est un graphe : c'est clairement un espace vectoriel, et si $(0, g) \in \mathrm{G}_{\mathbf{A}}$, on aura $\int_x^y g = 0$ pour tous $x < y$, ce qui signifie que g est orthogonale à toutes les fonctions en escalier, qui sont denses dans $\mathrm{L}_2(\mathbb{R})$, donc $g = 0$, ce qu'il fallait démontrer.

Exercice 10.1.4. Montrer que l'adhérence du graphe de T est *égale* à $\mathrm{G}_{\mathbf{A}}$.

On appelle $\mathrm{H}_1(\mathbb{R})$ (espace de Sobolev) l'espace des fonctions $f \in \mathrm{L}_2(\mathbb{R})$ telles qu'il existe $g \in \mathrm{L}_2(\mathbb{R})$ telle que $(f, g) \in \mathrm{G}_{\mathbf{A}}$. On dit que g est la *dérivée généralisée* de f, et on note simplement $g = f'$. La fermeture de T de l'exemple 1, **A** est donc l'opérateur $\overline{\mathrm{T}}$ de $\mathrm{L}_2(\mathbb{R})$ dans lui-même dont le domaine est $\mathrm{H}_1(\mathbb{R})$ et qui est défini par $\overline{\mathrm{T}}(f) = f'$ pour $f \in \mathrm{H}_1(\mathbb{R})$.

On définit de même l'espace $\mathrm{H}_1([0, 1])$ des fonctions $f \in \mathrm{L}_2([0, 1])$ (en fait f sera continue) telles qu'il existe une fonction $g \in \mathrm{L}_2([0, 1])$ telle que $f(x) = f(0) + \int_0^x g(t)\, dt$, pour tout $x \in [0, 1]$. Si on se rappelle l'opérateur-exemple V de $\mathrm{L}_2([0, 1])$ dans lui-même qui associe à chaque $g \in \mathrm{L}_2([0, 1])$ sa “primitive” nulle en zéro, on voit que $\mathrm{H}_1([0, 1])$ est égal à $\mathrm{im(V)} + \mathbb{K}1$. On peut vérifier que les fermetures des variantes B1, B2, B3 sont définies sur les domaines

– 1 : $f \in \mathrm{H}_1([0, 1])$
– 2 : $f \in \mathrm{H}_1([0, 1])$ et $f(0) = f(1) = 0$
– 3 : $f \in \mathrm{H}_1([0, 1])$ et $f(0) = f(1)$.

Dans les trois cas $j = 1, 2, 3$ la valeur de l'extension $\overline{\mathrm{T}}_j(f)$ est égale à f', la dérivée généralisée de f, quand f est dans le domaine de $\overline{\mathrm{T}}_j$.

Proposition 10.1.2. *L'inverse d'un opérateur injectif fermé est fermé.*

Démonstration. Soient E et F deux espaces de Banach et T un opérateur fermé de E dans F ; on a $\mathrm{Gr}(\mathrm{T}^{-1}) = \rho(\mathrm{Gr}(\mathrm{T}))$ où $\rho : \mathrm{E} \times \mathrm{F} \to \mathrm{F} \times \mathrm{E}$ est l'homéomorphisme $(x, y) \to (y, x)$, donc $\mathrm{Gr}(\mathrm{T}^{-1})$ est fermé.

//

Si S est une application linéaire continue de E dans F, elle définit un opérateur de la façon la plus évidente : on pose $\mathrm{dom(S)} = \mathrm{E}$ et $\mathrm{S}(x)$ aura le sens habituel pour tout $x \in \mathrm{E}$; si T est un opérateur de E dans F, le domaine de $\mathrm{S}+\mathrm{T}$ sera égal à celui de T.

Exemple 10.1.5. Revenons à l'opérateur borné V de $\mathrm{L}_2([0,1])$ dans lui-même ; on se rappelle que V est injectif ; on peut donc définir un V^{-1} au sens des non bornés, dont le domaine est im(V) ; puisque V est continu, son graphe est fermé, donc V^{-1} est fermé. On peut vérifier que V^{-1} est densément défini. Les opérateurs fermés et densément définis forment la classe la plus intéressante dans cette théorie.

Proposition 10.1.3. *Soient* E, F *et* G *des espaces de Banach ;*

(i) soient $\mathrm{S} : \mathrm{E} \to \mathrm{F}$ *une application linéaire continue et* T *un opérateur de* E *dans* F *; pour que* $\mathrm{S}+\mathrm{T}$ *soit fermé il faut et il suffit que* T *le soit ; pour que* $\mathrm{S}+\mathrm{T}$ *soit fermable il faut et il suffit que* T *le soit et, dans ce cas* $\overline{\mathrm{S}+\mathrm{T}} = \mathrm{S}+\overline{\mathrm{T}}$ *;*

(ii) soient $\mathrm{S} : \mathrm{E} \to \mathrm{F}$ *une application linéaire continue et* T *un opérateur fermé (resp. fermable) de* F *dans* G *; l'opérateur* TS *est fermé (resp. fermable) ;*

(iii) soient S *un opérateur fermé (resp. fermable) de* E *dans* F *et* $\mathrm{T} : \mathrm{G} \to \mathrm{F}$ *une application linéaire continue injective ; l'opérateur* $\mathrm{T}^{-1}\mathrm{S}$ *est fermé (resp. fermable).*

Démonstration. Le graphe de $\mathrm{S}+\mathrm{T}$ est $\{(x, y) \in \mathrm{E} \times \mathrm{F} : (x, y - \mathrm{S}(x)) \in \mathrm{Gr(T)}\}$; si T est fermé, cet ensemble est fermé. Si T est fermable, $\mathrm{S}+\mathrm{T}$ admet l'extension $\mathrm{S}+\overline{\mathrm{T}}$ qui est fermée par ce qui précède ; il en résulte que $\mathrm{S}+\mathrm{T}$ est fermable et $\overline{\mathrm{S}+\mathrm{T}} \subset \mathrm{S}+\overline{\mathrm{T}}$.

Remplaçant T par $\mathrm{T}+\mathrm{S}$ et S par $-\mathrm{S}$ on trouve que si $\mathrm{T}+\mathrm{S}$ est fermé ou fermable il en va de même pour T et que, dans ce cas $\overline{\mathrm{T}} \subset \overline{\mathrm{S}+\mathrm{T}} - \mathrm{S}$; donc $\mathrm{S}+\overline{\mathrm{T}} \subset \mathrm{S}+\overline{\mathrm{S}+\mathrm{T}}-\mathrm{S} = \overline{\mathrm{S}+\mathrm{T}}$.

Si T est fermé, le graphe de TS est $\{(x, y) \in \mathrm{E} \times \mathrm{G} : (\mathrm{S}(x), y) \in \mathrm{Gr(T)}\}$, donc est fermé. Si T est fermable, alors TS admet l'extension fermée $\overline{\mathrm{T}}\mathrm{S}$.

Si S est fermé, le graphe de $\mathrm{T}^{-1}\mathrm{S}$ est $\{(x, y) \in \mathrm{E} \times \mathrm{G} : (x, \mathrm{T}(y)) \in \mathrm{Gr(S)}\}$, donc est fermé. Si S est fermable, alors $\mathrm{T}^{-1}\mathrm{S}$ admet l'extension fermée $\mathrm{T}^{-1}\overline{\mathrm{S}}$.

//

Soient E et F deux espaces de Banach et soit T un opérateur de E dans F ; on appelle *topologie du graphe* associée à l'opérateur T, la topologie sur dom(T) induite par l'injection $x \to (x, \mathrm{T}(x))$ de dom(T) dans $\mathrm{E} \times \mathrm{F}$. Cette topologie est plus fine que celle induite par E et fait de dom(T) un espace complet si T est fermé.

Soient E et F deux espaces de Banach et T un opérateur fermable de E dans F ; un sous-espace D du domaine de T est appelé un *domaine essentiel* pour T si les opérateurs T et $\mathrm{T}_{|\mathrm{D}}$ ont même fermeture. Cela revient à dire que D est dense dans dom(T) pour la topologie du graphe.

10.2. Spectre des opérateurs fermés

Définition 10.2.1. Soient T un opérateur d'un espace de Banach complexe E dans lui même et $\lambda \in \mathbb{C}$; on dit que λ est une *valeur régulière* de T si $\mathrm{T} - \lambda\,\mathrm{Id}_\mathrm{E}$ est une application linéaire bijective de $\mathrm{dom}(\mathrm{T})$ sur E et si l'application linéaire réciproque définit une application linéaire continue $\mathrm{R}_\lambda(\mathrm{T})$ de E dans lui même. On appelle *spectre* de T le complémentaire $\mathrm{Sp}(\mathrm{T})$ dans $\mathbb{C}$ de l'ensemble des valeurs régulières de T.

Si λ est une valeur régulière de T, alors $(\mathrm{T} - \lambda\,\mathrm{Id}_\mathrm{E})^{-1}$ est continue, donc fermée ; donc $\mathrm{T} - \lambda\,\mathrm{Id}_\mathrm{E}$ est fermé, donc T est fermé. On en déduit que, si T n'est pas fermé, $\mathrm{Sp}(\mathrm{T}) = \mathbb{C}$.

Soit T un opérateur fermé d'un espace de Banach E dans lui-même ; remarquons que pour tout $\lambda \in \mathbb{C}$, l'opérateur $\mathrm{T} - \lambda\,\mathrm{Id}_\mathrm{E}$ est fermé (par la proposition 1.3). Si $\mathrm{T} - \lambda\,\mathrm{Id}_\mathrm{E}$ est bijectif de $\mathrm{dom}(\mathrm{T})$ sur E, alors λ est une valeur régulière car $(\mathrm{T} - \lambda\,\mathrm{Id}_\mathrm{E})^{-1}$ est fermé (proposition 1.2), donc continu par le théorème du graphe fermé (théorème 3.1.6).

Exemples 10.2.2.

1. Soit μ une mesure sur $\mathbb{C}$, positive et non nulle, donnant une mesure finie à tout compact. On considère dans $\mathrm{L}_2(\mu) = \mathrm{L}_2(\mathbb{C}, \mu)$ l'application de multiplication par z, définie sur le domaine

$$\mathrm{D} = \{f \in \mathrm{L}_2(\mu) : \int_\mathbb{C} |z|^2\, |f(z)|^2\, d\mu(z) < +\infty\},$$

ce qui donne un opérateur, en général non borné, qu'on notera M, qui agit sur $f \in \mathrm{D}$ par $(\mathrm{M}f)(z) = zf(z)$, et $\mathrm{M}f \in \mathrm{L}_2(\mu)$. On peut décrire l'appartenance de f au domaine D en une seule formule,

$$\int_\mathbb{C} (1 + |z|^2)\, |f(z)|^2\, d\mu(z) < +\infty.$$

On suppose d'abord que $\lambda \in \mathbb{C}$ est tel que pour tout $\varepsilon > 0$, on ait que $\mathrm{B} = \mathrm{B}(\lambda, \varepsilon)$ vérifie $\mu(\mathrm{B}) > 0$. On peut considérer la fonction $f = \mathbf{1}_\mathrm{B}$, qui est dans le domaine D, et qui n'est pas dans la classe nulle de $\mathrm{L}_2(\mu)$ puisque $\mu(\mathrm{B}) > 0$. On a $|(\mathrm{M} - \lambda)f| = |z - \lambda|\, \mathbf{1}_\mathrm{B} \le \varepsilon\, \mathbf{1}_\mathrm{B}$. Ceci montre que $\|(\mathrm{M} - \lambda)f\|_2 \le \varepsilon\, \|f\|_2$; si l'inverse $\mathrm{R}_\lambda(\mathrm{M})$ de $\mathrm{M} - \lambda\,\mathrm{Id}$ existait, il devrait vérifier $\|\mathrm{R}_\lambda(\mathrm{M})\| \ge 1/\varepsilon$, pour tout $\varepsilon > 0$, ce qui est impossible. Il en résulte que $\lambda \in \mathrm{Sp}(\mathrm{T})$.

Remarquons en passant : si μ est la mesure de Lebesgue de $\mathbb{C}$, toutes les boules ouvertes sont de mesure > 0, donc le spectre de M est $\mathbb{C}$ tout entier dans ce cas (on a dit qu'on obtient aussi que $\mathrm{Sp}(\mathrm{T}) = \mathbb{C}$ dans le cas dégénéré où T n'est pas fermé).

On suppose inversement que $\lambda \in \mathbb{C}$ est tel qu'il existe $\varepsilon_0 > 0$ tel que $\mu(\mathrm{B}(\lambda, \varepsilon_0)) = 0$. Considérons la fonction mesurable bornée g définie sur $\mathbb{C}$ par $g(z) = (z - \lambda)^{-1}$ si $|z - \lambda| \ge \varepsilon_0$ et $g(z) = 0$ sinon. La multiplication M_g est bornée sur $\mathrm{L}_2(\mu)$ puisque g est bornée, et on va voir que $\mathrm{M}_g = \mathrm{R}_\lambda(\mathrm{M})$. Si $f \in \mathrm{dom}(\mathrm{M})$, on voit que $\mathrm{M}_g(\mathrm{M}(f) - \lambda f) = g(z - \lambda)f$ est égale à f en dehors de B, et à 0 dans B ; mais puisque $\mu(\mathrm{B}) = 0$, on a bien $\mathrm{M}_g(\mathrm{M}(f) - \lambda f) = f$ en tant que classe. Inversement, si $h \in \mathrm{L}_2(\mu)$, on vérifie que $\mathrm{M}_g(h) \in \mathrm{dom}(\mathrm{M})$ (en effet,

$$\int_\mathbb{C} |z|^2\, |(\mathrm{M}_g h)(z)|^2\, d\mu(z) = \int_\mathbb{C} |z|^2\, |g(z)h(z)|^2\, d\mu(z) = \int_\mathbb{C} |zg(z)|^2\, |h(z)|^2\, d\mu(z) < +\infty$$

parce que $zg(z)$ est bornée sur $\mathbb{C}$) et ensuite $(\mathrm{M} - \lambda\,\mathrm{Id})(\mathrm{M}_g(h)) = h$. On a bien montré que $\mathrm{M}_g = \mathrm{R}_\lambda(\mathrm{M})$.

En bref, le spectre de M est exactement l'ensemble des $\lambda \in \mathbb{C}$ décrit précédemment, c'est à dire les λ dont tout voisinage a une μ-mesure > 0.

Si $\mu = \sum_{n=0}^{+\infty} 2^{-n}\delta_{z_n}$, où (z_n) est une suite quelconque de points de $\mathbb{C}$, on déduit de ce qui précède que le spectre de M est l'adhérence F de l'ensemble des points de la suite. Cela nous permet de dire que tout fermé non vide de $\mathbb{C}$ est le spectre d'un opérateur. C'est vrai aussi pour l'ensemble vide, comme on le verra avec l'exemple qui suit.

2. On désigne par V l'opérateur borné de $\mathrm{E} = \mathrm{L}_2([0,1])$ dans lui-même défini par $(\mathrm{V}f)(t) = \int_0^t f(s)\,ds$. L'image de V est l'ensemble D des fonctions F de $\mathrm{H}^1([0,1])$ qui sont nulles en 0, et on a vu que V est injectif. On peut donc considérer l'opérateur $\mathrm{T} = \mathrm{V}^{-1}$, dont le domaine est D ; le spectre de T est vide :

évidemment, 0 est valeur régulière de T et $\mathrm{R}_0(\mathrm{T}) = \mathrm{V}$. Pour $\lambda \neq 0$, cherchons à résoudre l'équation $\mathrm{T}x - \lambda x = y$, pour $y \in \mathrm{E}$ donné (on cherche $x \in \mathrm{D}$). Puisque T est surjectif, on peut écrire $y = \mathrm{T}z$, avec $z = \mathrm{V}(y) \in \mathrm{D}$. En appliquant V on trouve $x - \lambda \mathrm{V}x = z$, soit $\mathrm{V}x - \lambda^{-1}x = -\lambda^{-1}z$. On sait que λ^{-1} n'est pas dans le spectre de V (qui est réduit à $\{0\}$) donc on peut résoudre,

$$x = \mathrm{R}_{\lambda^{-1}}(\mathrm{V})(-\lambda^{-1}z) = -\lambda^{-1}\mathrm{R}_{\lambda^{-1}}(\mathrm{V})\,(\mathrm{V}y).$$

On vient donc d'identifier $\mathrm{R}_\lambda(\mathrm{T}) = -\lambda^{-1}\mathrm{R}_{\lambda^{-1}}(\mathrm{V})\,\mathrm{V}$. Finalement, on constate que tout nombre complexe est valeur régulière de T, donc le spectre de $\mathrm{T} = \mathrm{V}^{-1}$ est vide.

Le spectre d'un opérateur T est réunion des trois ensembles disjoints suivants :

– le *spectre ponctuel* $\mathrm{Sp}_p(\mathrm{T})$ de T est l'ensemble de ses valeurs propres ;

– le *spectre résiduel* $\mathrm{Sp}_r(\mathrm{T})$ de T formé des $\lambda \in \mathbb{C}$ qui ne sont pas valeur propre de l'opérateur T et tels que l'image de $\mathrm{T} - \lambda\,\mathrm{Id}_\mathrm{E}$ ne soit pas dense dans E ;

– le complémentaire $\mathrm{Sp}_c(\mathrm{T})$ dans $\mathrm{Sp}(\mathrm{T})$ de la réunion de ces deux ensembles appelé *spectre continu* de T.

Remarque 10.2.3. Pour $\lambda \in \mathrm{Sp}_c(\mathrm{T})$, l'opérateur $\mathrm{T} - \lambda\,\mathrm{Id}_\mathrm{E}$ est injectif d'image dense, mais $(\mathrm{T} - \lambda\,\mathrm{Id}_\mathrm{E})^{-1}$ n'est pas continu.

Lemme 10.2.1. *Soient* T *un opérateur injectif fermé d'un espace de Banach* E *dans lui même et* λ *une valeur régulière de* T *non nulle ; alors* λ^{-1} *est une valeur régulière de* T^{-1} *et on a*

$$\mathrm{R}_{\lambda^{-1}}(\mathrm{T}^{-1}) = -\lambda\mathrm{T}\,\mathrm{R}_\lambda(\mathrm{T}) = -\lambda\,\mathrm{Id}_\mathrm{E} - \lambda^2\mathrm{R}_\lambda(\mathrm{T}).$$

Démonstration. On veut résoudre pour tout $y \in \mathrm{E}$ l'équation $(\mathrm{T}^{-1} - \lambda^{-1}\,\mathrm{Id}_\mathrm{E})(x) = y$ (on cherche $x \in \mathrm{dom}(\mathrm{T}^{-1}) = \mathrm{im}(\mathrm{T})$). Puisque λ est régulière pour T, on peut écrire $y = (\mathrm{T} - \lambda\,\mathrm{Id})(z)$, avec $z = \mathrm{R}_\lambda(\mathrm{T})(y)$. On sait alors que $z \in \mathrm{dom}(\mathrm{T}) = \mathrm{im}(\mathrm{T}^{-1})$, donc il existe $u \in \mathrm{im}(\mathrm{T})$ tel que $z = \mathrm{T}^{-1}(u)$. L'équation proposée est donc

$$\mathrm{T}^{-1}(x) - \lambda^{-1}x = (\mathrm{T} - \lambda\,\mathrm{Id})(\mathrm{T}^{-1}(u)) = u - \lambda\mathrm{T}^{-1}(u) = \mathrm{T}^{-1}(-\lambda u) - \lambda^{-1}(-\lambda u).$$

Il en résulte que $x_0 = -\lambda u$ convient. Par ailleurs, $\mathrm{T}^{-1} - \lambda^{-1}\,\mathrm{Id}$ est injectif (donc la solution x_0 est unique) : si $x \in \mathrm{dom}(\mathrm{T}^{-1})$ et $\mathrm{T}^{-1}(x) - \lambda^{-1}x = 0$, alors $x = \lambda\mathrm{T}^{-1}(x) \in \mathrm{dom}(\mathrm{T})$ et $\mathrm{T}(x) = \lambda x$ implique $x = 0$ puisque λ est régulière pour T. Si $\mathrm{R}_{\lambda^{-1}}(\mathrm{T}^{-1})$ existe, on a donc

$$x = \mathrm{R}_{\lambda^{-1}}(\mathrm{T}^{-1})(y) = -\lambda u = -\lambda\mathrm{T}(z) = -\lambda\mathrm{T}\,\mathrm{R}_\lambda(\mathrm{T})(y).$$

Il reste à expliquer pourquoi l'opérateur $\mathrm{T}\,\mathrm{R}_\lambda(\mathrm{T})$ est borné. Cela provient de l'égalité $(\mathrm{T}-\lambda\,\mathrm{Id})\mathrm{R}_\lambda(\mathrm{T}) = \mathrm{Id}$, qui donne $\mathrm{T}\,\mathrm{R}_\lambda(\mathrm{T}) = \lambda\mathrm{R}_\lambda(\mathrm{T}) + \mathrm{Id}$, qui est bien continu.

//

Proposition 10.2.2.

(*i*) *Le spectre d'un opérateur fermé* T *d'un espace de Banach complexe* E *dans lui même est une partie fermée de* $\mathbb{C}$.

(*ii*) *L'application* $\lambda \to \mathrm{R}_\lambda(\mathrm{T})$ *est continue et holomorphe du complémentaire du spectre dans* $\mathcal{L}(\mathrm{E})$.

Démonstration. Montrons le premier point ; si $\mathrm{Sp}(\mathrm{T}) = \mathbb{C}$, le spectre est bien fermé ; nous supposons donc maintenant $\mathrm{Sp}(\mathrm{T}) \neq \mathbb{C}$; quitte à remplacer T par $\mathrm{T}-\lambda\,\mathrm{Id}_\mathrm{E}$, on peut supposer que 0 est valeur régulière de T. Posons $\mathrm{S} = \mathrm{R}_0(\mathrm{T}) = \mathrm{T}^{-1}$. Il résulte alors du lemme 1 que $\mathrm{Sp}(\mathrm{T}) = \{\lambda \in \mathbb{C} \setminus \{0\} : \lambda^{-1} \in \mathrm{Sp}(\mathrm{S})\}$ et comme $\mathrm{Sp}(\mathrm{S})$ est une partie compacte de $\mathbb{C}$, $\mathrm{Sp}(\mathrm{T})$ est une partie fermée de $\mathbb{C}$.

Passons au point (*ii*). Par le lemme 1, on a $\mathrm{R}_\lambda(\mathrm{T}) = -\lambda^{-1}\mathrm{S}\,\mathrm{R}_{\lambda^{-1}}(\mathrm{S})$ pour toute valeur régulière non nulle de T. Donc l'application $\lambda \to \mathrm{R}_\lambda(\mathrm{T})$ est holomorphe sur l'ouvert $\mathbb{C} \setminus (\mathrm{Sp}(\mathrm{T}) \cup \{0\})$. Comme $\mathrm{Sp}(\mathrm{T})$ est fermé et que $0 \notin \mathrm{Sp}(\mathrm{T})$, il existe une valeur régulière non nulle λ_0 de T. Appliquant ce qui précède à $\mathrm{T} - \lambda_0\,\mathrm{Id}_\mathrm{E}$ on en déduit que $\lambda \to \mathrm{R}_\lambda(\mathrm{T})$ est holomorphe dans $\mathbb{C} \setminus (\mathrm{Sp}(\mathrm{T}) \cup \{\lambda_0\})$, ce qui couvre le cas qui manquait (au voisinage de 0).

//

On a vu que pour toute partie fermée non vide F de $\mathbb{C}$, on peut construire un opérateur T d'un espace de Hilbert H dont le spectre $\mathrm{Sp}(\mathrm{T})$ soit égal à F. Considérons l'opérateur V de l'exemple 6.2.2. Posons $\mathrm{T} = \mathrm{V}^{-1}$. Comme $\mathrm{V} \in \mathcal{L}(\mathrm{H})$, on sait que $0 \notin \mathrm{Sp}(\mathrm{T})$; en appliquant le lemme 1, on voit que $\mathrm{Sp}(\mathrm{T}) = \{\lambda \in \mathbb{C} \setminus \{0\} : \lambda^{-1} \in \mathrm{Sp}(\mathrm{V})\} = \emptyset$, vu que $\mathrm{Sp}(\mathrm{V}) = \{0\}$.

10.3. Transposés et adjoints

Soient E et F deux espaces de Banach et T un opérateur densément défini de E dans F ; on définit *le transposé* de T, qui est un opérateur de F^* dans E^*, de la façon suivante : le domaine de ${}^t\mathrm{T}$ est l'ensemble des $y^* \in \mathrm{F}^*$ telles que la forme linéaire $x \in \mathrm{dom}(\mathrm{T}) \to y^*(\mathrm{T}(x))$ soit continue.

Dans le cas où $y^* \in \mathrm{dom}({}^t\mathrm{T})$, cette forme linéaire continue, définie sur le sous-espace dense $\mathrm{dom}(\mathrm{T}) \subset \mathrm{E}$, se prolonge de façon unique en une forme linéaire $x^* \in \mathrm{E}^*$ continue sur E. On pose alors ${}^t\mathrm{T}(y^*) = x^*$. On a donc

$$({}^t\mathrm{T})(y^*)(x) = y^*(\mathrm{T}(x))$$

pour tous $x \in \mathrm{dom}(\mathrm{T})$ et $y^* \in \mathrm{dom}({}^t\mathrm{T})$. Lorsque E et F sont deux espaces de Hilbert et T un opérateur densément défini de E dans F, on définit un opérateur T^* de F dans E de la façon suivante : on définit $\mathrm{T}^*(y) = x$ si la forme linéaire ℓ_y associée à $y \in \mathrm{H}$ est dans $\mathrm{dom}({}^t\mathrm{T})$, et si $\ell_x = x^* = {}^t\mathrm{T}(\ell_y)$. Le vecteur y est donc dans le domaine de T^* si et seulement si la forme linéaire $\ell : u \in \mathrm{E} \to \langle \mathrm{T}(u), y\rangle$ est continue sur $\mathrm{dom}(\mathrm{T})$ (muni de la norme de E), et le couple $(y, x) \in \mathrm{F} \times \mathrm{E}$ est dans le graphe de T^* si et seulement si

$$\langle \mathrm{T}(u), y\rangle = \langle u, x\rangle \qquad (*)$$

pour tout $u \in \mathrm{dom}(\mathrm{T})$, ce qui signifie que x représente la forme linéaire ℓ (et son prolongement continu à E). On a donc

$$\mathrm{Gr}(\mathrm{T}^*) = \{(y, x) \in \mathrm{F} \times \mathrm{E} : \forall z \in \mathrm{dom}(\mathrm{T}),\ \langle x, z\rangle = \langle y, \mathrm{T}(z)\rangle\}.$$

En effet, la forme linéaire $u \to \langle \mathrm{T}(u), y\rangle$ est alors continue puisqu'elle est égale à $u \to \langle x, u\rangle$ et dans ce cas on a $x = \mathrm{T}^*(y)$ par définition de l'adjoint. Il est clair que la condition $(*)$ définit un ensemble fermé de couples (y, x), ce qui montre que T^* est toujours un opérateur fermé.

On dit que T (densément défini sur un Hilbert) est *symétrique* si

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

pour tous $x, y \in \mathrm{dom}(\mathrm{T})$. Cela revient à dire que $\mathrm{T} \subset \mathrm{T}^*$. Un opérateur T de E dans lui même est dit *autoadjoint* si $\mathrm{T} = \mathrm{T}^*$. Tout autoadjoint est symétrique mais l'inverse n'est pas vrai.

Exercice 10.3.1. Adjoints des fermetures des exemples B1, B2, B3 de 1.1. Posons $\mathrm{E} = \mathrm{L}_2([0, 1])$. La fermeture de l'exemple B3 est l'opérateur T_c dont le domaine est

$$\mathrm{D}_c = \{f \in \mathrm{H}^1([0, 1]) : f(0) = f(1)\}$$

et qui est défini par $\mathrm{T}_c(f) = f'$ pour toute $f \in \mathrm{D}_c$; on va montrer que $\mathrm{S}_c = i\mathrm{T}_c$ est autoadjoint.

On montre d'abord que S_c est symétrique, c'est à dire que $\langle f_1, \mathrm{S}_c(f_2)\rangle = \langle \mathrm{S}_c(f_1), f_2\rangle$ pour toutes $f_1, f_2 \in \mathrm{D}_c$. On a en effet

$$\langle f_1, \mathrm{S}_c(f_2)\rangle = \int_0^1 f_1\, \overline{(if_2')} = \big[f_1\, \overline{(if_2)}\big]_0^1 - \int_0^1 f_1'\, \overline{(if_2)} = \int_0^1 (if_1')\, \overline{f}_2 = \langle \mathrm{S}_c(f_1), f_2\rangle$$

(le terme $\big[\,.\,\big]_0^1$ est nul parce que toutes les fonctions ont la même valeur en 0 et en 1 par définition de D_c). On montrerait de la même façon que l'exemple S_b correspondant à B2, défini sur $\mathrm{D}_b = \{f \in \mathrm{H}^1([0, 1]) : f(0) = f(1) = 0\}$ est symétrique : c'est évident puisque $\mathrm{S}_b \subset \mathrm{S}_c$.

On sait donc déjà que $\mathrm{D}_c \subset \mathrm{dom}(\mathrm{S}_c^*)$, et que $\mathrm{S}_c^*(f) = \mathrm{T}(f) = if'$ pour $f \in \mathrm{D}_c$. Il reste à voir que $\mathrm{dom}(\mathrm{S}_c^*) \subset \mathrm{D}_c$. Dire que (g, h) est dans le graphe de S_c^* signifie que $g \in \mathrm{E}$ est dans le domaine de S_c^* et que $h = \mathrm{S}_c^*(g) \in \mathrm{E}$ vérifie

$$\langle f, h\rangle = \langle \mathrm{T}_c(f), g\rangle$$

pour toute fonction $f \in \mathrm{D}_c$. On a si $(g, h) \in \mathrm{Gr}(\mathrm{S}_c^*)$

$$\int_0^1 f\, \overline{h} = \int_0^1 (if')\, \overline{g}$$

pour toute $f \in \mathrm{D}_c$. Posons $\mathrm{H}(t) = \int_0^t h(s)\, ds$. On obtient par intégration par parties

$$\int_0^1 f\, \overline{h} = \big[f\, \overline{\mathrm{H}}\big]_0^1 - \int_0^1 f'\, \overline{\mathrm{H}} = f(1)\, \overline{\mathrm{H}}(1) - \int_0^1 f'\, \overline{\mathrm{H}},$$

ce qui donne

$$f(1)\, \overline{\mathrm{H}}(1) - \int_0^1 f'\, \overline{\mathrm{H}} = \int_0^1 (if')\, \overline{g} = -\int_0^1 f'\, \overline{(ig)}$$

ou encore $f(1)\,\overline{\mathrm{H}}(1) = \int_0^1 f'\,\overline{(\mathrm{H} - ig)}$ pour toute $f \in \mathrm{D}_c$. Puisque la fonction $f_0 = 1$ est dans D_c, on obtient puisque $f_0' = 0$ que $\mathrm{H}(1) = 0$. On remarque que l'ensemble des f', lorsque $f \in \mathrm{D}_c$, est exactement l'ensemble de toutes les fonctions k de $\mathrm{E} = \mathrm{L}_2$ qui sont d'intégrale nulle sur $[0,1]$. Cet ensemble des fonctions d'intégrale nulle est égal à $(\mathbb{C}1)^\perp$, et l'équation précédente indique que $\mathrm{H} - ig$ est orthogonale à $(\mathbb{C}1)^\perp$, donc $\mathrm{H} - ig \in (\mathbb{C}1)^{\perp\perp} = \mathbb{C}1$. On obtient que $\mathrm{H} - ig$ est une fonction constante, donc $g = -i\mathrm{H} + \mathrm{C}te$; comme $\mathrm{H}(0) = \mathrm{H}(1)$ et que H est une fonction de $\mathrm{H}^1([0,1])$, il en résulte que $g \in \mathrm{D}_c$. On a déjà vu que $\mathrm{D}_c \subset \mathrm{dom}(\mathrm{S}_c^*)$, et on a maintenant $\mathrm{dom}(\mathrm{S}_c^*) \subset \mathrm{D}_c$, donc $\mathrm{dom}(\mathrm{S}_c^*) = \mathrm{D}_c$ et pour $g \in \mathrm{dom}(\mathrm{S}_c^*)$ on a $\mathrm{S}_c^*(g) = ig' = \mathrm{S}_c(g)$, ce qui montre que S_c est autoadjoint.

En suivant la même méthode, on vérifie que l'adjoint de S_a, défini sur $\mathrm{D}_a = \mathrm{H}^1([0,1])$ par $\mathrm{S}_a(f) = if'$, est l'opérateur S_b défini sur $\mathrm{D}_b = \{f \in \mathrm{D}_a : f(0) = f(1) = 0\}$ par $\mathrm{S}_b(f) = if'$. L'adjoint de S_b est S_a. Il en résulte que S_b est un exemple d'opérateur symétrique qui n'est pas autoadjoint.

Proposition 10.3.1. *Soient* E *et* F *deux espaces de Hilbert et* T *un opérateur densément défini de* E *dans* F *; alors* T^* *est fermé. Pour que* T *soit fermable, il faut et il suffit que* T^* *soit densément défini. Dans ce cas, on a* $\overline{\mathrm{T}} = (\mathrm{T}^*)^*$.

Démonstration. Soit $\mathrm{U}_0 \in \mathcal{L}(\mathrm{F} \oplus \mathrm{E}, \mathrm{E} \oplus \mathrm{F})$ l'opérateur unitaire qui à $(y,x) \in \mathrm{F} \oplus \mathrm{E}$ associe $(x,-y)$; le graphe $\mathrm{Gr}(\mathrm{T}^*)$ de T^* est l'orthogonal dans l'espace de Hilbert $\mathrm{F} \oplus \mathrm{E}$ de $\mathrm{U}_0^*(\mathrm{Gr}(\mathrm{T}))$. Donc $\mathrm{Gr}(\mathrm{T}^*)$ est fermé.

Si T^* est densément défini, alors $\mathrm{T} \subset (\mathrm{T}^*)^*$ donc T est fermable. Dans ce cas, $\mathrm{Gr}(\overline{\mathrm{T}}) = \overline{\mathrm{Gr}(\mathrm{T})} = \mathrm{Gr}(\mathrm{T})^{\perp\perp} = \mathrm{Gr}((\mathrm{T}^*)^*)$. Soient $y \in \mathrm{dom}(\mathrm{T}^*)^\perp$ et $(x,z) \in \mathrm{Gr}(\mathrm{T})^\perp$; alors $z \in \mathrm{dom}(\mathrm{T}^*)$ donc $\langle 0,x\rangle + \langle y,z\rangle = 0$, donc $(0,y) \in \mathrm{Gr}(\mathrm{T})^{\perp\perp} = \overline{\mathrm{Gr}(\mathrm{T})}$. Si T est fermable, $y = 0$, donc $\mathrm{dom}(\mathrm{T}^*)$ est dense.

//

Proposition 10.3.2. *Soit* T *un opérateur densément défini d'un espace de Hilbert* E *dans un espace de Hilbert* F *; alors* $\ker(\mathrm{T}^*) = \mathrm{im}(\mathrm{T})^\perp$.

Démonstration. Soit $y \in \mathrm{F}$; on a $y \in \ker(\mathrm{T}^*)$ si et seulement si, pour tout $x \in \mathrm{dom}(\mathrm{T})$, on a $\langle 0,x\rangle = \langle y,\mathrm{T}(x)\rangle$; cela a lieu si et seulement si $y \in \mathrm{im}(\mathrm{T})^\perp$.

//

Proposition 10.3.3. *Soient* E *et* F *deux espaces de Hilbert et* T *un opérateur densément défini, fermé de* E *dans* F *;*

(i) *pour tout* $\mathrm{S} \in \mathcal{L}(\mathrm{E},\mathrm{F})$ *on a* $(\mathrm{S}+\mathrm{T})^* = \mathrm{S}^* + \mathrm{T}^*$ *;*
(ii) *si* R *est une extension de* T*, alors* $\mathrm{R}^* \subset \mathrm{T}^*$ *;*
(iii) *si* T *est injectif et d'image dense, alors* $(\mathrm{T}^{-1})^* = (\mathrm{T}^*)^{-1}$.

Démonstration. Pour $x \in \mathrm{dom}(\mathrm{T})$, $y \in \mathrm{dom}(\mathrm{T}^*)$, on a

$$\langle(\mathrm{S}+\mathrm{T})(x),y\rangle = \langle\mathrm{S}(x),y\rangle + \langle\mathrm{T}(x),y\rangle = \langle x,(\mathrm{S}^*+\mathrm{T}^*)(y)\rangle,$$

donc $\mathrm{S}^* + \mathrm{T}^* \subset (\mathrm{S}+\mathrm{T})^*$. Pour $x \in \mathrm{dom}(\mathrm{T})$, $y \in \mathrm{dom}(\mathrm{S}+\mathrm{T})^*$, on a

$$\langle\mathrm{T}(x),y\rangle = \langle(\mathrm{S}+\mathrm{T})(x),y\rangle - \langle\mathrm{S}(x),y\rangle = \langle x,(\mathrm{S}+\mathrm{T})^*(y)\rangle - \langle x,\mathrm{S}^*(y)\rangle,$$

donc $y \in \mathrm{dom}(\mathrm{T}^*)$ et $\mathrm{T}^*(y) = (\mathrm{S}+\mathrm{T})^*(y) - \mathrm{S}^*(y)$, donc $(\mathrm{S}+\mathrm{T})^* = \mathrm{S}^* + \mathrm{T}^*$, ce qui termine le point (i). Si $\mathrm{Gr}(\mathrm{T}) \subset \mathrm{Gr}(\mathrm{R})$, alors $\mathrm{Gr}(\mathrm{R})^\perp \subset \mathrm{Gr}(\mathrm{T})^\perp$, d'où le deuxième point (ii). Enfin, pour $(x,y) \in \mathrm{E}\times\mathrm{F}$ on a $(x,y) \in \mathrm{Gr}((\mathrm{T}^{-1})^*)$ si et seulement si $(y,-x) \in \mathrm{Gr}(\mathrm{T}^{-1})^\perp$, ce qui équivaut à $(-x,y) \in \mathrm{Gr}(\mathrm{T})^\perp$ ou encore à $(y,x) \in \mathrm{Gr}(\mathrm{T}^*)$.

//

Proposition 10.3.4. *Soient* E *et* F *deux espaces de Hilbert et* T *un opérateur fermé densément défini de* E *dans* F *; l'opérateur* $(\mathrm{Id}_\mathrm{E} + \mathrm{T}^*\mathrm{T})$ *est injectif, son image est égale à* E *et* $(\mathrm{Id}_\mathrm{E} + \mathrm{T}^*\mathrm{T})^{-1}$ *est un élément positif de* $\mathcal{L}(\mathrm{E})$*. L'opérateur* $\mathrm{T}^*\mathrm{T}$ *est autoadjoint et son spectre est contenu dans* $[0,+\infty[$.

Démonstration. Soit $x \in \mathrm{E}$; comme $\mathrm{U}_0(\mathrm{G}(\mathrm{T}^*))$ est l'orthogonal de $\mathrm{Gr}(\mathrm{T})$, il existe deux vecteurs $\xi \in \mathrm{Gr}(\mathrm{T})$ et $\eta \in \mathrm{Gr}(\mathrm{T}^*)$ tels que $(x,0) = \xi + \mathrm{U}_0(\eta)$; en d'autres termes, il existe $z \in \mathrm{dom}(\mathrm{T})$ et $y \in \mathrm{dom}(\mathrm{T}^*)$ tels que $(x,0) = (z,\mathrm{T}(z)) + (\mathrm{T}^*(y),-y)$. Alors $y = \mathrm{T}(z)$ donc $z \in \mathrm{dom}(\mathrm{T}^*\mathrm{T}) = \mathrm{dom}(\mathrm{Id}_\mathrm{E} + \mathrm{T}^*\mathrm{T})$ et $x = (\mathrm{Id}_\mathrm{E} + \mathrm{T}^*\mathrm{T})(z)$. Donc $(\mathrm{Id}_\mathrm{E} + \mathrm{T}^*\mathrm{T})$ est surjectif. Soit $z \in \mathrm{dom}(\mathrm{Id}_\mathrm{E} + \mathrm{T}^*\mathrm{T})$; comme $z \in \mathrm{dom}(\mathrm{T})$ et $\mathrm{T}(z) \in \mathrm{dom}(\mathrm{T}^*)$, on a
$$\langle \mathrm{T}^*(\mathrm{T}(z)), z\rangle = \langle \mathrm{T}(z), \mathrm{T}(z)\rangle.$$
On a $\langle (\mathrm{Id}_\mathrm{E} + \mathrm{T}^*\mathrm{T})(z), z\rangle = \langle z + \mathrm{T}^*\mathrm{T}(z), z\rangle = \|z\|^2 + \|\mathrm{T}(z)\|^2$. Alors
$$\|z\|^2 \le \|z\|^2 + \|\mathrm{T}(z)\|^2 = \langle (\mathrm{Id}_\mathrm{E} + \mathrm{T}^*\mathrm{T})(z), z\rangle \le \|z\|\,\|(\mathrm{Id}_\mathrm{E} + \mathrm{T}^*\mathrm{T})(z)\|\,;$$
donc $\|z\| \le \|(\mathrm{Id}_\mathrm{E} + \mathrm{T}^*\mathrm{T})(z)\|$, il en résulte que $\mathrm{Id}_\mathrm{E} + \mathrm{T}^*\mathrm{T}$ est injectif, que l'inverse $(\mathrm{Id}_\mathrm{E} + \mathrm{T}^*\mathrm{T})^{-1}$ est continue et que $\|(\mathrm{Id}_\mathrm{E} + \mathrm{T}^*\mathrm{T})^{-1}\| \le 1$. Enfin, considérons $z = (\mathrm{Id}_\mathrm{E} + \mathrm{T}^*\mathrm{T})^{-1}(x)$ avec $x \in \mathrm{E}$; on a
$$\langle x, (\mathrm{Id}_\mathrm{E} + \mathrm{T}^*\mathrm{T})^{-1}(x)\rangle = \langle (\mathrm{Id}_\mathrm{E} + \mathrm{T}^*\mathrm{T})(z), z\rangle = \|z\|^2 + \|\mathrm{T}(z)\|^2 \ge 0.$$
Donc $(\mathrm{Id}_\mathrm{E} + \mathrm{T}^*\mathrm{T})^{-1}$ est un élément positif de $\mathcal{L}(\mathrm{E})$. Comme $(\mathrm{Id}_\mathrm{E} + \mathrm{T}^*\mathrm{T})^{-1}$ est injectif et autoadjoint, son image est dense (d'après la proposition 5.3.3). Par la proposition 3, l'opérateur $\mathrm{Id}_\mathrm{E} + \mathrm{T}^*\mathrm{T}$ est autoadjoint, donc $\mathrm{T}^*\mathrm{T}$ est autoadjoint. Comme l'opérateur $(\mathrm{Id}_\mathrm{E} + \mathrm{T}^*\mathrm{T})^{-1}$ est positif de norme ≤ 1, on a $\mathrm{Sp}(\mathrm{Id}_\mathrm{E} + \mathrm{T}^*\mathrm{T})^{-1} \subset [0,1]$; il en résulte que $\mathrm{Sp}(\mathrm{Id}_\mathrm{E} + \mathrm{T}^*\mathrm{T}) \subset [1,+\infty[$ (par le lemme 2.1) et $\mathrm{Sp}(\mathrm{T}^*\mathrm{T}) \subset [0,+\infty[$.

//

10.4. Décomposition spectrale

Exemple 10.4.1. Soit (X,μ) un espace mesuré et soit $\mathrm{L}_0(\mathrm{X},\mu)$ l'ensemble des classes d'équivalence de fonctions mesurables complexes pour la relation d'égalité μ-presque partout ; si $f,g \in \mathrm{L}_0(\mathrm{X},\mu)$, on voit facilement que la classe de $\widetilde{f}\widetilde{g}$ ne dépend pas des représentants $\widetilde{f}$, $\widetilde{g}$ des deux classes, ce qui permet de parler du produit ponctuel de deux classes. On peut alors définir un opérateur M_f dont le domaine est l'ensemble des $\xi \in \mathrm{L}_2(\mathrm{X},\mu)$ telles que $f\xi \in \mathrm{L}_2(\mathrm{X},\mu)$ et tel que $\mathrm{M}_f(\xi) = f\xi$ pour tout $\xi \in \mathrm{dom}(\mathrm{M}_f)$.

Pour tout $\xi \in \mathrm{L}_2(\mathrm{X},\mu)$ et tout $n > 0$, la fonction $\xi_n = (n+|f|)^{-1}n\xi$ est dans le domaine de M_f ; de plus la suite (ξ_n) converge partout vers ξ et est dominée par $|\xi|$. Il résulte alors du théorème de convergence dominée que le domaine de M_f est dense dans L_2. Soient $\xi,\eta \in \mathrm{L}_2(\mathrm{X},\mu)$ tels que $f\xi \in \mathrm{L}_2(\mathrm{X},\mu)$ et $\overline{f}\eta \in \mathrm{L}_2(\mathrm{X},\mu)$; on a
$$\langle f\xi,\eta\rangle = \int f(t)\xi(t)\overline{\eta(t)}\,d\mu(t) = \langle \xi, \overline{f}\eta\rangle.$$

On en déduit que $\mathrm{M}_{\overline{f}} \subset \mathrm{M}_f^*$. Enfin, soient $\xi, \eta \in \mathrm{L}_2(\mathrm{X}, \mu)$; posons $\xi_1 = (\xi + \overline{f}\eta)(1+|f|^2)^{-1}$ et $\eta_1 = (\eta - f\xi)(1+|f|^2)^{-1}$; clairement $f\xi_1 \in \mathrm{L}_2(\mathrm{X}, \mu)$, $\overline{f}\eta_1 \in \mathrm{L}_2(\mathrm{X}, \mu)$ et $(\xi, \eta) = (\xi_1, f\xi_1) + (-\overline{f}\eta_1, \eta_1)$. On en déduit que $\mathrm{L}_2(\mathrm{X}, \mu) \times \mathrm{L}_2(\mathrm{X}, \mu) = \mathrm{Gr}(\mathrm{M}_f) + \{(-z, y) : (y, z) \in \mathrm{Gr}(\mathrm{M}_{\overline{f}})\}$. Comme $\{(-z, y)\,,\ (y, z) \in \mathrm{Gr}(\mathrm{M}_{\overline{f}})\} \subset \mathrm{Gr}(\mathrm{M}_f)^\perp$, on en déduit que M_f et $\mathrm{M}_{\overline{f}}$ sont fermés et adjoint l'un de l'autre.

Posons $g = (1 + |f|)^{-1}$. Remarquons que $\mathrm{M}_g \in \mathcal{L}(\mathrm{L}_2(\mathrm{X}, \mu))$ est injective, que le domaine de M_f est l'image de M_g et que $\mathrm{M}_f \mathrm{M}_g = \mathrm{M}_{fg}$. On en déduit que $\mathrm{M}_f - \lambda\, \mathrm{Id}_{\mathrm{L}_2(\mathrm{X},\mu)}$ et $\mathrm{M}_{fg-\lambda g}$ ont même image. Le noyau de M_f est l'ensemble des fonctions $\xi \in \mathrm{L}_2(\mathrm{X}, \mu)$ telles que $f\xi$ soit μ-négligeable. Il coïncide avec celui de M_{fg}. Par la proposition 9.2.1, le spectre de M_f est l'ensemble des $\lambda \in \mathbb{C}$ tels que, pour tout $\varepsilon > 0$ l'ensemble $\{s \in \mathrm{X} : |f(s) - \lambda| < \varepsilon\}$ ne soit pas μ-négligeable.

Remarquons que M_f est injectif si et seulement si l'ensemble $\{s \in \mathrm{X} : f(s) = 0\}$ est μ-négligeable ; dans ce cas $\mathrm{im}(\mathrm{M}_f)$ est dense et $\mathrm{M}_f^{-1} = \mathrm{M}_{f^{-1}}$.

Théorème 10.4.1. *Soient* H *un espace de Hilbert séparable et* T *un opérateur densément défini autoadjoint de* H *dans* H *;*

(*i*) *le spectre de* T *est réel :* $\mathrm{Sp}(\mathrm{T}) \subset \mathbb{R}$ *;*

(*ii*) *l'opérateur* $\mathrm{U} = (\mathrm{T} - i\,\mathrm{Id}_\mathrm{H})(\mathrm{T} + i\,\mathrm{Id}_\mathrm{H})^{-1}$ *est un élément unitaire de* $\mathcal{L}(\mathrm{H})$ *;*

(*iii*) *il existe un espace mesuré* (X, μ)*, une fonction* $f : \mathrm{X} \to \mathbb{R}$ *mesurable et un isomorphisme* $u : \mathrm{L}_2(\mathrm{X}, \mu) \to \mathrm{H}$ *d'espaces de Hilbert tels que* $\mathrm{T} = u\, \mathrm{M}_f\, u^*$.

Démonstration. Soit $\lambda \in \mathbb{C} \setminus \mathbb{R}$; notons b sa partie imaginaire. Pour tout $x \in \mathrm{dom}(\mathrm{T})$, on a $\langle \mathrm{T}(x), x\rangle = \langle x, \mathrm{T}(x)\rangle$ donc $\langle \mathrm{T}(x), x\rangle \in \mathbb{R}$; la partie imaginaire de $\langle (\mathrm{T} - \lambda\, \mathrm{Id}_\mathrm{H})(x), x\rangle$ est donc $-b\|x\|^2$. On en déduit que

$$|b|\, \|x\|^2 \leq |\langle (\mathrm{T} - \lambda\, \mathrm{Id}_\mathrm{H})(x), x\rangle| \leq \|(\mathrm{T} - \lambda\, \mathrm{Id}_\mathrm{H})(x)\|\, \|x\|,$$

donc $\|(\mathrm{T} - \lambda\, \mathrm{Id}_\mathrm{H})(x)\| \geq |b|\, \|x\|$. On voit que pour tout $(x, y) \in \mathrm{Gr}(\mathrm{T} - \lambda\, \mathrm{Id}_\mathrm{H})$, on a $\|y\| \geq |b|\, \|x\|$, donc $(1 + |b|^2)\|y\|^2 \geq |b|^2(\|x\|^2 + \|y\|^2$. Par la proposition 6.3.1, l'application $(x, y) \to y$ de $\mathrm{Gr}(\mathrm{T} - \lambda\, \mathrm{Id}_\mathrm{H})$ dans H est injective d'image fermée ; donc $\mathrm{T} - \lambda\, \mathrm{Id}_\mathrm{H}$ est injective d'image fermée. Comme $(\mathrm{T} - \lambda\, \mathrm{Id}_\mathrm{H})^* = \mathrm{T} - \overline{\lambda}\, \mathrm{Id}_\mathrm{H}$ (proposition 3.3) est aussi injective, on en déduit que l'image de $\mathrm{T} - \lambda\, \mathrm{Id}_\mathrm{H}$ est dense (par la proposition 3.2), d'où le premier point.

Remarquons que l'image de $(\mathrm{T} + i\,\mathrm{Id}_\mathrm{H})^{-1}$ est le domaine de $\mathrm{T} - i\,\mathrm{Id}_\mathrm{H}$, ce qui implique que $\mathrm{U} = (\mathrm{T} - i\,\mathrm{Id}_\mathrm{H})(\mathrm{T} + i\,\mathrm{Id}_\mathrm{H})^{-1}$ est partout défini et bijectif (par (*i*)). Pour $x \in \mathrm{dom}(\mathrm{T})$, on a

$$\|(\mathrm{T} - i\,\mathrm{Id}_\mathrm{H})(x)\|^2 = \|\mathrm{T}(x)\|^2 + \|x\|^2 - i\langle x, \mathrm{T}(x)\rangle + i\langle \mathrm{T}(x), x\rangle =$$

$$= \|\mathrm{T}(x)\|^2 + \|x\|^2 = \|(\mathrm{T} + i\,\mathrm{Id}_\mathrm{H})(x)\|^2.$$

Soit $y \in \mathrm{H}$, posons $x = (\mathrm{T} + i\,\mathrm{Id}_\mathrm{H})^{-1}(y)$; on a $\|\mathrm{U}(y)\| = \|(\mathrm{T} - i\,\mathrm{Id}_\mathrm{H})(x)\|$ et aussi $\|\mathrm{U}(y)\| = \|(\mathrm{T} + i\,\mathrm{Id}_\mathrm{H})(x)\| = \|y\|$, donc U est isométrique, d'où (*ii*).

Passons au dernier point. Soit $y \in \mathrm{H}$, posons $x = (\mathrm{T} + i\,\mathrm{Id}_\mathrm{H})^{-1}(y)$; on écrit $\mathrm{U}(y) = (\mathrm{T} - i\,\mathrm{Id}_\mathrm{H})(x) = (\mathrm{T} + i\,\mathrm{Id}_\mathrm{H})(x) - 2ix = y - 2ix$; donc $x = i/2(\mathrm{U}(y) - y)$; on a donc $(\mathrm{T} + i\,\mathrm{Id}_\mathrm{H})^{-1} = i/2(\mathrm{U} - \mathrm{Id}_\mathrm{H})$, donc $\mathrm{T} = 2/i(\mathrm{U} - \mathrm{Id}_\mathrm{H})^{-1} - i\,\mathrm{Id}_\mathrm{H}$. Par le théorème 9.3.3, il existe un espace mesuré (X, μ), une fonction $g : \mathrm{X} \to \mathbb{C}$ mesurable de module 1 et un isomorphisme $u : \mathrm{L}_2(\mathrm{X}, \mu) \to \mathrm{H}$ d'espaces de Hilbert tels que $\mathrm{U} = u\, \mathrm{M}_g\, u^*$. L'application $\mathrm{U} - \mathrm{Id}_\mathrm{H}$ est injective, donc M_{g-1} est injective, d'où l'on déduit que $\{t \in \mathrm{X} : g(t) = 1\}$ est μ-négligeable. L'inverse de $\mathrm{U} - \mathrm{Id}_\mathrm{H}$ est donc

$(u\mathrm{M}_{g-1}u^*)^{-1} = (u^*)^{-1}\mathrm{M}_{(g-1)^{-1}}u^{-1} = u\,\mathrm{M}_{(g-1)^{-1}}\,u^*$, donc $\mathrm{T} = u\,\mathrm{M}_f\,u^*$ où on a posé $f = 2/i(g-1)^{-1} - i = -i(g+1)(g-1)^{-1}$.

//

Soient H un espace de Hilbert et T un opérateur densément défini autoadjoint de H dans H ; on peut, comme dans le cas borné, définir un calcul fonctionnel borélien pour T : écrivons $\mathrm{T} = u\,\mathrm{M}_f\,u^*$. Si g est une fonction borélienne bornée sur Sp(T), on pose $g(\mathrm{T}) = u\,\mathrm{M}_{g\circ f}\,u^*$. Si l'unitaire $\mathrm{U} = (\mathrm{T}-i\,\mathrm{Id_H})(\mathrm{T}+i\,\mathrm{Id_H})^{-1}$ s'écrit $u\,\mathrm{M}_h\,u^*$, on a $f = \varphi\circ h$, où $\varphi : t \to -i(t+1)(t-1)^{-1}$; pour toute fonction borélienne bornée g sur $\mathbb{R}$ on a donc

$$g(\mathrm{T}) = u\,\mathrm{M}_{g\circ f}\,u^* = u\,\mathrm{M}_{g\circ\varphi\circ h}\,u^* = g\circ\varphi(\mathrm{U}).$$

On en déduit que $g(\mathrm{T})$ ne dépend pas de l'écriture $\mathrm{T} = u\,\mathrm{M}_f\,u^*$.

10.5. Le théorème de Stone

Soit H un espace de Hilbert ; on appelle *groupe à un paramètre d'unitaires* une famille $(v_t)_{t\in\mathbb{R}}$ d'éléments unitaires de $\mathcal{L}(\mathrm{H})$ telle que :

(*i*) pour tous $s, t \in \mathbb{R}$ on a $v_{s+t} = v_s v_t$;

(*ii*) pour tout $x \in \mathrm{H}$ l'application $t \to v_t(x)$ est continue.

Lemme 10.5.1. *Soient* H *un espace de Hilbert,* $(v_t)_{t\in\mathbb{R}}$ *un groupe à un paramètre d'unitaires,* D *un sous-ensemble dense de* H *tel que pour tout* $z \in \mathrm{D}$ *et tout* $t \in \mathbb{R}$ *on ait* $v_t(z) \in \mathrm{D}$ *; soient* $x, y \in \mathrm{H}$ *tels que, pour tout* $z \in \mathrm{D}$*, l'application* $t \to \langle v_t(x), z\rangle$ *soit dérivable en* 0 *de dérivée* $i\langle y, z\rangle$ *; alors* $t \to v_t(x)$ *est continument dérivable de* $\mathbb{R}$ *dans* H *et sa dérivée en* t *est* $iv_t(y)$*.*

Démonstration. L'application $t \to v_t(y)$ est continue de $\mathbb{R}$ dans H. Pour $t \in \mathbb{R}$, posons $x_t = x + \int_0^t iv_s(y)\,ds$. L'application $t \to x_t$ est continument dérivable et sa dérivée en t est $iv_t(y)$. Soit $z \in \mathrm{D}$; posons $\varphi(t) = \langle x_t, z\rangle$ et $\psi(t) = \langle v_t(x), z\rangle$. La fonction φ est continument dérivable et sa dérivée en t est $\langle iv_t(y), z\rangle$. Soient $t, s \in \mathbb{R}$; on a $\psi(t+s) = \langle v_t v_s(x), v_t v_{-t}(z)\rangle = \langle v_s(x), v_{-t}(z)\rangle$; par hypothèse, cette fonction de s est dérivable en 0 et sa dérivée est $i\langle y, v_{-t}(z)\rangle = i\langle v_t(y), z\rangle$. On a montré que ψ est dérivable en t et $\psi' = \varphi'$. Comme $\varphi(0) = \psi(0)$, on trouve $\varphi = \psi$; donc, pour tout $z \in \mathrm{D}$ et tout $t \in \mathbb{R}$, on a $\langle (x_t - v_t(x)), z\rangle = 0$. Alors $x_t - v_t(x) \in \mathrm{D}^\perp$; donc $v_t(x) = x_t$, d'où le résultat.

//

Théorème 10.5.2 : Théorème de Stone. *Soit* H *un espace de Hilbert séparable ;*

(*i*) *soit* $(v_t)_{t\in\mathbb{R}}$ *un groupe à un paramètre d'opérateurs d'unitaires. Il existe un opérateur autoadjoint* T *sur* H *dont le graphe est l'ensemble des couples* $(x, y) \in \mathrm{H}\times\mathrm{H}$ *tels que la fonction* $t \to v_t(x)$ *soit dérivable en* 0*, de dérivée* iy*. Pour* $x \in \mathrm{dom(T)}$ *et* t *réel, on a* $v_t(x) \in \mathrm{dom(T)}$ *et* $\mathrm{T}(v_t(x)) = v_t(\mathrm{T}(x))$*.*

On dit que T *est le générateur infinitésimal de* $(v_t)_{t\in\mathbb{R}}$ *;*

(*ii*) *tout opérateur autoadjoint* T *est le générateur infinitésimal d'un unique groupe à un paramètre d'unitaires* $(v_t)_{t\in\mathbb{R}}$*.*

Démonstration. Montrons que le domaine de T est dense. Soient $x \in \mathrm{H}$ et $f : \mathbb{R} \to \mathbb{R}$ une fonction de classe C^1 à support compact ; posons $x_f = \int f(t) v_t(x)\, dt$. On a

$$v_t(x_f) = \int f(s) v_t v_s(x)\, ds = \int f(s-t) v_t(x)\, ds.$$

Donc $(v_t(x_f) - x_f)/t = \int t^{-1}[f(s-t) - f(s)] v_s(x)\, ds$. Quand t tend vers 0, $(v_t(x_f) - x_f)/t$ tend vers $-\int f'(s) v_s(x)\, ds$. On en déduit que $x_f \in \mathrm{dom(T)}$. Soit f_n une suite de fonctions positives de classe C^1 telles que $\int f_n(t)\, dt = 1$ et telles que f_n soit nulle en dehors de $]-1/n, 1/n[$; posons $y_n = \int f_n(t) v_t(x)\, dt$; on a $y_n \in \mathrm{dom(T)}$ et $\lim y_n = x$. Donc dom(T) est dense.

Si $x \in \mathrm{dom(T)}$, pour tout $z \in \mathrm{H}$ la fonction $t \to \langle v_t(x), z\rangle$ est dérivable en 0. Par le lemme 1, $t \to v_t(x)$ est de classe C^1. En particulier, pour tout $t \in \mathbb{R}$, l'application $s \to v_s(v_t(x))$ est dérivable en 0 ; donc $v_t(x) \in \mathrm{dom(T)}$.

Soient $x, y \in \mathrm{dom(T)}$; on a $\langle v_t(x), y\rangle = \langle x, v_{-t}(y)\rangle = \overline{\langle v_{-t}(y), x\rangle}$. La dérivée en 0 de $t \to \langle v_{-t}(y), x\rangle$ est $-i\langle \mathrm{T}(y), x\rangle$. On a donc $i\langle \mathrm{T}(x), y\rangle = \overline{-i\langle \mathrm{T}(y), x\rangle} = i\langle x, \mathrm{T}(y)\rangle$. On en déduit que $\mathrm{T} \subset \mathrm{T}^*$.

Enfin, soit $x \in \mathrm{dom(T^*)}$; pour tout $z \in \mathrm{dom(T)}$, on a $\langle v_t(x), z\rangle = \langle x, v_{-t}(z)\rangle$; donc $t \to \langle v_t(x), z\rangle$ est dérivable en 0 et sa dérivée est $i\langle x, \mathrm{T}(z)\rangle = i\langle \mathrm{T}^*(x), z\rangle$. Par le lemme 1, $x \in \mathrm{dom(T)}$ et $\mathrm{T}(x) = \mathrm{T}^*(x)$.

Passons au deuxième point. Soient (X, μ) un espace mesuré, $f : \mathrm{X} \to \mathbb{R}$ une fonction mesurable et $u : \mathrm{L}_2(\mathrm{X}, \mu) \to \mathrm{H}$ un isomorphisme d'espaces de Hilbert tels que l'on ait $\mathrm{T} = u\, \mathrm{M}_f\, u^*$ (théorème 4.1) ; pour $t \in \mathbb{R}$, notons $g_t : \mathrm{X} \to \mathbb{C}$ l'application $\exp(itf)$; posons $v_t = u\, \mathrm{M}_{g_t}\, u^*$. Comme $|g_t| = 1$, M_{g_t} est unitaire, donc v_t est unitaire ; comme $g_s g_t = g_{s+t}$, on a $v_s v_t = v_{s+t}$. Si $t_n \to t$, pour tout $\xi \in \mathrm{L}_2(\mathrm{X}, \mu)$ la suite de fonctions $g_{t_n}\xi$ converge partout vers $g_t\xi$, son module est constant égal à $|\xi|$; par le théorème de convergence dominée, $g_{t_n}\xi$ converge vers $g_t\xi$ dans L_2. Il en résulte que $t \to v_t$ est fortement continu ; c'est donc un groupe à un paramètre d'unitaires. Notons S son générateur infinitésimal. Pour tout $t \in \mathbb{R}^*$, on a $|g_t - 1| \leq |tf|$; si $\xi \in \mathrm{dom}(\mathrm{M}_f)$, pour toute suite (t_n) tendant vers 0, la suite de fonctions $t_n^{-1}(g_{t_n}\xi - \xi)$ converge partout vers $if\xi$, son module est majoré par $|f\xi|$; par le théorème de convergence dominée, $t_n^{-1}(g_{t_n}\xi - \xi)$ converge vers $if\xi$ dans L_2. Donc $t \to g_t\xi$ est dérivable en 0 et sa dérivée est $if\xi$. On en déduit que $u\xi \in \mathrm{dom(S)}$ et $i\mathrm{S}(u\xi) = u(if\xi)$. On a montré que $\mathrm{T} = u\, \mathrm{M}_f\, u^* \subset \mathrm{S}$. On en déduit que $\mathrm{S} = \mathrm{S}^* \subset \mathrm{T}^* = \mathrm{T}$. Donc $\mathrm{T} = \mathrm{S}$.

Enfin, soit $(v_t)_{t\in\mathbb{R}}$ et $(w_t)_{t\in\mathbb{R}}$ deux groupes à un paramètre d'unitaires ayant même générateur infinitésimal T ; pour $x \in \mathrm{dom(T)}$ et $t \in \mathbb{R}$, on a $v_t(x) \in \mathrm{dom(T)}$, $\mathrm{T}(v_t(x)) = v_t(\mathrm{T}(x))$, et $w_t(x) \in \mathrm{dom(T)}$, $\mathrm{T}(w_t(x)) = w_t(\mathrm{T}(x))$. Posons $\varphi(t) = \|v_t(x) - w_t(x)\|^2$ pour tout t réel ; on a

$$\varphi(t) = \|v_t(x)\|^2 + \|w_t(x)\|^2 - 2\,\mathrm{Re}\langle v_t(x), w_t(x)\rangle = 2(\|x\|^2 - \mathrm{Re}\langle v_t(x), w_t(x)\rangle)\,;$$

La fonction φ est alors dérivable en tout $t \in \mathbb{R}$ et l'on a

$$\varphi'(t) = -2\,\mathrm{Re}(\langle v_t(x), i\mathrm{T}(w_t(x))\rangle + \langle i\mathrm{T}(v_t(x)), w_t(x)\rangle).$$

Comme $\mathrm{T} = \mathrm{T}^*$, on a $\langle \mathrm{T}(v_t(x)), w_t(x)\rangle = \langle v_t(x), \mathrm{T}(w_t(x))\rangle$, donc $\varphi' = 0$. Comme $\varphi(0) = 0$, on trouve $\varphi(t) = 0$. Donc $v_t(x) = w_t(x)$. Comme dom(T) est dense, on en déduit $v_t = w_t$.

//

On a en fait montré que, si T est le générateur infinitésimal d'un groupe à un paramètre d'unitaires $(v_t)_{t\in\mathbb{R}}$, alors $v_t = \exp(it\mathrm{T})$.

Index des deux parties

Index des notations