MT404, Cours n^o 9, Lundi 22 octobre 2001.

Compléments sur la transformée de Laplace d'une fonction bornée sur $[0,+\infty[$

On considère une fonction mesurable bornée g sur $(0,+\infty)$ et l'ouvert du plan complexe $\mathrm{U}=\{z\in\mathbb{C} : \operatorname{Re} z>0\}$. En écrivant $z=s+it\in\mathrm{U}$, $s>0$ et t réel, on aura

$$(\mathcal{L}g)(z)=f(s,t)=\int_0^{+\infty}\mathrm{e}^{-sx-itx}\,g(x)\,dx$$

et les théorèmes de dérivation à la Lebesgue montrent que f est de classe C^1 sur U, avec

$$\frac{\partial f}{\partial s}(s_0,t_0)=\int_0^{+\infty}-x\,\mathrm{e}^{-s_0x-it_0x}\,g(x)\,dx,$$

$$\frac{\partial f}{\partial t}(s_0,t_0)=\int_0^{+\infty}-ix\,\mathrm{e}^{-s_0x-it_0x}\,g(x)\,dx=i\,\frac{\partial f}{\partial s}(s_0,t_0)$$

ce qui implique que $\mathcal{L}g$ est holomorphe dans l'ouvert U. Si $(\mathcal{L}g)(y)=0$ pour tout y réel dans un intervalle ouvert non vide de $]0,+\infty[$, le principe des zéros isolés entraîne que $(\mathcal{L}g)(z)=0$ pour tout $z\in\mathrm{U}$. En particulier, si on fixe un réel $s>0$ on aura

$$\int_{\mathbb{R}}\bigl(\mathbf{1}_{[0,+\infty)}(x)\,\mathrm{e}^{-sx}\,g(x)\bigr)\,\mathrm{e}^{-itx}\;dx=0$$

pour tout t réel, c'est à dire que la transformée de Fourier de la fonction h définie sur $\mathbb{R}$ par $h(x)=\mathbf{1}_{(0,+\infty)}(x)\,\mathrm{e}^{-sx}\,g(x)$ est nulle ; comme $h\in\mathrm{L}_2(\mathbb{R})$, les propriétés de la transformation de Fourier impliquent $h=0$, donc $g(x)=0$ pour presque tout $x>0$.

Dans le même ordre d'idées que l'exercice de la dernière fois sur la totalité de la suite de fonctions exponentielles, mentionnons le résultat suivant, “pour la culture”.

Théorème de Müntz-Szasz. Si (a_n) est une suite strictement croissante de nombres >0 telle que $\sum 1/a_n=+\infty$, la famille formée de la fonction $\mathbf{1}$ et des fonctions $x\to x^{a_n}$ est totale dans $\mathrm{C}([0,1])$.

Comme cas particulier, on peut prendre la suite $a_n=n$ qui correspond au théorème de Weierstrass ; mais en fait, la démonstration de Müntz-Szasz s'appuie sur le théorème de Weierstrass.

En posant $x=\mathrm{e}^{-t}$ pour $t\ge 0$, on obtient à partir de Müntz-Szasz que si (a_n) est une suite strictement croissante de nombres >0 telle que $\sum 1/a_n=+\infty$, la famille formée des fonctions $t\to\mathrm{e}^{-a_nt}$ est totale dans $\mathrm{C}_0([0,+\infty[)$ (l'espace des fonctions continues sur $[0,+\infty[$ qui tendent vers 0 à l'infini).

Hahn-Banach et transposition

Définition 1.6.2. Soient X et Y deux espaces normés et $\mathrm{T}\in\mathcal{L}(\mathrm{X},\mathrm{Y})$; on appelle *transposée* de T l'application ${}^t\mathrm{T} : y^*\to y^*\circ\mathrm{T}$ de Y^* dans X^*.

Regroupons dans la proposition suivante des propriétés élémentaires de la transposition :

Proposition 1.6.2. *Soient* X, Y *et* Z *des espaces normés ;*

(i) *pour tout* $\mathrm{T}\in\mathcal{L}(\mathrm{X},\mathrm{Y})$, *l'application* ${}^t\mathrm{T}$ *est linéaire et continue et* $\|{}^t\mathrm{T}\|\le\|\mathrm{T}\|$ *;*

(ii) *l'application* $\mathrm{T}\to{}^t\mathrm{T}$ *est linéaire de* $\mathcal{L}(\mathrm{X},\mathrm{Y})$ *dans* $\mathcal{L}(\mathrm{Y}^*,\mathrm{X}^*)$ *;*

(iii) pour tout $S \in \mathcal{L}(X, Y)$ *et tout* $T \in \mathcal{L}(Y, Z)$*, on a* ${}^t(T \circ S) = {}^tS \circ {}^tT$ *(bien noter l'interversion de* S *et* T*).*

Vérifions que $\|{}^tT\| \le \|T\|$. Soit $y^* \in Y^*$; on a

$$\forall x \in X, \quad |{}^tT(y^*)(x)| = |y^*(T(x))| \le \|y^*\| \, \|T(x)\| \le \|y^*\| \, \|T\| \, \|x\|,$$

d'où il résulte que $\|{}^tT(y^*)\| \le \|T\| \, \|y^*\|$ pour tout $y^* \in Y^*$, donc $\|{}^tT\| \le \|T\|$. La démonstration des autres points est laissée en exercice.

On vient de montrer que la transposée ${}^tT$ d'une application linéaire continue T a une norme inférieure ou égale à celle de T. Grâce au théorème de Hahn-Banach on peut compléter ce résultat.

Proposition 4.2.10. *Soient* X *et* Y *deux espaces normés ; pour tout* $T \in \mathcal{L}(X, Y)$*, on a* $\|{}^tT\| = \|T\|$.

Démonstration. On sait déjà que $\|{}^tT\| \le \|T\|$ par la proposition 1.6.2, nous allons montrer l'égalité. Pour tout $\varepsilon > 0$, on peut trouver un vecteur $x \in X$ tel que $\|x\| \le 1$ et tel que $\|T(x)\| > \|T\| - \varepsilon$, puis par Hahn-Banach, une forme linéaire $y^* \in F^*$ telle que $\|y^*\| \le 1$ et $y^*(T(x)) = \|T(x)\|$. Alors

$$\|{}^tT\| \ge \|{}^tT(y^*)\| \ge |{}^tT(y^*)(x)| = y^*(T(x)) = \|T(x)\| > \|T\| - \varepsilon.$$

Exemple. Pour toute fonction f intégrable sur $[0, 1]$ posons

$$\forall x \in [0, 1], \quad (Vf)(x) = \int_0^x f(t)\, dt, \quad (Wf)(x) = \int_x^1 f(t)\, dt.$$

On a déjà vu que V envoie continûment L_1 dans les fonctions continues sur $[0, 1]$. Il en résulte que pour tout p tel que $1 \le p \le +\infty$, on définit un endomorphisme continu V_p de $L_p([0, 1])$ en posant $V_p(f) = V(f)$ pour toute $f \in L_p$. De même, en désignant par q l'exposant conjugué de p, on définit un endomorphisme W_q de $L_q([0, 1])$, en posant $W_q(g) = W(g)$ pour toute $g \in L_q$. Posons

$$\forall f \in L_p, \ \forall g \in L_q, \quad (f, g) = \int_0^1 f(t) g(t)\, dt$$

et remarquons en employant Fubini que

$$(V_p(f), g) = \int_0^1 (Vf)(t) g(t)\, dt = \int \mathbf{1}_{0 \le t \le 1} \Big(\int \mathbf{1}_{0 \le s \le t} f(s)\, ds \Big) g(t)\, dt =$$

$$= \int \mathbf{1}_{0 \le s \le t \le 1} \, f(s) g(t)\, ds dt = \int_0^1 f(s) \Big(\int_s^1 g(t)\, dt \Big)\, ds = (f, W_q(g)).$$

Supposons que $1 \le p < +\infty$ et que x^* soit une forme linéaire continue sur L_p ; on sait que x^* peut se représenter par une fonction $g \in L_q$ sous la forme $x^* = j_q(g)$, c'est à dire

$$\forall f \in L_p, \quad x^*(f) = \int_0^1 f(t) g(t)\, dt = (f, g).$$

Calculons ${}^tV(x^*)$; c'est une forme linéaire, qui est donc connue par son action sur L_p,

$$\forall f \in L_p, ({}^tVx^*)(f) = x^*(V_p(f)) = (V_p(f), g) = (f, W_q(g)) = j_q(W_q(g))(f)$$

ce qui montre que ${}^t\mathrm{V}(x^*) = j_q(\mathrm{W}_q(g))$; ainsi, si $x^* \in (\mathrm{L}_p)^*$ est représentée par $g \in \mathrm{L}_q$, la forme linéaire ${}^t\mathrm{V}(x^*)$ est représentée par $\mathrm{W}_q(g)$. Il est donc naturel de dire que W_q "est" la transposée de V_p, comme on a dit que L_q "est" le dual de L_p.

Dans le cas $1 < p < +\infty$, on a aussi $1 < q < +\infty$ et on peut voir de la même façon que W_q "est" la transposée de V_p.

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

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

Lemme 4.2.11. *Soient* X, Y *deux espaces normés et* $\mathrm{T} \in \mathcal{L}(\mathrm{X}, \mathrm{Y})$ *; l'application* ${}^t\mathrm{T}$ *est injective si et seulement si* $\mathrm{im}(\mathrm{T})$ *est dense dans* Y. *De plus, si* $\mathrm{im}({}^t\mathrm{T})$ *est dense dans* X^*, *l'application* T *est injective.*

Démonstration. Si $\mathrm{im}(\mathrm{T})$ n'est pas dense, son adhérence Z est un sous-espace vectoriel fermé de Y, distinct de Y. D'après le corollaire 6, il existe une forme linéaire y^* non nulle sur Y, mais dont la restriction à Z est nulle ; en particulier, $y^*(\mathrm{T}(x)) = 0$ pour tout $x \in \mathrm{X}$ puisque Z contient l'image de T. On a donc ${}^t\mathrm{T}(y^*)(x) = 0$ pour tout $x \in \mathrm{X}$, 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 $y^* \in \mathrm{Y}^*$ non nulle telle que ${}^t\mathrm{T}(y^*) = 0$, ce qui signifie que $y^*(\mathrm{T}(x)) = 0$ pour tout $x \in \mathrm{X}$. On voit alors que l'image de T est contenue dans le noyau de y^*, qui est un sous-espace fermé de Y, distinct de Y. Il en résulte que $\mathrm{im}(\mathrm{T})$ n'est pas dense dans Y.

Si $\mathrm{T}(x) = 0$, on a ${}^t\mathrm{T}(y^*)(x) = y^*(\mathrm{T}(x)) = 0$ pour tout $y^* \in \mathrm{Y}^*$, 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 X^*, on en déduit par continuité que $x^*(x) = 0$ pour tout $x^* \in \mathrm{X}^*$, donc $x = 0$ par Hahn-Banach ; il en résulte que T est injective.

Exemple (suite) : opérateurs V et W.

Supposons que $1 < p < +\infty$ et montrons que V_p est injectif. Si $\mathrm{V}_p(f) = 0$, la fonction $f \in \mathrm{L}_p([0,1])$ vérifie

$$\forall x \in [0,1], \quad 0 = \int_0^x f(t)\,dt = j_p(f)(\mathbf{1}_{[0,x]})$$

ce qui entraîne, en passant aux combinaisons linéaires de fonctions $\mathbf{1}_{[0,x]}$, que la forme linéaire $j_p(f)$ est nulle sur toutes les fonctions *en escalier* ; mais il résulte de la densité des fonctions continues dans L_q que les fonctions en escalier forment aussi un sous-ensemble dense dans L_q, donc $j_p(f)$ est la forme linéaire nulle sur L_q et il en résulte que $f = 0$, donc V_p est injectif.

On montre de la même façon que W_q est injectif. Comme $1 < p < +\infty$, l'opérateur W_q est la transposée de V_p et V_p est la transposée de W_q, ce qui permet d'appliquer le lemme précédent dans les deux sens, et de déduire que V_p et W_q sont à image dense.

En fait dans le cas présent, c'est une question d'humeur : on peut aussi bien montrer directement que les images sont denses (ça n'est pas plus dur) et en déduire l'injectivité. Par exemple, on voit que l'image de V_p contient toutes les fonctions continues, linéaires par morceaux et nulles en 0, et on montre facilement que cet ensemble est dense dans L_p quand $p < +\infty$.

Lorsque $p = 1$, on peut encore montrer que V_1 est injectif, mais le transposé W_∞ n'est pas à image dense : l'image $\mathrm{W}_\infty(\mathrm{L}_\infty)$ est formée de fonctions continues sur $[0,1]$, et $\mathrm{C}([0,1])$ est un sous-espace fermé de L_∞, distinct de L_∞.

4.3. Bidual d'un espace normé. Espaces de Banach réflexifs

Soit X un espace normé ; le dual du dual X^* de X s'appelle le *bidual* de X et se note $\mathrm{X}^{**} = (\mathrm{X}^*)^*$. Pour $x \in \mathrm{X}$ notons $\mathrm{J_X}(x) : \mathrm{X}^* \to \mathbb{K}$ la forme linéaire sur X^* qui à $x^* \in \mathrm{X}^*$ associe $x^*(x)$,

$$\forall x^* \in \mathrm{X}^*, \quad \mathrm{J_X}(x^*) = x^*(x).$$

Pour tout $x^* \in \mathrm{X}^*$, on a $|\mathrm{J_X}(x)(x^*)| = |x^*(x)| \le \|x^*\|\,\|x\|$, donc $\mathrm{J_X}(x) \in \mathrm{X}^{**}$ et $\|\mathrm{J_X}(x)\| \le \|x\|$. On dit que $\mathrm{J_X} \in \mathcal{L}(\mathrm{X}, \mathrm{X}^{**})$ est *l'application canonique* de X dans son bidual.

Proposition 4.3.1. *L'application canonique* $\mathrm{J_X} : \mathrm{X} \to \mathrm{X}^{**}$ *est isométrique.*

Démonstration. Soit $x \in \mathrm{X}$; par le corollaire 2.7, il existe $x^* \in \mathrm{X}^*$ tel que $\|x^*\| \le 1$ et $x^*(x) = \|x\|$. Alors

$$\|x\| = |x^*(x)| = |\mathrm{J_X}(x)(x^*)| \le \|x^*\|\,\|\mathrm{J_X}(x)\| \le \|\mathrm{J_X}(x)\|,$$

vu que $\|x^*\| \le 1$; donc $\|\mathrm{J_X}(x)\| = \|x\|$.

Remarque. Puisque X^{**} est toujours complet et que l'espace normé X s'injecte isométriquement dans X^{**}, on obtient une description d'un complété de l'espace X en considérant $\widehat{\mathrm{X}} = \overline{\mathrm{J_X}(\mathrm{X})}$: l'adhérence de l'image de X dans l'espace complet X^{**} est complète.

Si X est un espace normé l'application $\mathrm{J_X}$ est injective par la proposition 1. Remarquons que si $\mathrm{J_X}$ est bijective alors $\mathrm{J_X}$ est une isométrie de X sur un espace de Banach (par la proposition 1.2.5). Il s'ensuit que si $\mathrm{J_X}$ est bijective, alors nécessairement X est un espace de Banach. Ceci explique que nous restreindrons la définition qui suit aux espaces de Banach.

Définition 4.3.2. Un espace de Banach E est dit *réflexif* si l'application canonique $\mathrm{J_E} : \mathrm{E} \to \mathrm{E}^{**}$ est bijective.

Autrement dit, un espace de Banach E est réflexif lorsque toute forme linéaire x^{**} continue sur le dual E^* provient d'un vecteur x de E de la façon expliquée précédemment,

$$\forall x^* \in \mathrm{E}^*, \quad x^{**}(x^*) = x^*(x).$$

Cours n$^{\rm o}$ 10, Mercredi 24 octobre 2001.

V_1 *est injectif*

Ce résultat se prouve par un raisonnement de théorie de la mesure : lorsque deux mesures sont égales sur une classe de parties $\mathcal{C}$ stable par intersection finie qui engendre la tribu $\mathcal{A}$, elles sont égales pour tout ensemble $\mathrm{A} \in \mathcal{A}$ (lemme des classes monotones) ; on y reviendra.

Dans ce qui suit, la plupart des résultats sur la réflexivité utilisent la même stratégie : on veut montrer qu'un espace de Banach E est réflexif, donc on se donne une forme linéaire continue $\ell : \mathrm{E}^* \to \mathbb{K}$ et on cherche à la représenter sous la forme voulue, c'est à dire trouver un vecteur $x_0 \in \mathrm{E}$ tel que $\ell(x^*) = x^*(x_0)$ pour tout $x^* \in \mathrm{E}^*$; dans plusieurs situations, on disposera d'une "paramétrisation" de E^* par une application linéaire continue surjective $\mathrm{P} : \mathrm{Z} \to \mathrm{E}^*$, où Z est un autre espace de Banach. On regardera alors la forme linéaire $f = \ell \circ \mathrm{P}$ sur Z ; puisque P est surjective, la propriété voulue pour x_0 s'écrira

$$\forall z \in \mathrm{Z}, \quad f(z) = \ell(\mathrm{P}(z)) = (\mathrm{P}z)(x_0)\,;$$

si on dispose d'informations supplémentaires sur le dual de Z, on arrive à avancer.

Proposition 4.3.2. *Tout espace de Hilbert est réflexif.*

Démonstration. Soient H un espace de Hilbert et $x^{**} \in \mathrm{H}^{**}$; pour tout vecteur $y \in \mathrm{H}$ soit $\ell_y \in \mathrm{H}^*$ la forme linéaire sur H définie par $\ell_y(x) = \langle x, y\rangle$; l'application $y \to \overline{x^{**}(\ell_y)}$ est une forme linéaire et continue sur H. Par la proposition 2.3.5, il existe $x \in \mathrm{H}$ tel que, pour tout $y \in \mathrm{H}$ on ait $\overline{x^{**}(\ell_y)} = \langle y, x\rangle$. D'après la proposition 2.3.5, toute $f \in \mathrm{H}^*$ est de la forme $f = \ell_y$ pour un certain $y \in \mathrm{H}$, donc on a $x^{**}(f) = x^{**}(\ell_y) = \langle x, y\rangle = f(x)$, c'est à dire que x^{**} est l'image de x par l'application canonique de H dans H^{**}, qui est donc surjective.

Ici on a paramétré le dual $\mathrm{E}^* = \mathrm{H}^*$ avec $\mathrm{Z} = \mathrm{H}$ et P l'application $y \to \ell_y$; dans cette variante P est *antilinéaire*.

Exemple 4.3.3. Les espaces $\mathrm{L}_p(\Omega, \mathcal{A}, \mu)$ sont réflexifs lorsque $1 < p < +\infty$; on pourrait dire un peu vite : le dual de L_p est L_q, et celui de L_q est L_p, donc ça marche ; c'est un peu trop rapide, parce que le dual de L_p n'**est** pas L_q, mais *s'identifie* à L_q au moyen d'une certaine bijection. Il faut donc prendre la peine, au moins une fois, de vérifier que tout colle bien.

Expliquons le cas de $\mathrm{X} = \mathrm{L}_p$; soit j_q l'application isométrique de L_q sur le dual X^* de L_p. Si x^{**} est une forme linéaire continue sur $\mathrm{X}^* = (\mathrm{L}_p)^*$, la composée $x^{**} \circ j_q$ est une forme linéaire continue sur L_q ; il existe donc une fonction $f \in \mathrm{L}_p = \mathrm{X}$ telle que $x^{**} \circ j_q = j_p(f)$, c'est à dire

$$\forall g \in \mathrm{L}_q, \quad x^{**}(j_q(g)) = \int_\Omega f g \, d\mu.$$

Soit $x^* \in \mathrm{X}^*$; il existe $g \in \mathrm{L}_q$ tel que $x^* = j_q(g)$, et alors $x^*(f) = \int_\Omega f g\, d\mu$. La ligne précédente signifie donc bien que l'on a trouvé un vecteur $f \in \mathrm{X} = \mathrm{L}_p$ tel que

$$\forall x^* \in \mathrm{X}^* = (\mathrm{L}_p)^*, \quad x^{**}(x^*) = x^*(f).$$

Ici on a paramétré le dual de L_p par $j_q : \mathrm{L}_q \to (\mathrm{L}_p)^*$.

Espaces normés isomorphes

On dit que deux espaces normés X et Y sont *isomorphes* (en tant qu'espaces normés) s'il existe une application linéaire continue $\mathrm{T} : \mathrm{X} \to \mathrm{Y}$ bijective telle que T^{-1} soit continue de Y dans X (si X et Y sont complets, cette dernière condition est automatique par le théorème des isomorphismes).

Si X et Y sont isomorphes, on dispose d'un dictionnaire qui permet de transporter toutes les notions topologico-algébriques de X à Y et inversement : au vecteur $x \in \mathrm{X}$ on associe $y = \mathrm{T}(x) \in \mathrm{Y}$, et alors $x = \mathrm{T}^{-1}(y)$; à une forme linéaire $x^* \in \mathrm{X}^*$ on associe $y^* = x^* \circ \mathrm{T}^{-1} = {}^t(\mathrm{T}^{-1})(x^*) \in \mathrm{Y}^*$, et inversement $x^* = y^* \circ \mathrm{T} = {}^t\mathrm{T}(y^*)$. Il n'est alors pas surprenant que :

Lemme 4.3.3. *Si* X *est réflexif et si* Y *est isomorphe à* X*, alors* Y *est réflexif.*

Démonstration. Soit y^{**} une forme linéaire continue sur Y^* ; alors $x^{**} = y^* \circ {}^t(\mathrm{T}^{-1})$ est dans X^{**} ; puisque X est réflexif il existe $x \in \mathrm{X}$ tel que $x^{**}(x^*) = x^*(x)$ pour tout $x^* \in \mathrm{X}^*$. On pose $y = \mathrm{T}(x)$ et on vérifie que y représente y^{**} : soit y^* quelconque dans Y^* et écrivons $y^* = {}^t(\mathrm{T}^{-1})(x^*)$; on a

$$y^{**}(y^*) = y^{**} \circ {}^t(\mathrm{T}^{-1})(x^*) = x^{**}(x^*) = x^*(x) = {}^t\mathrm{T}(y^*)(x) = y^*(\mathrm{T}(x)) = y^*(y),$$

ce qu'il fallait démontrer (on a paramétré Y^* avec ${}^t(\mathrm{T}^{-1})$).

Proposition 4.3.4. *Si* X *est réflexif, alors* X^* *est réflexif.*

Démonstration. Posons $\mathrm{E} = \mathrm{X}^*$. Si $z^{**} = x^{***}$ est une forme linéaire sur le dual $\mathrm{E}^* = \mathrm{X}^{**}$ de $\mathrm{E} = \mathrm{X}^*$, elle définit une forme linéaire continue $z = x^* = x^{***} \circ \mathrm{J}_\mathrm{X}$ sur X. Il reste seulement à vérifier que z définit la forme z^{**}, au sens précédent. Soit $z^* \in \mathrm{E}^* = \mathrm{X}^{**}$; puisque X est réflexif il existe $x \in \mathrm{X}$ tel que $z^* = \mathrm{J}_\mathrm{X}(x)$. Alors

$$z^{**}(z^*) = x^{***}(\mathrm{J}_\mathrm{X}(x)) = x^*(x) = \mathrm{J}_\mathrm{X}(x)(x^*) = z^*(z),$$

ce qui montre bien que z^{**} provient du vecteur $z \in \mathrm{E}$. Ici on a paramétré E^* avec J_X.

Proposition 4.3.5. *Si* X *est réflexif, tout sous-espace fermé* Y *de* X *est réflexif.*

Démonstration. Soit π l'application de restriction définie de X^* sur Y^* (surjective par le théorème de Hahn-Banach) ; soit y^{**} une forme linéaire continue sur Y^*. Alors $x^{**} = y^{**} \circ \pi$ est une forme linéaire continue sur X^*, donc il existe $x \in \mathrm{X}$ tel que $x^{**}(x^*) = x^*(x)$ pour tout x^* in X^*. Il suffit de voir que $x \in \mathrm{Y}$ pour pouvoir conclure assez facilement ; si on avait $x \notin \mathrm{Y}$, on pourrait trouver d'après le corollaire 2.6 une forme linéaire $x^* \in \mathrm{X}^*$ telle que $x^*(x) = 1$ mais $x^*(y) = 0$ pour tout $y \in \mathrm{Y}$. On aurait alors $\pi(x^*) = 0$, donc $x^{**}(x^*) = y^{**}(\pi(x^*)) = 0$, ce qui contredit $x^{**}(x^*) = x^*(x) = 1$. Ici, on a paramétré Y^* par $\pi : \mathrm{X}^* \to \mathrm{Y}^*$.

Corollaire 4.3.6. *Si* E *est un espace de Banach et si* E^* *est réflexif, alors* E *est réflexif.*
En effet E^{**} est alors réflexif et E est isomorphe à un sous-espace fermé de E^{**}.

Pourquoi s'intéresser aux espaces réflexifs ?

Les espaces réflexifs ont une sorte de compacité : on verra que si (C_n) est une suite décroissante de convexes fermés bornés non vides d'un espace réflexif E, l'intersection $\bigcap_n \mathrm{C}_n$ est non vide. On en déduit que si f est une fonction convexe continue sur un convexe fermé borné non vide C d'un espace réflexif E, alors f atteint son minimum sur C. Cela permet de montrer que certains problèmes de minimisation ont une solution, quand on travaille avec un espace réflexif.

Hahn-Banach comme théorème de séparation

Dans un espace normé X, on suppose donnés un convexe fermé non vide C et un point $x \notin \mathrm{C}$. On veut trouver une forme linéaire x^* continue sur X et un nombre t tels que C soit contenu dans le demi-espace affine $\{y \in \mathrm{X} : x^*(y) < t\}$, et x dans l'autre demi-espace, c'est à dire $x^*(x) > t$.

On dit que $\mathrm{C} \subset \mathrm{X}$ est un *cône* si pour tout $c \in \mathrm{C}$ et tout $\lambda > 0$ on a $\lambda\, c \in \mathrm{C}$.

Si C est un cône convexe, on a $c_1 + c_2 \in \mathrm{C}$ lorsque $c_1, c_2 \in \mathrm{C}$. En effet,

$$c_1 + c_2 = 2\Big(\frac{1}{2}\,c_1 + \frac{1}{2}\,c_2\Big).$$

Lemme. *Soit* X *un espace normé ; la distance à un cône convexe non vide* $\mathrm{C} \subset \mathrm{X}$ *est une fonction sous-linéaire sur* X*.*

Démonstration. On pose

$$\forall x \in \mathrm{C}, \quad q(x) = \mathrm{dist}(x, \mathrm{C}) = \inf\{\|x - c\| : c \in \mathrm{C}\}.$$

Pour $\varepsilon > 0$ donné on peut trouver $c \in \mathrm{C}$ tel que $\|x - c\| < q(x) + \varepsilon$; pour $\lambda > 0$ on aura $\lambda c \in \mathrm{C}$ donc $\mathrm{dist}(\lambda x, \mathrm{C}) \le \|\lambda x - \lambda c\| < \lambda(q(x) + \varepsilon)$, donc $q(\lambda x) \le \lambda q(x)$ puisque $\varepsilon > 0$ est arbitraire. En répétant avec $x' = \lambda x$ et $\lambda' = \lambda^{-1}$ on aura $q(x) = q(\lambda' x') \le \lambda^{-1} q(\lambda x)$, donc

$$\forall \lambda > 0, \quad q(\lambda x) = \lambda q(x).$$

Par continuité on obtient aussi le cas $\lambda = 0$. On montre que $q(x_1 + x_2) \le q(x_1) + q(x_2)$ de la façon suivante : pour $i = 1, 2$ on trouve $c_i \in \mathrm{C}$ tel que $\|x_i - c_i\| < q(x_i) + \varepsilon/2$, et puisque $c_1 + c_2 \in \mathrm{C}$ on a

$$q(x_1 + x_2) \le \|(x_1 + x_2) - (c_1 + c_2)\| \le \|x_1 - c_1\| + \|x_2 - c_2\| < q(x_1) + q(x_2) + \varepsilon.$$

Remarquons que $q(x) = \mathrm{dist}(x, \mathrm{C}) = \mathrm{dist}(x, \overline{\mathrm{C}})$, et $0_{\mathrm{X}} \in \overline{\mathrm{C}}$ (si C est non vide ; prendre $c \in \mathrm{C}$ puis λc avec $\lambda \to 0$). On a donc $q(x) \le \|x - 0_{\mathrm{X}}\| = \|x\|$.

Proposition. *Si* $\mathrm{C} \subset \mathrm{X}$ *est un cône convexe non vide et si* $\mathrm{dist}(x, \mathrm{C}) > 0$*, il existe une forme linéaire continue* x^* *sur* X*, non nulle, telle que* $x^* \le 0$ *sur* C *et* $x^*(x) > 0$*.*

Démonstration. Sur le sous-espace $\mathrm{Y} = \mathbb{R}x$, on définit une forme linéaire ℓ par

$$\ell(tx) = tq(x)$$

pour tout t réel. On vérifie que $\ell \le q$, ce qui permet de prolonger par Hahn-Banach en une forme linéaire $m \le q$. On note que $q(x) \le \|x\|$, donc $m = x^*$ est continue sur X. Puisque $q(y) = 0$ quand $y \in \mathrm{C}$, on a bien $x^* \le 0$ sur C, et comme x^* prolonge ℓ on a $x^*(x) = \ell(x) = q(x) > 0$.

MT404, Cours n° 11, lundi 29 octobre 2001.

V_1 *est injectif : suite et fin*

Supposons que $f \in \mathrm{L}_1(0,1)$ et $\mathrm{V}_1(f) = 0$. On sait que cela implique que

$$\int_0^1 \varphi(t) f(t)\, dt = 0$$

pour toute fonction φ en escalier sur $[0,1]$. Considérons un borélien B quelconque de $[0,1]$ et sa fonction indicatrice $\mathbf{1}_\mathrm{B}$. Puisque le sous-espace des fonctions en escalier est dense dans $\mathrm{L}_1(0,1)$, on peut trouver une suite (φ_n) de fonctions en escalier qui converge dans L_1 vers $\mathbf{1}_\mathrm{B}$. On peut supposer que $|\varphi_n| \le 1$, en remplaçant φ_n par $\psi_n = \max(0, \min(\varphi_n, 1))$, fonction à valeurs dans $[0,1]$ et qui donne une meilleure approximation de $\mathbf{1}_\mathrm{B}$ parce que

$$|s - \max(0, \min(t,1))| \le |s - t|$$

lorsque $0 \le s \le 1$ (appliquer avec $s = \mathbf{1}_\mathrm{B}(x)$ et $t = \varphi_n(x)$). D'après un théorème d'intégration, on peut trouver une sous-suite (φ_{n_k}) qui converge presque partout vers $\mathbf{1}_\mathrm{B}$. Alors $(\varphi_{n_k} f)$ converge presque partout vers f en étant dominée en module par la fonction intégrable fixe $|f|$. D'après le théorème de convergence dominée de Lebesgue, on aura

$$\int_0^1 \mathbf{1}_\mathrm{B}(t) f(t)\, dt = \lim_k \int_0^1 \varphi_{n_k}(t) f(t)\, dt = 0.$$

Ceci ayant lieu pour tout $\mathrm{B} \in \mathcal{B}$, il en résulte que $f = 0$ (voir le lemme 3.4.1).

Séparation : suite

Rappel. *Si* $\mathrm{C} \subset \mathrm{X}$ *est un cône convexe non vide et si* $\mathrm{dist}(x, \mathrm{C}) > 0$, *il existe une forme linéaire continue* x^* *sur* X, *non nulle, telle que* $x^* \le 0$ *sur* C *et* $x^*(x) > 0$.

Remarque : le cas d'un Hilbert (réel) ; dans ce cas on obtient le résultat précédent en utilisant le théorème de projection dans un Hilbert : si y est le point de $\overline{\mathrm{C}}$ le plus proche de x et si $v = x - y$, la forme linéaire $x^* : z \to \langle z, v \rangle$ vérifie les conclusions voulues.

Théorème. *Soit* X *un espace normé réel ; si* C *est un convexe ouvert non vide de* X *qui ne contient pas* x, *il existe une forme linéaire continue* x^* *sur* X *qui vérifie*

$$\forall a \in \mathrm{A}, \quad x^*(a) < x^*(x).$$

Démonstration. On commence par translater A en

$$\mathrm{A}_1 = \mathrm{A} - x = \{a - x : a \in \mathrm{A}\}\,;$$

ce nouvel ensemble A_1 est toujours convexe, ouvert, et maintenant $0 \notin \mathrm{A}_1$. On fabrique le cône $\mathrm{C} = \bigcup_{\lambda > 0} \lambda \mathrm{A}_1$. Cet ensemble est clairement ouvert, et il est aussi un cône (facile), convexe : si $c_1 = \lambda_1 a_1$, $c_2 = \lambda_2 a_2$, avec $\lambda_1, \lambda_2 > 0$, $a_1, a_2 \in \mathrm{A}_1$ et si $0 \le t \le 1$,

$$(1-t)c_1 + tc_2 = \big((1-t)\lambda_1 + t\lambda_2\big) \Big(\frac{(1-t)\lambda_1}{(1-t)\lambda_1 + t\lambda_2}\, a_1 + \frac{t\lambda_2}{(1-t)\lambda_1 + t\lambda_2}\, a_2 \Big)$$

est bien de la forme $\mu\, b$, avec $\mu = (1-t)\lambda_1 + t\lambda_2 > 0$ et $b \in \mathrm{A}_1$.

Puisque A_1 est non vide, on peut choisir $x_0 \in \mathrm{A}_1$ et puisque A_1 est ouvert on peut trouver $r > 0$ tel que $\mathrm{B}(x_0, r) \subset \mathrm{A}_1$; on va vérifier que $\mathrm{B}(-x_0, r)$ est disjoint de C, ce qui signifie que $\mathrm{dist}(-x_0, \mathrm{C}) \geq r > 0$: si on avait $y \in \mathrm{C} \cap \mathrm{B}(-x_0, r)$, on pourrait écrire $y = -x_0 + v \in \mathrm{C}$ avec $\|v\| < r$; mais $z = x_0 - v \in \mathrm{B}(x_0, r) \subset \mathrm{C}$, et on aurait $0_\mathrm{X} = y + z \in \mathrm{C}$ ce qui n'est pas vrai (parce que $0_\mathrm{X} \notin \mathrm{A}_1$).

On peut donc appliquer le *rappel* ci-dessus et trouver $x^* \in \mathrm{X}^*$ telle que $x^* \leq 0$ sur C et $x^*(-x_0) > 0$. Soit $a_1 \in \mathrm{A}_1$; puisque A_1 est ouvert, il existe $t > 0$ assez petit tel que $a_1 - tx_0 \in \mathrm{A}_1 \subset \mathrm{C}$, donc $x^*(a_1 - tx_0) = x^*(a_1) + tx^*(-x_0) \leq 0$, donc $x^*(a_1) < 0$. Pour tout $a \in \mathrm{A}$, on obtient avec $a_1 = a - x \in \mathrm{A}_1$ l'inégalité $x^*(a - x) < 0$, ce qui donne le résultat.

6. Opérateurs bornés sur les espaces de Hilbert

On introduira dans ce chapitre les principales classes d'opérateurs bornés entre espaces de Hilbert.

6.1. Applications linéaires continues entre espaces de Hilbert

On commence avec la notion essentielle d'application linéaire *adjointe* associée à une application linéaire continue T entre deux espaces de Hilbert ; plusieurs des classes particulières d'opérateurs bornés sur les espaces de Hilbert seront ensuite définies au moyen de cette notion.

Proposition 6.1.1. *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 $\langle \mathrm{T}(x), y \rangle = \langle x, \mathrm{T}^*(y) \rangle$. On vérifie facilement que $\mathrm{T}^*(y) + \lambda \mathrm{T}^*(z)$ vérifie la propriété caractéristique de $\mathrm{T}^*(y + \lambda z)$, pour tous $y, z \in \mathrm{F}$ et $\lambda \in \mathbb{K}$, d'où l'on déduit que T^* est linéaire. On a, par définition de $\|\mathrm{T}\|$ et par la proposition 2.2.1

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

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 ${}^t\mathrm{T}$ 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 6.1.1. Soient E et F deux espaces de Hilbert et $\mathrm{T} \in \mathcal{L}(\mathrm{E}, \mathrm{F})$; l'unique application linéaire $\mathrm{T}^* \in \mathcal{L}(\mathrm{F}, \mathrm{E})$ telle 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 appelée *adjointe* de T.

Regroupons dans la proposition suivante quelques propriétés des adjoints.

Proposition 6.1.2. *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 de Hilbert* 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 $\|\mathrm{T}^* \circ \mathrm{T}\| \le \|\mathrm{T}^*\| \, \|\mathrm{T}\| = \|\mathrm{T}\|^2$. De plus, pour tout $x \in \mathrm{E}$ tel que $\|x\| \le 1$ on a

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

grâce à l'inégalité de Cauchy-Schwarz, d'où résulte $\|\mathrm{T}\|^2 \le \|\mathrm{T}^* \circ \mathrm{T}\|$ et l'égalité cherchée. Les autres propriétés sont laissées en exercice.

Exemples 6.1.2.

1. Opérateur diagonal dans une base hilbertienne (h_n) de H : soit $\alpha = (\alpha_n)$ une suite bornée de scalaires et définissons Δ_α sur H par

$$\forall c = (c_n) \in \ell_2, \quad \Delta_\alpha\Bigl(\sum_{n=0}^{+\infty} c_n h_n\Bigr) = \sum_{n=0}^{+\infty} \alpha_n c_n h_n.$$

On voit que Δ_α est continu et que $\|\Delta_\alpha\| = \|\alpha\|_\infty$. On vérifie que l'adjoint de Δ_α est l'opérateur diagonal associé à la suite complexe conjuguée, $\Delta_\alpha^* = \Delta_{\overline{\alpha}}$.

Si f est un élément de $\mathrm{L}_\infty(\Omega, \mathcal{A}, \mu)$, on définit l'opérateur de multiplication M_f sur $\mathrm{L}_2(\Omega, \mathcal{A}, \mu)$ par $\mathrm{M}_f(g) = fg$ pour toute $g \in \mathrm{L}_2(\Omega, \mathcal{A}, \mu)$. On vérifie que M_f est borné sur $\mathrm{L}_2(\Omega, \mathcal{A}, \mu)$, avec $\|\mathrm{M}_f\| \le \|f\|_\infty$, et on montre que $\mathrm{M}_f^* = \mathrm{M}_{\overline{f}}$.

Exercice proposé. Montrer que $\|\mathrm{M}_f\|_{\mathcal{L}(\mathrm{L}_2)} = \|f\|_\infty$.

– Opérateur à noyau mesurable borné. Si $\mathrm{K}(s,t)$ est une fonction mesurable bornée sur le carré $[0,1]^2$, on peut poser pour $f \in \mathrm{L}_2(0,1)$ et pour tout $s \in [0,1]$

$$(\mathrm{T_K} f)(s) = \int_0^1 \mathrm{K}(s,t)\, f(t)\, dt\,;$$

en effet la fonction $t \to \mathrm{K}(s,t) f(t)$ est mesurable et bornée en module par f, donc elle est dans $\mathrm{L}_2(0,1)$, donc intégrable. De plus l'inégalité $|(\mathrm{T_K} f)(s)| \le \|\mathrm{K}\|_\infty \|f\|_2$ montre que la fonction image $\mathrm{T_K} f$ est bornée sur $[0,1]$, mesurable par Fubini, donc $\mathrm{T_K} f \in \mathrm{L}_2$. On a ainsi défini un opérateur linéaire $\mathrm{T_K}$ borné sur $\mathrm{L}_2(0,1)$. Avec Fubini on montre que $(\mathrm{T_K})^*$ est l'opérateur défini par le noyau K^* donné par la formule

$$\forall s,t \in [0,1], \quad \mathrm{K}^*(s,t) = \overline{\mathrm{K}(t,s)}.$$

L'opérateur V_2 étudié précédemment entre dans ce cadre : il faut prendre pour noyau K l'indicatrice du triangle $\mathbf{1}_{\{0 \le t \le s \le 1\}}$; on retrouve W_2 comme adjoint de V_2.

Définition. Soit H un espace de Hilbert ; un élément $\mathrm{T} \in \mathcal{L}(\mathrm{H})$ est appelé *hermitien* ou *autoadjoint* si $\mathrm{T} = \mathrm{T}^*$, et *positif* s'il est hermitien et si $\langle \mathrm{T}(x), x \rangle$ est réel ≥ 0 pour tout $x \in \mathrm{H}$.

Remarque. Si A est hermitien, $\langle \mathrm{A}x, x \rangle$ est réel pour tout $x \in \mathrm{H}$.

Exemples.

— L'opérateur diagonal Δ_α est hermitien si et seulement si la suite α est réelle. L'opérateur de multiplication M_f est hermitien si et seulement si f est réelle presque partout. Un opérateur à noyau T_K est hermitien quand le noyau K a la *symétrie hermitienne*,

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

— Pour tout opérateur borné $\mathrm{T} \in \mathcal{L}(\mathrm{E},\mathrm{F})$ entre deux Hilbert, $\mathrm{T}^*\mathrm{T}$ est hermitien : en effet, $(\mathrm{T}^*\mathrm{T})^* = \mathrm{T}^*\mathrm{T}^{**} = \mathrm{T}^*\mathrm{T}$; de plus $\mathrm{T}^*\mathrm{T}$ est positif, puisque $\langle \mathrm{T}^*\mathrm{T}(x), x\rangle = \|\mathrm{T}(x)\|^2$ pour tout $x \in \mathrm{E}$.

Cours n° 12, mercredi 31 octobre 2001.

Exemple. Soient H un espace de Hilbert, F un sous-espace vectoriel fermé de H et $\mathrm{P} = \mathrm{P}_\mathrm{F} \in \mathcal{L}(\mathrm{H})$ la projection orthogonale sur F. On va voir que P est hermitien, et positif.

Pour tous vecteurs $x, y \in \mathrm{H}$ on peut dire que $\mathrm{P}x \in \mathrm{F}$ et $y - \mathrm{P}y \in \mathrm{F}^\perp$. On a donc $\langle \mathrm{P}x, y - \mathrm{P}y\rangle = 0$, ce qui donne $\langle \mathrm{P}x, y\rangle = \langle \mathrm{P}x, \mathrm{P}y\rangle$. En recommençant avec $x - \mathrm{P}x \in \mathrm{F}^\perp$ et $\mathrm{P}y \in \mathrm{F}$ on obtient finalement que

$$\langle \mathrm{P}x, y\rangle = \langle \mathrm{P}x, \mathrm{P}y\rangle = \langle x, \mathrm{P}y\rangle\ ;$$

l'égalité entre les extrêmes signifie que $\mathrm{P}^* = \mathrm{P}$; en utilisant le terme central et $y = x$ on montre la positivité de P.

Lemme de polarisation. *Si* $\mathrm{A} \in \mathcal{L}(\mathrm{H})$ *est hermitien et si* $\langle \mathrm{A}x, x\rangle = 0$ *pour tout* $x \in \mathrm{H}$, *alors* $\mathrm{A} = 0$.

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

$$0 = \langle \mathrm{A}(x+y), x+y\rangle = \langle \mathrm{A}x, x\rangle + \langle \mathrm{A}x, y\rangle + \langle \mathrm{A}y, x\rangle + \langle \mathrm{A}y, y\rangle =$$
$$\langle \mathrm{A}x, y\rangle + \langle y, \mathrm{A}x\rangle = 2\,\mathrm{Re}\langle \mathrm{A}x, y\rangle.$$

En appliquant avec iy on trouve aussi $\mathrm{Im}\langle \mathrm{A}x, y\rangle = 0$, donc $\langle \mathrm{A}x, y\rangle = 0$ pour tous x, y ; en appliquant avec $y = \mathrm{A}x$ on obtient $\|\mathrm{A}x\|^2 = 0$, donc $\mathrm{A}x = 0$ pour tout x, donc $\mathrm{A} = 0$.

Définition. Soit H un espace de Hilbert ; un élément $\mathrm{T} \in \mathcal{L}(\mathrm{H})$ est appelé *normal* si $\mathrm{T}^* \circ \mathrm{T} = \mathrm{T} \circ \mathrm{T}^*$.

Exemples.

Les opérateurs diagonaux Δ_α de l'exemple 2 sont normaux car on voit facilement que

$$\Delta_\alpha \Delta_\alpha^* = \Delta_\alpha \Delta_{\overline{\alpha}} = \Delta_{\alpha\overline{\alpha}} = \Delta_{\overline{\alpha}} \Delta_\alpha = \Delta_\alpha^* \Delta_\alpha.$$

Pour la même raison les opérateurs de multiplication M_f sont normaux.

Définition. On dit que $\mathrm{U} \in \mathcal{L}(\mathrm{E},\mathrm{F})$ est *isométrique* si $\|\mathrm{U}(x)\| = \|x\|$ pour tout $x \in \mathrm{E}$.

Exemple. On définit le *shift à droite* (ou *décalage à droite*) $\mathrm{S} \in \mathcal{L}(\ell_2(\mathbb{N}))$ par

$$\mathrm{S}(x_0, x_1, \ldots) = (0, x_0, x_1, \ldots)$$

pour tout $x = (x_0, x_1, \ldots) \in \ell_2(\mathbb{N})$. Il est clair que S est linéaire et isométrique. Son adjoint S^* est le décalage à gauche,

$$\mathrm{S}^*(x_0, x_1, \ldots) = (x_1, x_2, \ldots).$$

Lemme. *L'opérateur* $\mathrm{U} \in \mathcal{L}(\mathrm{E}, \mathrm{F})$ *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 le lemme de polarisation : si on sait que U est isométrique, on a $\langle \mathrm{U}^*\mathrm{U}x, x \rangle = \langle x, x \rangle$ pour tout x, donc l'opérateur hermitien $\mathrm{A} = \mathrm{U}^*\mathrm{U} - \mathrm{Id}_\mathrm{E}$ vérifie $\langle \mathrm{A}x, x \rangle = 0$ pour tout x, donc $\mathrm{A} = 0$ par le lemme de polarisation.

Définition. 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 $\mathrm{U} \circ \mathrm{U}^* = \mathrm{Id}_\mathrm{F}$.

Le shift dans le cas de $\ell_2(\mathbb{Z})$ est unitaire. On définit maintenant la suite $\mathrm{S}x$, si $x = (x_n)_{n \in \mathbb{Z}}$ est dans $\ell_2(\mathbb{Z})$, par la formule

$$(\mathrm{S}x)_n = x_{n-1}$$

pour tout $n \in \mathbb{Z}$; l'adjoint (et inverse) de S est le décalage à gauche, donné par $(\mathrm{S}^*x)_n = x_{n+1}$ pour tout $n \in \mathbb{Z}$.

Proposition 6.1.3. *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*) *l'opérateur* T *est une isométrie de* E *sur* F.

Démonstration. Voir poly.

On verra plus loin un théorème de représentation des normaux : tout endomorphisme normal d'un espace de Hilbert complexe *peut être représenté* comme un opérateur M_f de multiplication par une fonction mesurable bornée. On va traiter un exemple à la main pour le shift sur $\ell_2(\mathbb{Z})$.

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 équivalent à l'opérateur de multiplication par la fonction $t \to \mathrm{e}^{it}$ sur H_2.

Suites faiblement convergentes

Définition. Soit H un espace de Hilbert ; une suite $(x_n) \subset \mathrm{H}$ est dite *faiblement convergente* vers un vecteur $x \in \mathrm{H}$ si

$$\lim_n \langle x_n, y \rangle = \langle x, y \rangle$$

pour tout $y \in \mathrm{H}$.

Cela revient à dire que pour toute forme linéaire continue ℓ sur H, on a $\ell(x) = \lim_n \ell(x_n)$. Sous cette forme la définition se généralise à tout espace normé X : une suite

$(x_n) \subset \mathrm{X}$ est faiblement convergente vers $x \in \mathrm{X}$ si $x^*(x) = \lim_n x^*(x_n)$ pour toute forme linéaire continue $x^* \in \mathrm{X}^*$.

On vérifie facilement que la limite faible est unique quand elle existe. Dans le cas hilbertien, on aura si x et x' sont deux candidats limite faible,

$$\langle x, y\rangle = \langle x', y\rangle$$

pour tout y, donc $\langle x - x', y\rangle = 0$ pour tout $y \in \mathrm{H}$, et on en déduit que $x - x' = 0$ en choisissant $y = x - x'$.

Exemple. Soit $(e_n)_{n\geq 0}$ une suite orthonormée dans un espace de Hilbert H ; alors la suite (e_n) converge faiblement vers le vecteur nul 0_{H}.

En effet, si H est un Hilbert et (e_n) une suite orthonormée dans H, on a

$$\sum_{n\geq 0} |\langle e_n, y\rangle|^2 \leq \|y\|^2$$

pour tout $y \in \mathrm{H}$ (inégalité de Bessel - théorème 3.2) ; la suite $(\langle e_n, y\rangle)$ est de carré sommable donc tend vers $0 = \langle 0_{\mathrm{H}}, y\rangle$.

Lemme 5.3.1 *Dans un espace normé toute suite faiblement convergente est bornée.*

Démonstration. Soit (x_n) une suite faiblement convergente ; en plongeant isométriquement X dans X^{**} on peut considérer (x_n) comme une suite d'applications linéaires de l'espace de Banach X^* dans $\mathbb{K}$ qui converge en tout point $x^* \in \mathrm{X}^*$; il résulte alors du théorème de Banach-Steinhaus (corollaire 4.1.9) que $\{\|x_n\| : n \geq 0\}$ est borné.

Lemme. *Si* C *est un convexe fermé d'un espace normé* X *et si* $(x_n) \subset \mathrm{C}$ *converge faiblement vers* x*, alors* $x \in \mathrm{C}$.

Démonstration. Raisonnons par l'absurde en supposant que $x \notin \mathrm{C}$. Supposons d'abord que $\mathbb{K} = \mathbb{R}$. D'après le théorème de séparation d'un convexe fermé et d'un point, on aurait si $x \notin \mathrm{C}$ une forme linéaire $x^* \in \mathrm{X}^*$ et un réel t tels que $x^* \leq t$ sur C et $x^*(x) > t$.

Si $\mathbb{K} = \mathbb{C}$, on considère d'abord X comme espace réel pour trouver x^* $\mathbb{R}$-linéaire comme ci-dessus, puis on trouve y^* $\mathbb{C}$-linéaire telle que $\mathrm{Re}\, y^* = x^*$. Dans tous les cas on aura une forme $\mathbb{K}$-linéaire continue y^* telle que $\mathrm{Re}\, y^* \leq t$ sur C et $\mathrm{Re}\, y^*(x) > t$.

En particulier, $\mathrm{Re}\, y^*(x_n) \leq t$ pour tout n devrait donner $\mathrm{Re}\, y^*(x) \leq t$ à la limite, contradiction.

Lemme. *Soit* (x_n) *une suite* **bornée** *dans un espace de Hilbert* H *; pour que cette suite soit faiblement convergente vers* x*, il suffit que* $\langle x, z\rangle = \lim_n \langle x_n, z\rangle$ *pour tout* z *d'un ensemble* $\mathcal{T}$ *total dans* H.

Démonstration. Pour fixer les idées, supposons $\|x_n\| \leq 1$ pour tout $n \geq 0$. On sait alors que $\|x\| \leq 1$, soit en appliquant le lemme précédent, soit en écrivant

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

Si la limite existe pour tout z d'un ensemble total $\mathcal{T}$, elle existe aussi, par linéarité, pour tout d de l'ensemble dense $\mathrm{D} = \mathrm{Vect}(\mathcal{T})$. Montrons que $(\langle x_n, y\rangle)_{n\geq 0}$ converge vers $\langle x, y\rangle$ pour tout $y \in \mathrm{H}$. Choisissons $\varepsilon > 0$ et $d \in \mathrm{D}$ tels que $\|y - d\| < \varepsilon/3$. Pour tout entier n assez grand, disons $n \geq \mathrm{N}_0(\varepsilon)$, on a $|\langle x_n, d\rangle - \langle x, d\rangle| < \varepsilon/3$, et pour tout entier n on a $|\langle x_n, d\rangle - \langle x_n, y\rangle| < \|x_n\|\,\|d - y\| < \varepsilon/3$, et $|\langle x, d\rangle - \langle x, y\rangle| < \|x\|\,\|d - y\| < \varepsilon/3$; il en résulte que pour $n \geq \mathrm{N}_0(\varepsilon)$, on aura $|\langle x_n, y\rangle - \langle x, y\rangle| < \varepsilon$, et la suite $(x_n)_{n\geq 0}$ converge faiblement vers x.

Exemple. Définissons une suite (f_n) de fonctions dans $\mathrm{L}_2(0,1)$ par $f_n(t) = (-1)^{[nt]}$ (la notation $[x]$ désigne la partie entière du nombre réel x). Cette suite est formée de fonctions de module un, donc de norme un dans L_2. On va montrer qu'elle tend faiblement vers 0 dans L_2 :

prenons $\mathcal{T} = \mathrm{C}([0,1])$, qui est dense dans L_2. On aura pour g continue

$$(*) \qquad \langle f_n, g\rangle = \int_0^1 f_n(t)\overline{g(t)}\,dt = \sum_{k=0}^{n-1} (-1)^k \int_{k/n}^{(k+1)/n} \overline{g(t)}\,dt.$$

Utilisons la continuité uniforme de g. Etant donné $\varepsilon > 0$, il existe un réel $\delta > 0$ tel que $|g(t) - g(s)| < \varepsilon$ si $|t - s| < \delta$. On prend n_0 tel que $n_0 > \delta^{-1}$. On voit que si $n \ge n_0$

$$\Bigl|\int_{k/n}^{(k+1)/n} g(t)\,dt - \int_{(k+1)/n}^{(k+2)/n} g(t)\,dt\Bigr| \le \frac{\varepsilon}{n}$$

et on en déduit en regroupant dans $(*)$ les morceaux d'intégrale deux par deux que $|(f_n, g)| \le \varepsilon + \|g\|_\infty/n$.

Remarque. En dimension finie, convergence faible et convergence en norme sont identiques.

Théorème. *Si* H *est un espace de Hilbert, toute suite bornée de* H *admet des sous-suites faiblement convergentes.*

Démonstration. On commence par le cas où H est séparable, et de dimension infinie (sinon, le résultat voulu est donné par le théorème de Bolzano-Weierstrass). Pour exprimer la démonstration, il est utile d'introduire une petite convention de notation. Si $\mathrm{M} = \{n_0 < \ldots < n_j < \ldots\}$ est un sous-ensemble infini de $\mathbb{N}$, convenons de noter la sous-suite (x_{n_j}) par $(x_n)_{n\in\mathrm{M}}$. Soit donc (e_k) une base hilbertienne de H, et (x_n) une suite bornée dans H, telle que par exemple $\|x_n\| \le 1$ pour tout entier $n \ge 0$; la suite de scalaires $(\langle x_n, e_0\rangle)$ est bornée, donc elle admet une sous-suite convergente $(\langle x_n, e_0\rangle)_{n\in\mathrm{M}_0}$, vers une limite $c_0 \in \mathbb{K}$. La suite $(\langle x_n, e_1\rangle)_{n\in\mathrm{M}_0}$ est encore bornée, donc on peut trouver un nouvel ensemble infini $\mathrm{M}_1 \subset \mathrm{M}_0$ tel que la sous-suite $(\langle x_n, e_1\rangle)_{n\in\mathrm{M}_1}$ soit convergente vers un scalaire c_1. En continuant ainsi, on construit une suite décroissante $\mathrm{M}_0 \supset \mathrm{M}_1 \supset \ldots \supset \mathrm{M}_j \supset \ldots$ telle que $(\langle x_n, e_j\rangle)_{n\in\mathrm{M}_j}$ soit convergente vers un scalaire c_j, pour tout $j \ge 0$.

On note que $\sum |c_j|^2 \le 1$. En effet, pour tout k, on remarque que $(\langle x_n, e_j\rangle)_{n\in\mathrm{M}_k}$ converge vers c_j pour tout $j \le k$, donc

$$\sum_{j=0}^{k} |c_j|^2 = \lim_{n\in\mathrm{M}_k} \sum_{j=0}^{k} |\langle x_n, e_j\rangle|^2 \le \limsup_n \|x_n\|^2 \le 1.$$

Il existe donc un vecteur $x \in \mathrm{H}$, à savoir $x = \sum_{k=0}^{+\infty} c_k\, e_k$, tel que $c_k = \langle x, e_k\rangle$ pour tout $k \ge 0$.

C'est ici qu'intervient le procédé de *la suite diagonale*. Construisons un ensemble infini M formé du premier élément n_0 de M_0, puis du premier élément n_1 de M_1 qui soit $> n_0$, etc... On constate que pour tout entier $k \ge 0$, la sous-suite $(\langle x_n, e_k\rangle)_{n\in\mathrm{M}}$ est convergente vers $c_k = \langle x, e_k\rangle$: en effet, l'ensemble M est contenu dans M_k à un ensemble fini près, pour tout $k \geq 0$. Puisque (x_n) est bornée, la convergence sur l'ensemble total (e_k) suffit à prouver la convergence faible : on a donc trouvé une sous-suite $(x_n)_{n \in \mathrm{M}}$ qui converge faiblement vers x.

Pour finir on suppose que H est un espace de Hilbert arbitraire. On désigne par H_0 le sous-espace vectoriel fermé engendré par la suite (x_n) ; c'est un sous-espace séparable, donc on peut trouver d'après la première partie une sous-suite (x_{n_j}) et un vecteur $x \in \mathrm{H}_0$ tels que

$$\lim_j \langle x_{n_j}, y' \rangle = \langle x, y' \rangle$$

pour tout $y' \in \mathrm{H}_0$. Soit maintenant $y \in \mathrm{H}$ quelconque ; on peut écrire $y = y' + y''$, avec $y' \in \mathrm{H}_0$ et y'' orthogonal à H_0. On aura $\langle x_n, y \rangle = \langle x_n, y' \rangle$ puisque y'' est orthogonal à x_n pour tout n, et le résultat en découle,

$$\lim_j \langle x_{n_j}, y \rangle = \lim_j \langle x_{n_j}, y' \rangle = \langle x, y' \rangle = \langle x, y \rangle.$$

Proposition. *Si* (C_n) *est une suite décroissante de convexes fermés bornés non vides d'un espace de Hilbert* H, *l'intersection* $\bigcap_n \mathrm{C}_n$ *est non vide.*

Démonstration. Puisque chaque ensemble C_n est non vide, on peut choisir $x_n \in \mathrm{C}_n$ pour tout n ; puisque les C_n sont décroissants, toute la suite (x_n) est dans l'ensemble borné C_0, donc la suite est bornée ; on peut alors trouver une sous-suite (x_{n_j}) faiblement convergente vers un vecteur $x \in \mathrm{H}$. On montre que x est dans tous les C_n. Pour chaque k fixé, on aura $n_j \geq k$ pour j assez grand, donc tous les termes de la sous-suite sont dans C_k pour j assez grand ; puisque C_k est convexe fermé, il en résulte que la limite faible x reste dans C_k ; ceci étant vrai pour tout $k \geq 0$, on en déduit que $x \in \bigcap_n \mathrm{C}_n$.

Corollaire. *Si* C *est un convexe fermé borné non vide d'un espace de Hilbert et si* f *est une fonction convexe continue sur* C, *elle atteint son minimum sur* C.

MT404, Cours n$^{\text{o}}$ 13, lundi 5 novembre 2001.

Rappels

Si $\mathrm{H}_1, \mathrm{H}_2$ sont deux espaces de Hilbert et $\mathrm{T} \in \mathcal{L}(\mathrm{H}_1, \mathrm{H}_2)$, l'opérateur linéaire borné *adjoint* $\mathrm{T}^* \in \mathcal{L}(\mathrm{H}_2, \mathrm{H}_1)$ est caractérisé par

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

Un opérateur linéaire borné $\mathrm{A} \in \mathcal{L}(\mathrm{H})$ est hermitien si et seulement si

$$\forall x, y \in \mathrm{H}, \quad \langle \mathrm{A}x, y\rangle = \langle x, \mathrm{A}y\rangle.$$

En effet, cette formule résume le fait que $\mathrm{A}^* = \mathrm{A}$.

Rappel : convergence faible

Soit X un espace normé ; une suite $(x_n) \subset \mathrm{X}$ converge faiblement vers $x \in \mathrm{X}$ si

$$x^*(x) = \lim_n x^*(x_n)$$

pour toute forme linéaire continue $x^* \in \mathrm{X}^*$. Dans le cas où X est un espace de Hilbert H, cela revient à dire que

$$\langle x, y\rangle = \lim_n \langle x_n, y\rangle$$

pour tout $y \in \mathrm{H}$.

Remarques.

— Remarque évidente ; la convergence forte d'une suite implique la convergence faible : si $\|x_n - x\|$ tend vers 0, il en résulte que $(x^*(x_n))$ tend vers $x^*(x)$, puisque x^* est une fonction continue sur X.

— Unicité de la limite faible : si x et x' sont deux candidats pour être limite faible d'une même suite $(x_n) \subset \mathrm{X}$, on a $x^*(x) = \lim_n x^*(x_n) = x^*(x')$ pour tout $x^* \in \mathrm{X}^*$, donc $x^*(x - x') = 0$ pour tout x^*, donc $x - x' = 0$ d'après une des formes de Hahn-Banach.

– Image linéaire continue d'une convergence faible. Si $(x_n) \subset \mathrm{X}$ converge faiblement vers $x \in \mathrm{X}$ et si $\mathrm{T} \in \mathcal{L}(\mathrm{X}, \mathrm{Y})$, alors $(\mathrm{T}(x_n))$ tend faiblement vers $\mathrm{T}(x)$: en effet, pour toute forme linéaire $y^* \in \mathrm{Y}^*$, on écrira

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

Rappel : si C est un convexe fermé d'un espace normé X, si $(x_n) \subset \mathrm{C}$ converge faiblement vers $x \in \mathrm{X}$, alors $x \in \mathrm{C}$.

Ce résultat découle du théorème de séparation suivant.

Théorème : séparation d'un convexe fermé et d'un point. *Si* C *est un convexe fermé non vide d'un espace normé réel* X *et si* $x \notin \mathrm{C}$, *il existe une forme linéaire continue* $x^* \in \mathrm{X}^*$ *telle que* $\sup x^*(\mathrm{C}) < x^*(x)$.

Démonstration. L'ensemble C est fermé et $x \notin \mathrm{C}$, donc $\varepsilon = \mathrm{dist}(x, \mathrm{C}) > 0$. On montre que la fonction $y \to \mathrm{dist}(y, \mathrm{C})$ est une fonction convexe, lipschitzienne donc continue sur X. On considère l'ensemble

$$\mathrm{A} = \{y \in \mathrm{X} : \mathrm{dist}(y, \mathrm{C}) < \varepsilon\}.$$

Cet ensemble est convexe ouvert, et $x \notin \mathrm{A}$, donc il existe une forme linéaire x^* telle que $x^*(a) < x^*(x)$ pour tout $a \in \mathrm{A}$. En particulier x^* est non nulle et on peut trouver un vecteur v tel que $\|v\| < 1$ et $x^*(v) > 0$. Pour tout $y \in \mathrm{C}$, on aura $y + \varepsilon v \in \mathrm{A}$, donc $x^*(y) \le x^*(x) - \varepsilon x^*(v)$. En laissant v fixé et en faisant varier y dans C on déduit

$$\sup x^*(\mathrm{C}) = \sup\{x^*(y) : y \in \mathrm{C}\} \le x^*(x) - \varepsilon\, x^*(v) < x^*(x).$$

Extraction de sous-suites faiblement convergentes

On rappelle que *toute suite bornée dans un espace de Hilbert admet des sous-suites faiblement convergentes.*

On va en tirer quelques conséquences.

Image de la boule unité de H *par un opérateur linéaire*

Proposition. *Soit* T *une application linéaire continue d'un espace de Hilbert* H *dans un espace normé* Y *; l'image* $\mathrm{T}(\mathrm{B_H})$ *de la boule unité fermée de* H *est fermée dans* Y.

Démonstration. Soit $y \in \mathrm{Y}$ un point adhérent à l'image de la boule ; il existe donc une suite $(x_n) \subset \mathrm{B_H}$ telle que $\|y - \mathrm{T}(x_n)\| \to 0$; il existe par ailleurs une sous-suite (x_{n_k}) qui converge faiblement vers $x \in \mathrm{B_H}$. Alors $(\mathrm{T}(x_{n_k}))$ converge faiblement vers $\mathrm{T}(x)$, et fortement, donc aussi faiblement, vers y, donc $y = \mathrm{T}(x) \in \mathrm{T}(\mathrm{B_E})$.

Compacité de V_2

L'ensemble C des fonctions continues F sur $[0,1]$ telles qu'il existe $f \in \mathrm{L}_2(0,1)$ avec $\|f\|_2 \le 1$ et

$$\forall x \in [0,1], \quad \mathrm{F}(x) = \int_0^x f(t)\,dt = \langle f, \mathbf{1}_{[0,x]} \rangle$$

est compact pour la norme L_2.

On voit que $\mathrm{F} = \mathrm{V}_2(f)$; l'ensemble C est donc égal à l'image $\mathrm{V}_2(\mathrm{B}_{\mathrm{L}_2})$, et il est par conséquent fermé dans $\mathrm{L}_2(0,1)$.

Soit (F_n) une suite dans cet ensemble, écrivons $\mathrm{F}_n = \mathrm{V}(f_n)$; soit (f_{n_k}) une sous-suite faiblement convergente vers une f de la boule unité de $\mathrm{L}_2(0,1)$; alors $\mathrm{V}(f_{n_k})$ converge simplement vers $\mathrm{V}(f)$ sur $[0,1]$, en étant dominée par 1, donc $|\mathrm{V}(f_{n_k}) - \mathrm{V}(f)|^2$ converge simplement vers 0 en étant dominée par la fonction constante intégrable 4, et Lebesgue donne la convergence L_2 vers $\mathrm{V}_2(f)$.

Exercice proposé. Montrer la compacité de $\mathrm{T_K}(\mathrm{B}_{\mathrm{L}_2})$, lorsque K est un noyau mesurable borné défini sur $[0,1]^2$.

Diagonalisation des hermitiens compacts

Définition. Soient E et F deux espaces de Banach ; un opérateur $\mathrm{T} \in \mathcal{L}(\mathrm{E},\mathrm{F})$ est dit *compact* si l'adhérence de l'image $\mathrm{T}(\mathrm{B_E})$ de la boule unité de E est compacte dans F.

D'après ce qui précède, on n'a pas besoin de l'adhérence lorsque E = H est un espace de Hilbert : un opérateur $\mathrm{T} \in \mathcal{L}(\mathrm{H},\mathrm{F})$ est compact si et seulement si $\mathrm{T}(\mathrm{B_H})$ est compacte dans F.

C'est ce qu'on a vu dans l'exemple de V_2.

Lemme. *Si* $\mathrm{B} \in \mathcal{L}(\mathrm{H})$ *est hermitien positif on a*

$$\forall x \in \mathrm{H}, \quad \|\mathrm{B}x\|^2 \le \|\mathrm{B}\|\,\langle \mathrm{B}x, x\rangle.$$

Démonstration. Posons $\varphi(x,y) = \langle \mathrm{B}x, y\rangle$ pour tous $x, y \in \mathrm{H}$; on définit de cette façon un produit scalaire sur H, auquel on peut appliquer l'inégalité de Cauchy-Schwarz

$$|\langle \mathrm{B}x, y\rangle|^2 \le \langle \mathrm{B}x, x\rangle\,\langle \mathrm{B}y, y\rangle.$$

Si $\|y\| \le 1$, on aura $\langle \mathrm{B}y, y\rangle \le \|\mathrm{B}y\|\,\|y\| \le \|\mathrm{B}\|\,\|y\|^2 \le \|\mathrm{B}\|$, donc

$$|\langle \mathrm{B}x, y\rangle|^2 \le \|\mathrm{B}\|\,\langle \mathrm{B}x, x\rangle$$

pour tout vecteur y de norme ≤ 1. En prenant le sup sur les y de norme 1 on obtient le résultat, puisque

$$\|v\| = \sup\{|\langle v, y\rangle| : \|y\| \le 1\}$$

pour tout vecteur $v \in \mathrm{H}$ (c'est évident si $v = 0_\mathrm{H}$; dans le cas contraire, choisir le vecteur $y = \|v\|^{-1} v$).

Lemme. *Si* $\mathrm{A} \in \mathcal{L}(\mathrm{H})$ *est hermitien compact et non nul, il admet au moins un vecteur propre, avec valeur propre réelle non nulle.*

Démonstration. Supposons A non nul ; par polarisation, on sait qu'il existe x tel que $\langle \mathrm{A}x, x\rangle \ne 0$, par exemple $\langle \mathrm{A}x, x\rangle > 0$; (dans le cas contraire on raisonnera sur $-\mathrm{A}$) ; posons dans ce cas

$$\sup\{\langle \mathrm{A}x, x\rangle : \|x\| \le 1\} = r > 0.$$

On va montrer que r est valeur propre de A. Par homogénéité on déduit de la définition de r que $\langle \mathrm{A}y, y\rangle \le r\,\|y\|^2$ pour tout vecteur $y \in \mathrm{H}$. On voit alors que l'opérateur hermitien $\mathrm{B} = r\,\mathrm{Id} - \mathrm{A}$ est positif, puisque pour tout $y \in \mathrm{H}$ on a

$$\langle \mathrm{B}y, y\rangle = \langle ry - \mathrm{A}y, y\rangle = r\|y\|^2 - \langle \mathrm{A}y, y\rangle \ge 0.$$

Par définition de r, il existe une suite $(x_n) \subset \mathrm{B_H}$ telle que $\langle \mathrm{A}x_n, x_n\rangle$ tende vers r ; ceci implique que $\lim_n \|x_n\| = 1$, puisque

$$0 \le r\,\|x_n\|^2 - \langle \mathrm{A}x_n, x_n\rangle \le r - \langle \mathrm{A}x_n, x_n\rangle$$

tend vers 0, ce qui entraîne que $r(1-\|x_n\|^2)$ tend vers 0. On voit donc que $\langle \mathrm{B}x_n, x_n\rangle$ tend vers 0 ; il en résulte que $\mathrm{B}x_n = rx_n - \mathrm{A}x_n$ tend vers 0, puisque $\|\mathrm{B}x_n\|^2 \le \|\mathrm{B}\|\,\langle \mathrm{B}x_n, x_n\rangle$; puisque A est compact on peut supposer, en passant à une sous-suite, que $\mathrm{A}x_n$ converge vers un vecteur $y \in \mathrm{H}$; il en résulte que rx_n tend aussi vers y, donc x_n converge vers $x = r^{-1}y$ et $\mathrm{A}x - rx = 0$ à la limite. De plus, $\|x\| = \lim_n \|x_n\| = 1$ garantit que x n'est pas nul.

Théorème. *Soit* H *un espace de Hilbert séparable, réel ou complexe, et soit* $\mathrm{A} \in \mathcal{L}(\mathrm{H})$ *hermitien et compact ;*

– *les valeurs propres de* A *sont réelles ;*
– *les sous-espaces propres sont deux à deux orthogonaux ;*
– *le sous-espace propre correspondant à chaque valeur propre* $\lambda \ne 0$ *est de dimension finie ;*
– *on peut ranger les valeurs propres non nulles dans une suite qui tend vers* 0*, à moins qu'il n'y ait qu'un nombre fini de valeurs propres non nulles ;*
– *il existe une base hilbertienne de* H *formée de vecteurs propres de* A.

Démonstration. Les deux premières informations se démontrent comme en DEUG pour les matrices symétriques réelles. Si $\mathrm{A}x = \lambda x$, avec $x \ne 0_\mathrm{H}$, on aura

$$\lambda\,\langle x, x\rangle = \langle \mathrm{A}x, x\rangle = \langle x, \mathrm{A}x\rangle = \langle x, \lambda x\rangle = \overline{\lambda}\,\langle x, x\rangle$$

ce qui implique $\lambda = \overline{\lambda}$ puisque $\langle x, x\rangle > 0$. Si $\mathrm{A}x = \lambda x$ et $\mathrm{A}y = \mu y$ avec $\lambda \ne \mu$, on a

$$\lambda\,\langle x, y\rangle = \langle \mathrm{A}x, y\rangle = \langle x, \mathrm{A}y\rangle = \langle x, \mu y\rangle = \overline{\mu}\,\langle x, y\rangle = \mu\,\langle x, y\rangle$$

puisque μ est réelle, et ceci entraîne $\langle x, y\rangle = 0$.

Si la dimension du sous-espace propre E_r était infinie, on raisonnerait ainsi : soit (e_n) une suite infinie orthonormée dans le sous-espace propre ; la suite $(\mathrm{A}e_n)$ doit avoir des sous-suites convergentes en norme, puisque A est compact, mais $\mathrm{A}e_n - \mathrm{A}e_m = r(e_n - e_m)$ est de norme $r\sqrt{2}$ pour tous $m \neq n$, contradiction.

Cours n$^\text{o}$ 14, mercredi 7 novembre 2001.

Exercice traité. Si C est un convexe non vide dans un espace normé X, la fonction f définie par $f(y) = \mathrm{dist}(y, \mathrm{C})$ est une fonction convexe sur X.

Supposons que $x_0, x_1 \in \mathrm{X}$ et $0 \le t \le 1$. Considérons le point $x_t = (1-t)x_0 + tx_1$; nous devons vérifier que $f(x_t) \le (1-t)f(x_0) + tf(x_1)$.

On choisit $y_i \in \mathrm{C}$, $i = 0, 1$ tel que $\|x_i - y_i\| \le \mathrm{dist}(x_i, \mathrm{C}) + \varepsilon$; puisque C est convexe, $y_t = (1-t)\, y_0 + t\, y_1 \in \mathrm{C}$, donc

$$f(x_t) \le \|x_t - y_t\| = \|(1-t)\,(x_0 - y_0) + t\,(x_1 - y_1)\| \le$$

$$(1-t)\,\|x_0 - y_0\| + t\,\|x_1 - y_1\| \le (1-t)\, f(x_0) + t\, f(x_1) + \varepsilon.$$

Démonstration du théorème de diagonalisation. On suppose A hermitien compact sur H, Hilbert réel ou complexe. On va trouver un sous-espace fermé $\mathrm{F} \subset \mathrm{H}$, engendré par une suite orthonormée de vecteurs propres de A (peut-être infinie, peut-être finie, ou même vide), tel que la restriction de A à $\mathrm{F}^\perp$ soit nulle.

On procède par récurrence. On commence avec $\mathrm{F}_0 = \{0\}$.

On suppose que F_{k-1} est de dimension $k-1$, engendré par $k-1$ vecteurs propres $e_1, \ldots, e_{k-1}$ de A, deux à deux orthogonaux, pour lesquels $\mathrm{A}(e_j) = \mu_j\, e_j$, $j = 1, \ldots, k-1$. On sait que chaque valeur propre μ_j est réelle (lorsque $k = 1$, l'espace $\mathrm{F}_0 = \{0\}$ est engendré par $k - 1 = 0$ vecteurs).

On pose $\mathrm{H}_k = \mathrm{F}_{k-1}^\perp$. Comme $\mathrm{A}(\mathrm{F}_{k-1}) \subset \mathrm{F}_{k-1}$ clairement, et que A est hermitien, il en résulte que H_k est stable par A et on peut considérer la restriction A_k de A à H_k ; alors $\mathrm{A}_k \in \mathcal{L}(\mathrm{H}_k)$ est hermitien et compact.

— Ou bien A_k est nul et le processus s'arrête ; on pose dans ce cas

$$\mathrm{F} = \mathrm{F}_{k-1} = \mathrm{Vect}(e_1, \ldots, e_{k-1})\,;$$

cet espace admet une base orthonormée formée de vecteurs propres de A, et la restriction de A à $\mathrm{F}^\perp$ est nulle.

— ou bien A_k n'est pas nul ; on sait alors par le lemme précédent qu'il existe des valeurs propres non nulles pour A_k. On désigne par ν_k le sup des $|\lambda|$, pour toutes les valeurs propres λ de A_k ; on a $\nu_k > 0$; on choisit une valeur propre $\mu_k = \lambda$ de A_k telle que $|\mu_k| > \frac{1}{2}\,\nu_k$, et un vecteur propre $e_k \in \mathrm{H}_k$ tel que $\mathrm{A}(e_k) = \mu_k\, e_k$ et $\|e_k\| = 1$. Puisque $e_k \in \mathrm{H}_k$, il est orthogonal aux vecteurs déjà choisis $e_1, \ldots, e_{k-1}$.

Si le processus continue jusqu'à l'infini, on a construit une suite orthonormée $(e_k)_{k\ge1}$ de vecteurs propres de A, et on va montrer maintenant que $\lim_k \mu_k = 0$. On montre en fait que $\lim_k \nu_k = 0$. Par construction, la suite (ν_k) est décroissante positive ; si elle reste $\ge \delta > 0$, on aura choisi une suite orthonormée infinie $(e_k)_{k\ge1}$ de vecteurs propres, telle que $\mathrm{A}(e_k) = \mu_k e_k$ avec $|\mu_k| \ge \frac{1}{2}\,\nu_k \ge \delta/2$ pour tout $k \ge 1$; ceci contredit la compacité de A. En effet, pour tout $m \neq n$,

$$\|\mathrm{A}e_m - \mathrm{A}e_n\|^2 = \|\mu_m e_m - \mu_n e_n\|^2 = |\mu_m|^2 + |\mu_n|^2 \ge \delta^2/2 > 0$$

montre que la suite $(\mathrm{A}e_n)$ ne peut pas avoir de sous-suite de Cauchy ; ceci est impossible puisque cette suite est contenue dans le compact $\mathrm{A}(\mathrm{B}_\mathrm{H})$.

Considérons maintenant le sous-espace fermé F engendré par la suite orthonormée $(e_k)_{k\in\mathbb{N}^*}$; on a encore $\mathrm{A}(\mathrm{F}) \subset \mathrm{F}$ (écrire un vecteur y quelconque de F dans la base hilbertienne $(e_k)_{k\geq 1}$ de F, et exprimer $\mathrm{A}(y)$), donc $\mathrm{H}_\infty = \mathrm{F}^\perp$ est stable par A ; désignons par A_∞ la restriction de A à F_∞. Puisque $\mathrm{F} \supset \mathrm{F}_{k-1}$ pour tout $k \geq 1$, on a $\mathrm{H}_\infty \subset \mathrm{H}_k$ pour tout k, et on en déduit que le sup des valeurs absolues des valeurs propres de A_∞ est $\leq \nu_k$ pour tout k, donc il est nul, donc A_∞ est nul.

Dans le cas où le processus s'arrête, comme dans le cas où il y a une infinité de pas, on a trouvé un sous-espace fermé F avec une base orthonormée (vide, ou finie, ou dénombrable) formée de vecteurs propres de A, et tel que A soit nul sur $\mathrm{F}^\perp$. Pour terminer il suffit de compléter, s'il y a lieu, la suite orthonormée (e_k) par une base orthonormée de $\mathrm{F}^\perp$ pour obtenir le résultat voulu.

Exemple de $\mathrm{V}^*\mathrm{V}$ *; calcul de la norme de* V_2

On a répété plusieurs fois que V_2 envoie $\mathrm{H} = \mathrm{L}_2(0,1)$ dans les fonctions continues ; voici une preuve rapide : si $0 \leq x < y \leq 1$,

$$\big|(\mathrm{V}f)(y) - (\mathrm{V}f)(x)\big| = \Big|\int_x^y f(t)\,dt\Big| = \Big|\langle f, \mathbf{1}_{(x,y)}\rangle\Big| \leq \|f\|_2\,\|\mathbf{1}_{(x,y)}\|_2 = \|f\|\,\sqrt{|y-x|}.$$

La même preuve s'applique à V_2^*.

Si f est une fonction continue, on sait depuis le DEUG que la fonction F définie par $\mathrm{F}(x) = (\mathrm{V}f)(x) = \int_0^x f(t)\,dt$ est une fonction de classe C^1 dont la dérivée est $\mathrm{F}'(x) = f(x)$. De même, $\mathrm{G}(x) = (\mathrm{V}^*f)(x) = \int_x^1 f(t)\,dt$ est dérivable, mais cette fois $\mathrm{G}'(x) = -f(x)$.

On a vu que V_2 et V_2^* sont injectifs, donc $\mathrm{A} = \mathrm{V}^*\mathrm{V}$ est hermitien, et injectif. Par ailleurs, $\mathrm{A}(\mathrm{B}_\mathrm{H}) = \mathrm{V}^*(\mathrm{V}(\mathrm{B}_\mathrm{H}))$ est l'image du compact $\mathrm{V}_2(\mathrm{B}_\mathrm{H})$ par l'application continue V^*, donc c'est un compact de H.

Puisque A est hermitien compact injectif, sa diagonalisation ne comprend que des valeurs propres $\neq 0$. Par ailleurs, les valeurs propres sont ≥ 0 parce que A est positif ; finalement, les valeurs propres de A sont des réels $\lambda > 0$.

Supposons que $f \in \mathrm{H}$ et $\mathrm{A}f = \lambda f$. Alors $f = \lambda^{-1}\mathrm{V}^*(\mathrm{V}(f))$ montre que f est continue, parce que $\mathrm{V}^*(g)$ est continue pour toute $g \in \mathrm{H}$. Alors $\mathrm{V}f$ est C^1, et on en déduit que $\mathrm{V}^*(\mathrm{V}f)$ est de classe C^2, donc f elle-même est de classe C^2. On a d'abord $f' = -\lambda^{-1}\mathrm{V}(f)$, puis $f'' = -\lambda^{-1}f$. Posons $\lambda^{-1} = \omega^2$ pour un $\omega > 0$. On sait que les solutions sont de la forme $f(x) = \mathrm{A}\cos(\omega x) + \mathrm{B}\sin(\omega x)$.

La relation $f = \lambda^{-1}\mathrm{V}^*(\mathrm{V}(f))$ montre que $f(1) = 0$, et $f' = -\lambda^{-1}\mathrm{V}(f)$ montre que $f'(0) = 0$. Puisque $f'(x) = -\mathrm{A}\,\omega\sin(\omega x) + \mathrm{B}\,\omega\cos(\omega x)$, on voit que $f'(0) = \mathrm{B}\,\omega = 0$, donc $\mathrm{B} = 0$. On cherche donc f de la forme $f(x) = \cos(\omega x)$, et $f(1) = 0$ donne $\cos(\omega) = 0$.

Résumons : si λ est valeur propre de A, alors $\lambda > 0$, et si on pose $\omega^2 = \lambda^{-1}$ on aura $\cos(\omega) = 0$. Les valeurs de ω sont donc de la forme $\omega_k = \pi/2 + k\,\pi$, avec $k = 0, 1, \ldots$. Réciproquement, on vérifie que chaque fonction $f_k(x) = \cos(\omega_k x)$ est vecteur propre de A, avec valeur propre $\lambda_k = \omega_k^{-2}$.

On a donc trouvé toutes les valeurs propres et directions propres possibles pour l'opérateur A. D'après le théorème de diagonalisation, on sait que les fonctions $(f_k)_{k\geq 0}$, une fois normalisées, donnent une base hilbertienne de H.

On en déduit que $\|A\|$ est égale à la plus grande valeur propre, à savoir $\lambda_0 = 4/\pi^2$. On en déduit $\|V_2\| = 2/\pi$, puisque $\|A\| = \|V\|^2$.

Le cas normal compact

Ici il **faut** supposer que $\mathbb{K} = \mathbb{C}$.

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

Si $S \in \mathcal{L}(H)$ est normal, on note que

$$\|Sx\|^2 = \langle Sx, Sx\rangle = \langle S^*Sx, x\rangle = \langle SS^*x, x\rangle = \langle S^*x, S^*x\rangle = \|S^*x\|^2,$$

ce qui montre en particulier que $\ker S = \ker S^*$. Si T est normal, si on applique à l'opérateur normal $S = T - \lambda \,\mathrm{Id}_H$, alors $S^* = T^* - \overline{\lambda}\,\mathrm{Id}_H$ ce qui montre que le sous-espace propre de T pour la valeur propre λ coïncide avec le sous-espace propre de T^* pour la valeur propre $\overline{\lambda}$.

Lemme. *Si* H *est un espace de Hilbert complexe,* $T \in \mathcal{L}(H)$ *un opérateur normal compact et non nul, il existe une valeur propre non nulle pour* T.

Démonstration. L'opérateur $A = T^*T = TT^*$ est hermitien et compact, non nul parce que $\|A\| = \|T\|^2 > 0$. D'après le cas hermitien compact, il existe une valeur propre non nulle μ pour A ; comme T commute avec A, le sous-espace propre $E_\mu = \ker(A - \mu\,\mathrm{Id}_H)$ de A est stable par T ; mais on a vu que les sous-espaces propres de A, pour les valeurs propres non nulles, sont de dimension finie. La restriction T_1 de T à E_μ est donc un endomorphisme d'un espace complexe de dimension finie, donc T_1 admet au moins une valeur propre λ.

Il existe donc $x \in E_\mu$, non nul, tel que $Tx = \lambda x$; comme T est normal, on a $T^*x = \overline{\lambda}x$, donc $\mu x = Ax = |\lambda|^2 x$, ce qui prouve que $|\lambda|^2 = \mu \neq 0$, donc $\lambda \neq 0$.

On peut maintenant reprendre la démonstration de la diagonalisation qui a été donnée dans le cas hermitien ; on produit de proche en proche des vecteurs propres $e_1, \ldots, e_{k-1}$ de T, on pose $F_{k-1} = \mathrm{Vect}(e_1, \ldots, e_{k-1})$; il faut juste un mot de plus pour justifier que $H_k = F_{k-1}^\perp$ est stable par T.

Si $U \in \mathcal{L}(H)$ et si $U(F) \subset F$, alors $U^*(F^\perp) \subset F^\perp$: si y est orthogonal à F et $x \in F$,

$$\langle U^*y, x\rangle = \langle y, Ux\rangle = 0$$

puisque $Ux \in F$. Il en résulte aussi que si $U^*(F) \subset F$, alors $U(F^\perp) \subset F^\perp$.

On voit que F_{k-1} est stable par T et T^*, parce que les vecteurs $e_1, \ldots, e_{k-1}$ sont propres pour T et pour T^*. Il en résulte que H_k est stable par T^* et T, ce qui permet de considérer la restriction T_k de T à H_k ; c'est à nouveau un opérateur normal et compact.

Aperçu du cas banachique pour la convergence faible ou $*$-faible

On a déjà défini la convergence faible des suites dans le cas des espaces normés ; on a montré que les convexes fermés sont "fermés pour les suites faiblement convergentes".

Il existe une autre notion, qu'on ne peut pas distinguer de la précédente dans le cas hilbertien, la *convergence $*$-faible* des suites de formes linéaires continues : si (x_n^*) est une suite de formes linéaires continues sur un espace normé X, on dit qu'elle converge $*$-faiblement vers $x^* \in X^*$ si

$$\forall x \in X, \quad x^*(x) = \lim_n x_n^*(x).$$

Le théorème d'extraction de sous-suites trouve sa généralisation avec cette notion de convergence $*$-faible.

Proposition 5.3.4. *Si* X *est un espace normé séparable, toute suite bornée de* X^* *admet des sous-suites* $*$*-faiblement convergentes.*

Dans le cas réflexif, on retrouve la situation des espaces de Hilbert, où il y a mélange entre les points de vue vecteur et forme linéaire :

Théorème 5.3.5. *Si* X *est réflexif, toute suite bornée dans* X *admet des sous-suites faiblement convergentes.*

5.2. Topologie faible sur un espace normé

On essaie de définir une topologie $\mathcal{T}$ sur un espace normé X, en y mettant *le moins possible* d'ouverts, mais de façon que toutes les applications $x \to x^*(x)$, pour $x^* \in \mathrm{X}^*$, restent continues pour $\mathcal{T}$. On supposera X réel pour l'instant, pour simplifier la description.

Si $\mathcal{T}$ vérifie cette condition et si $x^* \in \mathrm{X}^*$, $a, b \in \mathbb{R}$, $a < b$, alors l'ensemble

$$\{x \in \mathrm{X} : a < x^*(x) < b\}$$

doit être un ouvert de $\mathcal{T}$, puisque nous voulons que x^* soit $\mathcal{T}$-continue. Puisque toute intersection finie d'ouverts est un ouvert, la famille d'ouverts $\mathcal{T}$ doit contenir tout ensemble B de la forme

$$\mathrm{B} = \bigcap_{i=1}^{k} \{x \in \mathrm{X} : a_i < x_i^*(x) < b_i\}$$

pour tout entier $k \geq 1$, toutes suites finies $x_1^*, \ldots, x_k^*$ dans X^* et $a_1, b_1, \ldots, a_k, b_k$ dans $\mathbb{R}$.

Pour finir $\mathcal{T}$ doit contenir toutes les réunions $\bigcup_{i \in \mathrm{I}} \mathrm{B}_i$ d'ensembles B_i de la forme précédente. Mais la famille de ces réunions est une topologie sur X ; c'est la topologie voulue, *la topologie la moins fine qui rend continues* toutes les applications $x \to x^*(x)$, quand x^* décrit X^*.

La topologie $\mathcal{T}$ sur X s'appelle la *topologie faible* sur X. La topologie faible est plus faible que la topologie de la norme.

Soit X un espace vectoriel normé ; on va maintenant s'intéresser à une topologie faible sur le dual X^*, la *topologie* $*$*-faible* sur X^*. C'est la topologie la moins fine sur X^* rendant continues toutes les applications $x^* \in \mathrm{X}^* \to x^*(x)$, où x décrit X.

Théorème 5.2.3. *Munie de la topologie* $*$*-faible, la boule unité de* X^* *est compacte.*

Ce théorème, appelé théorème de *Banach-Alaoglu*, peut s'obtenir comme un corollaire du théorème de Tykhonov.

Théorème 5.2.4 : théorème de Tykhonov. *Tout produit d'espaces compacts (muni de la topologie produit) est compact.*

MT404, Cours n° 15, lundi 12 novembre 2001.

7. Algèbres de Banach et théorie spectrale

Un certain nombre de résultats de ce chapitre et des suivants n'a 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 aussi bien dans le cas réel que complexe.

7.1. 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 dans A 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 $a, b \in \mathrm{A}$. On en déduit immédiatement que l'application $(a, b) \to ab$ est 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. En fait, si $a_n \to a$ et $b_n \to b$, il en résulte que $a_n b_n \to ab$.

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 7.1.1.

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 $\mathrm{E} \ne \{0\}$.

Un cas particulier très important est celui où E est un espace de Hilbert. Dans ce cas on dispose d'une opération supplémentaire, l'opération $\mathrm{T} \to \mathrm{T}^*$ d'adjoint. On appelle *C^*-algèbre* une algèbre de Banach unitaire complexe A muni d'une involution antilinéaire $a \in \mathrm{A} \to a^* \in \mathrm{A}$ telle que $\|a^* a\| = \|a\|^2$ pour tout $a \in \mathrm{A}$ et $(ab)^* = b^* a^*$.

2. Soit K un espace compact non vide ; considérons l'espace de Banach $\mathrm{A} = \mathrm{C}(\mathrm{K})$ des fonctions continues sur K à valeurs complexes, muni du produit usuel et de la norme de convergence uniforme (exemples 1.1.6) ; c'est une algèbre de Banach unitaire. L'élément 1_A est la fonction constante égale à 1. Cet exemple donne une algèbre commutative. C'est aussi une C^*-algèbre, si on définit $f^* = \overline{f}$, la fonction complexe conjuguée.

Définition 7.1.2. 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 $ab = ba = 1_\mathrm{A}$.

Il suffit d'avoir $b_1, b_2 \in \mathrm{A}$ tels que $b_1 a = ab_2 = 1_\mathrm{A}$ pour en déduire que $b_1 = b_1(ab_2) = (b_1 a) b_2 = b_2$ est bien l'inverse de a.

Exemples 7.1.3.

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.

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.

Lemme 7.1.1. *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

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

Si on sait simplement que la série $\sum a^k$ *converge dans* A*, sa somme* $\sum_{k=0}^{+\infty} a^k$ *est l'inverse de* $1_A - a$*.*

Démonstration. Commençons par la remarque finale. Si $\sum a^n$ converge dans A, posons $S = \sum_{n=0}^{+\infty} a^n$. Puisque l'opération G_a de multiplication à gauche par a, définie par $G_a(x) = ax$ pour tout $x \in A$ est linéaire et continue de A dans A, on en déduit que

$$aS = G_a(S) = \sum_{n=0}^{+\infty} G_a(a^n) = \sum_{n=0}^{+\infty} a^{n+1} = S - 1_A,$$

d'où $S - aS = 1_A$. On a le même résultat avec la multiplication à droite, ce qui implique que $S(1_A - a) = (1_A - a)S = 1_A$.

Supposons $\|a\| < 1$. 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 vectoriel complet A. En majorant la norme de la série par la série des normes, on obtient la majoration $\|(1_A - a)^{-1}\| \le \sum_{k=0}^{+\infty} \|a\|^k = (1 - \|a\|)^{-1}$.

Proposition 7.1.2. *Soit* A *une algèbre de Banach ; l'ensemble des éléments inversibles dans* A *est un ouvert non vide* U *de* A*. Plus précisément, pour tout* $u \in A$ *inversible, et tout* $b \in A$ *tel que* $\|b\| < \|u^{-1}\|^{-1}$*, on a*

$$(u - b)^{-1} = u^{-1} + u^{-1}b\,u^{-1} + u^{-1}b\,u^{-1}b\,u^{-1} + \cdots = \Big(\sum_{k=0}^{+\infty}(u^{-1}b)^k\Big)u^{-1}.$$

Démonstration. Soit $u \in A$ inversible et soit $b \in A$ tel que $\|b\| < \|u^{-1}\|^{-1}$; on écrit $u - b = u(1_A - u^{-1}b)$, et si on pose $a = u^{-1}b$ on aura $\|a\| = \|u^{-1}b\| \le \|u^{-1}\|\,\|b\| < 1$, ce qui implique que $1_A - a = 1_A - u^{-1}b$ est inversible dans A, donc $u - b$ aussi, et $(u-b)^{-1} = (1_A - a)^{-1}u^{-1}$. En utilisant le développement en série vu au lemme 1, on obtient que lorsque $\|b\| < \|u^{-1}\|^{-1}$, on a $(u-b)^{-1} = \left(\sum_{k=0}^{+\infty} a^k\right) u^{-1}$, ce qui peut s'écrire

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

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

$$R_\lambda(a) = (a - \lambda 1_A)^{-1}.$$

Si $|\lambda| > \|a\|$, on peut écrire $a - \lambda 1_A = -\lambda(1_A - a/\lambda)$, et $\|a/\lambda\| < 1$, ce qui montre que $a - \lambda 1_A$ 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

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

Exemples 7.1.8.

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

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

Théorème 7.1.4. *Soient* A *une algèbre de Banach unitaire complexe et* $a \in A$ *; le spectre de* a *est une partie compacte non vide de* $\mathbb{C}$.

Démonstration. Si λ n'est pas dans $\mathrm{Sp}(a)$, l'élément $u = a - \lambda 1_A$ est inversible ; d'après la proposition 2, $a - (\lambda + h)1_A$ sera encore inversible pour tout $h \in \mathbb{C}$ assez petit (précisément, si $|h| < \|u^{-1}\|^{-1}$) et on aura

$$R_{\lambda+h}(a) = (u - h1_A)^{-1} = u^{-1} + hu^{-2} + h^2u^{-3} + h^3u^{-4} + \cdots$$

Ceci montre en particulier que le complémentaire du spectre de a 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 compact.

Soit $\lambda \in \mathbb{C} \setminus \mathrm{Sp}(a)$; posons $u = a - \lambda 1_A$; alors u est inversible, $u^{-1} = R_\lambda(a)$ et on a dit que pour z assez petit, $a - (\lambda + z)1_A = u - z1_A$ est inversible et

$$R_{\lambda+z}(a) = u^{-1} + zu^{-2} + z^2u^{-3} + z^3u^{-4} + \cdots$$

Si x^* est une forme linéaire continue sur A, la fonction scalaire $g(\lambda) = x^*(\mathrm{R}_\lambda(a))$, définie sur l'ouvert $\Omega = \mathbb{C} \setminus \mathrm{Sp}(a)$, est holomorphe ; en effet, pour tout $\lambda \in \Omega$ la relation précédente montre que pour $|z|$ assez petit, on peut écrire

$$g(\lambda + z) = \sum_{n=0}^{+\infty} x^*(u^{-n-1}) z^n == \sum_{n=0}^{+\infty} c_n(\lambda) z^n,$$

donc g est développable en série entière au voisinage de chaque point λ de Ω.

Il reste à montrer que $\mathrm{Sp}(a) \neq \emptyset$. Choisissons λ_0 hors du spectre ; alors $\mathrm{R}_{\lambda_0}(a)$ est non nul puisqu'inversible et on peut trouver une forme linéaire x^* continue sur A telle que $x^*(\mathrm{R}_{\lambda_0}(a)) \neq 0$ (corollaire 4.2.7) ; l'application $g : \lambda \to x^*(\mathrm{R}_\lambda(a))$ est une fonction holomorphe scalaire définie sur $\mathbb{C} \setminus \mathrm{Sp}(a)$, telle que $g(\lambda_0) \neq 0$. 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 $g(\lambda) = x^*(\mathrm{R}_\lambda(a))$ 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 $\mathrm{Sp}(a) \neq \emptyset$.

Exemples.

Soit M l'opérateur sur $\mathrm{H} = \mathrm{L}_2(0, 1)$ (complexe) défini par

$$\forall f \in \mathrm{H}, \quad (\mathrm{M}f)(x) = x f(x)$$

pour tout $x \in [0, 1]$. Le spectre de M est $[0, 1]$, et M n'a aucune valeur propre.

Quand $\lambda \notin [0, 1]$, on voit facilement que l'opérateur borné de multiplication par $1/(x - \lambda)$, $0 \leq x \leq 1$, est l'inverse de $\mathrm{M} - \lambda\,\mathrm{Id}$; quand $\lambda \in [0, 1]$, on considère un petit intervalle $\mathrm{I}_\varepsilon \subset [0, 1]$ de longueur $\varepsilon > 0$ contenant λ et la fonction de norme 1 donnée par $f_\varepsilon = \varepsilon^{-1/2} \mathbf{1}_{\mathrm{I}_\varepsilon}$; on voit que $|(x - \lambda) f_\varepsilon| \leq \varepsilon\, f_\varepsilon$, donc $\|(\mathrm{M} - \lambda\,\mathrm{Id})(f_\varepsilon)\|_2 \leq \varepsilon$; ceci montre que $(\mathrm{M} - \lambda\,\mathrm{Id})$ n'est pas inversible dans ce cas.

Résolvante de V

On considère ici que V définit un endomorphisme $\mathrm{V_C}$ de $\mathrm{E} = \mathrm{C}(0, 1)$, par la formule habituelle $(\mathrm{V}f)(x) = \int_0^x f(t)\, dt$. On est sûr pour diverses raisons que 0 est dans le spectre de V ; la plus simple est que V n'est pas surjectif, puisque les fonctions de l'image sont toutes de classe C^1 (et nulles en 0 : une autre raison).

On cherche à résoudre l'équation $(\mathrm{V} - \lambda) f = g$, où g est une fonction continue donnée. On va commencer par le cas où g est de classe C^1, pour deviner la formule. Dans ce cas, et puisqu'on cherche f continue de toute façon, on aura $\lambda f = \mathrm{V}f - g$ qui est de classe C^1, donc f est C^1 quand $\lambda \neq 0$.

L'équation devient alors une équation différentielle bien simple,

$$f' - \frac{1}{\lambda} f = -\frac{1}{\lambda} g',$$

qui se résout par des méthodes classiques,

$$f(x) = \mathrm{e}^{x/\lambda} \Big(\int_0^x \mathrm{e}^{-t/\lambda} \big(-\frac{g'(t)}{\lambda} \big)\, dt + \mathrm{C} \Big).$$

On transforme par intégration par parties,

$$f(x) = \mathrm{e}^{x/\lambda} \Big(-\mathrm{e}^{-x/\lambda} g(x)/\lambda - \int_0^x \mathrm{e}^{-t/\lambda} \big(\frac{g(t)}{\lambda^2} \big)\, dt + \mathrm{C}_2 \Big)$$

et on élimine la constante inconnue C_2 en revenant à la véritable équation qui nous intéresse, à savoir $\mathrm{V}f - \lambda f = g$, qui donne la condition $\lambda f(0) = g(0)$, donc

$$f(x) = -\frac{1}{\lambda}\, g(x) - \frac{1}{\lambda^2} \int_0^x \mathrm{e}^{(x-t)/\lambda}\, g(t)\, dt.$$

On vérifie sans peine que le membre de droite définit un endomorphisme continu S_λ de $\mathrm{E} = \mathrm{C}(0,1)$. On vérifie aussi que S_λ est l'inverse de $\mathrm{V} - \lambda$. Ainsi, pour tout $\lambda \neq 0$ et toute fonction $g \in \mathrm{C}([0,1])$

$$\forall x \in [0,1], \quad (\mathrm{R}_\lambda \mathrm{V})(g)(x) = -\frac{1}{\lambda}\, g(x) - \frac{1}{\lambda^2} \int_0^x \mathrm{e}^{(x-t)/\lambda}\, g(t)\, dt.$$

On a ainsi montré que

$$\mathrm{Sp}(\mathrm{V}) = \{0\}.$$

On a vu aussi que la formule de R_λ fournit une façon de résoudre l'équation $(\mathrm{V}-\lambda)f = g$; c'est ce qui explique la terminologie historique de *résolvante*.

Cours n° 16, mercredi 14 novembre 2001.

Suite d'exercice

L'opérateur M de multiplication par $x \to x$ n'a pas de vecteur propre ; en effet, si $(\mathrm{M} - \lambda\,\mathrm{Id})(f) = 0$, on a $(x-\lambda)f(x) = 0$ presque partout, donc $f = 0$ presque partout c'est à dire que $f = 0_{\mathrm{L}_2}$.

Spectre des normaux compacts

Si T est un opérateur normal compact sur un espace de Hilbert complexe, séparable et de dimension infinie, il existe une base hilbertienne $(e_n)_{n\in\mathbb{N}}$ de H telle que T soit diagonal dans cette base, $\mathrm{T}(e_n) = \lambda_n\, e_n$ pour tout $n \geq 0$, avec $\lim_n \lambda_n = 0$. Alors

$$\mathrm{Sp}(\mathrm{T}) = \{\lambda_n : n \in \mathbb{N}\} \cup \{0\}.$$

En effet, il est clair que chaque λ_n est dans le spectre, et le spectre est fermé, donc

$$\mathrm{K} = \{\lambda_n : n \in \mathbb{N}\} \cup \{0\} \subset \mathrm{Sp}(\mathrm{T}).$$

Inversement pour tout $\mu \notin \mathrm{K}$, on voit que l'opérateur diagonal de coefficients $(\lambda_n - \mu)^{-1}$ est borné, et est l'inverse de $\mathrm{T} - \mu\,\mathrm{Id}_{\mathrm{H}}$.

7.2. 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 $a \in \mathrm{A}$. 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

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

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

$$\rho(a) = \lim_{n \to \infty} \|a^n\|^{1/n}.$$

Démonstration. Pour la convergence de la suite $(\|a^n\|^{1/n})$, voir le poly, lemme 7.2.2.

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.

Pour que $a - \lambda = -\lambda(1 - \lambda^{-1}a)$ soit inversible pour un $\lambda \ne 0$, il suffit d'après le lemme 1.1 que la série géométrique $\sum \lambda^{-n} a^n$ soit convergente dans A ; pour cela il suffit que $\sum |\lambda|^{-n} \|a^n\| < +\infty$; d'après le critère de Cauchy pour les séries numériques, il suffit que

$$|\lambda|^{-1} \limsup_n \|a^n\|^{1/n} < 1.$$

Ceci signifie qu'aucun nombre complexe λ tel que $|\lambda| > \limsup_n \|a^n\|^{1/n}$ ne peut être dans le spectre de a, c'est à dire que $\rho(a) \le \limsup_n \|a^n\|^{1/n}$.

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

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

ce qui fournit les inégalités de Cauchy

$$|c_n|\, \mathrm{R}^n \le \mathrm{M}(\mathrm{R}, g) = \max\{|g(z)| : |z| = \mathrm{R}\}.$$

Considérons la fonction vectorielle $f(z) = (1_\mathrm{A} - za)^{-1}$; elle est définie pour tout complexe z tel que $1/z$ ne soit pas dans le spectre de a, ce qui est le cas lorsque $|z| < r = \rho(a)^{-1}$.

Si on pose $u = 1_\mathrm{A} - za$ avec $|z| < r$, on sait que pour $|h| < h_0 = \|u^{-1}a\|^{-1}$, l'élément $f(z+h) = (u - ha)^{-1} = (1_\mathrm{A} - hu^{-1}a)^{-1}u^{-1}$ est la somme de la série de vecteurs

$$f(z+h) = f(z) + ha_1 + h^2 a_2 + \cdots + h^n a_n + \cdots$$

où $a_n = (u^{-1}a)^n u^{-1}$ pour tout $n \ge 1$ (en fait, a commute avec u^{-1} ici, donc $a_n = a^n u^{-n-1}$). On en déduit en particulier la continuité de l'application $z \to f(z)$ pour $|z| < r$; pour tout R tel que $0 < \mathrm{R} < r$, la fonction $z \to \|f(z)\|$ est donc bornée par un certain $\mathrm{M}_0(\mathrm{R})$ sur le cercle de rayon R (qui est compact).

Soit x^* une forme linéaire continue sur A, $\|x^*\| \le 1$; posons $g(z) = x^*(f(z))$ lorsque $|z| < r$; en appliquant x^* à la série précédente, on voit que $x^*(f(z+h))$ est pour h assez petit (dépendant de z) la somme d'une série entière en h, donc g est holomorphe dans $\mathrm{B}(0, r)$ et par conséquent, développable en série entière $\sum_{k=0}^{+\infty} c_k z^k$, convergente lorsque $|z| < r$. Par ailleurs, pour z assez petit on sait que $f(z) = \sum_{k=0}^{+\infty} z^k a^k$ (lemme 1.1), donc $g(z) = \sum_k z^k x^*(a^k)$; par l'unicité des coefficients de Taylor il résulte que $c_n = x^*(a^n)$ pour tout n.

Puisque $\|x^*\| \le 1$, on a $|g(z)| \le \|f(z)\|$, ce qui entraîne que $\mathrm{M}(\mathrm{R}, g) \le \mathrm{M}_0(\mathrm{R})$; les inégalités de Cauchy, appliquées à g, donnent $|x^*(a^n)| \le \mathrm{M}_0(\mathrm{R})/\mathrm{R}^n$ pour tout n et toute

$x^* \in \mathrm{A}^*$ telle que $\|x^*\| = 1$; pour chaque $n \ge 1$ donné on peut choisir par Hahn-Banach (corollaire 4.2.7) une forme linéaire x^* telle que $\|x^*\| = 1$ et $x^*(a^n) = \|a^n\|$; on obtient ainsi $\|a^n\| \le \mathrm{M}_0(\mathrm{R})/\mathrm{R}^n$ pour tout $n \ge 1$, ce qui implique $\limsup_n \|a^n\|^{1/n} \le 1/\mathrm{R}$, d'où $\limsup_n \|a^n\|^{1/n} \le \rho(a)$ en faisant tendre R vers $r = 1/\rho(a)$.

Proposition 7.2.3. *Soit* H *un espace de Hilbert complexe ; le rayon spectral de tout élément normal* T *de* $\mathcal{L}(\mathrm{H})$ *est égal à sa norme,* $\rho(\mathrm{T}) = \|\mathrm{T}\|$.

Démonstration. Soit d'abord A un élément hermitien ; on a $\|\mathrm{A}^2\| = \|\mathrm{A}^*\mathrm{A}\| = \|\mathrm{A}\|^2$ (proposition 6.1.2) ; on en déduit par récurrence que $\|\mathrm{A}^{2^n}\| = \|\mathrm{A}\|^{2^n}$ pour tout $n \ge 0$, 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 hermitien, donc $\rho(\mathrm{T})^2 = \rho(\mathrm{T}^*\mathrm{T}) = \|\mathrm{T}^*\mathrm{T}\| = \|\mathrm{T}\|^2$.

Exemple 7.2.2. Posons $\mathrm{E} = \mathrm{C}([0,1])$ et $\mathrm{A} = \mathcal{L}(\mathrm{E})$; pour toute fonction $f \in \mathrm{E}$ et $s \in [0,1]$, on pose $\mathrm{V}(f)(s) = \int_0^s f(t)\,dt$. Si $\|f\|_\infty \le 1$ on voit que $|\mathrm{V}(f)(s)| \le s$, ce qui implique que

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

puis, par récurrence sur n, que $|\mathrm{V}^n(f)(s)| \le s^n/n!$; en effet

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

ce qui donne $\|\mathrm{V}^n\| \le (n!)^{-1}$. Comme $\lim_n (n!)^{-1/n} = 0$, il s'ensuit que le rayon spectral de V est nul, donc $\mathrm{Sp}(\mathrm{V}) = \{0\}$, comme on l'avait déjà vu par le calcul de la résolvante.

7.3. Décomposition du spectre d'un opérateur borné

Si un opérateur borné T de E dans F est inversible, il possède les deux propriétés suivantes :

A. Il existe une constante $c > 0$ telle que $\|\mathrm{T}x\| \ge c\,\|x\|$ pour tout $x \in \mathrm{E}$.
B. On a $\overline{\mathrm{T}(\mathrm{E})} = \mathrm{F}$.

La deuxième propriété est une forme faible de surjectivité : l'image de T est dense dans F ; c'est évidemment vrai quand T est inversible, puisqu'alors T est surjectif. De plus, lorsque T est inversible, la propriété **A** est vraie avec $c = \|\mathrm{T}^{-1}\|^{-1} > 0$: en effet, on a pour tout $x \in \mathrm{E}$, lorsque T^{-1} existe dans $\mathcal{L}(\mathrm{F}, \mathrm{E})$

$$\|x\| = \|\mathrm{T}^{-1}(\mathrm{T}(x))\| \le \|\mathrm{T}^{-1}\|\,\|\mathrm{T}(x)\|.$$

Lemme 7.3.2. *Soient* E *et* F *deux espaces de Banach ; un opérateur* $\mathrm{T} \in \mathcal{L}(\mathrm{E}, \mathrm{F})$ *est inversible si et seulement s'il vérifie* **A et B**.

Démonstration. On a déjà vu une des directions : si T est inversible, il vérifie les deux conditions. Inversement, supposons que **A** et **B** soient vraies ; on voit alors que T(E) est fermé : si $(\mathrm{T}x_n)$ tend vers $y \in \mathrm{F}$, la suite $(\mathrm{T}x_n)$ est de Cauchy, et l'inégalité $\|x_m - x_n\| \le c^{-1}\|\mathrm{T}x_m - \mathrm{T}x_n\|$ montre que (x_n) est de Cauchy, donc convergente dans E complet vers un vecteur x, et $\mathrm{T}(x) = \lim_n \mathrm{T}(x_n) = y$.

D'autre part T(E) est dense d'après **B**, donc $\mathrm{T(E)} = \mathrm{F}$. Si $\mathrm{T}(x) = \mathrm{T}(x')$ on aura $x = x'$ puisque $0 = \|\mathrm{T}(x - x')\| \geq c\,\|x - x'\|$ d'après **A**. Cela permet de définir une application (linéaire) S de $\mathrm{F} = \mathrm{T(E)}$ sur E en posant $\mathrm{S}(y) = x \in \mathrm{E}$ si et seulement si $y \in \mathrm{F}$ et $\mathrm{T}(x) = y$. En traduisant **A**, on obtient $\|\mathrm{S}(y)\| \leq c^{-1}\,\|y\|$ pour tout $y \in \mathrm{F}$, ce qui montre que S est continue. Pour finir il est clair que S est l'inverse de T.

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

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

Proposition 7.3.3. *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 $({}^t\mathrm{T})^{-1} = {}^t(\mathrm{T}^{-1})$. Supposons inversement que T ne soit pas inversible. On sait que, ou bien T ne vérifie pas la condition **B**, ou bien il ne vérifie pas **A**.

Si T ne vérifie pas **B**, l'image T(E) n'est pas dense, donc ${}^t\mathrm{T}$ n'est pas injective par le lemme 4.2.11, ce qui implique que ${}^t\mathrm{T}$ n'est pas inversible. Si T ne vérifie pas **A**, il existe d'après la proposition 1 une suite $(x_n) \subset \mathrm{E}$ de vecteurs de norme un telle que $\mathrm{T}(x_n) \to 0$. Considérons pour tout entier n l'opérateur R_n de $\mathbb{K}$ dans E, défini par $\mathrm{R}_n(\lambda) = \lambda x_n$. Sa norme est égale à $\|x_n\| = 1$, et $\mathrm{T} \circ \mathrm{R}_n$ tend vers 0. En transposant, ${}^t\mathrm{R}_n \circ {}^t\mathrm{T}$ tend vers 0, alors que $\|{}^t\mathrm{R}_n\| = \|\mathrm{R}_n\| = 1$ pour tout n, ce qui entraîne encore que ${}^t\mathrm{T}$ ne peut être inversible.

On en déduit immédiatement :

Corollaire 7.3.4. *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 nombre complexe λ.

Dans le cas hilbertien, on préfère le plus 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 7.3.5. *Soient* H *un espace de Hilbert complexe et* $\mathrm{T} \in \mathcal{L}(\mathrm{H})$ *; le spectre de l'adjoint* T^* *est formé des complexes conjugués des éléments du spectre de* T,

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

Exemple. Le shift à droite sur $\mathrm{H} = \ell_2(\mathbb{N})$.

On sait déjà que $\|\mathrm{S}\| = 1$, donc le spectre de S est contenu dans le disque unité de $\mathbb{C}$. Soit $\lambda \in \mathbb{C}$; si $|\lambda| < 1$, posons $y = (\lambda^n)_{n\geq 0}$; c'est un élément non nul de ℓ_2 et $\mathrm{S}^*(y) = \lambda y$. Il en résulte que le spectre de S^* contient le disque unité ouvert, et il est contenu dans le disque unité fermé ; puisque le spectre est fermé,

$$\mathrm{Sp}(\mathrm{S}^*) = \{\lambda \in \mathbb{C} : |\lambda| \leq 1\}$$

donc $\mathrm{Sp}(\mathrm{S}) = \mathrm{Sp}(\mathrm{S}^*)$ d'après le corollaire précédent. On a vu que tout λ tel que $|\lambda| < 1$ est valeur propre de S^*. En revanche, S n'a pas de vecteur propre ; comme S est isométrique, la seule possibilité de valeur propre est $|\lambda| = 1$; soit $\eta = (x_n)_{n\geq 0} \in \ell_2$ tel que $\mathrm{S}(\eta) = \lambda\eta$; il s'ensuit alors $0 = \lambda\, x_0$, donc $x_0 = 0$; ensuite, pour tout $n \geq 0$, on a $x_n = \lambda x_{n+1}$ qui implique que $x_{n+1} = 0$ par récurrence sur n, et enfin, $\eta = 0_{\mathrm{H}}$.

Soit λ de module 1 un point quelconque de $\partial\,\mathrm{Sp}(\mathrm{S})$; 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 $\lambda \in \mathbb{T}$. Le même principe fonctionne pour S^*.

MT404, Cours n° 17, lundi 19 novembre 2001.

Commentaires sur le deuxième exercice du partiel

Le but de cet exercice était d'obtenir le résultat classique suivant.

Toute fonction lipschitzienne F *sur* $[0,1]$ *est "primitive" d'une fonction* $f \in \mathrm{L}_\infty(0,1)$, au sens que $\mathrm{F}(x) - \mathrm{F}(0) = \int_0^x f(t)\,dt$ pour tout $x \in [0,1]$. Avec la convergence faible, ce résultat est plus facile à obtenir (mais la convergence faible n'était pas au programme du partiel) : à partir de la suite (f_N) de la question II.2.b, bornée dans $\mathrm{L}_2(0,1)$, il suffisait d'extraire une sous-suite faiblement convergente vers une fonction f ; mais en fait, les (f_N) sont dans la boule unité de $\mathrm{L}_\infty(0,1)$, qui est un ensemble convexe, et fermé dans $\mathrm{L}_2(0,1)$ (problème 1) ; on a donc $|f| \le 1$ à la limite, et la convergence faible implique que $\mathrm{F}(x) - \mathrm{F}(0) = \int_0^x f(t)\,dt$ pour tout $x \in [0,1]$.

Le théorème de dérivabilité de Lebesgue dit que $\mathrm{F}(x) = \int_0^x f(t)\,dt$ est dérivable pour presque tout $x \in [0,1]$, et que cette dérivée est égale à $f(x)$. Ce résultat est sensiblement plus difficile à obtenir.

Proposition. *Si* $\mathrm{S} \in \mathcal{L}(\mathrm{H})$ *est normal, on a*

$\mathrm{H} = \ker(\mathrm{S}) \oplus \overline{\mathrm{im}(\mathrm{S})}$, *et la somme est une somme orthogonale.*

Si S *est injectif, il est à image dense.*

Si S *vérifie la propriété* **A**, S *est inversible*

Démonstration. On sait que $\ker \mathrm{S} = \ker \mathrm{S}^*$ quand S est normal. Il suffit donc de montrer que $\ker \mathrm{S} = \ker \mathrm{S}^*$ est l'orthogonal de $\overline{\mathrm{im}(\mathrm{S})}$, puisqu'on a $\mathrm{H} = \mathrm{F} \oplus \mathrm{F}^\perp$ pour tout sous-espace vectoriel fermé de H.

Pour tout couple (x,y) on a $\langle \mathrm{S}x, y\rangle = \langle x, \mathrm{S}^* y\rangle$. Par conséquent la propriété

$$(1) \qquad \forall x, \quad \langle \mathrm{S}x, y\rangle = 0$$

équivaut à

$$(2) \qquad \forall x, \quad \langle x, \mathrm{S}^* y\rangle = 0.$$

La propriété (1) équivaut à dire que y est orthogonal à l'image de S, et la deuxième signifie que $\mathrm{S}^* y = 0$. Pour finir, être orthogonal à im(S) équivaut par continuité du produit scalaire à être orthogonal à $\overline{\mathrm{im}(\mathrm{S})}$. On a donc

$$\mathrm{H} = \ker(\mathrm{S}) \oplus \overline{\mathrm{im}(\mathrm{S})}.$$

Si S est normal et injectif, on a $\overline{\mathrm{S}(\mathrm{H})} = \mathrm{H}$ d'après ce qui précède, c'est à dire que l'image est dense. Si S vérifie la propriété **A**, il est injectif, donc on a **A** et **B**, par conséquent S est inversible.

Corollaire. *Soient* E *un espace de Hilbert,* T *un opérateur normal et* $\lambda \in \mathrm{Sp}(\mathrm{T})$ *; il existe une suite* $(x_n) \subset \mathrm{H}$ *de vecteurs de norme un telle que* $\lim_n(\mathrm{T}x_n - \lambda x_n) = 0_\mathrm{H}$. *On a de plus* $\lim_n(\mathrm{T}^* x_n - \overline{\lambda} x_n) = 0_\mathrm{H}$ *pour les mêmes vecteurs.*

Démonstration. Si $\lambda \in \mathrm{Sp}(\mathrm{T})$, l'opérateur $\mathrm{S} = \mathrm{T} - \lambda\,\mathrm{Id}_\mathrm{E}$ est normal non inversible, donc S ne vérifie pas **A** : pour tout $c > 0$, la propriété de définition de **A** est fausse. Si $c = 2^{-n}$, il existe donc un vecteur y_n tel que $\|\mathrm{S}y_n\| < 2^{-n}\|y_n\|$; ceci exige $y_n \ne 0_\mathrm{H}$; en posant $x_n = \|y_n\|^{-1} y_n$, on aura $\|x_n\| = 1$ et $\|\mathrm{S}x_n\| < 2^{-n}$. On a obtenu la suite (x_n) souhaitée.

Le dernier résultat vient de la relation $\|\mathrm{S}x\|^2 = \|\mathrm{S}^* x\|^2$, valable pour tout opérateur normal S, appliquée à $\mathrm{S} = \mathrm{T} - \lambda\,\mathrm{Id}_\mathrm{E}$: on a $\|\mathrm{T}x_n - \lambda x_n\| = \|\mathrm{T}^* x_n - \overline{\lambda}\, x_n\|$ pour tout n.

Spectre des hermitiens et unitaires

Proposition 7.3.11. *Soit* H *un espace de Hilbert complexe ; le spectre de tout élément hermitien de* $\mathcal{L}(\mathrm{H})$ *est réel ; le spectre de tout élément hermitien positif de* $\mathcal{L}(\mathrm{H})$ *est contenu dans* $[0,+\infty[$. *Le spectre de tout élément unitaire de* $\mathcal{L}(\mathrm{H})$ *est contenu dans le cercle unité.*

Démonstration. Dans les deux cas T est normal ; si $\lambda \in \mathrm{Sp}(\mathrm{T})$ il existe une suite (x_n) de vecteurs de norme 1 telle que $\mathrm{T}x_n - \lambda x_n \to 0$.

Dans le cas où T est hermitien, on écrit alors

$$\bigl|\langle \mathrm{T}x_n - \lambda x_n, x_n\rangle\bigr| \le \|\mathrm{T}x_n - \lambda x_n\|\,\|x_n\| \to 0$$

d'où résulte que $\langle \mathrm{T}_n, x_n\rangle - \lambda\langle x_n, x_n\rangle$ tend vers 0. On a donc

$$\lambda = \lim_n \langle \mathrm{T}x_n, x_n\rangle$$

qui est réel comme limite d'une suite de réels. Si l'opérateur T est de plus positif, la limite est réelle ≥ 0 puisque $\langle \mathrm{T}x_n, x_n\rangle$ est ≥ 0 pour tout n.

Quand U est unitaire on écrit simplement que $\bigl|\|\mathrm{U}x_n\| - \|\lambda x_n\|\bigr| \le \|\mathrm{U}x_n - \lambda x_n\|$ tend vers 0, donc $|\lambda| = \lim_n \|\mathrm{U}x_n\| = 1$.

9. 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 linéaire borné et *normal* sur un espace de Hilbert complexe H (non nul).

C'est très facile quand T est de plus compact sur un espace de Hilbert séparable de dimension infinie. Il existe alors une base hilbertienne $(e_n)_{n\in\mathbb{N}}$ de H telle que $\mathrm{T}e_n = \lambda_n e_n$ pour tout $n \ge 0$. Le spectre de T est $\mathrm{K} = \{\lambda_n : n \in \mathbb{N}\} \cup \{0\}$. Dans ce cas on pose simplement $(f(\mathrm{T}))(e_n) = f(\lambda_n)e_n$ pour tout $n \ge 0$. Si f est continu sur K, on vient de définir un opérateur diagonal correspondant à une suite $(f(\lambda_n))$ bornée, donc $f(\mathrm{T})$ existe comme opérateur borné sur H. On vérifie facilement que l'application $f \to f(\mathrm{T})$ est un homomorphisme d'algèbres de Banach unitaires, de $\mathrm{C}(\mathrm{K})$ dans $\mathcal{L}(\mathrm{H})$.

9.1. 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 aux questions de 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 9.1.1. *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 montre facilement que $\mathrm{P}(a)b = b\mathrm{P}(a)$ (commencer avec $a^k b = ba^k$ pour tout $k \geq 0$). On déduit ensuite $\mathrm{P}(a)\mathrm{Q}(b) = \mathrm{Q}(b)\mathrm{P}(a)$.

Lemme 9.1.2. *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 9.1.3. *Si* $c, \mu_1, \ldots, \mu_k \in \mathbb{K}$ *et si l'élément* $c\,(a - \mu_1 1_\mathrm{A}) \ldots (a - \mu_k 1_\mathrm{A})$ *est inversible dans* A*, alors chaque* $a - \mu_j 1_\mathrm{A}$ *est inversible, pour* $j = 1, \ldots, k$*.*

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_1a_2 est inversible, donc a_1 est inversible.

Théorème 9.1.4 : Petit théorème spectral. *Soit* A *une algèbre de Banach unitaire* complexe *; pour tout* $a \in \mathrm{A}$*, on a*

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

Démonstration. Posons $\mathrm{K} = \mathrm{Sp}(a)$. Montrons d'abord que $\mathrm{Q}(a)$ est inversible si et seulement si Q ne s'annule pas sur K, c'est à dire si et seulement si 0 n'est pas dans l'ensemble $\mathrm{Q}(\mathrm{K})$ des valeurs de Q sur K. Le cas où Q est un polynôme constant est évident, supposons donc $k = \deg \mathrm{Q} \geq 1$. Puisqu'on est sur $\mathbb{C}$, on peut factoriser le polynôme Q sous la forme

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

avec $c \neq 0$ et $k = \deg \mathrm{P} \geq 1$. Alors

$$\mathrm{Q}(a) = c\prod_{i=1}^{k}(a - \mu_i 1_\mathrm{A})$$

est inversible dans A si et seulement si tous les μ_i sont en dehors de K, ce qui revient exactement à dire que Q ne s'annule pas sur K.

En considérant $\mathrm{Q} = \mathrm{P} - \lambda$, on voit que λ est dans le spectre de $\mathrm{P}(a)$ si et seulement si $\mathrm{P} - \lambda$ s'annule sur K, c'est à dire si et seulement si $\lambda \in \mathrm{P}(\mathrm{K})$.

9.2. 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

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

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

Lemme 9.2.1. *Si* H *est un espace de Hilbert complexe et si* $\mathrm{T} \in \mathcal{L}(\mathrm{H})$ *est normal, on a*

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

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

Démonstration. Soit $\mathrm{K} = \mathrm{Sp}(\mathrm{T})$; on a vu dans le théorème 1.4 que le spectre de $\mathrm{P}(\mathrm{T})$ est $\mathrm{P}(\mathrm{K})$. Par ailleurs $\mathrm{P}(\mathrm{T})$ est normal, donc

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

d'après la proposition 7.2.3.

Passons au théorème sur le calcul fonctionnel continu pour les *hermitiens*. Ce théorème est une étape sur la route du résultat général, valable pour les opérateurs normaux. Nous développerons donc les nombreux corollaires du calcul fonctionnel après l'obtention du résultat général.

Si K est un compact de $\mathbb{C}$, on notera i_{K} la fonction $z \in \mathrm{K} \to z \in \mathbb{C}$.

Théorème 9.2.4. *Soient* H *un espace de Hilbert complexe (non nul) et* $\mathrm{T} \in \mathcal{L}(\mathrm{H})$ *hermitien ; posons* $\mathrm{K} = \mathrm{Sp}(\mathrm{T})$. *Il existe un 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* $f(\mathrm{T}) = \varphi_{\mathrm{T}}(f)$, $f(\mathrm{T})$ *commute avec tout opérateur* S *qui commute avec* T.

Démonstration. On a vu que $\mathrm{K} = \mathrm{Sp}(\mathrm{T})$ est contenu dans $\mathbb{R}$. Désignons par $\mathcal{A}$ l'ensemble (c'est une algèbre) des fonctions continues f sur K de la forme $f : s \to \mathrm{P}(s)$ pour un $\mathrm{P} \in \mathbb{C}[\mathrm{X}]$ (fonctions polynomiales). D'après une version du théorème de Weierstrass, ces fonctions polynomiales sont uniformément denses dans l'espace $\mathrm{C}(\mathrm{K})$.

Montrons l'existence de φ_{T}, en commençant par la définition d'un homomorphisme ψ sur $\mathcal{A}$: pour toute $f \in \mathcal{A}$ l'élément $\psi(f)$ peut être défini de façon unique puisque si $f = \mathrm{P}_1 = \mathrm{P}_2$ sur K,

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

d'après le lemme 1. On posera donc $\psi(f) = \mathrm{P(T)}$, où P est n'importe quel polynôme qui représente la fonction f sur K. De plus, on a $\|\psi(f)\| = \|f\|_\infty$.

L'ensemble $\mathcal{A}$ est dense dans C(K), et on a un homomorphisme isométrique ψ de $\mathcal{A}$ 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 $f \in \mathrm{C(K)}$. 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

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

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

Si ST = 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.

Cours n° 18, mercredi 21 novembre 2001.

Remarque. L'adjoint de $f(\mathrm{T})$, T hermitien. Si (P_n) est une suite de polynômes telle que les fonctions polynomiales correspondantes convergent uniformément vers f sur $\mathrm{K} = \mathrm{Sp(T)} \subset \mathbb{R}$, on a $f(\mathrm{T}) = \lim_n \mathrm{P}_n(\mathrm{T})$ dans $\mathcal{L}(\mathrm{H})$, donc $(f(\mathrm{T}))^* = \lim(\mathrm{P}_n(\mathrm{T}))^* = \lim \widetilde{\mathrm{P}}_n(\mathrm{T}^*)$. Mais on a ici $\mathrm{T}^* = \mathrm{T}$ et de plus, puisque le spectre K est réel, les fonctions polynomiales associées aux polynômes $\widetilde{\mathrm{P}}_n$ convergent vers la fonction $\overline{f}$ sur K, donc

$$(f(\mathrm{T}))^* = \overline{f}(\mathrm{T}).$$

En particulier, $f(\mathrm{T})$ est hermitien quand f est réelle.

Exemples.

Racine carrée, un début. Si T est hermitien positif, il existe un hermitien positif S tel que $\mathrm{T} = \mathrm{S}^2$.

L'ensemble $\mathrm{K} = \mathrm{Sp(T)}$ est dans $[0, +\infty[$. On considère la fonction continue $f : t \in \mathrm{K} \to t^{1/4}$; cette fonction est réelle, donc $f(\mathrm{T})$ est hermitien et $\mathrm{S} = f(\mathrm{T})f(\mathrm{T})$ est hermitien positif ; de plus $\mathrm{S}^2 = f^4(\mathrm{T}) = i_{\mathrm{K}}(\mathrm{T}) = \mathrm{T}$.

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

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

Démonstration. Posons $\mathrm{K} = \mathrm{Sp(T)}$ comme toujours. 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 $\lambda \in \mathrm{K}$, donc $\|g\|_{\mathrm{C(K)}} \ge |g(\lambda)| = 1$ et clairement $\|g\|_\infty \le 1$, donc $\|g(\mathrm{T})\| = \|g\|_{\mathrm{C(K)}} = 1$, 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 $|f(t) - t| \le \varepsilon$ 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 $\|f(\mathrm{T}) - \mathrm{T}\| \le \varepsilon$ puisque $\|f - i_\mathrm{K}\|_{\mathrm{C(K)}} \le \varepsilon$. Pour tout $x \in \mathrm{F}$, on aura $f(\mathrm{T})(x) = \lambda x$ et $\|f(\mathrm{T})(x) - \mathrm{T}(x)\| \le \varepsilon \|x\|$, et c'est fini.

9.4. 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 $\mathrm{K} = \mathrm{Sp(T)}$; on voit clairement de quelle fonction de C(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 $\overline{i_\mathrm{K}}$ sur T^*, alors l'image de $i_\mathrm{K}\overline{i_\mathrm{K}}$ doit être TT^* ; dans le cas hermitien ou unitaire, la fonction $i_\mathrm{K}\overline{i_\mathrm{K}}$ 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 $i_\mathrm{K}\overline{i_\mathrm{K}}$ 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 $i_\mathrm{K}^p\,\overline{i_\mathrm{K}}^q$ pour tous entiers $p, q \ge 0$. Nous allons donc considérer l'algèbre $\mathbb{C}[\mathrm{X}, \mathrm{Y}]$ des polynômes en deux variables, puis prendre l'ensemble des fonctions sur $\mathrm{K} = \mathrm{Sp(T)}$ obtenues en remplaçant X par i_K et Y par $\overline{i_\mathrm{K}}$. Notre algèbre de base $\mathcal{A}$ qui remplacera l'algèbre des polynômes sera l'algèbre de toutes les fonctions f sur K de la forme

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

avec $c_{p,q} \in \mathbb{C}$, et où N varie dans $\mathbb{N}$. On a envie de poser ensuite

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

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 $\mathcal{A}$ considérée est dense dans C(K) par Stone-Weierstrass (facile), et l'application que nous avons en tête sera isométrique.

Lemme crucial. *Soient* H *un espace de Hilbert complexe non nul et* T *un opérateur normal sur* H *; posons* $\mathrm{K} = \mathrm{Sp}(\mathrm{T}) \subset \mathbb{C}$ *; soient par ailleurs* $f \in \mathcal{A}$ *définie par*

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

et

$$\mathrm{V} = \sum_{p,q=0}^{\mathrm{N}} c_{p,q} \mathrm{T}^p \, (\mathrm{T}^*)^q \in \mathcal{L}(\mathrm{H}).$$

On a

$$\|\mathrm{V}\|_{\mathcal{L}(\mathrm{H})} = \|f\|_{\mathrm{C(K)}}.$$

Démonstration. Commençons par montrer que $\|\mathrm{V}\| \geq \|f\|$. Soit $\lambda \in \mathrm{K}$, quelconque ; on sait qu'il existe une suite (x_n) de vecteurs de norme un telle que $\lim_n(\mathrm{T}(x_n) - \lambda x_n) = 0_{\mathrm{H}}$ et $\lim_n(\mathrm{T}^*(x_n) - \overline{\lambda} x_n) = 0_{\mathrm{H}}$ (amphi précédent) ; on en déduit que $\lim_n(\mathrm{T}^p(x_n) - \lambda^p x_n) = 0_{\mathrm{H}}$ et $\lim_n((\mathrm{T}^*)^q(x_n) - \overline{\lambda}^q x_n) = 0_{\mathrm{H}}$ pour tous $p, q \in \mathbb{N}$, puis $\lim_n(\mathrm{V}(x_n) - f(\lambda) x_n) = 0_{\mathrm{H}}$, donc $\|\mathrm{V}\| \geq |f(\lambda)|$, pour tout $\lambda \in \mathrm{K}$, ce qui donne $\|\mathrm{V}\| \geq \|f\|_{\mathrm{C(K)}}$.

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 2.7, 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 (comme ci-dessus) 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 $\|\mathrm{V}\| \leq \|f\|_{\mathrm{C(K)}}$.

Théorème 9.4.1. *Soient* H *un espace de Hilbert complexe et* $\mathrm{T} \in \mathcal{L}(\mathrm{H})$ *normal ; posons* $\mathrm{K} = \mathrm{Sp}(\mathrm{T})$ *; il existe un et un seul homomorphisme d'algèbres de Banach unitaires complexes* $\varphi_{\mathrm{T}} : \mathrm{C(K)} \to \mathcal{L}(\mathrm{H})$ *tel que* $\varphi_{\mathrm{T}}(i_{\mathrm{K}}) = \mathrm{T}$ *et* $\varphi_{\mathrm{T}}(\overline{i_{\mathrm{K}}}) = \mathrm{T}^*$.

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

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

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

Démonstration. A partir du *lemme crucial*, la démonstration est très semblable à celle du cas hermitien. On applique tout d'abord le théorème de Stone-Weierstrass pour dire que la nouvelle algèbre $\mathcal{A}$ est dense dans C(K), ce qui permet d'étendre la correspondance de $\mathcal{A}$ à C(K), et on vérifie comme avant la propriété d'homomorphisme, et le caractère isométrique. A suivre.