MT404, Cours n^o 9, Lundi 25 Octobre 1999.

Exercice proposé. Soient $(\Omega_1, \mathcal{A}_1, \mu_1)$ et $(\Omega_2, \mathcal{A}_2, \mu_2)$ deux espaces mesurés, avec μ_1, μ_2 mesures positives finies. Soit K une fonction réelle mesurable bornée sur $(\Omega_1 \times \Omega_2)$. Pour toute $f \in \mathrm{L}_p(\Omega_2, \mu_2)$ (où $1 \le p < +\infty$) on pose

$$(\mathrm{T_K} f)(s) = \int_{\Omega_2} \mathrm{K}(s,t) f(t)\, d\mu_2(t)$$

pour tout $s \in \Omega_1$. Montrer que $\mathrm{T_K}$ définit une application linéaire continue de $\mathrm{L}_p(\Omega_2, \mu_2)$ dans $\mathrm{L}_p(\Omega_1, \mu_1)$. Déterminer la transposée ${}^t\mathrm{T_K}$, considérée comme application linéaire de $\mathrm{L}_q(\Omega_1, \mu_1)$ dans $\mathrm{L}_q(\Omega_2, \mu_2)$, avec $1/q + 1/p = 1$ (d'une certaine façon, K donne l'analogue continu d'une matrice, et la formule pour $\mathrm{T_K}$ l'analogue continu du calcul de l'image d'un vecteur par une matrice).

Si (K, d) *est un espace métrique compact, l'espace* $\mathrm{C(K)}$ *est séparable.*

Partition de l'unité ; on donne un recouvrement fini de K par des ouverts $\omega_1, \ldots, \omega_\mathrm{N}$; il existe des fonctions continues $\varphi_1, \ldots, \varphi_\mathrm{N}$ telles que $0 \le \varphi_j \le 1$ sur K, $\varphi_j(s) = 0$ si $s \notin \omega_j$ et $\sum_{j=1}^{\mathrm{N}} \varphi_j(s) = 1$ pour tout $s \in \mathrm{K}$.

Pour chaque $j = 1, \ldots, \mathrm{N}$ et $s \in \mathrm{K}$ on pose $\psi_j(s) = \mathrm{dist}(s, \omega_j^c)$. On remarque que $\psi_j(s) > 0$ équivaut à $s \in \omega_j$; la fonction $\psi = \sum_{j=1}^{\mathrm{N}} \psi_j$ est continue et > 0 sur K, donc $1/\psi$ est continue. On pose alors $\varphi_j = \psi_j/\psi$, et on a bien

$$\sum_{j=1}^{\mathrm{N}} \varphi_j(s) = 1, \quad 0 \le \varphi_j \le 1$$

et φ_j nulle hors de ω_j.

Pour toute fonction continue f sur K, on introduit le *module de continuité de* f, défini pour tout $\delta > 0$ par

$$\omega_f(\delta) = \sup\{|f(s) - f(t)| : s, t \in \mathrm{K},\ d(s,t) \le \delta\}.$$

Dire que f est uniformément continue revient à dire que $\lim_{\delta \to 0} \omega_f(\delta) = 0$.

Fixons $\delta_0 > 0$ et considérons un recouvrement fini de K par des boules ouvertes $(\mathrm{B}(s_j, \delta_0))_{j=1,\ldots,\mathrm{N}}$ (existence par Borel-Lebesgue). Soit $\varphi_1, \ldots, \varphi_\mathrm{N}$ une partition de l'unité associée au recouvrement de K par les ouverts $\omega_j = \mathrm{B}(s_j, \delta_0)$; soit $\mathrm{F} = \mathrm{F}_{\delta_0}$ le sous-espace de dimension finie de C(K) engendré par $\varphi_1, \ldots, \varphi_\mathrm{N}$.

Pour toute fonction continue f *sur* K, *on a* $\mathrm{dist}(f, \mathrm{F}_{\delta_0}) \le \omega_f(\delta_0)$.

On pose $g = \sum_{j=1}^{\mathrm{N}} f(s_j)\, \varphi_j \in \mathrm{F} = \mathrm{F}_{\delta_0}$; on voit que

$$f(s) - g(s) = \sum_{j=1}^{\mathrm{N}} \varphi_j(s)(f(s) - f(s_j)).$$

Si $\varphi_j(s)(f(s) - f(s_j)) \ne 0$, on a $\varphi_j(s) \ne 0$, donc $s \in \mathrm{B}(s_j, \delta_0)$, donc $|f(s) - f(s_j)| \le \omega_f(\delta_0)$, donc pour tout $j = 1, \ldots, \mathrm{N}$ on a

$$\varphi_j(s)|f(s) - f(s_j)| \le \omega_f(\delta_0)\, \varphi_j(s),$$

d'où le résultat en sommant en j.

Si on prend $\delta = 2^{-n}$ pour $n = 0, 1, \ldots$ et si (F_n) sont des sous-espaces de dimension finie correspondants, on aura pour toute fonction continue f

$$\mathrm{dist}(f, \mathrm{F}_n) \leq \omega_f(2^{-n}) \to 0.$$

Il en résulte que $\bigcup_n \mathrm{F}_n$ est dense dans C(K), donc C(K) est séparable.

Attention : ça ne marche que si K est métrisable.

Autre conséquence : Ascoli. Soit A un sous-ensemble de C(K) ; on suppose que A est borné dans C(K) et que A est uniformément équicontinu, c'est à dire qu'il existe une fonction ω tendant vers 0 en 0 telle que $\omega_f \leq \omega$ pour toute $f \in \mathrm{A}$. Alors toute suite $(f_n) \subset \mathrm{A}$ admet une sous-suite uniformément convergente. Autrement dit, l'adhérence de A est compacte dans C(K).

Début de démonstration : on montre que pour tout $\varepsilon > 0$, il existe une sous-suite (f_{n_k}) telle que $\|f_{n_k} - f_{n_\ell}\|_\infty < \varepsilon$ pour tous k, ℓ. Il faut ensuite répéter des extractions et utiliser une sous-suite diagonale.

Choisissons $\delta > 0$ tel que $\omega(\delta) < \varepsilon/4$, et soit $\mathrm{F} = \mathrm{F}_\delta$ le sous-espace de dimension finie associé. Pour chaque n, il existe une fonction $g_n \in \mathrm{F}$ telle que $\|f_n - g_n\|_\infty \leq \omega_f(\delta) \leq \omega(\delta) < \varepsilon/4$. La suite (g_n) est bornée puisque $(f_n) \subset \mathrm{A}$ est bornée et que $\|g_n\| \leq \|f_n\| + \|g_n - f_n\|$. Puisque (g_n) est bornée dans un espace vectoriel de dimension finie, on peut utiliser Bolzano-Weierstrass : il existe une sous-suite (g_{n_k}) qui converge uniformément. On a pour $k, \ell \geq \mathrm{K}$ assez grand $\|g_{n_k} - g_{n_\ell}\|_\infty < \varepsilon/2$, donc $\|f_{n_k} - f_{n_\ell}\|_\infty < \varepsilon$ par l'inégalité triangulaire.

Exemple. Pour toute fonction f intégrable sur $[0, 1]$ définissons la fonction continue $\mathrm{V}(f)$ comme la semaine précédente, en posant pour tout $t \in [0, 1]$

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

Considérons l'opérateur V_2 de $\mathrm{L}_2 = \mathrm{L}_2(0, 1)$ dans $\mathrm{C}([0, 1])$ qui associe à chaque fonction $f \in \mathrm{L}_2(0, 1)$ la fonction continue $\mathrm{V}(f)$; considérons l'ensemble $\mathrm{A} = \mathrm{V}_2(\mathrm{B}_{\mathrm{L}_2}) \subset \mathrm{C}([0, 1])$. Si $f \in \mathrm{B}_{\mathrm{L}_2}$, on voit que $|\mathrm{V}(f)(t)| \leq \bigl(\int_0^1 |f(s)|^2\, ds\bigr)^{1/2} \leq 1$, donc l'ensemble A est borné dans $\mathrm{C}([0, 1])$; d'autre part, par Cauchy-Schwarz (cas particulier de Hölder) on a (en supposant $0 \leq s \leq t \leq 1$ pour fixer les idées)

$$|\mathrm{V}(f)(s) - \mathrm{V}(f)(t)| = \Bigl| \int_0^1 1_{[s,t]} f \Bigr| \leq \Bigl(\int_0^1 1_{[s,t]}^2 \Bigr)^{1/2} \Bigl(\int_0^1 |f(u)|^2\, du \Bigr)^{1/2} \leq \sqrt{|s - t|}$$

pour toute $f \in \mathrm{B}_{\mathrm{L}_2}$, donc A est borné et équicontinu, avec pour module de continuité uniforme $\omega(\delta) = \sqrt{\delta}$; l'ensemble A est donc relativement compact dans $\mathrm{C}([0, 1])$ par Ascoli. On dit que l'opérateur V_2 est un *opérateur compact* de L_2 dans $\mathrm{C}([0, 1])$ (l'image de la boule unité du premier espace est relativement compacte dans le second).

Topologie faible sur un espace normé

Soit X un espace vectoriel normé ; la *topologie faible* sur X est la topologie la moins fine rendant continues toutes les applications $x \in \mathrm{X} \to x^*(x)$, où x^* décrit l'ensemble de toutes les formes linéaires continues (en norme) sur X ; on la note $\sigma(\mathrm{X}, \mathrm{X}^*)$; bien entendu la topologie de la norme rend déjà continues toutes ces applications, donc la topologie faible $\sigma(\mathrm{X}, \mathrm{X}^*)$ est plus faible que la topologie de la norme.

On peut décrire un système fondamental de voisinages pour la topologie faible : pour que $W \subset X$ soit un voisinage faible du point $x_0 \in X$, il faut et il suffit qu'il existe un nombre fini de formes linéaires continues $x_1^*, \ldots, x_n^* \in X^*$ et un nombre $\varepsilon > 0$ tels que

$$x_0 \in W(x_0; x_1^*, \ldots, x_n^*, \varepsilon) := \{x \in X : \forall j = 1, \ldots, n, \quad |x_j^*(x) - x_j^*(x_0)| < \varepsilon\} \subset W.$$

Notons que la topologie faible est séparée (Hahn-Banach).

Proposition 3.2.1. *Soient* X *et* Y *deux espaces normés et* $T \in \mathcal{L}(X, Y)$ *; alors* T *est continue de* X *muni de la topologie* $\sigma(X, X^*)$ *dans* Y *muni de la topologie* $\sigma(Y, Y^*)$.

Démonstration. Soient $x_0 \in X$, $y_0 = T(x_0) \in Y$ et $W = W(y_0; y_1^*, \ldots, y_n^*, \varepsilon)$ un voisinage faible élémentaire de $T(x_0)$. On pose $V = W(x_0; {}^tT(y_1^*), \ldots, {}^tT(y_n^*), \varepsilon)$, qui est un voisinage faible de x_0, et on constate que $T(V) \subset W$.

■

Théorème 3.2.1. *Soit* X *un espace normé ; si* C *est un sous-ensemble convexe de* X*, il est fermé en norme si et seulement s'il est faiblement fermé.*

Démonstration. Comme la topologie de la norme est plus fine que la topologie faible, toute partie fermée pour la topologie faible est fermée pour la topologie de la norme.

Supposons $\mathbb{K} = \mathbb{R}$ (voir poly pour le cas complexe). Les demi-espaces affines fermés de la forme $\{x \in X : x^*(x) \geq c\}$, où $x^* \in X^*$ sont faiblement fermés puisque x^* est faiblement continue ; or on a vu que C est intersection de tels ensembles faiblement fermés (corollaire 2.2.3).

■

Proposition. *Tout convexe fermé borné* C *d'un espace réflexif* E *est faiblement compact.*

Démonstration. Puisque $C \subset M\,B_E$ et puisque C est faiblement fermé, il suffit de montrer que la boule unité B_E est faiblement compacte. On vérifie que I_E est un homéomorphisme de $\sigma(E, E^*)$ sur $\sigma(E^{**}, E^*)$, et puisque I_E est isométrique l'image de B_E par cet homéomorphisme est exactement la boule unité fermée de E^{**}, qui est $*$-faiblement compacte d'après Banach-Alaoglu, appliqué à l'espace $Z = E^*$ et à son dual $Z^* = E^{**}$.

Pour expliquer l'homéomorphisme il vaut mieux adopter la convention de notation suivante (je ne l'ai pas adoptée à l'amphi) : pour toute lettre x désignons par x^{**}, non pas un élément quelconque de E^{**}, mais l'élément précis $x^{**} = I_E(x)$, où on suppose que $x \in E$; dire que E est réflexif signifie que tout élément de E^{**} est le x^{**} d'un certain $x \in E$.

Posons $Z = E^*$ et considérons $z_0^* = x_0^{**} \in Z^* = E^{**}$. Un voisinage $*$-faible élémentaire W de x_0^{**} est déterminé par la donnée de $z_1, \ldots, z_n \in Z$, c'est à dire $x_1^*, \ldots, x_n^* \in E^*$ et d'un $\varepsilon > 0$, et alors

$$\begin{aligned} W &= \{z^* : \forall j = 1, \ldots, n, \quad |z^*(z_j) - z_0^*(z_j)| < \varepsilon\} = \\ &= \{x^{**} : \forall j = 1, \ldots, n, \quad |x^{**}(x_j^*) - x_0^{**}(x_j^*)| < \varepsilon\}. \end{aligned}$$

Mais $x^{**}(x_j^*) = x_j^*(x)$, donc W est exactement l'ensemble des x^{**} tels que x soit dans V donné par

$$V = \{x \in E : \forall j = 1, \ldots, n, \quad |x_j^*(x) - x_j^*(x_0)| < \varepsilon\}.$$

Avec ces informations on verra que les voisinages élémentaires dans $\sigma(E^{**}, E^*)$ sont exactement les ensembles $I_E(V)$, où V est un ouvert faible élémentaire dans E.

Remarque. On en déduit que dans un réflexif, on a la propriété d'intersection des familles de convexes fermés bornés non vides, qui a été mentionnée en exercice précédemment.

3.3. Suites faiblement convergentes

Une suite $(x_n) \subset \mathrm{X}$ est faiblement convergente vers $x \in \mathrm{X}$ si $x^*(x) = \lim_n x^*(x_n)$ pour tout $x^* \in \mathrm{X}^*$.

Toute **suite** faiblement convergente dans X est bornée ; ce résultat vient du corollaire 2.1.2. En effet, si (x_n) est une suite dans X telle que $\lim_n x^*(x_n)$ existe pour tout $x^* \in \mathrm{X}^*$, la suite $(\mathrm{I}_\mathrm{X}(x_n))$ est une suite d'applications linéaires définies sur l'espace de Banach X^*, telle que $(\mathrm{I}_\mathrm{X}(x_n)(x^*)$ soit bornée pour chaque $x^* \in \mathrm{X}^*$. D'après Banach-Steinhaus, on a $\sup_n \|\mathrm{I}_\mathrm{X}(x_n)\| < +\infty$, et on sait que $\|x_n\| = \|\mathrm{I}_\mathrm{X}(x_n)\|$.

Théorème 3.3.1. *Si* E *est réflexif, toute suite bornée dans* E *admet des sous-suites faiblement convergentes.*

Démonstration. Supposons d'abord que E^* soit séparable, et considérons $\mathrm{Z} = \mathrm{E}^*$. Si (x_n) est une suite bornée dans E, elle définit la suite bornée $(\mathrm{I}_\mathrm{E}(x_n)) \subset \mathrm{E}^{**} = \mathrm{Z}^*$. Puisque Z est séparable, on sait qu'on peut extraire une sous-suite $(\mathrm{I}_\mathrm{E}(x_{n_k}))$, qui soit $*$-faiblement convergente vers un certain $x^{**} = \mathrm{I}_\mathrm{X}(x)$, c'est à dire que pour tout vecteur $x^* \in \mathrm{E}^* = \mathrm{Z}$, on ait $\lim_k \mathrm{I}_\mathrm{E}(x_{n_k})(x^*) = \mathrm{I}_\mathrm{E}(x)(x^*)$. Mais ceci se réécrit : pour tout $x^* \in \mathrm{E}^*$, on a $\lim_k x^*(x_{n_k}) = x^*(x)$, ce qui signifie que la sous-suite (x_{n_k}) converge faiblement vers x.

Pour le cas général on aura besoin du résultat suivant, qui sera vu plus loin : si X est normé et X^* séparable, alors X est séparable.

Si F est réflexif et séparable, alors $\mathrm{F}^{**} \simeq \mathrm{F}$ est séparable, donc F^* est séparable. Soit (x_n) une suite bornée dans E. Le sous-espace fermé F engendré par la suite (x_n) est un espace réflexif séparable, donc son dual F^* est séparable et réflexif. D'après la première partie on sait qu'on peut extraire une sous-suite (x_{n_k}) qui converge faiblement dans F vers $x \in \mathrm{F}$; pour vérifier qu'il y a aussi convergence faible dans E, on raisonne ainsi : si x^* est une forme linéaire continue sur E, elle n'agira sur les (x_n) et sur $x \in \mathrm{F}$ que par sa restriction $y^* \in \mathrm{F}^*$ à l'espace F, et on aura encore $x^*(x) = y^*(x) = \lim_j y^*(x_{n_j}) = \lim_j x^*(x_{n_j})$. On a donc montré que la sous-suite (x_{n_j}) converge faiblement dans E vers le vecteur x.

Cours n$^\mathrm{o}$ 10, Mercredi 27 Octobre 1999.

Exercice.

– Montrer que si $m \neq n$

$$\int_0^{2\pi} |\,\mathrm{e}^{imt} - \mathrm{e}^{int}\,|^2 \, dt = 4\pi.$$

– On désigne par S le cercle unité du plan complexe et on considère le produit $\mathrm{K} = \mathrm{S}^{[0,2\pi]}$ qui est compact pour la topologie produit (convergence simple sur $[0, 2\pi]$) par Tykhonov. On définit une suite $(f_n)_{n\geq 0}$ d'éléments de K par $f_n(s) = \mathrm{e}^{ins}$; montrer que la suite (f_n) n'admet pas de sous-suite simplement convergente (c'est à dire pas de sous-suite convergente dans K ; on pensera au théorème de convergence dominée).

– On considère l'espace de Banach $\mathrm{X} = \mathrm{M}([0, 2\pi])$ (espace des mesures sur $[0, 2\pi]$). Montrer que la suite (f_n) précédente définit une suite dans la boule unité de X^*, qui n'admet aucune sous-suite $*$-faiblement convergente.

Lemme. *Si* X* *est séparable,* X *est séparable.*

Démonstration. Soit $(x_n^*) \subset X^*$ une suite dense dans X^*. Pour tout n, on peut trouver une suite $(x_{n,k})_k$ dans B_X telle que $\|x_n^*\| = \sup_k |x_n^*(x_{n,k})|$. Posons

$$p(x^*) = \sup_{n,k} |x^*(x_{n,k})|.$$

Alors p est une semi-norme continue ($p(x^*) \leq \|x^*\|$, donc $|p(x^*) - p(y^*)| \leq p(x^* - y^*) \leq \|x^* - y^*\|$) et $p(x_n^*) = \|x_n^*\|$ pour tout n ; cette égalité de fonctions continues se prolonge donc à l'adhérence de l'ensemble (x_n^*), c'est à dire à X^*. Considérons le sous-espace séparable F de X engendré par les vecteurs $(x_{n,k})$; s'il n'était pas dense, il existerait une forme linéaire x^* non nulle sur X, donc $\|x^*\| > 0$, mais nulle sur F, donc $p(x^*) = 0$, ce qui est impossible puisque $p(x^*) = \|x^*\|$.

On dit qu'un ensemble $\Delta \subset X$ est *total* dans X si le sous-espace vectoriel engendré par Δ est dense dans X. Pour que X soit séparable, il faut et il suffit qu'il existe un ensemble dénombrable total dans X.

Théorème 3.3.2. *Si* C *est un convexe fermé borné non vide d'un espace réflexif et si* f *est une fonction réelle convexe continue sur* C*, elle atteint son minimum sur* C*.*

Démonstration. On peut sélectionner une suite (c_n) de réels strictement décroissante telle que $\lim_n c_n = m = \inf f(C)$ (peut-être $-\infty$) ; la suite

$$C_n = \{x \in C : f(x) \leq c_n\}$$

est une suite décroissante de convexes fermés bornés non vides (parce que $m < c_n$), donc son intersection $A = \bigcap_n C_n$ est non vide par compacité faible. Si $x_0 \in A$, on a $f(x_0) \leq c_n$ pour tout n, donc $f(x_0) \leq m$, donc $f(x_0) = m$, ce qui montre que f atteint son minimum au point x_0.

Corollaire 3.3.1. *Soit* C *un convexe fermé non vide d'un espace réflexif* E *; si* f *est convexe continue sur* C *et si* $f(x)$ *tend vers* $+\infty$ *lorsque* $x \in C$ *et* $\|x\| \to +\infty$*, la fonction* f *atteint son minimum sur* C*.*

Démonstration. Soit x_0 un point de C, et soit $C_0 = \{x \in E : f(x) \leq f(x_0)\}$. L'ensemble C_0 est convexe fermé, non vide, et il est borné parce que $f(x) \to +\infty$ lorsque $\|x\| \to +\infty$. La fonction f atteint donc son minimum sur C_0, et il est facile de voir que ce minimum est aussi le minimum sur C tout entier.

Application. Si on trouve une forme linéaire sur un Banach E qui n'atteint pas son maximum sur la boule unité fermée, on est sûr que E n'est pas réflexif. Ainsi, $C([0,1])$ n'est pas réflexif ; exercice : considérer $\ell(f) = \int_0^1 g(t)f(t)\,dt$ sur $C([0,1])$, où $g(t) = 1$ si $0 \leq t \leq 1/2$ et $g(t) = -1$ si $t > 1/2$. Montrer que $\|\ell\| = 1$ mais $|\ell(f)| < 1$ pour toute fonction continue f.

Exercice traité. Sachant que L_4 est réflexif, on va montrer que le dual s'identifie à $L_{4/3}$ (on considère le cas réel). Considérons la fonction F définie sur L_4 par

$$F(f) = \int_\Omega |f(s)|^4 \,d\mu(s) = \|f\|_4^4.$$

Alors

$$F(f+h) = F(f) + 4\int_\Omega (f(s))^3 h(s)\,d\mu(s) + R$$

avec $\mathrm{R} = \mathrm{O}(\|h\|^2)$ quand $h \to 0$. Il en résulte que la fonction F est différentiable sur L_4, et que la différentielle $(d\mathrm{F})_f$ est la forme linéaire associée à la fonction $g = 4f^3 \in \mathrm{L}_{4/3}$.

Soit ℓ une forme linéaire continue sur L_4 ; on considère la fonction convexe continue $f \to \frac{1}{4}\,\mathrm{F}(f) - \ell(f)$ sur L_4 ; cette fonction tend vers $+\infty$ lorsque $\|f\| \to +\infty$, donc elle admet un minimum en un certain point f_0, en lequel la différentielle va s'annuler. On aura donc $(d\mathrm{F})_{f_0} = \ell$ en ce point, donc ℓ provient d'une fonction de $\mathrm{L}_{4/3}$.

4. Espaces hilbertiens

4.1. Produits scalaires

Soit E un espace vectoriel complexe ; on appelle *forme hermitienne* sur E une fonction de $\mathrm{E} \times \mathrm{E}$ dans $\mathbb{C}$ telle que :

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

Attention : on a $\varphi(x, \lambda y) = \overline{\lambda}\,\varphi(x, y)$.

On note que

$$\varphi(x + y, x + y) = \varphi(x, x) + 2\,\mathrm{Re}\,\varphi(x, y) + \varphi(y, y).$$

Une forme hermitienne B sur un espace vectoriel complexe E est dite *positive* si, pour tout $x \in \mathrm{E}$, le nombre $\mathrm{B}(x, x)$ est réel ≥ 0.

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

Proposition 4.1.2 : inégalité de Cauchy-Schwarz. *Soit* E *un espace vectoriel muni d'un produit scalaire ; pour tous* $x, y \in \mathrm{E}$ *on a*

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

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

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

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

∎

Corollaire 4.1.2. *Soit* E *un espace vectoriel muni d'un produit scalaire ; l'application* $x \to \langle x, x\rangle^{1/2}$ *est une semi-norme sur* E.

Démonstration. Voir poly.

4.2. Espaces hilbertiens

On appellera *espace préhilbertien* un espace vectoriel normé E (réel ou complexe) tel qu'il existe un produit scalaire sur E pour lequel $\|x\| = \sqrt{\langle x, x\rangle}$ pour tout $x \in \mathrm{E}$.

Définition 4.2.1. Soit E un espace préhilbertien ; on dit que les éléments x et y de E sont *orthogonaux* si $\langle x, y\rangle = 0$.

Définition 4.2.2. Un *espace hilbertien* est un espace préhilbertien complet.

Exemple 4.2.1. L'espace $\mathrm{L}_2(\Omega, \mu)$ est un espace hilbertien pour le produit scalaire

$$\langle f, g\rangle = \int_\Omega f(s)\overline{g(s)}\, d\mu(s).$$

Redonnons en termes de normes la relation

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

Lorsque x et y sont orthogonaux, il en résulte que $\|x+y\|^2 = \|x\|^2+\|y\|^2$, et cette relation s'étend par récurrence à toute famille finie de vecteurs deux à deux orthogonaux. On a par ailleurs la *relation du parallélogramme*,

$$\|\frac{x+y}{2}\|^2 + \|\frac{x-y}{2}\|^2 = \frac{\|x\|^2 + \|y\|^2}{2}.$$

4.3. Le théorème de projection

Pour ne pas refaire strictement la même chose qu'en licence, on va élargir légèrement le cadre du théorème de projection. On dit qu'un espace de Banach E est *uniformément convexe* s'il existe une fonction $\delta : \varepsilon > 0 \to \delta(\varepsilon) > 0$ telle que : pour tous $x, y \in \mathrm{B_E}$, l'hypothèse $\|(x - y)/2\| \geq \varepsilon$ entraîne que $\|(x + y)/2\| \leq 1 - \delta(\varepsilon)$.

On peut supposer que δ est croissante en prenant pour $\delta(\varepsilon)$ le plus grand nombre possible pour lequel la propriété est vraie. Notons aussi que $\delta(\varepsilon) \leq 1$.

La relation du parallélogramme implique que les espaces de Hilbert sont uniformément convexes : si $\|x\| \leq 1$ et $\|y\| \leq 1$, il en résulte que $\|(x+y)/2\|^2 \leq 1-\|(x-y)/2\|^2$. On peut prendre $\delta(\varepsilon) = 1 - \sqrt{1 - \varepsilon^2} > 0$.

Théorème 4.3.1 : théorème de projection. *Soient* E *un espace de Banach uniformément convexe et* C *une partie convexe fermée non vide de* E *; pour tout* $x \in \mathrm{E}$*, il existe un et un seul point* y_0 *de* C *en lequel la fonction* $y \to \|y - x\|$ *atteint son minimum.*

Démonstration. Par translation on se ramène au cas où $x = 0$; le seul cas intéressant est celui où $0 \notin \mathrm{C}$. Notons $d = \inf\{\|0 - z\| : z \in \mathrm{C}\} = \inf\{\|z\| : z \in \mathrm{C}\} > 0$ la distance de 0 à C. On pose $d_\varepsilon = d(1 + \delta(\varepsilon))$ et

$$\mathrm{C}_\varepsilon = \{z \in \mathrm{C} : \|z\| \leq d_\varepsilon\}.$$

l'ensemble C_ε est convexe fermé, et il est non vide parce que $d < d_\varepsilon$; on va voir que l'uniforme convexité implique que le diamètre de C_ε tend vers 0 quand $\varepsilon \to 0$.

Soient donc z_1 et z_2 deux points de C_ε ; posons $x_j = z_j/d_\varepsilon$, ce qui entraîne $\|x_j\| \leq 1$. Si on avait $\|(z_1-z_2)/2\| \geq \varepsilon d_\varepsilon$, on en déduirait que $\|(x_1-x_2)/2\| \geq \varepsilon$, donc par uniforme convexité $\|(x_1 + x_2)/2\| \leq (1 - \delta(\varepsilon))$. On aurait alors

$$\|(z_1 + z_2)/2\| \leq d_\varepsilon(1 - \delta(\varepsilon)) = d(1 - \delta(\varepsilon)^2) < d,$$

ce qui est impossible puisque $(z_1 + z_2)/2 \in \mathrm{C}$ par convexité, et que d est l'infimum des normes des éléments de C.

On a donc $\|z_1 - z_2\| < 2\varepsilon d(1 + \delta(\varepsilon)) \leq 4d\varepsilon$, donc le diamètre de C_ε est $\leq 4d\varepsilon$; si on prend $\varepsilon = 1/n$ pour $n = 1, \ldots$ on obtient une suite décroissante de fermés non vides dont les diamètres tendent vers 0 ; puisque E est complet, l'intersection est non vide et contient un point unique y. Ce point vérifie $\|y\| = d$ (il faut remarquer que $\delta(\varepsilon)$ tend vers 0 avec ε : j'ai oublié de le dire et de le faire à l'amphi ; on peut voir simplement que $\delta(\varepsilon) \leq \varepsilon$ en considérant x de norme 1, $y = (1 - 2\varepsilon)\,x$; alors $\|(x - y)/2\| = \varepsilon$ et $\|(x + y)/2\| = 1 - \varepsilon$).

∎

Rappel sur la convexité uniforme : soit E un espace de Banach ; on pose pour $\varepsilon > 0$

$$\delta_{\mathrm{E}}(\varepsilon) = \inf\{1 - \|\frac{x+y}{2}\| : x, y \in \mathrm{B}_{\mathrm{E}},\ \|\frac{x-y}{2}\| \ge \varepsilon\}.$$

On dit que E est uniformément convexe si $\delta_{\mathrm{E}}(\varepsilon) > 0$ pour tout $\varepsilon > 0$. En prenant $\|x\| = 1$ et $y = (1-2\varepsilon)x$, on voit que $\delta_{\mathrm{E}}(\varepsilon) \le \varepsilon$, donc $\delta(\varepsilon)$ tend vers 0 lorsque $\varepsilon \to 0$: c'est le point qui manquait la semaine dernière.

Dans la littérature, ce que nous appelons $\delta(\varepsilon)$ s'appellerait $\delta(2\varepsilon)$. Si on suit les conventions mathématiques usuelles, la fonction δ_{E} vaut $+\infty$ lorsque $\varepsilon > 1$ (l'ensemble des choix de x, y est vide dans ce cas ; cette remarque n'a pas été faite à l'amphi). Rappelons le théorème de projection :

Soient E *un espace de Banach uniformément convexe et* C *une partie convexe fermée non vide de* E *; pour tout* $x \in \mathrm{E}$, *il existe un et un seul point* y *de* C *en lequel la fonction* $z \in \mathrm{C} \to \|z - x\|$ *atteint son minimum.*

Le théorème de projection : cas d'un sous-espace vectoriel fermé de H

Un cas particulier important est celui où C est un sous-espace vectoriel fermé F d'un espace de Hilbert H. Dans ce cas, la projection y de x sur F est entièrement caractérisée par les deux conditions suivantes :

– le vecteur y appartient à F ;
– le vecteur $x - y$ est orthogonal à F.

Supposons d'abord que y soit la projection de x sur F ; soit v un vecteur de F ; on a $\|x-(y+tv)\|^2 \ge \|x-y\|^2$ pour tout t réel ; en développant $\|(x-y)-tv\|^2$ et en prenant $t \to 0$ on déduit que $\mathrm{Re}\langle x-y, v\rangle = 0$; soit w quelconque dans F, posons $\langle x-y, w\rangle = r\,\mathrm{e}^{i\theta}$ et utilisons $v = \mathrm{e}^{i\theta}\,w \in \mathrm{F}$; on a $\langle x-y, v\rangle = r$ réel, donc nul par ce qui précède, donc $\langle x-y, w\rangle = 0$ pour tout $w \in \mathrm{F}$.

Inversement, si les deux conditions sont vérifiées et si z est un élément quelconque de F, on aura

$$(*) \qquad \|x-z\|^2 = \|(x-y)+(y-z)\|^2 = \|x-y\|^2 + \|y-z\|^2$$

parce que $y - z \in \mathrm{F}$ est orthogonal à $x-y$. Cette relation montre que $\|x-z\|^2 \ge \|x-y\|^2$ pour tout $z \in \mathrm{F}$, c'est à dire que y est bien le point de F le plus proche de x.

On notera $\mathrm{P}_{\mathrm{F}}(x) = y$ la projection orthogonale de x sur F. La caractérisation ci-dessus montre que $\mu \mathrm{P}_{\mathrm{E}}(x) + \mu' \mathrm{P}_{\mathrm{E}}(x')$ est la projection de $\mu x + \mu' x'$, autrement dit l'application P_{F} est une application linéaire. L'égalité $(*)$ ci-dessus donne aussi $\|x-z\| \ge \|\mathrm{P}_{\mathrm{F}}(x) - z\|$ pour tout $z \in \mathrm{F}$, donc $\|x\| \ge \|\mathrm{P}_{\mathrm{F}}(x)\|$ en prenant $z = 0$; on a donc $\|\mathrm{P}_{\mathrm{F}}\| \le 1$.

On vérifie que $\mathrm{P}_{\mathrm{F}}(\mathrm{P}_{\mathrm{F}}(x)) = x$ pour tout $x \in \mathrm{H}$, c'est à dire que $\mathrm{P}_{\mathrm{F}}^2 = \mathrm{P}_{\mathrm{F}}$ (c'est la relation qui caractérise les *projecteurs* en algèbre linéaire). Si $\mathrm{F} \ne \{0\}$ on a $\|\mathrm{P}_{\mathrm{F}}\| = 1$. Si F est un sous-espace vectoriel fermé d'un espace hilbertien H on appelle *projecteur orthogonal* sur F l'opérateur $\mathrm{P}_{\mathrm{F}} : \mathrm{H} \to \mathrm{H}$ qui associe à tout vecteur $x \in \mathrm{H}$ sa projection sur F.

Projection sur un sous-espace de dimension finie engendré par un système orthonormé

Soit d'abord $f_1, \ldots, f_{\mathrm{N}}$ une suite orthonormée finie de vecteurs de H. On a par orthogonalité

$$\Big\|\sum_{j=1}^{\mathrm{N}} c_j f_j\Big\|^2 = \sum_{j=1}^{\mathrm{N}} |c_j|^2$$

pour tous scalaires $c_1, \ldots, c_{\mathrm{N}}$. Soit $\mathrm{F} = \mathrm{Vect}(f_1, \ldots, f_{\mathrm{N}})$; la projection orthogonale sur F est donnée par

$$\mathrm{P}_{\mathrm{F}}(x) = \sum_{j=1}^{\mathrm{N}} \langle x, f_j\rangle\, f_j.$$

en effet, le vecteur $y = \sum_{j=1}^{\mathrm{N}} \langle x, f_j\rangle\, f_j$ est un vecteur de F et on voit que $\langle y, f_\ell\rangle = \langle x, f_\ell\rangle$ pour tout $\ell = 1, \ldots, \mathrm{N}$, donc $x - y$ est orthogonal à chaque f_ℓ, donc à F en passant aux combinaisons linéaires. On a de plus

$$\sum_{j=1}^{\mathrm{N}} |\langle x, f_j\rangle|^2 = \|\mathrm{P}_{\mathrm{F}}(x)\|^2 \le \|x\|^2$$

en utilisant le fait que $\|\mathrm{P}_{\mathrm{F}}\| \le 1$.

Passons maintenant au cas où on a une suite infinie $(f_k)_{k\ge 0}$ orthonormée dans H. Dans ce cas on voit d'abord que si $\sum_k |c_k|^2 < +\infty$, la série de vecteurs $\sum_{k=0}^{+\infty} c_k f_k$ converge dans H (on vérifie que les sommes partielles $\mathrm{S}_n = \sum_{k=0}^{n} c_k f_k$ sont de Cauchy : si $\mathrm{N} \le m \le n$, on a

$$\|\mathrm{S}_n - \mathrm{S}_m\|^2 = \Big\|\sum_{k=m+1}^{n} c_k f_k\Big\|^2 = \sum_{k=m+1}^{n} |c_k|^2 \le \sum_{k>\mathrm{N}} |c_k|^2$$

qui peut être rendu petit si N est grand). De plus si F est l'adhérence de $\mathrm{Vect}(f_k, k \ge 0)$, on aura maintenant

$$\mathrm{P}_{\mathrm{F}}(x) = \sum_{k=0}^{+\infty} \langle x, f_k\rangle\, f_k.$$

On vérifie d'abord que

$$\sum_{k=0}^{+\infty} |\langle x, f_k\rangle|^2 \le \|x\|^2$$

ce qui est facile puisque toutes les sommes finies vérifient cette majoration. On peut donc considérer $y = \sum_{k=0}^{+\infty} c_k f_k$, avec $c_k = \langle x, f_k\rangle$ pour tout entier $k \ge 0$. Les sommes partielles $\mathrm{S}_n = \sum_{k=0}^{n} c_k f_k$ vérifient $\langle \mathrm{S}_n, f_\ell\rangle = \langle x, f_\ell\rangle$ pour tout $n \ge \ell$, donc à la limite $\langle y, f_\ell\rangle = \langle x, f_\ell\rangle$ pour tout ℓ (l'application $z \to \langle z, f_\ell\rangle$ est continue par Cauchy-Schwarz). Le vecteur $y - x$ est donc orthogonal à tous les f_ℓ, donc à $\mathrm{Vect}(f_\ell, \ell \ge 0)$, donc à F par continuité.

Exemple 4.3.1.

Considérons l'injection isométrique i de $\mathrm{L}_2(0,1)$ dans $\mathrm{L}_2([0,1]^2)$ définie par

$$i(g)(s,t) = g(s).$$

Soit F l'image de i ; c'est un sous-espace vectoriel fermé de $\mathrm{L}_2([0,1]^2)$ (F est formé des fonctions qui ne dépendent que de la première variable). On vérifiera que la projection orthogonale sur F est donnée par

$$\mathrm{P}_\mathrm{F} f(s,t) = \int_0^1 f(s,u)\,du$$

pour toute fonction $f \in \mathrm{L}_2([0,1]^2)$.

Plus généralement, si $(\Omega, \mathcal{A}, \mu)$ est un espace de probabilité et si $\mathcal{F}$ est une sous-tribu de $\mathcal{A}$, on peut considérer le sous-espace vectoriel F de L_2 formé de toutes les fonctions qui sont $\mathcal{F}$-mesurables. Le sous-espace F est fermé, et la projection orthogonale de L_2 sur F s'appelle *l'espérance conditionnelle*.

Proposition 4.3.2. *Soit* H *un espace hilbertien ; pour toute forme linéaire continue* ℓ *sur* H*, il existe un vecteur* $y_\ell \in \mathrm{H}$ *unique qui représente la forme linéaire* ℓ *au sens suivant :*

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

Démonstration. Soit $\ell \in \mathrm{H}^*$; notons F son noyau. Si $\ell = 0$ on prend $y_\ell = 0$; sinon, on a $\ell \neq 0$, on peut choisir un vecteur x_0 tel que $\ell(x_0) = 1$. Soit $y_0 = \mathrm{P}_\mathrm{F}(x_0)$, et posons $x_1 = x_0 - y_0$; alors x_1 est orthogonal à F et $\ell(x_1) = \ell(x_0) = 1$. Pour tout vecteur $x \in \mathrm{H}$ on écrit $x = (x - \ell(x)x_1) + \ell(x)x_1$; on a $\ell(x - \ell(x)x_1) = 0$ donc $x - \ell(x)x_1 \in \mathrm{F}$, ce qui implique que ce vecteur est orthogonal à x_1. On a donc

$$\langle x, x_1 \rangle = \langle \ell(x)x_1, x_1 \rangle = \|x_1\|^2\, \ell(x).$$

Il en résulte que ℓ est représentée par $y = \|x_1\|^{-2}\, x_1$.

■

Exemple 4.3.2. Pour toute forme linéaire continue ℓ sur $\mathrm{L}_2(\Omega, \mu)$, il existe une fonction $g \in \mathrm{L}_2$ telle que

$$\forall f \in \mathrm{L}_2, \quad \ell(f) = \int_\Omega f(s)\overline{g(s)}\, d\mu(s).$$

Une application très utile est le "petit" théorème de Radon-Nikodym. Soit μ la mesure de Lebesgue sur $\mathbb{R}^n$ (mais ça pourrait aussi bien être n'importe quelle mesure positive σ-finie) ; soit ν une mesure positive finie sur $(\mathbb{R}^n, \mathcal{B})$, où $\mathcal{B}$ est la tribu borélienne de $\mathbb{R}^n$; on suppose que

$$\nu(\mathrm{A}) \leq \mu(\mathrm{A})$$

pour tout $\mathrm{A} \in \mathcal{B}$; il existe alors une fonction mesurable bornée f telle que

$$\nu(\mathrm{A}) = \int_\mathrm{A} f(s)\, d\mu(s) = \int_\Omega 1_\mathrm{A}(s) f(s)\, d\mu(s)$$

pour tout ensemble $\mathrm{A} \in \mathcal{B}$, c'est à dire que la mesure ν peut se représenter comme la mesure de densité f par rapport à μ.

La démonstration fonctionne ainsi : il résulte de l'hypothèse que $\int h\, d\nu \leq \int h\, d\mu$ pour toute fonction étagée positive, puis pour toute fonction mesurable positive h, et ceci implique que $\|g\|_{\mathrm{L}_2(\nu)} \leq \|g\|_{\mathrm{L}_2(\mu)}$ pour toute fonction $g \in \mathrm{L}_2(\mu)$, ce qui implique

$L_2(\mu) \subset L_2(\nu)$. Comme ν est finie, la forme linéaire $g \to \int g \, d\nu$ est définie et continue sur $L_2(\nu)$, donc sur $L_2(\mu)$. On peut donc la représenter par une fonction $f \in L_2(\mu)$, c'est à dire que

$$\int g \, d\nu = \int g f \, d\mu$$

pour toute fonction $g \in L_2(\mu)$. En appliquant avec $g = 1_A$ où A est un borélien de mesure de Lebesgue finie on obtient le résultat annoncé. Je n'ai pas montré que f est bornée ; j'ai oublié à l'amphi de dire que je pouvais travailler dans le cas réel, et j'ai trainé une barre de conjugaison sur f ; je n'ai pas non plus montré que f est en fait réelle et ≥ 0.

4.5. Familles sommables dans un espace de Banach

Famille sommable de réels positifs : soit I un ensemble d'indices quelconque et soit $(a_i)_{i \in I}$ une famille de nombres réels ≥ 0 indexée par I ; on pose

$$\sum_{i \in I} a_i = \sup \{ \sum_{j \in J} a_j : J \text{ fini } \subset I \}.$$

Par exemple, si les a_i sont tous nuls on est en train de dire que leur "somme" est égale à 0, quelle que soit la cardinalité de l'ensemble d'indices. Lorsque la somme S est finie on peut pour tout $\varepsilon > 0$ trouver J_0 fini tel que $S - \varepsilon < \sum_{j \in J_0} a_j \leq S$. Pour tout sous-ensemble J fini contenant J_0 on aura alors $S - \sum_{j \in J} a_j < \varepsilon$. Ceci suggère la définition suivante :

Définition 4.5.1. Soient X un espace normé, I un ensemble et $(x_i)_{i \in I}$ une famille d'éléments de X ; on dit que la famille $(x_i)_{i \in I}$ est *sommable* de somme $S \in X$ et on écrit $S = \sum_{i \in I} x_i$ si pour tout $\varepsilon > 0$, il existe une partie finie J de I telle que, pour toute partie finie K de I contenant J on ait $\left\| S - \sum_{i \in K} x_i \right\| < \varepsilon$.

Si $(x_i)_{i \in I}$ est sommable dans un espace normé X, alors

$$L = \{ i \in I : x_i \neq 0_X \}$$

est un ensemble fini ou dénombrable (voir poly) ; expliquons le cas réel positif : si $\sum_{i \in I} a_i \leq 1$ (par exemple) on voit que pour tout $n \geq 1$ l'ensemble

$$L_n = \{ i \in I : a_i \geq 1/n \}$$

a au plus n éléments, donc $L = \bigcup_n L_n$ est fini ou dénombrable.

Si $(f_i)_{i \in I}$ est une famille orthonormée dans un Hilbert H, alors la famille $(c_i f_i)_{i \in I}$ est sommable dans H si et seulement si $\sum_{i \in I} |c_i|^2 < +\infty$. Il faudra que j'y revienne la prochaine fois.

MT404, Cours n° 12, Lundi 8 Novembre 1999.

Rappels : famille sommable de réels ≥ 0, famille sommable de vecteurs dans un espace normé ;

Remarque. Si la famille $(x_i)_{i\in \mathrm{I}}$ est sommable, les sommes finies $\sum_{j\in \mathrm{J}} x_j$ sont bornées, lorsque J varie dans l'ensemble des parties finies de I.

(très mal traité à l'amphi ; j'aurais dû dire : étant donné $\varepsilon = 1/2$ par exemple, il existe J_0 fini tel que $\|\mathrm{S} - \sum_{j\in \mathrm{J}} x_j\| < \varepsilon$ pour tout ensemble $\mathrm{J} \subset \mathrm{I}$ fini qui contient J_0, donc $\| \sum_{j\in \mathrm{J}} x_j - \sum_{j\in \mathrm{J}_0} x_j\| < 2\varepsilon = 1$; en prenant L fini disjoint de J_0 puis $\mathrm{J} = \mathrm{J}_0 \cup \mathrm{L}$ on en déduit $\| \sum_{j\in \mathrm{L}} x_j\| < 1$ pour tout ensemble fini L disjoint de J_0. Posons d'autre part

$$\mathrm{M}_0 = \sup\{\big\|\sum_{j\in \mathrm{J}_1} x_j\big\| \colon \mathrm{J}_1 \subset \mathrm{J}\} < +\infty$$

—c'est un sup fini— et posons $\mathrm{M} = \mathrm{M}_0 + 1$. On voit alors que $\| \sum_{j\in \mathrm{J}} x_j\| \leq \mathrm{M}$ pour tout J fini : on découpe J en $\mathrm{J} \cap \mathrm{J}_0$ et $\mathrm{J} \setminus \mathrm{J}_0$, on applique l'inégalité triangulaire et les estimations précédentes aux deux sommes.)

Image linéaire continue d'une famille sommable

Si T est linéaire continue de X dans Y et si $(x_i)_{i\in \mathrm{I}}$ est sommable dans X, alors $(\mathrm{T}(x_i))_{i\in \mathrm{I}}$ est sommable dans Y et

$$\mathrm{T}\big(\sum_{i\in \mathrm{I}} x_i\big) = \sum_{i\in \mathrm{I}} \mathrm{T}(x_i).$$

Soient H *un espace de Hilbert,* $(f_i)_{i\in \mathrm{I}}$ *une famille orthonormée dans* H, *et* $(c_i)_{i\in \mathrm{I}}$ *une famille de scalaires ; alors la famille* $(c_i f_i)_{i\in \mathrm{I}}$ *est sommable dans l'espace de Hilbert* H *si et seulement si* $\sum_{i\in \mathrm{I}} |c_i|^2 < +\infty$. *On a dans ce cas*

$$\big\|\sum_{i\in \mathrm{I}} c_i f_i\big\|^2 = \sum_{i\in \mathrm{I}} |c_i|^2.$$

Démonstration. Supposons la famille sommable. Il existe M tel que $\| \sum_{j\in \mathrm{J}} c_j f_j\| \leq \mathrm{M}$ pour tout ensemble fini $\mathrm{J} \subset \mathrm{I}$; mais

$$\sum_{j\in \mathrm{J}} |c_j|^2 = \big\|\sum_{j\in \mathrm{J}} c_j f_j\big\|^2 \leq \mathrm{M}^2$$

ce qui montre que $\sum_{i\in \mathrm{I}} |c_i|^2 \leq \mathrm{M}^2 < +\infty$. Inversement quand cette condition est réalisée, on peut trouver une suite croissante d'ensembles finis $\mathrm{J}_n \subset \mathrm{I}$ telle que

$$\sum_{i\notin \mathrm{J}_n} |c_i|^2 < 2^{-n}$$

pour tout $n \geq 0$. Posons $x_n = \sum_{j\in \mathrm{J}_n} c_j f_j$; on vérifie que la suite (x_n) est de Cauchy : en effet, si $m \leq n$ on aura

$$\|x_n - x_m\|^2 = \big\| \sum_{j\in \mathrm{J}_n\setminus \mathrm{J}_m} c_j f_j\big\|^2 = \sum_{j\in \mathrm{J}_n\setminus \mathrm{J}_m} |c_j|^2 \leq \sum_{i\notin \mathrm{J}_m} |c_i|^2 < 2^{-m}.$$

La suite (x_n) converge donc vers un vecteur x, et $\|x - x_m\|^2 \le 2^{-m}$ en passant à la limite quand $n \to +\infty$. Si J est n'importe quel ensemble fini contenant J_m, on aura par le même calcul que ci-dessus $\|x_m - \sum_{j\in \mathrm{J}} c_j f_j\|^2 < 2^{-m}$, donc $\|x - \sum_{j\in \mathrm{J}} c_j f_j\|^2 \le 2.2^{-m}$. On a donc $x = \sum_{i\in \mathrm{I}} c_i f_i$. Pour finir,

$$\|x\|^2 = \lim_n \|x_n\|^2 = \lim_n \sum_{j\in \mathrm{J}_n} |c_j|^2 = \sum_{i\in \mathrm{I}} |c_i|^2.$$

Soient H *un espace de Hilbert,* $(f_i)_{i\in \mathrm{I}}$ *une famille orthonormée dans* H *; soit* F *le sous-espace vectoriel fermé engendré par la famille* $(f_i)_{i\in \mathrm{I}}$ *; pour tout* $x \in \mathrm{H}$ *on a*

$$\mathrm{P_F}(x) = \sum_{i\in \mathrm{I}} \langle x, f_i\rangle\, f_i.$$

Il en résulte que

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

Démonstration. Posons $c_i = \langle x, f_i\rangle$ pour tout $i \in \mathrm{I}$. Pour tout sous-ensemble fini $\mathrm{J} \subset \mathrm{I}$ on sait que $\sum_{j\in \mathrm{J}} |c_j|^2 \le \|x\|^2$, donc la famile $(|c_i|^2)_{i\in \mathrm{I}}$ est sommable et on peut considérer $y = \sum_{i\in \mathrm{I}} c_i f_i$. Ce vecteur est limite de sommes finies de vecteurs de F, donc $y \in \mathrm{F}$. En prenant l'image linéaire continue dans $\mathbb{K}$ par produit scalaire avec un vecteur fixé f_ℓ on obtient

$$\langle y, f_\ell\rangle = \sum_{i\in \mathrm{I}} c_i \,\langle f_i, f_\ell\rangle = c_\ell = \langle x, f_\ell\rangle$$

donc $x - y$ est orthogonal à tous les vecteurs (f_ℓ), donc à leurs combinaisons linéaires, puis à l'adhérence ; finalement $x - y$ est orthogonal à F et y est bien la projection orthogonale de x sur F.

Bases hilbertiennes

On dit que $(f_i)_{i\in \mathrm{I}}$ est une base hilbertienne de H si c'est un système orthonormé de vecteurs et si de plus l'espace vectoriel fermé F engendré par les $(f_i)_{i\in \mathrm{I}}$ est égal à H.

Puisque l'espace fermé F engendré par $(f_i)_{i\in \mathrm{I}}$ est H, la projection sur ce sous-espace est l'identité, et par ailleurs on l'a calculée précédemment. On a donc la propriété de décomposition des vecteurs dans la base :

pour tout $x \in \mathrm{H}$,

$$x = \sum_{i\in \mathrm{I}} \langle x, f_i\rangle\, f_i \quad \text{et} \quad \|x\|^2 = \sum_{i\in \mathrm{I}} |\langle x, f_i\rangle|^2.$$

Exemple du système de Fourier. Considérons dans $\mathrm{H} = \mathrm{L}_2([0, 2\pi], ds/2\pi)$ (complexe) la famille $(f_n)_{n\in\mathbb{Z}}$ indexée par $\mathbb{Z}$ définie par $f_n(s) = \mathrm{e}^{ins}$ pour tout $n \in \mathbb{Z}$; Avec la normalisation choisie pour la mesure on a $\int_0^{2\pi} |f_n(s)|^2\, ds/2\pi = 1$, et on vérifie que les fonctions sont deux à deux orthogonales. Par ailleurs l'espace engendré est dense (on peut invoquer Stone-Weierstrass, qui reviendra de toute façon dans la chapitre suivant).

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

Démonstration. Pour exprimer la démonstration, il est utile de remplacer l'indexation $(f_i)_{i\in \mathrm{I}}$ des familles orthonormées par l'auto-indexation : on dit qu'une partie $\mathrm{B} \subset \mathrm{H}$ est orthonormée si $\|b\| = 1$ pour tout $b \in \mathrm{B}$ et si $\langle b, b'\rangle = 0$ pour tous $b, b' \in \mathrm{B}$ tels que $b \neq b'$. Pour revenir au point de vue de l'indexation, on peut ensuite définir à partir de $\mathrm{B} \subset \mathrm{H}$ orthonormée une famille $(f_b)_{b\in\mathrm{B}}$ indexée par B en posant $f_b = b$ pour tout $b \in \mathrm{B}$!

Le lemme de Zorn donne l'existence d'une partie orthonormée B maximale dans H ; soit F le sous-espace vectoriel fermé qu'elle engendre ; si on avait $\mathrm{F} \neq \mathrm{H}$, on pourrait trouver $x \notin \mathrm{F}$, puis $\beta = x - \mathrm{P_F}(x)$, qui est non nul et orthogonal à F, donc orthogonal à tout élément de B (puisque $\mathrm{B} \subset \mathrm{F}$). Si on pose $\widetilde{b} = \|\beta\|^{-1}\beta$, on obtient un nouveau vecteur de norme 1 qui est orthogonal à tous les éléments de B, donc $\widetilde{\mathrm{B}} = \mathrm{B} \cup \{\widetilde{b}\}$ est une partie orthonormée strictement plus grande que B, ce qui contredit la maximalité. On doit donc avoir $\mathrm{F} = \mathrm{H}$.

5. Théorie spectrale et calcul fonctionnel

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 ; on suppose qu'il existe un élément neutre 1_A pour la multiplication ($1_\mathrm{A} a = a 1_\mathrm{A} = a$ pour tout $a \in \mathrm{A}$) et que

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

pour tous $a, b \in \mathrm{A}$. Les applications $b \to ab$ et $b \to ba$ sont donc continues de A dans A. On remarquera que notre définition exclut $\mathrm{A} = \{0\}$, puisqu'on ne pourra pas y trouver 1_A de norme 1 !

Exemples.

– L'exemple de loin le plus important sera $\mathrm{A} = \mathcal{L}(\mathrm{E})$, où E est un espace de Banach ; si $\mathrm{E} \neq \{0\}$, 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 $1_\mathrm{A} = \mathrm{Id_E}$ est l'élément neutre du produit ; il est de norme 1 quand $\mathrm{E} \neq \{0\}$.

– Soit K un espace compact métrique non vide ; considérons l'espace de Banach $\mathrm{A} = \mathrm{C(K)}$ des fonctions continues sur K à valeurs complexes, muni du produit usuel et de la norme de convergence uniforme (exemples 1.1.1) ; c'est une algèbre de Banach unitaire. L'élément 1_A est la fonction constante égale à 1. Cet exemple donne une algèbre commutative.

– L'espace $\mathrm{L}_\infty([0, 1], dx)$ donne un autre exemple d'algèbre de Banach commutative.

5.1. Spectre et résolvante

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

Exemples.

– Soit E un espace de Banach et considérons $\mathrm{A} = \mathcal{L}(\mathrm{E})$; une application linéaire continue $\mathrm{T} \in \mathrm{A}$ est inversible dans A s'il existe $\mathrm{S} \in \mathcal{L}(\mathrm{E})$ telle que $\mathrm{ST} = \mathrm{Id_E} = 1_\mathrm{A}$ et $\mathrm{TS} = \mathrm{Id_E}$. 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.

– Soit $f \in \mathrm{A} = \mathrm{C}(\mathrm{K})$; si f est inversible il existe une fonction continue g telle que $f(s)g(s) = 1$ pour tout $s \in \mathrm{K}$ donc $f(s) \neq 0$ pour tout $s \in \mathrm{K}$. 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 $\mathrm{A} = \mathrm{C}(\mathrm{K})$. On voit donc que f est inversible dans C(K) si et seulement si elle ne s'annule pas sur K.

Exercice proposé : exprimer l'inversibilité dans L_∞ de $f \in \mathrm{L}_\infty$.

Lemme 5.1.1. *Soient* A *une algèbre de Banach unitaire et* $a \in \mathrm{A}$ *tel que* $\|a\| < 1$ *; alors, la série* $\sum_k a^k$ *est convergente dans* A *et sa somme est l'inverse de* $1_\mathrm{A} - a$,

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

On a de plus l'estimation

$$\left\|(1_\mathrm{A} - a)^{-1}\right\| \leq \frac{1}{1 - \|a\|}.$$

Démonstration. Comme $\|a^k\| \leq \|a\|^k$ pour tout entier $k \geq 0$, la série $\sum_{k=0}^{+\infty} a^k$ est normalement convergente, donc convergente dans l'espace complet A. Notons S sa somme. On vérifie facilement en utilisant l'image de la série par l'application linéaire continue de produit par a (à droite et à gauche) que

$$\mathrm{S}a = a\mathrm{S} = \sum_{k=0}^{+\infty} a^{k+1} = \mathrm{S} - 1_\mathrm{A}$$

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

Conséquence. L'ensemble des éléments inversibles est ouvert. En effet, si $x \in \mathrm{A}$ est inversible et si $h \in \mathrm{A}$ est tel que $\|h\| < \|x^{-1}\|^{-1}$, on aura $x + h = x(1_\mathrm{A} + x^{-1}h)$, et si on pose $a = -x^{-1}h$ on aura $\|a\| = \|x^{-1}h\| \leq \|x^{-1}\|\,\|h\| < 1$, ce qui implique que $1_\mathrm{A} - a = 1_\mathrm{A} + x^{-1}h$ est inversible dans A, donc $x + h$ aussi.

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 5.1.1. Soient A une algèbre de Banach unitaire complexe et $a \in \mathrm{A}$; on appelle *spectre* de a et l'on note $\mathrm{Sp}(a)$ l'ensemble des $\lambda \in \mathbb{C}$ tels que $a - \lambda 1_\mathrm{A}$ ne soit pas inversible.

Exemples.

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

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

Une application f d'un ouvert U de $\mathbb{C}$ dans un espace de Banach F est dite holomorphe si elle est $\mathbb{C}$-dérivable en tout point $\lambda \in \mathrm{U}$,

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

Théorème 5.1.1. *Soient* A *une algèbre de Banach unitaire complexe et* $a \in \mathrm{A}$ *; le spectre de* a *est une partie compacte non vide de* $\mathbb{C}$*, l'application* $\lambda \to (a - \lambda 1_{\mathrm{A}})^{-1}$ *est holomorphe sur* $\mathbb{C} \setminus \mathrm{Sp}(a)$ *(et sa dérivée est continue).*

Démonstration. La prochaine fois.

Cours n° 13, Mercredi 10 Novembre 1999.

On appelle *résolvante* de a l'application qui à chaque $\lambda \in \mathbb{C} \setminus \mathrm{Sp}(a)$ associe l'inverse $(a - \lambda 1_{\mathrm{A}})^{-1}$. On notera si $\lambda \notin \mathrm{Sp}(a)$

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

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

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

Démonstration du théorème annoncé la dernière fois : on a vu que l'ensemble des éléments inversibles est ouvert dans A ; si $\lambda \notin \mathrm{Sp}(a)$, l'élément $a - \lambda 1_{\mathrm{A}}$ est inversible dans A, donc $(a - \lambda 1_{\mathrm{A}}) - z 1_{\mathrm{A}}$ est encore inversible si $|z| < \varepsilon$ suffisamment petit, ce qui signifie que $\mathrm{B}(\lambda, \varepsilon) \subset \mathbb{C} \setminus \mathrm{Sp}(a)$, donc $\mathbb{C} \setminus \mathrm{Sp}(a)$ est ouvert et $\mathrm{Sp}(a)$ est fermé dans $\mathbb{C}$. On a vu ci-dessus qu'il est borné.

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

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

En écrivant $\mathrm{R}_{\lambda+z}(a) = \mathrm{R}_\lambda(a) + z(\mathrm{R}_\lambda(a))^2 + c(z)$ on montre que $\|c(z)\| = \mathrm{O}(|z|^2)$ lorsque $z \to 0$ (majorer la norme de la série par la série des normes), ce qui entraîne que l'application résolvante est dérivable (complexe) au point λ, avec $(\mathrm{R}_\lambda(a))^2$ pour dérivée en ce point. On pourra même remarquer que la dérivée est continue : en effet $\mathrm{R}_\lambda(a)$ est continue en λ puisque dérivable, et la dérivée est le carré de cette fonction continue.

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

Exemple.

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

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

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

Avec cette notation on va chercher à exprimer la transposée de S, considérée comme endomorphisme de ℓ_q. Soit T l'opération de décalage à gauche, définie par $\mathrm{T}((y_n)_{n\geq 0}) = (y_{n+1})_{n\geq 0}$. On constate sans peine que $(\mathrm{T}(y), x) = (y, \mathrm{S}(x))$ pour tous $x \in \ell_p$, $y \in \ell_q$. L'application T "est" donc la transposée de S. Quand un opérateur V sur ℓ_p est inversible, il est clair que sa transposée est inversible dans $\mathcal{L}(\ell_q)$, ce qui entraîne que $\mathrm{Sp}({}^t\mathrm{S}) \subset \mathrm{Sp(S)}$. Soit $\lambda \in \mathbb{C}$; si $|\lambda| < 1$, posons $y = (\lambda^n)_{n\geq 0}$; c'est un élément non nul de ℓ_q et ${}^t\mathrm{S}(y) = \lambda y$. Il en résulte que ${}^t\mathrm{S} - \lambda\,\mathrm{Id}_{\ell_q}$ n'est pas inversible, donc le spectre de ${}^t\mathrm{S}$ contient le disque unité ouvert, et il est contenu dans Sp(S) qui est contenu dans le disque unité fermé ; puisque le spectre est fermé,

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

5.2. Rayon spectral

La quantité

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

s'appelle le *rayon spectral de a*.

Théorème 5.2.1. *Soient* A *une algèbre de Banach unitaire complexe et $a \in \mathrm{A}$; on a*

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

En fait, la suite $\|a^n\|^{1/n}$ tend vers une vraie limite, donc $\rho(a) = \lim_n \|a^n\|^{1/n}$.

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

La deuxième inégalité demande de se rappeler le cours de fonctions holomorphes ; si $g(z) : \mathrm{B}(0, \mathrm{R}) \to \mathbb{C}$ est holomorphe ($\mathrm{R} = +\infty$ admis), alors elle est développable en série entière $\sum_{k=0}^{+\infty} b_k z^k$ dans cette boule ouverte ; il en résulte que pour tout r tel que $0 < r < \mathrm{R}$ et tout $n \geq 0$

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

ce qui donne les inégalités de Cauchy

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

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

Pour le fait que la limite existe, voir le poly.

■

Exemple 5.2.1.

Posons $\mathrm{H} = \mathrm{L}_2([0,1])$; pour $f \in \mathrm{H}$ et $s \in [0,1]$, on pose $\mathrm{V}(f)(s) = \int_0^s f(t)\,dt$. En appliquant Cauchy-Schwarz au produit $1_{[0,s]} f$ on voit que $|\mathrm{V}(f)(s)| \leq \sqrt{s}\,\|f\|_2$. Si $\|f\|_2 \leq 1$ on a donc $|\mathrm{V}(f)(s)| \leq \sqrt{s}\,\|f\|_2 \leq 1$ pour tout réel $s \in [0,1]$; on en déduit que $|\mathrm{V}(\mathrm{V}(f))(s)| = |\int_0^s \mathrm{V}(f)(t)\,dt| \leq s$, puis par récurrence sur n, que $|\mathrm{V}^{n+1}(f)(s)| \leq s^n/n! \leq 1/n!$ pour tout $s \in [0,1]$ donc $\|\mathrm{V}^{n+1}(f)\|_2 \leq 1/n!$; comme $\lim_n (n!)^{-1/n} = 0$, il s'ensuit que le rayon spectral de V est nul, donc $\mathrm{Sp}(\mathrm{V}) = \{0\}$.

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

Changement d'algèbre, homomorphismes

Un homomorphisme d'algèbres de Banach unitaires $\varphi : \mathrm{A} \to \mathrm{B}$ est une application linéaire continue telle que $\varphi(ab) = \varphi(a)\varphi(b)$ pour tous $a, b \in \mathrm{A}$ et telle que l'image de 1_{A} soit 1_{B}. Si a est inversible dans A son image est inversible dans B et l'inverse est l'image de l'inverse. De plus $\varphi(a - \lambda 1_{\mathrm{A}}) = \varphi(a) - \lambda 1_{\mathrm{B}}$. Il en résulte que

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

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

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

Rappel. Un homomorphisme d'algèbres de Banach unitaires $\varphi : \mathrm{A} \to \mathrm{B}$ est une application linéaire continue telle que $\varphi(ab) = \varphi(a)\varphi(b)$ pour tous $a, b \in \mathrm{A}$ et telle que l'image de 1_A soit 1_B. On a vu que si $\mathbb{K} = \mathbb{C}$, on a $\mathrm{Sp}(\varphi(a)) \subset \mathrm{Sp}(a)$ pour tout $a \in \mathrm{A}$. On dit qu'on a un homomorphisme isométrique de A dans B si φ est de plus isométrique.

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

Alors $\mathrm{K} = \mathrm{Sp}(\mathrm{T}) = \{\lambda_1, \lambda_2, \lambda_3\}$. Pour toute fonction f (continue !) sur K on considérera la matrice

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

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

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

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

Conséquence : *pour tout homomorphisme isométrique φ de* C(K) *(complexe) dans une algèbre de Banach unitaire complexe* B, *on aura*

$$\mathrm{Sp}(\varphi(f)) = \mathrm{Sp}(f) = f(\mathrm{K})$$

pour toute $f \in \mathrm{C}(\mathrm{K})$.

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

Dans notre exemple matriciel ci-dessus, on avait bien

$$\mathrm{Sp}(\varphi_\mathrm{M}(f)) = f(\mathrm{K}) = \{f(\lambda_1), f(\lambda_2), f(\lambda_3)\}.$$

Calcul fonctionnel polynomial

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

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

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

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

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

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

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

Lemme. *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 $c(a_1 a_2) = 1_\mathrm{A} = (a_1 a_2)c$; on voit que a_1 est inversible à gauche et à droite : $1_\mathrm{A} = a_1(a_2 c)$ et $1_\mathrm{A} = c(a_1 a_2) = (c a_2) a_1$; il en résulte que $a_2 c = c a_2$ est l'inverse de a_1 :

$$a_2 c = (a_2 c)(a_1 a_2 c) = a_2 (c a_1 a_2) c = a_2 c.$$

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

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

Petit Théorème spectral. ($\mathbb{K} = \mathbb{C}$) Pour tout $a \in \mathrm{A}$, on a

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

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

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

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

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

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

Retour à $\mathcal{L}(\mathrm{E})$

Soient E un espace de Banach et T un endomorphisme continu de E ; on a l'alternative suivante :

1. Il existe une constante $c > 0$ telle que $\|\mathrm{T}(x)\| \geq c\,\|x\|$ pour tout $x \in \mathrm{E}$.

2. Il n'existe pas de telle constante $c > 0$.

On va traduire chacune de ces deux possibilités.

Le cas **1** se produit si et seulement si T définit un isomorphisme de E sur l'image T(E) : si on a isomorphisme, il existe une application linéaire continue S de T(E) dans E telle que $\mathrm{T}(x) = y$ soit équivalent à $\mathrm{S}(y) = x$ pour tous $x \in \mathrm{E}$, $y \in \mathrm{T}(\mathrm{E})$. Alors $\|\mathrm{S}(y)\| \leq \|\mathrm{S}\|\,\|y\|$ pour tout $y \in \mathrm{T}(\mathrm{E})$ revient à dire que $\|\mathrm{T}(x)\| \geq \|\mathrm{S}\|^{-1}\|x\|$ pour tout $x \in \mathrm{E}$. Inversement, si la constante $c > 0$ existe, on remarque d'abord que T est injectif ; on peut donc définir algébriquement une application linéaire S de T(E) dans E telle que $\mathrm{T}(x) = y$ si et seulement si $x = \mathrm{S}(y)$; de plus, l'inégalité de norme $\|\mathrm{T}(x)\| \geq c\,\|x\|$ signifie que $\|\mathrm{S}\| \leq c^{-1}$, donc S est continu, et T est un isomorphisme de E sur T(E).

Dans le cas **1** on obtient comme conséquence que T(E) est un sous-espace fermé de E, puisqu'il est complet.

Dans le cas **2**, on peut dire pour tout $n \geq 0$ que la constante $c = 2^{-n}$ ne convient pas ; il existe donc un vecteur x_n tel que $\|\mathrm{T}(x_n)\| < 2^{-n}\|x_n\|$, ce qui implique que $x_n \neq 0_{\mathrm{E}}$; en multipliant par un scalaire, on se ramène à :

il existe une suite $(x_n) \subset \mathrm{E}$ *telle que* $\|x_n\| = 1$ *pour tout* n *mais* $\|\mathrm{T}(x_n)\| \to 0$.

Bien entendu cette propriété implique en retour qu'il n'existe pas de constante $c > 0$ telle que… donc cette propriété caractérise le cas **2**.

Supposons maintenant que T soit non inversible ; si on est dans le cas **1**, on sait que T définit un isomorphisme de E sur T(E) qui est de plus fermé ; si on avait $\mathrm{T}(\mathrm{E}) = \mathrm{E}$, l'opérateur T serait inversible. On a donc $\mathrm{T}(\mathrm{E}) \neq \mathrm{E}$.

On peut donc dire : si T n'est pas inversible dans $\mathcal{L}(\mathrm{E})$, ou bien T définit un isomorphisme de E sur T(E) **et** $\mathrm{T}(\mathrm{E}) \neq \mathrm{E}$, ou bien il existe une suite $(x_n) \subset \mathrm{E}$ de vecteurs de norme 1 telle que $\mathrm{T}(x_n) \to 0$.

Proposition. *Soient* E *un espace de Banach et* $\mathrm{T} \in \mathcal{L}(\mathrm{E})$ *; alors* T *est inversible si et seulement* ${}^t\mathrm{T}$ *est inversible dans* $\mathcal{L}(\mathrm{E}^*)$.

Démonstration. On a déjà dit que si T est inversible, alors ${}^t\mathrm{T}$ est inversible (transposer la relation $\mathrm{T}\mathrm{T}^{-1} = \mathrm{T}^{-1}\mathrm{T} = \mathrm{Id}_{\mathrm{E}}$) ; supposons donc T non inversible.

Dans le cas **1**, on a T(E) fermé et $\mathrm{T}(\mathrm{E}) \neq \mathrm{E}$; d'après Hahn-Banach, on peut trouver une forme linéaire y^* non nulle telle que $y^* = 0$ sur T(E), c'est à dire $y^*(\mathrm{T}(x)) = 0$ pour tout $x \in \mathrm{E}$; ceci donne $({}^t\mathrm{T})(y^*)(x) = 0$ pour tout x, c'est à dire $({}^t\mathrm{T})(y^*) = 0$, donc ${}^t\mathrm{T}$ n'est pas injective, donc pas inversible.

Dans le cas **2**, on trouve une suite $(x_n) \subset \mathrm{E}$ telle que $\|x_n\| = 1$ et (par exemple) $\|\mathrm{T}(x_n)\| \leq 2^{-n}$; par Hahn-Banach on peut trouver une forme linéaire x_n^* de norme 1 telle que $x_n^*(x_n) = \|x_n\| = 1$; si ${}^t\mathrm{T}$ était inversible, il existerait $y_n^* = ({}^t\mathrm{T})^{-1}(x_n^*)$ telle que $x_n^* = {}^t\mathrm{T}(y_n^*)$ et alors

$$1 = x_n^*(x_n) = ({}^t\mathrm{T})(y_n^*)(x_n) = y_n^*(\mathrm{T}(x_n)) \leq \|y_n^*\|\,\|\mathrm{T}(x_n)\| \leq 2^{-n}\|y_n^*\|$$

montre que $\|y_n^*\| \geq 2^n$, donc $\|({}^t\mathrm{T})^{-1}\| \geq 2^n$ pour tout $n \geq 0$, ce qui est impossible ; on en déduit que ${}^t\mathrm{T}$ est non inversible.

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

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

Démonstration. Il suffit de remarquer que ${}^t\mathrm{T} - \lambda \,\mathrm{Id}_{\mathrm{E}^*} = {}^t(\mathrm{T} - \lambda \,\mathrm{Id}_\mathrm{E})$ est inversible si et seulement si $\mathrm{T} - \lambda \,\mathrm{Id}_\mathrm{E}$ est inversible.

Valeurs propres approchées

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

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

Cours n° 15, Mercredi 17 Novembre 1999.

Rappels de quelques éléments qui seront utilisés aujourd'hui :

– Si φ est un homomorphisme isométrique d'algèbres de Banach unitaires complexes de C(K) dans B, on a $\mathrm{Sp}(\varphi(f)) = \mathrm{Sp}(f) = f(\mathrm{K})$.

– Petit théorème spectral : $\mathrm{Sp}(\mathrm{P}(a)) = \mathrm{P}(\mathrm{Sp}(a))$.

– Lemme du bord du spectre : $(\mathbb{K} = \mathbb{C})$ Si $\lambda \in \partial \,\mathrm{Sp}(\mathrm{T})$, il existe une suite de vecteurs propres approchés $(x_n) \subset \mathrm{E}$, vecteurs de norme un tels que $\mathrm{T}(x_n) - \lambda x_n \to 0$.

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

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

et on note que $\|\mathrm{S}(x_n) - \lambda^{-1} x_n\| \leq 2n^{-1/2} \to 0$, ce qui donne des presque vecteurs propres pour $\lambda^{-1} \in \mathbb{T}$ (ça aurait été plus malin de construire le vecteur avec des λ^{-1}, mais je n'ai pas été malin).

Adjoint d'une application linéaire continue sur un Hilbert

On prend $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$.

Proposition. *Soient* E *et* F *deux espaces hilbertiens 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* $\langle \mathrm{T}(x), y \rangle = \langle x, \mathrm{T}^*(y) \rangle$.

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

$\langle x, T^*(y)\rangle$. Pour tout $x \in E$, l'application $y \to \langle T^*(y), x\rangle = \langle y, T(x)\rangle$ est linéaire d'où l'on déduit que T^* est linéaire. Pour tout $x \in E$ et tout $y \in F$ on a

$$|\langle x, T^*(y)\rangle| = \langle T(x), y\rangle| \leq \|T(x)\| \, \|y\| \leq \|T\| \, \|x\| \, \|y\|;$$

or, pour $y \in F$ on a (par la proposition 4.2.1)

$$\|T^*(y)\| = \sup\{\langle x, T^*(y)\rangle : x \in E, \ \|x\| \leq 1\};$$

donc $\|T^*(y)\| \leq \|T\| \, \|y\|$. Donc T^* est continue et $\|T^*\| \leq \|T\|$.

■

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

$$T^* = h_E^{-1} \, ({}^t T) \, h_F.$$

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

Exemples.

– cas matriciel : si on munit $E = F = \mathbb{C}^n$ de sa norme euclidienne usuelle et si $T \in \mathcal{L}(\mathbb{C}^n)$ est représenté dans la base canonique par une matrice $M = (m_{i,j})$, alors l'adjoint de T est représenté par la matrice M^* dont les coefficients sont $m^*_{i,j} = \overline{m_{j,i}}$.

– opérateur diagonal dans une base orthonormée, à coefficients (λ_n) : l'adjoint est l'opérateur diagonal de coefficients $(\overline{\lambda}_n)$;

– multiplication par une fonction continue f sur $L_2(0,1)$: l'opérateur $M_f : g \in L_2 \to fg \in L_2$ admet pour adjoint l'opérateur de multiplication par la fonction complexe conjuguée $\overline{f}$.

– exercice proposé : opérateur $T_K \in \mathcal{L}(L_2(0,1))$ défini par un noyau $K(s,t) \in L_2([0,1]^2)$; l'adjoint est l'opérateur défini par $K^*(t,s) = \overline{K(s,t)}$.

Regroupons dans la proposition suivante quelques propriétés des adjoints :

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

Démonstration. Montrons que $\|T^* \circ T\| = \|T\|^2$. On a $\|T^* \circ T\| \leq \|T^*\| \, \|T\| \leq \|T\|^2$. De plus, pour tout $x \in E$, $\|x\| \leq 1$ on a $\|T(x)\|^2 = \langle T(x), T(x)\rangle = \langle x, T^* \circ T(x)\rangle \leq \|T^* \circ T\|$ grâce à l'inégalité de Cauchy-Schwarz. Donc

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

Les autres propriétés sont laissées en exercice.

■

Définition. Soient E et F deux espaces hilbertiens ; un élément $\mathrm{U} \in \mathcal{L}(\mathrm{E}, \mathrm{F})$ est appelé *unitaire* si $\mathrm{U}^* \circ \mathrm{U} = \mathrm{Id}_\mathrm{E}$ et $\mathrm{U} \circ \mathrm{U}^* = \mathrm{Id}_\mathrm{F}$. Un élément $\mathrm{T} \in \mathcal{L}(\mathrm{E})$ est appelé *normal* si $\mathrm{T}^* \circ \mathrm{T} = \mathrm{T} \circ \mathrm{T}^*$, *autoadjoint* si $\mathrm{T} = \mathrm{T}^*$. On dit aussi *hermitien* comme synonyme d'autoadjoint.

Exemples.

– matrice hermitienne : on retrouve la condition de premier cycle $m_{i,j} = \overline{m_{j,i}}$; en particulier si M est réelle symétrique. Matrice normale, unitaire : pas grand chose à ajouter.

– opérateur diagonal à coefficients (λ_n) : il est hermitien si et seulement si $\lambda_n \in \mathbb{R}$ pour tout n, unitaire si $|\lambda_n| = 1$; l'opérateur est toujours normal.

– multiplication par une fonction réelle f sur $\mathrm{L}_2(0,1)$: elle est hermitienne si et seulement si f est réelle, unitaire si et seulement si $|f| = 1$; la multiplication est toujours normale.

– Un opérateur à noyau est hermitien si $\mathrm{K}(s,t) = \overline{\mathrm{K}(t,s)}$ (en particulier quand il est réel symétrique). On verra plus tard qu'un opérateur à noyau n'est jamais unitaire.

– $\mathrm{T}^*\mathrm{T}$ est toujours hermitien.

Polynômes et adjoints

Il résulte des propriétés des adjoints que $(\mathrm{T}^k)^* = (\mathrm{T}^*)^k$ pour tout entier $k \geq 0$. Si $\mathrm{P} = \sum c_k \mathrm{X}^k$ est un polynôme, on posera $\widetilde{\mathrm{P}} = \sum \overline{c}_k \mathrm{X}^k$, et 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 $\widetilde{\mathrm{P}}(\overline{z}) = \overline{\mathrm{P}(z)}$.

On voit que si P est un polynôme à coefficients réels, alors $\mathrm{P}(\mathrm{T})$ est hermitien. 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.

Spectres

Proposition 5.2.1. *Soit* H *un espace hilbertien complexe ; le rayon spectral de tout élément normal de* $\mathcal{L}(\mathrm{H})$ *est égal à sa norme.*

Démonstration. Soit d'abord S hermitien. On a

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

En prenant le sup sur x dans la boule unité on voit que $\|\mathrm{S}\|^2 \leq \|\mathrm{S}^2\|$, donc $\|\mathrm{S}^2\| = \|\mathrm{S}\|^2$; on remarque que S^k est hermitien pour tout k et on en déduit par récurrence $\|\mathrm{S}^{2^n}\| = \|\mathrm{S}\|^{2^n}$. La formule du rayon spectral termine l'histoire.

Supposons maintenant T normal. Alors $\mathrm{T}^*\mathrm{T} = \mathrm{S}$ est hermitien et $\mathrm{S}^n = (\mathrm{T}^*\mathrm{T})^n = (\mathrm{T}^*)^n \mathrm{T}^n$ parce que T est normal,

$$\|\mathrm{T}^n\|^2 = \|(\mathrm{T}^*)^n \mathrm{T}^n\| = \|(\mathrm{T}^*\mathrm{T})^n\| = \|\mathrm{S}^{2n}\| = \|\mathrm{S}\|^{2n} = \|\mathrm{T}\|^{2n},$$

d'où le résultat.

Proposition. *Soit* H *un espace hilbertien complexe ; le spectre de tout élément hermitien de* $\mathcal{L}(\mathrm{H})$ *est réel.*

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

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

tend vers 0, et $\langle \mathrm{T}(x_n), x_n\rangle = \langle x_n, \mathrm{T}(x_n)\rangle$ est réel, donc λ est réel, $b = \beta = 0$ et tout le spectre est réel.

Le résultat essentiel est le suivant : si T est normal,

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

Soit $\mathrm{K} = \mathrm{Sp(T)}$. On a vu que le spectre de P(T) est P(K). Par ailleurs P(T) est normal, donc

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

Calcul fonctionnel continu pour les opérateurs hermitiens

Soit K un compact de $\mathbb{C}$; on notera z_{K} la fonction $z \in \mathrm{K} \to z \in \mathbb{C}$.

Théorème 5.4.2. *Soient* H *un espace hilbertien complexe et* $\mathrm{T} \in \mathcal{L}(\mathrm{H})$ *autoadjoint ; il existe un et un seul homomorphisme d'algèbres de Banach unitaires complexes* φ_{T} : $\mathrm{C(Sp(T))} \to \mathcal{L}(\mathrm{H})$ *tel que* $\varphi_{\mathrm{T}}(z_{\mathrm{Sp(T)}}) = \mathrm{T}$. *On notera* $f(\mathrm{T}) = \varphi_{\mathrm{T}}(f)$.

De plus, φ_{T} *est isométrique,* $(f(\mathrm{T}))^* = \overline{f}(\mathrm{T})$ *et* $f(\mathrm{T})$ *commute avec tout opérateur* S *qui commute avec* T *(donc* $f(\mathrm{T})$ *est normal). On a*

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

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

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

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

On a une application isométrique de A dans $\mathcal{L}(\mathrm{H})$, qui se prolonge par densité en application linéaire isométrique. La suite la prochaine fois.