2006-12-03 by BM

Notes de Cours

THÉORIE SPECTRALE

B. Maurey

Décembre 2004
révision 11/2006

Ces Notes sont issues d'un enseignement donné à Paris 7 pendant les deux années 2003 et 2004, dans le cadre du magistère de Cachan. Leur contenu provient en partie d'un polycopié de titre voisin *Analyse Fonctionnelle et Théorie Spectrale*, que j'avais rédigé pour un cours de maîtrise à Paris 7 ; on peut trouver ce poly à l'adresse

`http://www.math.jussieu.fr/~maurey/ts012/poly/index.html`

Comparées à ce poly de maîtrise, ces Notes adoptent un rythme plus soutenu, et se placent à un niveau un peu plus ambitieux, entre maîtrise et début de troisième cycle. Elles contiennent aussi un certain nombre d'éléments qui n'ont jamais été réellement enseignés devant les étudiants, mais que je ne me suis pas résolu à supprimer.

Le lecteur se rendra compte facilement que j'ai forcément étudié le livre de Rudin, *Functional Analysis*, qui constitue une lecture très recommandable. Le contenu de ces Notes doit aussi beaucoup à un autre poly de maîtrise de Paris 7, que m'avait légué Georges Skandalis, et sur lequel j'avais construit le poly mentionné ci-dessus ; on peut encore trouver dans le document qui suit des morceaux qui ont été tapés par Skandalis ! L'esprit de ce poly de Skandalis provenait en partie du livre de Reed et Simon, *Methods of modern mathematical physics*, volume 1, une autre excellente lecture. Les quelques passages plus spécifiquement liés à la théorie des espaces de Banach ont été inspirés par le livre de Lindenstrauss et Tzafriri, *Classical Banach Spaces*, volume I. J'ai consulté le livre de G. Pedersen, C^*-*algebras and their automorphism groups*, pour plusieurs points concernant les C^*-algèbres, et absorbé inconsciemment un grand nombre d'informations et de façons de procéder, qui me viennent de cours anciens que j'ai pu suivre ou de conversations avec des collègues.

La numérotation des énoncés et des formules est assez primitive : en général, seuls les éléments qui servent de référence ailleurs dans le texte sont numérotés. Les fins de démonstration sont indiquées par le signe « /// » placé comme ci-dessous.

///

Chevaleret, décembre 2004
Révisé en novembre 2006

Sommaire

1. Opérateurs compacts sur les espaces de Hilbert

1.1. Généralités

Quelques notations

Si X est un ensemble, on note Id_X *l'application identité* de X, c'est-à-dire l'application de X dans X telle que $\mathrm{Id}_\mathrm{X}(x) = x$ pour tout $x \in \mathrm{X}$. Si A est un sous-ensemble de X, on notera A^c le complémentaire de A. On notera $\mathbf{1}_\mathrm{A}$ la *fonction indicatrice* de A, qui est égale à 1 en tout point de A et à 0 en tout point de A^c. On notera parfois 0_X le vecteur nul d'un espace vectoriel X, quand il semblera que cette notation lourde lève toute ambiguïté.

Opérateurs linéaires

Si X est un espace normé sur $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$, on notera B_X sa boule unité fermée,

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

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

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

la définition de la norme de T est en général utilisée pour pouvoir écrire la majoration $\|\mathrm{T}(x)\| \le \|\mathrm{T}\|\,\|x\|$ pour tout vecteur x ; de fait, $\|\mathrm{T}\|$ est la plus petite constante C telle qu'on puisse écrire $\|\mathrm{T}(x)\| \le \mathrm{C}\,\|x\|$ pour tout $x \in \mathrm{X}$. Pour alléger la notation, on écrira $\mathrm{T}x$ pour l'image $\mathrm{T}(x)$ d'un vecteur $x \in \mathrm{X}$ par $\mathrm{T} \in \mathcal{L}(\mathrm{X},\mathrm{Y})$. Lorsque Y est de plus *complet* (espace de Banach), l'espace $\mathcal{L}(\mathrm{X},\mathrm{Y})$ est complet pour la norme précédente. On notera $\mathcal{L}(\mathrm{X})$ l'espace des endomorphismes continus de X.

Une forme du théorème de Hahn-Banach

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

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

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

Séries de vecteurs

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

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

et on dit que s est la *somme de la série de vecteurs*. La convergence de la série des normes $\sum \|u_n\|$ implique que la suite (s_n) des sommes partielles est de Cauchy dans X ; par conséquent, lorsque X est complet, la convergence de la série des normes $\sum \|u_n\|$ implique que la série de vecteurs converge dans X. Si T est une application linéaire continue, elle transformera par linéarité les sommes partielles de la série $\sum u_n$ en sommes partielles de la série $\sum \mathrm{T}u_n$, et donnera la convergence des images par continuité. On aura par conséquent

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

Ce principe simple est utilisé très souvent.

Opérateurs compacts

L'argument diagonal est utilisé souvent dans les questions de compacité : il permet d'affirmer la chose suivante : si on a une famille indexée par $m \in \mathbb{N}$ formée de suites numériques bornées, chacune d'entre elles $(x_{m,n})$ étant indexée par $n \in \mathbb{N}$ et telle que $\sup_n |x_{m,n}| < +\infty$, on peut trouver une sous-suite d'indices n_k telle que

$$\lim_k x_{m,n_k}$$

existe, pour tout indice m. Cette affirmation revient à dire que $[0,1]^{\mathbb{N}}$ est compact pour la topologie produit ; elle se généralise au cas où chaque suite $(x_{m,n})_n$ est contenue dans un espace métrique compact X_m, $m \in \mathbb{N}$. On verra un exemple d'argument de sous-suite diagonale un peu plus loin.

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

Si X_0 est un sous-espace vectoriel fermé de X et si $\mathrm{T} \in \mathcal{L}(\mathrm{X}, \mathrm{Y})$ est compact, sa restriction $\mathrm{T}_{|\mathrm{X}_0}$ est un opérateur compact de X_0 dans Y, puisque l'image de la boule unité de X_0 est contenue dans l'image de la boule de l'espace entier. Si T est un endomorphisme compact de X et si $\mathrm{Y} \subset \mathrm{X}$ est stable par T, la restriction $\mathrm{T}_{|\mathrm{Y}}$, vue comme application de Y dans Y, est un endomorphisme compact de Y.

Dire que $\mathrm{T} \in \mathcal{L}(\mathrm{X}, \mathrm{Y})$ est compact revient exactement à dire ceci : *pour toute suite bornée* (x_n) *dans* X, *la suite image* $(\mathrm{T}x_n)$ *admet des sous-suites convergentes dans* Y.

Si $T_1, T_2 \in \mathcal{L}(X, Y)$ sont compacts, il est clair que $a_1 T_1 + a_2 T_2$ est compact pour tous scalaires a_1, a_2 : on peut en effet trouver une première sous-suite d'indices (n_k) telle que $(T_1 x_{n_k})$ soit convergente, puis une sous-sous-suite (n_{k_j}) telle que $T_2 x_{n_{k_j}}$ converge ; alors, la sous-suite $(a_1 T_1 + a_2 T_2)(x_{n_{k_j}})$ converge dans Y. Ainsi, *les opérateurs compacts de* X *dans* Y *forment un sous-espace vectoriel de* $\mathcal{L}(X, Y)$. Cet espace des opérateurs compacts sera noté $\mathcal{K}(X, Y)$. On a de plus le résultat important qui suit.

Théorème 1.1.1. *Le sous-espace vectoriel* $\mathcal{K}(X, Y)$ *est fermé dans* $\mathcal{L}(X, Y)$.

Preuve. — Commençons par une remarque simple :
si (K, d) *est un espace métrique compact,* $r > 0$ *et* $\varphi : I \to K$ *une application quelconque définie sur un ensemble* I infini*, il existe un point* $x \in K$ *tel que l'image inverse* $\varphi^{-1}\big(B(x, r)\big) \subset I$ *soit infinie.*
En effet, le compact K peut être couvert par un nombre fini de boules ouvertes $B(x, r)$, donc l'ensemble infini I est couvert par un nombre fini d'ensembles $\varphi^{-1}\big(B(x, r)\big)$ et il faut bien que l'un d'entre eux soit infini.

Supposons que $T \in \mathcal{L}(X, Y)$ soit adhérent au sous-ensemble $\mathcal{K}(X, Y)$ des opérateurs compacts ; pour tout entier $j \geq 0$, on peut trouver un opérateur compact T_j de X dans Y, tel que $\|T - T_j\| < 2^{-j}$. Considérons par ailleurs une suite quelconque $(x_n) \subset B_X$ et montrons que la suite image $(T x_n)$ admet des sous-suites convergentes dans Y. Ceci prouvera que T est compact de X dans Y.

On peut définir une suite décroissante de sous-ensembles infinis $M_j \subset \mathbb{N}$ tels que pour tout $j \geq 0$:

$$(m, n \in M_j) \Rightarrow \|T_j\, x_m - T_j\, x_n\| < 2^{-j}\,; \tag{1}$$

supposons que $M_{j-1} \subset \mathbb{N}$ soit défini (pour $j = 0$, convenons que $M_{-1} = \mathbb{N}$). L'application de la remarque préliminaire à la correspondance $m \in M_{j-1} \to T_j\, x_m \in \overline{T_j(B_X)}$ fournit un point $y_j \in Y$ et un ensemble infini $M_j \subset M_{j-1}$ tels que $\|T_j\, x_m - y_j\| < 2^{-j-1}$ pour tout $m \in M_j$, ce qui implique $\|T_j\, x_m - T_j\, x_n\| < 2^{-j}$ quand $m, n \in M_j$. On obtient ainsi la condition (1) par récurrence.

On définit ensuite une sous-suite diagonale ainsi : désignons par n_j le plus petit élément n de M_j tel que $n > n_{j-1}$. Alors la sous-suite $(T x_{n_j})$ est de Cauchy dans Y, donc convergente : en effet, pour tout entier $\ell > j$, on a $n_\ell \in M_\ell \subset M_j$, par conséquent $\|T_j\, x_{n_j} - T_j\, x_{n_\ell}\| < 2^{-j}$, ce qui donne $\|T x_{n_j} - T x_{n_\ell}\| < 3\,.\,2^{-j}$ par l'inégalité triangulaire, puisque $\|T x_n - T_j\, x_n\| \leq \|T - T_j\|\, \|x_n\| \leq \|T - T_j\| < 2^{-j}$.

///

Remarque. Les opérateurs de rang fini sont évidemment compacts (par application de Bolzano-Weierstrass dans $\mathbb{K}^n$). Toute limite en norme d'opérateur d'une suite d'opérateurs de rang fini est donc un opérateur compact([a]).

Les opérateurs compacts ont une propriété *d'idéal* : *si* S *et* T *sont composables et si l'un des deux est compact, la composition* $T \circ S$ *est compacte* ; en effet, si (x_n) est bornée dans l'espace de départ, on trouvera une sous-suite d'indices (n_k) telle que, ou bien $S\, x_{n_k}$ converge, si c'est S qui est compact, ou bien telle que $T(S\, x_{n_k})$ converge, si c'est T qui est compact (en appliquant la propriété de définition à la suite bornée $(S\, x_n)$) ; dans les deux cas on aura trouvé une sous-suite telle que $T(S\, x_{n_k})$ converge.

Exemple. L'espace $\mathrm{C}([0,1])$ des fonctions scalaires continues sur $[0,1]$ est un espace de Banach pour la norme uniforme, définie par

$$\|f\|_\infty = \max\{|f(t)| : t \in [0,1]\}.$$

L'espace $\mathrm{C}^1([0,1])$ des fonctions scalaires f continûment dérivables sur $[0,1]$ (en entendant par $f'(0)$ la dérivée à droite en 0 et par $f'(1)$ la dérivée à gauche en 1) est un espace de Banach pour la norme

$$\|f\|_{\mathrm{C}^1} = \|f'\|_\infty + \|f\|_\infty.$$

On peut considérer l'application linéaire d'injection $j : \mathrm{C}^1([0,1]) \to \mathrm{C}([0,1])$; elle est compacte d'après le théorème d'Arzela-Ascoli : l'image $\mathrm{A} \subset \mathrm{C}([0,1])$ de la boule unité de $\mathrm{C}^1([0,1])$ est formée de fonctions uniformément équicontinues, puisque

$$\forall f \in \mathrm{A},\ \forall s,t \in [0,1], \quad |f(t) - f(s)| \le |t-s|$$

par le théorème des accroissements finis (on a $\|f'\|_\infty \le 1$), et ces fonctions dans A sont uniformément bornées puisque $\|f\|_\infty \le 1$ pour toute $f \in \mathrm{A}$. L'ensemble A est donc relativement compact dans $\mathrm{C}([0,1])$. Il est facile de voir que cet ensemble image A n'est pas fermé en norme uniforme : en effet, son adhérence en norme uniforme contient des fonctions non dérivables, telles que la fonction $t \to |t - 1/2|$ par exemple. On a ici un exemple qui justifie la présence de *l'adhérence* de l'image dans la définition des opérateurs compacts.

Espaces de Hilbert : rappels

Si E est un espace vectoriel sur $\mathbb{C}$, une *forme sesquilinéaire* φ sur E est une application de $\mathrm{E} \times \mathrm{E}$ dans $\mathbb{K} = \mathbb{C}$ telle que

$\varphi(x_1 + \lambda x_2, y) = \varphi(x_1, y) + \lambda\varphi(x_2, y)$

et

$\varphi(x, y_1 + \lambda y_2) = \varphi(x, y_1) + \overline{\lambda}\varphi(x, y_2)$

pour tous vecteurs x, x_1, x_2, y, y_1, y_2 dans E et tout $\lambda \in \mathbb{C}$. Pour ne pas être obligé de faire une distinction entre espaces réels et espaces complexes, on peut considérer qu'une forme bilinéaire réelle sur un espace vectoriel réel est une forme sesquilinéaire. On dit que la forme sesquilinéaire φ est *hermitienne* si

$$\forall x, y \in \mathrm{E}, \quad \varphi(y, x) = \overline{\varphi(x, y)},$$

ce qui implique que $\varphi(x, x)$ est réel pour tout vecteur $x \in \mathrm{E}$; on dit qu'une forme sesquilinéaire φ hermitienne est *positive* si $\varphi(x, x) \ge 0$ pour tout $x \in \mathrm{E}$; on a alors l'inégalité de Cauchy-Schwarz

$$\forall x, y \in \mathrm{E}, \quad \bigl|\varphi(x, y)\bigr| \le \varphi(x, x)^{1/2}\, \varphi(y, y)^{1/2}.$$

Pour prouver cette inégalité, il suffit d'écrire $\varphi(\lambda x - \mu y, \lambda x - \mu y) \ge 0$, en choisissant $\lambda = \varphi(y, y)^{1/2}$ et $\mu = \mathrm{e}^{\mathrm{i}\theta}\, \varphi(x, x)^{1/2}$, où θ réel est choisi de façon que le nombre $\varphi(x, \mathrm{e}^{\mathrm{i}\theta}\, y)$ soit réel ≥ 0.

Un *espace de Hilbert* H sur $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$ est muni d'un produit scalaire

$$(x, y) \to \langle x, y\rangle \in \mathbb{K}$$

qui est une forme sesquilinéaire hermitienne positive sur H. L'application $x \to \langle x, y\rangle \in \mathbb{K}$ est donc linéaire pour tout $y \in \mathrm{H}$, on a aussi

$$\langle y, x\rangle = \overline{\langle x, y\rangle},$$

et $\langle x, x\rangle \geq 0$ pour tout vecteur $x \in \mathrm{H}$; on suppose de plus que $\langle x, x\rangle > 0$ pour tout $x \neq 0_{\mathrm{H}}$; la fonction $x \to \langle x, x\rangle^{1/2}$ est alors une norme $\|\cdot\|_{\mathrm{H}}$ sur H, pour laquelle H est supposé complet. L'inégalité de Cauchy-Schwarz s'écrit

$$\forall x, y \in \mathrm{H}, \quad \bigl|\langle x, y\rangle\bigr| \leq \langle x, x\rangle^{1/2} \, \langle y, y\rangle^{1/2}$$

ou bien $|\langle x, y\rangle| \leq \|x\|_{\mathrm{H}} \, \|y\|_{\mathrm{H}}$.

On dit que x *est orthogonal à* y, ce qui se note $x \perp y$, lorsque $\langle x, y\rangle = 0$; on dit que x est orthogonal à une partie A de H si x est orthogonal à tous les vecteurs de A. Pour toute partie $\mathrm{A} \subset \mathrm{H}$, on définit *l'orthogonal*

$$\mathrm{A}^{\perp} = \{x \in \mathrm{H} : x \perp \mathrm{A}\}$$

de cette partie. L'orthogonal $\mathrm{A}^{\perp}$ est toujours un sous-espace vectoriel fermé de H.

Soit Y un sous-espace vectoriel fermé de l'espace de Hilbert H ; pour tout $x \in \mathrm{H}$, il existe un unique vecteur $\mathrm{P_Y} x$ qui vérifie les deux propriétés suivantes : $\mathrm{P_Y} x \in \mathrm{Y}$ et $x - \mathrm{P_Y} x \perp \mathrm{Y}$. L'application $x \to \mathrm{P_Y} x$ est linéaire ; on l'appelle la *projection orthogonale* de H sur Y. On a $\|\mathrm{P_Y}\| \leq 1$. On voit que

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

Si Z est un sous-espace vectoriel (non nécessairement fermé) de H, on voit que Z est dense dans H si (et seulement si) 0_{H} est le seul vecteur de H orthogonal à Z.

Grâce à ces considérations, on montre que toute forme linéaire continue sur un espace de Hilbert H provient du produit scalaire avec un vecteur de H : on peut définir pour tout vecteur $x \in \mathrm{H}$ la forme linéaire continue $i_{\mathrm{H}} \, x$ sur H par

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

On voit facilement que $\|i_{\mathrm{H}} x\| \leq \|x\|$ par Cauchy-Schwarz, et on voit que $\|i_{\mathrm{H}} x\| = \|x\|$ en appliquant $i_{\mathrm{H}} x$ à x lui-même.

Proposition. *Toute forme linéaire continue ℓ sur l'espace de Hilbert* H *est de la forme $\ell = i_{\mathrm{H}} x$ pour un certain $x \in \mathrm{H}$.*

Si la forme linéaire continue ℓ n'est pas nulle, on trouve facilement([b]) le vecteur x tel que $\ell = i_{\mathrm{H}} x$ en considérant le noyau (de dimension 1) de la projection orthogonale sur l'hyperplan fermé $\ker \ell$. L'application $i_{\mathrm{H}} : x \to i_{\mathrm{H}} \, x$ est donc une bijection antilinéaire isométrique de H sur son dual topologique H^* (on dit qu'une application u est *antilinéaire* quand elle vérifie $u(x + \lambda y) = ux + \overline{\lambda} \, uy$ pour tous les vecteurs x, y et tout scalaire λ).

Lorsque des vecteurs $x_1, \ldots, x_n \in \mathrm{H}$ sont deux à deux orthogonaux, on a

$$\|x_1 + \cdots + x_n\|^2 = \|x_1\|^2 + \cdots + \|x_n\|^2 \, ;$$

lorsque $(x_n)_{n\in\mathbb{N}}$ est une suite de vecteurs deux à deux orthogonaux, il en résulte que la série de vecteurs $\sum x_n$ converge dans H si et seulement si la série numérique $\sum \|x_n\|^2$ converge.

Une famille de vecteurs $(e_i)_{i\in\mathrm{I}}$ est une *base hilbertienne* de H si ces vecteurs sont de norme 1, deux à deux orthogonaux, et si le sous-espace vectoriel $\mathrm{Vect}(e_i, i \in \mathrm{I})$ qu'ils engendrent est dense dans H. Il en résulte que tout vecteur x admet une représentation unique de la forme

$$x = \sum_{i\in\mathrm{I}} c_i e_i.$$

Cette notation signifie que pour tout $\varepsilon > 0$, il existe un sous-ensemble fini $\mathrm{J}_0 \subset \mathrm{I}$ tel que pour tout sous-ensemble fini $\mathrm{J} \supset \mathrm{J}_0$ de I, on ait

$$\Big\|x - \sum_{j\in\mathrm{J}} c_j e_j\Big\| < \varepsilon\,;$$

on dit que $\sum_{i\in\mathrm{I}} c_i e_i$ est la somme de la *famille sommable* $(c_i e_i)_{i\in\mathrm{I}}$. Les coefficients sont donnés par $c_i = \langle x, e_i\rangle$ pour tout $i \in \mathrm{I}$. Tout espace de Hilbert admet une base hilbertienne. Lorsque H est de dimension infinie et séparable, toutes ses bases hilbertiennes sont dénombrables. Dans ce cas on peut se passer de la notion de famille sommable et écrire simplement des séries convergentes : pour toute énumération $(e_n)_{n\in\mathbb{N}}$ d'une base hilbertienne de H, on a pour tout vecteur $x \in \mathrm{H}$

$$x = \sum_{n=0}^{+\infty} \langle x, e_n\rangle\, e_n = \lim_n \sum_{k=0}^{n} \langle x, e_k\rangle\, e_k.$$

L'espace $\mathrm{L}_2(0,1)$ est un espace de Hilbert pour le produit scalaire

$$\langle f, g\rangle_{\mathrm{L}_2} = \int_0^1 f(t)\,\overline{g(t)}\,\mathrm{d}t.$$

De même $\mathrm{L}_2(\mathbb{R})$ ou $\mathrm{L}_2(\mathbb{R}^d)$ sont des espaces de Hilbert.

Questions de densité

Pour pouvoir travailler avec $\mathrm{L}_2(\mathbb{R}^d)$, ou plus généralement avec les espaces de la forme $\mathrm{L}_p(\mathbb{R}^d)$, il est important de connaître la densité de fonctions plus simples : les fonctions en escalier (de la théorie de l'intégrale de Riemann) sont denses dans $\mathrm{L}_p(\mathbb{R})$ pour tout $p \in [1, +\infty[$; les fonctions continues à support compact, ou bien l'espace $\mathcal{D}$ des fonctions C^∞ à support compact, sont denses dans $\mathrm{L}_p(\mathbb{R})$ ou $\mathrm{L}_p(\mathbb{R}^d)$ pour tout $p \in [1, +\infty[$.

La densité des fonctions continues provient de la définition de la tribu borélienne, engendrée par les ensembles ouverts ; on peut procéder en montrant que les sous-ensembles A de $[0,1]$ (par exemple) dont la fonction indicatrice $\mathbf{1}_\mathrm{A}$ est limite dans $\mathrm{L}_1(0,1)$ d'une suite de fonctions continues (φ_n) telles que $0 \le \varphi_n \le 1$, forment une tribu de parties de $[0,1]$ qui contient les intervalles. La densité des fonctions de classe C^∞ se montre en général par la technique de *régularisation par convolution*.

1.2. Opérateurs sur un espace de Hilbert

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

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

On voit que $\mathrm{T}^* y = (i_{\mathrm{H}})^{-1}\big(i_{\mathrm{K}}\, y \circ \mathrm{T}\big)$. On vérifie facilement les propriétés suivantes :

$$(\mathrm{Id})^* = \mathrm{Id},\ (\lambda \mathrm{T})^* = \overline{\lambda}\, \mathrm{T}^*,\ (\mathrm{T}_1 + \mathrm{T}_2)^* = \mathrm{T}_1^* + \mathrm{T}_2^*,\ (\mathrm{T}^*)^* = \mathrm{T},\ (\mathrm{S}\mathrm{T})^* = \mathrm{T}^*\mathrm{S}^*.$$

De plus, on a $\|\mathrm{T}^*\| = \|\mathrm{T}\|$; l'application $\mathrm{T} \to \mathrm{T}^*$ est une bijection isométrique antilinéaire de $\mathcal{L}(\mathrm{H}, \mathrm{K})$ sur $\mathcal{L}(\mathrm{K}, \mathrm{H})$.

Exemples.

1. Considérons l'espace de Hilbert $\mathrm{L}_2(0,1)$. La multiplication par une fonction mesurable bornée $\varphi \in \mathrm{L}_\infty([0,1])$ définit un opérateur M_φ borné sur $\mathrm{L}_2(0,1)$,

$$\forall t \in (0,1), \quad (\mathrm{M}_\varphi f)(t) = \varphi(t) f(t)\,;$$

l'adjoint de M_φ est la multiplication par la fonction complexe conjuguée $\overline{\varphi}$.

2. Le *shift à droite* S sur l'espace $\mathrm{H} = \ell_2(\mathbb{N})$ est défini en posant pour tout vecteur $x = (x_n)_{n \ge 0} \in \mathrm{H}$

$$(\mathrm{S}x)_n = x_{n-1} \ \text{ si } \ n > 0, \quad (\mathrm{S}x)_0 = 0\,;$$

le vecteur $\mathrm{S}x$ est obtenu en décalant toutes les coordonnées d'un cran vers la droite, et en introduisant 0 à la place ouverte à la coordonnée 0,

$$\mathrm{S}(x_0, x_1, x_2, \ldots) = (0, x_0, x_1, x_2, \ldots)\,;$$

l'opérateur S est une isométrie ; son adjoint S^* est le shift à gauche,

$$\forall n \ge 0, \quad (\mathrm{S}^* x)_n = x_{n+1}.$$

On a $\mathrm{S}^*\mathrm{S} = \mathrm{Id}_{\mathrm{H}}$, relation qui traduit le caractère isométrique de S.

On peut aussi envisager le *shift à droite bilatéral* $\mathrm{S}_{\mathbb{Z}}$ sur l'espace $\mathrm{H} = \ell_2(\mathbb{Z})$, qui est défini en posant pour tout vecteur $x = (x_n)_{n \in \mathbb{Z}} \in \mathrm{H}$ et tout $n \in \mathbb{Z}$

$$(\mathrm{S}_{\mathbb{Z}} x)_n = x_{n-1}\,;$$

il s'agit d'un *opérateur unitaire*, vérifiant $\mathrm{S}_{\mathbb{Z}}^*\mathrm{S}_{\mathbb{Z}} = \mathrm{S}_{\mathbb{Z}}\mathrm{S}_{\mathbb{Z}}^* = \mathrm{Id}_{\mathrm{H}}$, bijection isométrique de $\ell_2(\mathbb{Z})$ sur lui-même. Évidemment, l'inverse de $\mathrm{S}_{\mathbb{Z}}$, qui est égal à $\mathrm{S}_{\mathbb{Z}}^*$, est le shift à gauche dans $\ell_2(\mathbb{Z})$.

Définition. Un endomorphisme continu T d'un espace de Hilbert réel ou complexe H est dit *hermitien* si $\mathrm{T}^* = \mathrm{T}$, ce qui revient très exactement à dire que

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

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

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

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

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

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

est réel pour tout $x \in \mathrm{H}$. Il en résulte immédiatement que les valeurs propres (éventuelles) d'un opérateur hermitien sont réelles.

On dit que $\mathrm{A} \in \mathcal{L}(\mathrm{H})$ hermitien est *positif* lorsque

$$\forall x \in \mathrm{H}, \quad \langle \mathrm{A}x, x\rangle \geq 0.$$

Exemples. Les opérateurs $\mathrm{T} + \mathrm{T}^*$, $\mathrm{i}(\mathrm{T} - \mathrm{T}^*)$ (si $\mathbb{K} = \mathbb{C}$) sont hermitiens, ce qui permet (dans le cas complexe) d'écrire tout opérateur $\mathrm{T} \in \mathcal{L}(\mathrm{H})$ sous la forme $\mathrm{T} = \mathrm{A} + \mathrm{i}\mathrm{B}$, A et B hermitiens. Tout projecteur orthogonal Q est hermitien. On voit que $\mathrm{T}^*\mathrm{T}$ et $\mathrm{T}\,\mathrm{T}^*$ sont hermitiens positifs pour tout $\mathrm{T} \in \mathcal{L}(\mathrm{H})$. Les projecteurs orthogonaux sont donc positifs, puisque $\mathrm{Q} = \mathrm{Q}^2 = \mathrm{Q}^*\mathrm{Q}$. Plus généralement, le carré d'un hermitien est hermitien positif.

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

Exemples. Les opérateurs unitaires ($\mathrm{U}^* = \mathrm{U}^{-1}$) sont évidemment normaux. La multiplication M_φ par une fonction mesurable bornée φ est un opérateur normal sur $\mathrm{L}_2(0, 1)$, car

$$(\mathrm{M}_\varphi)^*\,\mathrm{M}_\varphi = \mathrm{M}_{\overline{\varphi}}\,\mathrm{M}_\varphi = \mathrm{M}_{\overline{\varphi}\varphi} = \mathrm{M}_\varphi\,(\mathrm{M}_\varphi)^*.$$

L'opérateur M_φ est hermitien quand la fonction φ est réelle.

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

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

Si $\mathrm{T}x = \lambda x$, il en résulte que $\mathrm{T}^*x = \overline{\lambda}x$: en appliquant ce qui précède à l'opérateur normal $\mathrm{S} = \mathrm{T} - \lambda\,\mathrm{Id}_\mathrm{H}$, et en notant que l'adjoint de $\lambda\,\mathrm{Id}_\mathrm{H}$ est $\overline{\lambda}\,\mathrm{Id}_\mathrm{H}$, on obtient

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

Lemme 1.2.1. *Si* $\mathrm{A} \in \mathcal{L}(\mathrm{H})$ *est hermitien positif, on a pour tout vecteur* $x \in \mathrm{H}$

$$\|\mathrm{A}x\| \le \|\mathrm{A}\|^{1/2} \langle \mathrm{A}x, x\rangle^{1/2}.$$

Preuve. — Si $\mathrm{A}x = 0$ c'est évident ; sinon, soit $y = \|\mathrm{A}x\|^{-1}\,\mathrm{A}x$; on a $\|y\| = 1$ et par Cauchy-Schwarz appliqué à la forme hermitienne positive $(u, v) \to \langle \mathrm{A}u, v\rangle$

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

///

Lemme 1.2.2. *Si* T *est hermitien compact sur un espace de Hilbert* $\mathrm{H} \ne \{0\}$, *et si le nombre*

$$\mu = \sup\{\langle \mathrm{T}x, x\rangle : x \in \mathrm{H},\ \|x\| = 1\}$$

est non nul, l'opérateur T *admet* μ *pour valeur propre.*

Preuve. — On peut trouver une suite de vecteurs $(x_n) \subset \mathrm{H}$ telle que $\|x_n\| = 1$ pour tout n et $\mu = \lim \langle \mathrm{T}x_n, x_n\rangle$; puisque T est compact on peut supposer, quitte à passer à une sous-suite, que $\mathrm{T}x_n$ converge vers un vecteur $y \in \mathrm{H}$; considérons l'opérateur

$$\mathrm{A} = \mu\,\mathrm{Id} - \mathrm{T}\,;$$

puisque μ est réel et T hermitien, A est hermitien et pour tout x de norme un on a

$$\langle \mathrm{A}x, x\rangle = \mu\,\langle x, x\rangle - \langle \mathrm{T}x, x\rangle = \mu - \langle \mathrm{T}x, x\rangle \ge 0\,;$$

il en résulte que $\langle \mathrm{A}x, x\rangle \ge 0$ pour tout x par homogénéité, donc A est hermitien positif ; de plus,

$$\langle \mathrm{A}x_n, x_n\rangle = \mu - \langle \mathrm{T}x_n, x_n\rangle \to 0.$$

Par le lemme 1.2.1, il en résulte que $\mathrm{A}x_n = \mu x_n - \mathrm{T}x_n$ tend vers 0 ; puisque $\mathrm{T}x_n$ tend vers y on déduit que la suite (μx_n) tend aussi vers y. Enfin, la condition $\mu \ne 0$ nous permet de conclure que la suite (x_n) elle-même tend vers le vecteur $x = \mu^{-1}y$. Par continuité, $\|x\| = 1$, donc $x \ne 0$ et $\mathrm{A}x = 0$, c'est-à-dire

$$\mathrm{T}x = \mu x.$$

///

Lemme 1.2.3. *Si* $\mathrm{T} \in \mathcal{L}(\mathrm{H})$ *est compact, le sous-espace propre* $\mathrm{E}_\lambda = \ker(\mathrm{T} - \lambda\,\mathrm{Id})$ *est de dimension finie pour tout* $\lambda \ne 0$.

Preuve. — Il s'agit en fait d'un résultat valable pour tout Banach, mais c'est plus facile pour un Hilbert. Si E_λ est de dimension infinie, on peut trouver dans E_λ une suite orthonormée infinie $(e_n)_{n\ge 0}$; les vecteurs image vérifient pour $m \ne n$

$$\|\mathrm{T}e_n - \mathrm{T}e_m\|^2 = \|\lambda e_n - \lambda e_m\|^2 = 2\,|\lambda|^2.$$

Si $\lambda \ne 0$, la suite $(\mathrm{T}e_n)$ ne peut pas avoir de sous-suite de Cauchy, donc pas de sous-suite convergente et T ne peut être compact.

///

Lemme 1.2.4. *Si* T *est un opérateur compact sur un espace de Hilbert* $\mathrm{H} \neq \{0\}$, *et si*

- *ou bien :* $\mathbb{K} = \mathbb{R}$ *ou* $\mathbb{C}$, *et* T *hermitien,*
- *ou bien :* $\mathbb{K} = \mathbb{C}$, *et* T *normal,*

l'opérateur T *admet une valeur propre de module égal à* $\|\mathrm{T}\|$.

Preuve. — Si $\mathrm{T} = 0$, tous les vecteurs non nuls conviennent comme vecteurs propres de la valeur propre $0 = \|\mathrm{T}\|$. Sinon, on pose $r = \|\mathrm{T}\| > 0$ et on introduit l'opérateur compact hermitien positif $\mathrm{A} = \mathrm{T}^*\mathrm{T}$; considérons

$$\begin{aligned} \mu &= \sup\{\langle \mathrm{A}x, x\rangle : x \in \mathrm{H},\ \|x\| = 1\} \\ &= \sup_{\|x\|=1} \langle \mathrm{T}^*\mathrm{T}x, x\rangle = \sup_{\|x\|=1} \langle \mathrm{T}x, \mathrm{T}x\rangle = r^2 > 0. \end{aligned}$$

D'après le lemme 1.2.2, l'espace propre $\mathrm{F} = \ker(\mathrm{A} - r^2\,\mathrm{Id})$ est non-nul. Pour continuer, on note dans le cas hermitien que $\mathrm{A} = \mathrm{T}^2$ et que pour $x_0 \in \mathrm{F}$ non nul,

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

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

Dans le cas où T est normal, on note que T commute avec A, donc le sous-espace propre F de A est stable par T, et F est de dimension finie par le lemme 1.2.3. La restriction de T à F définit un endomorphisme en dimension finie complexe, qui admet une valeur propre $\lambda \in \mathbb{C}$ d'après le théorème de d'Alembert ; on sait si $\mathrm{T}x = \lambda x$ que $\mathrm{T}^*x = \overline{\lambda}x$ (d'après la remarque qui suit la définition des opérateurs normaux) donc $\mathrm{A}x = |\lambda|^2 x$ et il en résulte que $|\lambda| = r$.

///

Remarque. Dans le cas réel, la diagonalisation des opérateurs normaux n'est pas toujours possible : dans $\mathbb{R}^2$ déjà, l'opérateur de rotation d'angle $\pi/2$ est normal mais n'a pas de vecteur propre.

Diagonalisation de T *compact hermitien ou normal*

Théorème. *Soient* H *un espace de Hilbert non nul, et* $\mathrm{T} \in \mathcal{L}(\mathrm{H})$ *un opérateur compact ; on suppose que*

- *ou bien :* $\mathbb{K} = \mathbb{R}$ *ou* $\mathbb{C}$, *et* T *hermitien,*
- *ou bien :* $\mathbb{K} = \mathbb{C}$, *et* T *normal.*

Il existe une base hilbertienne de H *formée de vecteurs propres de l'opérateur* T *; pour tout* $\varepsilon > 0$, *il n'existe qu'un nombre fini de vecteurs de cette base pour lesquels la valeur propre correspondante est de module* $\geq \varepsilon$ *(les valeurs propres sont réelles dans le cas hermitien).*

Preuve. — Pour gagner de la place on écrira le cas hermitien comme cas particulier du cas normal. Considérons l'ensemble $\mathcal{U}$ de toutes les parties $\mathrm{U} \subset \mathrm{H}$ qui vérifient les trois propriétés suivantes :

- si $x \in \mathrm{U}$, $\|x\| = 1$;
- si $x, y \in \mathrm{U}$ et $x \neq y$, alors $x \perp y$;
- si $x \in \mathrm{U}$, $\mathrm{T}x \in \mathbb{K}x$.

On ordonne $\mathcal{U}$ par l'inclusion des parties de H ; muni de cet ordre, l'ensemble $\mathcal{U}$ est inductif : pour toute famille $(\mathrm{U}_i)_{i\in\mathrm{I}}$ d'éléments de $\mathcal{U}$ qui est totalement ordonnée par inclusion, on peut trouver dans $\mathcal{U}$ un majorant pour cette famille, à savoir le sous-ensemble

$$\mathrm{U} = \bigcup_{i\in\mathrm{I}} \mathrm{U}_i \subset \mathrm{H},$$

pour lequel on montre facilement que $\mathrm{U} \in \mathcal{U}$. D'après le lemme de Zorn, l'ensemble $\mathcal{U}$ admet des éléments maximaux ; soit U_0 un tel élément maximal ; les deux premières propriétés montrent que les éléments de U_0 constituent un système orthonormé ; on va de plus montrer que U_0 engendre un espace dense dans H, c'est-à-dire que U_0 est une base hilbertienne de H. Enfin, la troisième propriété dit que les vecteurs de cette base sont vecteurs propres de T.

Considérons le sous-espace vectoriel $\mathrm{H}_0 = \mathrm{Vect}(\mathrm{U}_0)$; puisque tous les vecteurs x de U_0 vérifient $\mathrm{T}x = \lambda x \in \mathrm{H}_0$ et $\mathrm{T}^*x = \overline{\lambda}x \in \mathrm{H}_0$ pour un certain $\lambda \in \mathbb{K}$, il est clair que H_0 est stable par T et T^* ; il en résulte que le sous-espace fermé $\mathrm{H}_1 = \mathrm{H}_0^\perp$ est stable par T^* et T ; en effet, si $y \perp \mathrm{H}_0$ on a pour tout $h_0 \in \mathrm{H}_0$

$$\langle \mathrm{T}^*y, h_0\rangle = \langle y, \mathrm{T}h_0\rangle = 0$$

parce que $\mathrm{T}h_0 \in \mathrm{H}_0$, donc $\mathrm{T}^*y \perp \mathrm{H}_0$ et de même on a $\mathrm{T}y \perp \mathrm{H}_0$. Considérons l'opérateur T_1 défini sur l'espace de Hilbert H_1 en posant

$$\forall h_1 \in \mathrm{H}_1, \quad \mathrm{T}_1 h_1 = \mathrm{T}h_1 \in \mathrm{H}_1.$$

On voit facilement que $\mathrm{T}_1 \in \mathcal{L}(\mathrm{H}_1)$ est normal et compact. D'après le lemme 1.2.4, on peut trouver si $\mathrm{H}_1 \neq \{0\}$ un vecteur $v_1 \in \mathrm{H}_1$ de norme un tel que $\mathrm{T}v_1 \in \mathbb{K}v_1$. Mais alors $\mathrm{U} = \mathrm{U}_0 \cup \{v_1\}$ est un élément de $\mathcal{U}$ qui contredit la maximalité de U_0. On en déduit que $\mathrm{H}_1 = \{0\}$, ce qui implique que H_0 est dense, et que U_0 fournit une base hilbertienne de H formée de vecteurs propres de T.

Réécrivons cette base de vecteurs propres sous la forme « indexée » plus habituelle $(e_i)_{i\in\mathrm{I}}$ (en prenant par exemple $\mathrm{I} = \mathrm{U}_0$ et $e_i = i$ pour tout i !), écrivons $\mathrm{T}e_i = \lambda_i e_i$ pour tout $i \in \mathrm{I}$ et considérons pour $\varepsilon > 0$ l'ensemble

$$\mathrm{I}_\varepsilon = \{i \in \mathrm{I} : |\lambda_i| \geq \varepsilon\}.$$

Si $i, j \in \mathrm{I}_\varepsilon$ et $i \neq j$ on a

$$\|\mathrm{T}e_i - \mathrm{T}e_j\|^2 = \|\lambda_i e_i - \lambda_j e_j\|^2 = |\lambda_i|^2 + |\lambda_j|^2 \geq 2\varepsilon^2 ;$$

la famille $(\mathrm{T}e_i)_{j\in\mathrm{I}_\varepsilon}$ est contenue dans le compact $\overline{\mathrm{T}(\mathrm{B_H})}$, et formée de points dont les distances mutuelles sont $\geq \sqrt{2}\varepsilon$: une telle famille est nécessairement finie.

Si on n'aime pas le lemme de Zorn on peut procéder comme suit. On supposera $\dim \mathrm{H} = +\infty$ pour ne pas avoir à distinguer plusieurs cas. On construit par récurrence une suite croissante (E_n) de sous-espaces de dimension finie et une suite décroissante (X_n) de sous-espaces fermés de dimension infinie, stables par T et T^*. On amorce l'induction avec $\mathrm{E}_0 = \{0\}$.

Si $n \geq 0$ et si $\mathrm{E}_n = \mathrm{Vect}(e_1, \ldots, e_n)$ est donné, engendré par des vecteurs propres de T, on raisonne ainsi : on a $\mathrm{E}_n \neq \mathrm{H}$ puisque $\dim \mathrm{H} = +\infty$, donc $\mathrm{X}_{n+1} = \mathrm{E}_n^\perp$ n'est pas réduit à $\{0\}$. Puisque E_n est stable par T et T^*, on a vu que X_{n+1} est stable par T^* et T, et la restriction $\mathrm{T}_{n+1} = \mathrm{T}_{|\mathrm{X}_{n+1}} \in \mathcal{L}(\mathrm{X}_{n+1})$ est normale et compacte ; elle admet donc par le lemme 1.2.4 une valeur propre λ_{n+1} telle que $|\lambda_{n+1}| = \|\mathrm{T}_{n+1}\|$, et il existe un vecteur $e_{n+1} \in \mathrm{X}_{n+1}$, qu'on peut choisir de norme 1, tel que $\mathrm{T}e_{n+1} = \mathrm{T}_{n+1}e_{n+1} = \lambda_{n+1}e_{n+1}$. On pose $\mathrm{E}_{n+1} = \mathrm{Vect}(e_1, \ldots, e_n, e_{n+1})$ et la boucle est bouclée.

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

On a $\mathrm{T}e_n = \lambda_n e_n$ avec $|\lambda_n| = \|\mathrm{T}_n\|$ pour tout $n \geq 1$. La suite $(\|\mathrm{T}_n\|)$ est décroissante (car la suite (X_n) est décroissante), donc converge vers un nombre $d \geq 0$. Par la compacité on doit avoir $d = 0$ (ceci résulte de la finitude de I_ε démontrée plus haut). L'espace fermé E_∞ engendré par cette suite de vecteurs propres (e_n) est stable par T et T^*, ainsi que son orthogonal X_∞ ; on voit que la restriction T_∞ de T à X_∞ est nulle, car $\|\mathrm{T}_\infty\| \leq \|\mathrm{T}_n\|$ pour tout n, ce qui entraîne $\|\mathrm{T}_\infty\| = 0$. On a diagonalisé T : il suffit de compléter la suite orthonormée (e_n) par une base hilbertienne $(f_j)_{j\in \mathrm{J}}$ de X_∞, où l'ensemble J peut être vide, dans le cas où $\mathrm{X}_\infty = \{0\}$; on aura $\mathrm{T}f_j = 0$ pour tout indice $j \in \mathrm{J}$.

///

Remarque. Si $\lambda_n \neq 0$ pour tout n, l'espace X_∞ ci-dessus est égal au noyau de T ; on aura au contraire $\lambda_n = 0$ à partir d'un certain rang quand T est de rang fini.

Exercice 1.2.5. Si P désigne l'opérateur de primitive sur $\mathrm{L}_2(0,1)$, défini par

$$\forall f \in \mathrm{L}_2(0,1), \quad (\mathrm{P}f)(t) = \int_0^t f(s)\,\mathrm{d}s,$$

diagonaliser l'opérateur hermitien compact $\mathrm{P}^*\mathrm{P}$.

1.3. Opérateurs de Hilbert-Schmidt

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

Lemme. *Soient* H *et* K *deux espaces de Hilbert,* B *une base hilbertienne de* H *et* C *une base hilbertienne de* K *; pour tout* $\mathrm{T} \in \mathcal{L}(\mathrm{H}, \mathrm{K})$ *on a :*

$$\sum_{b\in \mathrm{B}} \|\mathrm{T}b\|^2 = \sum_{c\in \mathrm{C}} \|\mathrm{T}^*c\|^2$$

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

Preuve.— Pour $x \in \mathrm{H}$ et $y \in \mathrm{K}$ on a

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

et on en déduit que

$$\sum_{b \in \mathrm{B}} \|\mathrm{T}b\|^2 = \sum_{b \in \mathrm{B},\, c \in \mathrm{C}} |\langle c, \mathrm{T}b\rangle|^2 = \sum_{b \in \mathrm{B},\, c \in \mathrm{C}} |\langle \mathrm{T}^* c, b\rangle|^2 = \sum_{c \in \mathrm{C}} \|\mathrm{T}^* c\|^2.$$

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

///

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

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

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

Remarques.

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

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

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

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

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

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

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

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

4. Les opérateurs de Hilbert-Schmidt sont compacts.

Il suffit de montrer que si $\mathrm{T} \in \mathcal{L}(\mathrm{H}, \mathrm{K})$ est de Hilbert-Schmidt, il est adhérent à l'ensemble des opérateurs de rang fini : pour tout $\varepsilon > 0$, il existe S de rang fini tel que $\|\mathrm{T} - \mathrm{S}\| < \varepsilon$. Il suffit de trouver S tel que $\|\mathrm{T} - \mathrm{S}\|_{\mathrm{HS}} < \varepsilon$, puisque $\|\mathrm{T} - \mathrm{S}\| \le \|\mathrm{T} - \mathrm{S}\|_{\mathrm{HS}}$.

Soit $(e_i)_{i \in \mathrm{I}}$ une base hilbertienne de H, et posons $r = \|\mathrm{T}\|_{\mathrm{HS}}$; on peut trouver un ensemble fini $\mathrm{J} \subset \mathrm{I}$ tel que

$$\sum_{j \in \mathrm{J}} \|\mathrm{T} e_j\|^2 > r^2 - \varepsilon^2, \quad \text{donc} \quad \sum_{i \notin \mathrm{J}} \|\mathrm{T} e_i\|^2 < \varepsilon^2.$$

Considérons l'espace de dimension finie $\mathrm{F}_1 = \mathrm{Vect}(e_j : j \in \mathrm{J})$ et son orthogonal $\mathrm{F}_2 = \mathrm{F}_1^{\perp}$, qui admet $(e_i)_{i \notin \mathrm{J}}$ pour base hilbertienne. Soient P_1 et P_2 les projections orthogonales sur F_1 et F_2 ; on a $\mathrm{P}_1 + \mathrm{P}_2 = \mathrm{Id}_{\mathrm{H}}$. Posons $\mathrm{S} = \mathrm{T}\mathrm{P}_1$; c'est un opérateur de rang fini, et $\mathrm{T} - \mathrm{S} = \mathrm{T}\mathrm{P}_2$. Puisque P_2 annule tous les vecteurs e_j, $j \in \mathrm{J}$, et que $\mathrm{P}_2\, e_i = e_i$ quand $i \notin \mathrm{J}$, on obtient

$$\|\mathrm{T} - \mathrm{S}\|_{\mathrm{HS}}^2 = \sum_{i \in \mathrm{I}} \|\mathrm{T}\mathrm{P}_2 e_i\|^2 = \sum_{i \notin \mathrm{J}} \|\mathrm{T} e_i\|^2 < \varepsilon^2.$$

Exemples 1.3.1.

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

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

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

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

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

Compléments sur les espaces mesurés produits

On suppose que $(\mathrm{S}, \mathcal{A})$ et $(\mathrm{T}, \mathcal{B})$ sont deux espaces mesurables ; on définit la tribu $\mathcal{A} \otimes \mathcal{B}$ de parties de $\mathrm{S} \times \mathrm{T}$ engendrée par les pavés mesurables $\mathrm{A} \times \mathrm{B}$, $\mathrm{A} \in \mathcal{A}$ et $\mathrm{B} \in \mathcal{B}$. On supposera que $(\mathrm{S}, \mathcal{A}, \mu)$ et $(\mathrm{T}, \mathcal{B}, \nu)$ sont deux espaces mesurés *σ-finis*, c'est-à-dire que $\mathrm{S} = \bigcup_n \mathrm{S}_n$, avec $\mathrm{S}_n \in \mathcal{A}$ et $\mu(\mathrm{S}_n) < +\infty$ pour tout n, et la même chose pour ν. On sait alors définir une mesure $\mu \otimes \nu$ sur cette tribu $\mathcal{A} \otimes \mathcal{B}$. L'outil fondamental pour travailler avec cette mesure produit tensoriel est le théorème de Fubini, qui permet de calculer l'intégrale de toute fonction $\mathcal{A} \otimes \mathcal{B}$-mesurable positive f définie sur $\mathrm{S} \times \mathrm{T}$, par la formule

$$\int_{\mathrm{S} \times \mathrm{T}} f \,\mathrm{d}(\mu \otimes \nu) = \int_{\mathrm{S}} \Bigl(\int_{\mathrm{T}} f(s,t) \,\mathrm{d}\nu(t)\Bigr) \mathrm{d}\mu(s),$$

(ou bien la formule symétrique obtenue en échangeant l'ordre des intégrations), formule où les valeurs infinies sont admises.

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

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

Avec Fubini on voit que $\|f \otimes g\|_{\mathrm{L}_2(\mathrm{S}\times\mathrm{T})} = \|f\|_{\mathrm{L}_2(\mathrm{S})} \|g\|_{\mathrm{L}_2(\mathrm{T})}$, et plus généralement,

$$\langle f_1 \otimes g_1, f_2 \otimes g_2 \rangle_{\mathrm{L}_2(\mathrm{S}\times\mathrm{T})} = \langle f_1, f_2 \rangle_{\mathrm{L}_2(\mathrm{S})} \langle g_1, g_2 \rangle_{\mathrm{L}_2(\mathrm{T})}.$$

On notera $\mathbf{1}_{\mathrm{A}}$ la fonction indicatrice d'un sous-ensemble A, égale à 1 sur l'ensemble A et à 0 en dehors de A.

Lemme 1.3.2. *Toute fonction $\mathbf{1}_{\mathrm{C}}$, $\mathrm{C} \in \mathcal{A} \otimes \mathcal{B}$ de mesure finie, peut être approchée dans $\mathrm{L}_2(\mathrm{S} \times \mathrm{T}, \mu \otimes \nu)$ par des combinaisons linéaires de fonctions $\mathbf{1}_{\mathrm{A}} \otimes \mathbf{1}_{\mathrm{B}}$, avec des ensembles $\mathrm{A} \in \mathcal{A}$, $\mathrm{B} \in \mathcal{B}$ de mesure finie.*

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

///

Opérateurs de Hilbert-Schmidt définis par un noyau $k(s,t)$

Soient (S, μ) et (T, ν) deux espaces mesurés σ-finis et $k(s,t)$ une fonction de carré intégrable sur $\mathrm{S} \times \mathrm{T}$; on désignera par $\|k\|_2$ la norme de k dans $\mathrm{L}_2(\mathrm{S} \times \mathrm{T}, \mu \otimes \nu)$; on va montrer pour commencer que pour toute fonction $f \in \mathrm{L}_2(\mathrm{T}, \nu)$,

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

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

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

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

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

ce qui donne en réintégrant

$$\int_{\mathrm{S}} \Big(\int_{\mathrm{T}} |k(s,t)|\,|f(t)|\,\mathrm{d}\nu(t) \Big)^2 \,\mathrm{d}\mu(s) \le \Big(\int_{\mathrm{S}\times\mathrm{T}} |k(s,t)|^2 \,\mathrm{d}\mu(s)\,\mathrm{d}\nu(t) \Big) \,\big(\int_{\mathrm{T}} |f(t)|^2 \,\mathrm{d}\nu(t) \big)$$

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

$$\int_{\mathrm{S}} |(\mathrm{T}_k f)(s)|^2 \,\mathrm{d}\mu(s) \le \Big(\int_{\mathrm{S}\times\mathrm{T}} |k(s,t)|^2 \,\mathrm{d}\mu(s)\,\mathrm{d}\nu(t) \Big) \Big(\int_{\mathrm{T}} |f(t)|^2 \,\mathrm{d}\nu(t) \Big) = \|k\|_2^2 \,\|f\|^2.$$

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

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

Utilisation de bases hilbertiennes

Supposons pour fixer les idées que les deux espaces $\mathrm{L}_2(\mathrm{S},\mu)$ et $\mathrm{L}_2(\mathrm{T},\nu)$ soient de dimension infinie et séparables (c'est de très loin le cas le plus important). On peut alors trouver deux bases hilbertiennes infinies dénombrables, que l'on peut indexer par $\mathbb{N}$.

Soit $k(s,t) \in \mathrm{L}_2(\mathrm{S}\times\mathrm{T}, \mu\otimes\nu)$ et soit $(\psi_n)_{n\ge 0}$ une base hilbertienne de $\mathrm{L}_2(\mathrm{T},\nu)$; l'étude précédente de l'opérateur de noyau k indique que la formule

$$k_n(s) = \int_{\mathrm{T}} k(s,t)\,\overline{\psi_n(t)}\,\mathrm{d}\nu(t) \tag{2}$$

définit une fonction $k_n \in \mathrm{L}_2(\mathrm{S},\mu)$ pour tout entier n, et

$$\|k_n\|_2 \le \|k\|_2 \,\|\psi_n\|_2 = \|k\|_2.$$

Les applications linéaires $k \to k_n$ sont donc toutes continues, et il en résulte que

$$k \to \mathrm{Q}_{\mathrm{N}} k = \sum_{n=0}^{\mathrm{N}} k_n \otimes \psi_n$$

est un endomorphisme continu de $\mathrm{L}_2(\mathrm{S}\times\mathrm{T}, \mu\otimes\nu)$; on voit facilement que Q_{N} est le projecteur orthogonal sur l'image $\mathrm{im}(\mathrm{Q}_{\mathrm{N}})$: si $f \in \mathrm{L}_2(\mathrm{S})$ et $n \le \mathrm{N}$, on vérifie que $\mathrm{Q}_{\mathrm{N}}(f\otimes\psi_n) = f\otimes\psi_n$; l'image de Q_{N} est donc l'espace de toutes les sommes $\sum_{n=0}^{\mathrm{N}} f_n\otimes\psi_n$, où les (f_n) sont quelconques dans $\mathrm{L}_2(\mathrm{S})$. Ensuite, on voit sans peine que $\mathrm{Q}_{\mathrm{N}}k - k$ est orthogonale à tous les générateurs $f\otimes\psi_n$, $n \le \mathrm{N}$, de $\mathrm{im}(\mathrm{Q}_{\mathrm{N}})$. Puisque Q_{N} est un projecteur orthogonal, on a

$$\|\mathrm{Q}_{\mathrm{N}}\| \le 1.$$

La famille d'opérateurs $(\mathrm{Q}_{\mathrm{N}})$ est équicontinue, et elle converge vers l'identité sur le sous-espace vectoriel $\mathrm{L}_2(\mathrm{S})\otimes\mathrm{L}_2(\mathrm{T})$ engendré par les fonctions de la forme $f\otimes g$; ce sous-espace est dense dans $\mathrm{L}_2(\mathrm{S}\times\mathrm{T})$ car il contient les fonctions $\mathbf{1}_{\mathrm{A}}\otimes\mathbf{1}_{\mathrm{B}}$, A, B de mesure finie, qui engendrent déjà un sous-espace dense d'après le lemme 1.3.2 ; cette convergence sur un sous-espace vectoriel dense implique la convergence sur l'espace entier, puisque la famille $(\mathrm{Q_N})$ est équicontinue. Montrons la convergence annoncée : il suffit de le faire pour les générateurs de $\mathrm{L}_2(\mathrm{S}) \otimes \mathrm{L}_2(\mathrm{T})$, les fonctions décomposées $f \otimes g$; lorsque $k = f \otimes g$, on a

$$k_n(s) = \int_{\mathrm{T}} f(s)g(t)\,\overline{\psi_n(t)}\,\mathrm{d}\nu(t) = f(s)\,\langle g, \psi_n\rangle,$$

donc

$$\mathrm{Q_N}(f \otimes g)(s,t) = f(s) \sum_{n=0}^{\mathrm{N}} \langle g, \psi_n\rangle\,\psi_n(t),$$

c'est-à-dire

$$\mathrm{Q_N}(f \otimes g) = f \otimes \mathrm{P_N} g,$$

où $\mathrm{P_N} g = \sum_{n=0}^{\mathrm{N}} \langle g, \psi_n\rangle\,\psi_n$ tend vers g dans $\mathrm{L}_2(\mathrm{T}, \nu)$, puisque (ψ_n) est une base hilbertienne de $\mathrm{L}_2(\mathrm{T})$; il en résulte que

$$\|f \otimes g - \mathrm{Q_N}(f \otimes g)\|_2 = \|f \otimes (g - \mathrm{P_N} g)\|_2 = \|f\|_2\,\|g - \mathrm{P_N} g\|_2 \to 0\,;$$

on en déduit que $k = \lim_{\mathrm{N}} \mathrm{Q_N} k$ pour toute $k \in \mathrm{L}_2(\mathrm{S} \times \mathrm{T})$, ou bien

$$k = \sum_{n=0}^{+\infty} k_n \otimes \psi_n\,;$$

comme les termes de la série sont deux à deux orthogonaux, on a aussi l'égalité

$$\|k\|_2^2 = \sum_{n=0}^{+\infty} \|k_n\|_2^2.$$

On a ainsi démontré le résultat qui suit.

Proposition 1.3.3. *Si $(\psi_n)_{n\in\mathbb{N}}$ est une base hilbertienne de $\mathrm{L}_2(\mathrm{T}, \nu)$, toute fonction $k \in \mathrm{L}_2(\mathrm{S} \times \mathrm{T}, \mu \otimes \nu)$ peut s'exprimer sous forme de la somme d'une série convergente de vecteurs deux à deux orthogonaux*

$$k = \sum_{n=0}^{+\infty} k_n \otimes \psi_n.$$

Les fonctions $(k_n) \subset \mathrm{L}_2(\mathrm{S}, \mu)$ sont uniques, définies par la formule (2).

Corollaire 1.3.4. *Si $(\varphi_n)_{n\in\mathbb{N}}$ est une base hilbertienne de $\mathrm{L}_2(\mathrm{S}, \mu)$ et $(\psi_n)_{n\in\mathbb{N}}$ une base hilbertienne de $\mathrm{L}_2(\mathrm{T}, \nu)$, la famille $(\varphi_m \otimes \psi_n)_{(m,n)\in\mathbb{N}^2}$ est une base hilbertienne pour l'espace $\mathrm{L}_2(\mathrm{S} \times \mathrm{T}, \mu \otimes \nu)$.*

Preuve. — Il est facile de vérifier que la famille $(\varphi_m \otimes \psi_n)_{(m,n)\in\mathbb{N}^2}$ est orthonormée dans $\mathrm{L}_2(\mathrm{S} \times \mathrm{T}, \mu \otimes \nu)$. Il suffit ensuite de montrer que l'espace vectoriel engendré par cette famille est dense. On a vu ci-dessus que tout élément $k \in \mathrm{L}_2(\mathrm{S} \times \mathrm{T})$ peut être approché arbitrairement par une somme finie de la forme $s_1 = \sum_{n=0}^{\mathrm{N}} k_n \otimes \psi_n$. Puisque la suite $(\varphi_m)_{m\in\mathbb{N}}$ est une base hilbertienne de $\mathrm{L}_2(\mathrm{S}, \mu)$, on peut trouver pour chaque indice $n = 0, \ldots, \mathrm{N}$ une combinaison linéaire h_n des vecteurs (φ_m) pour laquelle on aura $\|k_n - h_n\|_2 < \varepsilon/(\mathrm{N}+1)$. La somme finie $s_2 = \sum_{n=0}^{\mathrm{N}} h_n \otimes \psi_n$ est alors une combinaison linéaire des fonctions $(\varphi_m \otimes \psi_n)_{(m,n)\in\mathbb{N}^2}$, et $\|s_1 - s_2\|_2 < \varepsilon$, donc $\|k - s_2\|_2 < 2\varepsilon$. Ceci montre que la famille orthonormée $(\varphi_m \otimes \psi_n)_{(m,n)\in\mathbb{N}^2}$ engendre un sous-espace vectoriel dense, donc c'est une base hilbertienne de $\mathrm{L}_2(\mathrm{S} \times \mathrm{T}, \mu \otimes \nu)$.

///

Hilbert-Schmidt, suite

Appliquons la proposition 1.3.3 au noyau k et à la base hilbertienne (ψ_n) ; avec les notations précédentes, on a pour tout $n \geq 0$

$$k_n(s) = \int_{\mathrm{T}} k(s,t)\, \overline{\psi_n(t)}\, \mathrm{d}\nu(t) = \mathrm{T}_k \overline{\psi_n},$$

et en notant que la famille $(\overline{\psi_n})$ est aussi une base hilbertienne de $\mathrm{L}_2(\mathrm{T}, \nu)$, on voit que

$$\|k\|_2^2 = \sum_{n=0}^{+\infty} \|\mathrm{T}_k \overline{\psi_n}\|^2 = \|\mathrm{T}_k\|_{\mathrm{HS}}^2.$$

On a donc montré que T_k est de Hilbert-Schmidt, et que la norme Hilbert-Schmidt de T_k est *égale* à la norme L_2 du noyau k dans l'espace $\mathrm{L}_2(\mathrm{S} \times \mathrm{T}, \mu \otimes \nu)$. En particulier, l'application $k \to \mathrm{T}_k$ est injective : si l'opérateur T_k est l'opérateur nul, le noyau k est nul $\mu \otimes \nu$-presque partout sur $\mathrm{S} \times \mathrm{T}$.

Reprenons les choses un peu différemment, avec le corollaire 1.3.4. Si $k = \varphi \otimes \overline{\psi}$, on voit que $\mathrm{T}_{\varphi\otimes\overline{\psi}}\, f = \langle f, \psi \rangle\, \varphi$, donc l'opérateur est de rang un (son image est contenue dans la droite engendrée par φ). Si $(\varphi_n)_{n\geq 0}$ est une base hilbertienne de $\mathrm{L}_2(\mathrm{S}, \mu)$ et $(\overline{\psi_n})_{n\geq 0}$ une base hilbertienne de $\mathrm{L}_2(\mathrm{T}, \nu)$, on a vu que les fonctions $(s,t) \to \varphi_m(s)\, \overline{\psi_n(t)}$ (où m, n prennent toutes les valeurs entières ≥ 0) donnent une base hilbertienne de l'espace $\mathrm{L}_2(\mathrm{S} \times \mathrm{T}, \mu \otimes \nu)$. Supposons donné un noyau $k \in \mathrm{L}_2(\mathrm{S} \times \mathrm{T}, \mu \otimes \nu)$; on peut le décomposer sur la base hilbertienne précédente,

$$k = \sum_{m,n=0}^{+\infty} c_{m,n}\, \varphi_m \otimes \overline{\psi_n}. \tag{3}$$

On sait que

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

et la série (3) converge dans l'espace de Hilbert $\mathrm{L}_2(\mu \otimes \nu)$, quel que soit l'ordre dans lequel on énumère les termes. Par image linéaire continue, on en déduit que

$$\mathrm{T}_k = \sum_{m,n=0}^{+\infty} c_{m,n}\, \mathrm{T}_{\varphi_m\otimes\overline{\psi_n}},$$

où la série converge en norme d'opérateur, quel que soit l'ordre dans lequel on énumère les termes. On retrouve bien sûr le fait que l'opérateur T_k est de Hilbert-Schmidt : on déduit de la convergence de la série ci-dessus que

$$\text{(M)} \qquad \mathrm{T}_k f = \sum_{m,n=0}^{+\infty} c_{m,n}\, \mathrm{T}_{\varphi_m \otimes \overline{\psi_n}}\, f = \sum_{m,n=0}^{+\infty} c_{m,n}\, \langle f, \psi_n \rangle\, \varphi_m.$$

Il en résulte que $\mathrm{T}_k \psi_p = \sum_m c_{m,p}\, \varphi_m$, donc $\|\mathrm{T}_k \psi_p\|^2 = \sum_m |c_{m,p}|^2$, et ensuite

$$\sum_{p=0}^{+\infty} \|\mathrm{T}_k \psi_p\|^2 = \sum_{m,p=0}^{+\infty} |c_{m,p}|^2 = \|k\|_2^2 < +\infty.$$

On notera la correspondance entre l'opérateur T_k et les coefficients $c_{m,n}$ du développement de k,

$$\text{(4)} \qquad c_{m,n} = \langle \mathrm{T}_k \psi_n, \varphi_m \rangle.$$

On voit que les coefficients $(c_{m,n})$ et la formule (M) s'interprètent très facilement comme un calcul matriciel utilisant des matrices infinies, tel qu'on l'a évoqué dans le point 1 des exemples 1.3.1, à propos des opérateurs sur $\ell_2(\mathbb{N})$.

Exercice. Montrer que la composition de deux opérateurs T_{k_1} et T_{k_2} de la forme précédente est un opérateur T_k, avec

$$k(s,t) = \int k_1(s,u) k_2(u,t)\, \mathrm{d}u.$$

Opérateurs hermitiens à noyau

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

$$k(t,s) = \overline{k(s,t)}$$

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

$$c_{m,n} = \langle \mathrm{T}_k \varphi_n, \varphi_m \rangle = \lambda_n \, \langle \varphi_n, \varphi_m \rangle = \lambda_n\, \delta_{m,n},$$

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

$$k(s,t) = \sum_{n=0}^{+\infty} \lambda_n\, \varphi_n \otimes \overline{\varphi_n}$$

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

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

$$k(s,t) = \sum_{n=0}^{+\infty} \lambda_n \varphi_n(s) \overline{\varphi_n(t)}.$$

Le *théorème de Mercer* donne un cas particulier intéressant : si $k(s,t)$ est continue sur $[0,1]^2$ et si $\langle \mathrm{T}_k f, f \rangle$ est ≥ 0 pour toute $f \in \mathrm{L}_2$, les fonctions propres de T_k correspondant aux valeurs propres non nulles sont continues, et la série ci-dessus converge uniformément.

Exercice. Traiter le cas $k(s,t) = \max(s,t)$, $0 \leq s, t \leq 1$.

Notes du chapitre 1

([a]) La réciproque de cette propriété est vraie dans les espaces de Hilbert : si T est un endomorphisme compact d'un espace de Hilbert H, il est limite en norme d'opérateur d'une suite d'opérateurs de rang fini ; en effet, si H est séparable, muni d'une base hilbertienne $(e_n)_{n \geq 0}$ et si P_n désigne la projection orthogonale sur $\mathrm{Vect}(e_0, \ldots, e_n)$, les opérateurs de rang fini $\mathrm{P}_n \circ \mathrm{T}$ ou bien $\mathrm{T} \circ \mathrm{P}_n$ tendent vers T en norme d'opérateur quand T est compact. On peut aussi vérifier cette réciproque pour tous les espaces de Banach « classiques ». La question de savoir si cette réciproque est vraie pour tout espace de Banach a été considérée dès le livre de Banach *Théorie des opérations linéaires* paru en 1932, mais elle n'a été résolue (par la négative) qu'en 1972 par le mathématicien suédois Per Enflo.

([b]) Si la forme linéaire ℓ n'est pas nulle, on peut trouver un vecteur x_0 non nul dans le noyau de la projection orthogonale P sur l'hyperplan $\ker \ell$ (ce vecteur x_0 est donc orthogonal au noyau). Posons

$$x = \frac{\ell(x_0)}{\langle x_0, x_0 \rangle} \, x_0.$$

Tout vecteur $v \in \mathrm{H}$ peut s'écrire $v = z + \lambda x_0$ avec $z \in \ker \ell$ et $\lambda \in \mathbb{K}$; on voit que

$$\langle v, x \rangle = \langle z + \lambda x_0, x \rangle = \lambda \langle x_0, x \rangle = \lambda \ell(x_0) = \ell(z + \lambda x_0) = \ell(v),$$

ce qui montre que la forme linéaire ℓ est représentée par le produit scalaire avec x.

2. Algèbres de Banach

2.1. Algèbres et C*-algèbres

Une *algèbre* sur $\mathbb{K}$, avec $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$, est un $\mathbb{K}$-espace vectoriel A, muni d'une multiplication associative $(a, b) \in \mathrm{A} \times \mathrm{A} \to ab \in \mathrm{A}$, telle que les applications $x \in \mathrm{A} \to ax$ et $x \to xa$ soit des endomorphismes de l'espace vectoriel A, pour tout $a \in \mathrm{A}$.

Une *algèbre de Banach unitaire* est un espace de Banach A, muni d'une multiplication d'algèbre possédant une unité 1_{A} et vérifiant

$$\forall a, b \in \mathrm{A}, \quad \|ab\| \leq \|a\| \, \|b\| \, ; \quad \|1_{\mathrm{A}}\| = 1. \tag{1}$$

Remarque. Si une algèbre A ne possède pas d'unité, on peut lui en adjoindre une en remplaçant A par $\mathrm{A}_1 = \mathrm{A} \oplus \mathbb{K}\mathbf{1}$, muni du produit

$$(a + \lambda\mathbf{1})(b + \mu\mathbf{1}) = ab + \mu a + \lambda b + \lambda\mu\mathbf{1}.$$

Si l'opération de produit d'une algèbre de Banach avec unité est continue de $\mathrm{A} \times \mathrm{A}$ dans A, on peut munir A d'une norme équivalente pour laquelle les axiomes (1) sont satisfaits, en posant

$$|a| = \sup\{\|ab\| : \|b\| \leq 1\}.$$

Exemples 2.1.1.

1. L'algèbre des matrices $\mathrm{M}_n(\mathbb{K})$, munie de la norme subordonnée à une norme donnée sur $\mathbb{K}^n$, est une algèbre de Banach (de dimension finie) ; l'exemple le plus important est celui où l'espace $\mathbb{K}^n$ est muni de la norme euclidienne.
2. L'algèbre des endomorphismes $\mathcal{L}(\mathrm{X})$, si X est un espace de Banach $\neq \{0\}$.
3. Les espaces $\mathrm{C}(\mathrm{K})$, K compact non vide, ou l'espace $\mathrm{L}_\infty(0, 1)$, munis du produit ponctuel des fonctions, sont des exemples d'algèbres de Banach commutatives.
4. L'espace $\mathrm{L}_1(\mathbb{R})$, muni de la convolution comme produit, est un exemple d'algèbre sans unité ; pour ajouter une unité, on peut agrandir l'algèbre en considérant l'ensemble de toutes les mesures complexes $\mathrm{d}\mu(x) = f(x)\,\mathrm{d}x + \lambda\,\mathrm{d}\delta_0(x)$, avec $f \in \mathrm{L}_1(\mathbb{R})$, λ scalaire et δ_0 la mesure de Dirac en 0.
5. *L'algèbre de Calkin* est le quotient de $\mathcal{L}(\mathrm{H})$ par l'idéal des opérateurs compacts sur un espace de Hilbert H de dimension infinie. Elle joue un rôle intéressant dans la théorie des opérateurs de Fredholm.

C-algèbres*

Une *C*-algèbre unitaire* est une algèbre de Banach unitaire complexe A, munie d'une application continue $a \in \mathrm{A} \to a^* \in \mathrm{A}$ *antilinéaire*, telle que $(ab)^* = b^*a^*$, $(a^*)^* = a$ (involution) et surtout

$$\forall a \in \mathrm{A}, \quad \|a^*a\| = \|a\|^2.$$

Il en résulte que $\|a^*\| = \|a\|$. L'exemple fondamental de C*-algèbre est l'espace $\mathcal{L}(\mathrm{H})$ des endomorphismes d'un espace de Hilbert complexe H, où on prend pour opération étoile l'application $\mathrm{T} \to \mathrm{T}^*$ qui associe à chaque $\mathrm{T} \in \mathcal{L}(\mathrm{H})$ l'opérateur adjoint hilbertien.

Inversement, on sait montrer que toute C*-algèbre unitaire est isomorphe à une sous-algèbre fermée d'un $\mathcal{L}(\mathrm{H})$, pour un certain espace de Hilbert H, sous-algèbre invariante par $\mathrm{T} \to \mathrm{T}^*$ et contenant l'application identique de H (théorème 7.4).

Un autre exemple est fourni par l'algèbre C(K) des fonctions complexes continues sur un compact K non vide, munie de l'étoile

$$f^* = \overline{f}.$$

Cet exemple est commutatif. Le théorème précédent de représentation des C*-algèbres se complète par le suivant : toute C*-algèbre unitaire commutative est isomorphe à l'algèbre C(K) des fonctions complexes continues sur un certain compact K (théorème 7.7). Ce théorème s'applique en particulier à la C*-algèbre commutative $\mathrm{L}_\infty(0,1)$ des classes de fonctions mesurables bornées.

On ne parlera pas de C*-algèbres sur $\mathbb{R}$; on se limitera aussi au cas unitaire, bien que certains exemples non unitaires soient importants, par exemple la C*-algèbre $\mathcal{K}(\mathrm{H})$ des endomorphismes compacts d'un espace de Hilbert H de dimension infinie. On peut d'ailleurs facilement la plonger dans une algèbre avec unité : il suffit de considérer l'algèbre avec unité de tous les opérateurs de la forme $\lambda \,\mathrm{Id}_\mathrm{H} + \mathrm{T}$, T compact et $\lambda \in \mathbb{C}$.

Définition. Éléments hermitiens, normaux dans une C*-algèbre : on dit que $a \in \mathrm{A}$ est *hermitien* si $a^* = a$. Si $b \in \mathrm{A}$, l'élément $a = b^*b$ est hermitien. On dit que a est *normal* si $a^*a = a\,a^*$. On dit que $u \in \mathrm{A}$ est *unitaire* si $u^*u = u\,u^* = 1_\mathrm{A}$. Ces définitions généralisent le cas de l'algèbre $\mathcal{L}(\mathrm{H})$.

Éléments inversibles

Soit A une algèbre de Banach unitaire ; on dit que $a \in \mathrm{A}$ est *inversible dans* A s'il existe un élément $u \in \mathrm{A}$ tel que $au = ua = 1_\mathrm{A}$. Dans ce cas on note $a^{-1} = u$.

On va montrer que les éléments proches de 1_A sont inversibles ; la démonstration présente l'inverse au moyen d'une série vectorielle. Dans une algèbre de Banach, la continuité du produit (dans l'exemple ci-dessous : opération de multiplication à droite par b) permet de voir que

$$\Bigl(\sum_{n=0}^{+\infty} a_n\Bigr) b = \sum_{n=0}^{+\infty} (a_n b).$$

On fait la convention $a^0 = 1_\mathrm{A}$; on remarque que $\|a^n\| \le \|a\|^n$ pour tout entier $n \ge 0$.

Proposition 2.1.2. *Soient* A *une algèbre de Banach unitaire et* $a \in \mathrm{A}$ *; la série*

$$1_\mathrm{A} + a + a^2 + \cdots + a^n + \cdots$$

converge dans A *si et seulement si* $\inf_n \|a^n\| < 1$*, et dans ce cas sa somme est l'inverse de* $1_\mathrm{A} - a$*. En particulier*

$$\bigl(\|a\| < 1\bigr) \Rightarrow \bigl(1_\mathrm{A} - a \ \text{ inversible dans } \ \mathrm{A}\bigr).$$

Preuve. — Si la série ci-dessus converge, son terme général a^n doit tendre vers 0_{A}, donc $\inf_n \|a^n\| = 0 < 1$; inversement, si on a $\|a^m\| < 1$ pour un certain m, on en déduit que $\sum \|a^n\| < +\infty$ de la façon suivante : si $a^m = 0$, on a $a^n = 0$ pour tout $n \ge m$ et la convergence de la série est évidente ; sinon, on note que $m > 0$ (car $\|a^0\| = \|1_{\mathrm{A}}\| = 1$), on pose $\tau = \|a^m\|^{1/m}$; on a $0 < \tau < 1$, et on écrit alors pour n entier quelconque la division euclidienne $n = mq + r$, avec $0 \le r < m$; on obtient

$$\|a^n\| \le \|a^{mq}\| \, \|a^r\| \le \tau^{mq} \, \|a\|^r = \tau^{mq+r} \tau^{-r} \|a\|^r \le \mathrm{M}\, \tau^n$$

avec $\mathrm{M} = \bigl(\|a\|/\tau\bigr)^m$ (noter que $\tau \le \|a\|$), donc la série $\sum a^n$ converge normalement puisque la norme de son terme général est majorée par un multiple du terme général de la série géométrique convergente $\sum \tau^n$.

Si la série $\sum_{n=0}^{+\infty} a^n$ converge dans A, désignons par s sa somme ; alors as est égal à

$$a\bigl(1_{\mathrm{A}} + a + a^2 + \cdots + a^n + \cdots\bigr) = a + a^2 + \cdots + a^{n+1} + \cdots = s - 1_{\mathrm{A}},$$

donc $as = s - 1_{\mathrm{A}}$, c'est-à-dire $(1_{\mathrm{A}} - a)s = 1_{\mathrm{A}}$; de même $s(1_{\mathrm{A}} - a) = 1_{\mathrm{A}}$.

///

Exercice. Montrer que $\lim_n \|a^n\|^{1/n}$ existe.

Inverse de $u - a$

Si u est inversible et si a est assez petit pour que la série ci-dessous converge dans A, l'élément

$$x = u^{-1} + u^{-1} a\, u^{-1} + u^{-1} a\, u^{-1} a\, u^{-1} + \cdots$$

vérifie

$$ux = 1_{\mathrm{A}} + a\, u^{-1} + a\, u^{-1} a\, u^{-1} + \cdots$$

de sorte que $ux - ax = 1_{\mathrm{A}}$, et de même $xu - xa = 1_{\mathrm{A}}$. On en déduit que x est l'inverse de $u - a$. Pour que la série ci-dessus converge, il suffit que son terme général, qu'on peut écrire sous la forme $u^{-1}(au^{-1})^n$, soit majoré en norme par une série géométrique convergente ; c'est le cas dès que $\|au^{-1}\| < 1$, ce qui est garanti par la condition classique et simple

$$\|a\| < \|u^{-1}\|^{-1}.$$

On vient de montrer que les voisins d'un inversible restent inversibles.

Proposition. *L'ensemble des éléments inversibles de l'algèbre unitaire* A *forme un sous-groupe multiplicatif* G(A), *ouvert dans l'espace normé* A.

Preuve. — Sous-groupe est bien connu, ouvert résulte de ce qui précède.

///

Cas de $\mathcal{L}(\mathrm{X}, \mathrm{Y})$

Si T est un isomorphisme de X sur Y, si V désigne l'inverse de T et si S est un « petit » opérateur, on aura encore

$$(\mathrm{T} - \mathrm{S})^{-1} = \mathrm{V} + \mathrm{VSV} + \mathrm{VSVSV} + \cdots$$

ce qui prouve la stabilité des isomorphismes par perturbation. La même formule permet de traiter aussi le cas de l'inversibilité à droite ou à gauche : si on a seulement $\mathrm{VT} = \mathrm{Id}_{\mathrm{X}}$, la somme W de la série précédente vérifie $\mathrm{W}(\mathrm{T} - \mathrm{S}) = \mathrm{Id}_{\mathrm{X}}$; si T est inversible à gauche (ou à droite), les voisins de T le sont aussi.

2.2. Spectre d'un élément a d'une algèbre de Banach

On suppose que A est une algèbre de Banach unitaire *complexe*. Si a est un élément de A, on désigne par $\rho(a)$ l'ensemble des $\lambda \in \mathbb{C}$ tels que $a - \lambda 1_{\mathrm{A}}$ soit inversible dans A. Par la stabilité des inversibles, on voit que $\rho(a)$ est un ouvert. On dit que $\rho(a)$ est *l'ensemble résolvant* de a. La *résolvante de a* est l'application définie sur l'ouvert $\rho(a)$, à valeurs dans A,

$$\forall \lambda \in \rho(a), \quad \mathrm{R}_\lambda(a) = (\lambda 1_{\mathrm{A}} - a)^{-1} \in \mathrm{A}.$$

Si $|\lambda| > \|a\|$, on a $\|a/\lambda\| < 1$ et $1_{\mathrm{A}} - a/\lambda$ est inversible par la proposition 2.1.2, donc $a - \lambda 1_{\mathrm{A}}$ aussi et par conséquent $\lambda \in \rho(a)$; on voit ainsi que l'ensemble $\rho(a)$ contient le complémentaire du disque fermé de rayon $\|a\|$.

Le *spectre* $\sigma(a)$ est le complémentaire dans $\mathbb{C}$ de $\rho(a)$. Il est donc fermé et borné par $\|a\|$. Désignons par $\mathrm{rs}(a)$ le sup des modules des éléments du spectre, le *rayon spectral* de a. On va montrer plus loin que le spectre est non vide, et que

$$\mathrm{rs}(a) = \lim_n \|a^n\|^{1/n}$$

(formule du rayon spectral).

Exemples 2.2.1.

1. Si $\mathrm{A} = \mathrm{M}_n(\mathbb{C})$ est l'algèbre des matrices carrées complexes de taille $n \times n$, le spectre d'une matrice $a \in \mathrm{A}$ est l'ensemble des valeurs propres de la matrice ; mais dans le cas plus général où $\mathrm{A} = \mathcal{L}(\mathrm{X})$ est l'algèbre des endomorphismes d'un espace de Banach X de dimension infinie, il est important de savoir tout de suite que $\sigma(\mathrm{T})$, pour $\mathrm{T} \in \mathcal{L}(\mathrm{X})$, est en général **plus grand** que l'ensemble des valeurs propres (qui peut([a]) être vide, alors que le spectre n'est jamais vide).

2. Considérons l'algèbre $\mathrm{A} = \mathrm{C}(\mathrm{K})$ des fonctions complexes continues sur un espace topologique compact K non vide. Il est clair qu'une fonction f est inversible dans A si et seulement si f ne s'annule pas sur K. On en déduit facilement que

$$\sigma(f) = f(\mathrm{K}) = \{f(t) : t \in \mathrm{K}\}.$$

3. Si A est une C^*-algèbre unitaire, il est clair que b est inversible dans A si et seulement si b^* est inversible, et $(b^*)^{-1} = (b^{-1})^*$. En appliquant cette observation à $b = a - \lambda 1_{\mathrm{A}}$, on voit que

$$\sigma(a^*) = \{\overline{\lambda} : \lambda \in \sigma(a)\}.$$

Si on considère le shift à droite S dans $\ell_2(\mathbb{N})$, on voit que pour tout λ tel que $|\lambda| < 1$, le vecteur $x_\lambda = (\lambda^n)_{n \ge 0}$ est un vecteur de $\ell_2(\mathbb{N})$ tel que $\mathrm{S}^* x_\lambda = \lambda x_\lambda$. Le spectre de S^* contient donc le disque unité ouvert. Comme $\|\mathrm{S}\| = 1$, le spectre est contenu dans le disque unité fermé, donc

$$\sigma(\mathrm{S}) = \{\lambda : |\lambda| \le 1\}.$$

Remarque. Changement d'algèbre : supposons que l'algèbre de Banach A soit sous-algèbre d'une algèbre de Banach B ; l'ensemble résolvant augmente quand on augmente l'algèbre (il y a plus de candidats dans B pour être inverse d'un élément de A), donc le spectre d'un élément $a \in \mathrm{A}$ *diminue* quand on augmente l'algèbre. Mais on verra que le *bord* du spectre initial reste dans le spectre après augmentation de l'algèbre.

Exemple 2.2.2. Fonctions de A(D) plongées dans C($\mathbb{T}$) : notons $\mathbb{T}$ le cercle unité du plan complexe ; *l'algèbre du disque* A(D) est formée des fonctions continues f sur $\mathbb{T}$ qui se prolongent en une fonction $\widetilde{f}$, continue dans le disque unité fermé et holomorphe dans le disque unité ouvert ; c'est une sous-algèbre unitaire de C($\mathbb{T}$). L'élément $z \to z$ n'est pas inversible([b]) dans la petite algèbre A(D), mais le devient dans la grosse C($\mathbb{T}$).

Fonctions holomorphes vectorielles

Définition. Si f est une fonction définie sur un ouvert Ω du plan complexe, à valeurs dans un espace de Banach X, on dit que f est *analytique* dans Ω si pour tout $z \in \Omega$ il existe $\delta_z > 0$ et des vecteurs $(x_n(z)) \subset \mathrm{X}$ tels que

$$(h \in \mathbb{C},\ |h| \le \delta_z) \Rightarrow f(z+h) = \sum_{n=0}^{+\infty} h^n\, x_n(z).$$

Si f est analytique dans Ω, il en résulte que f est continue de Ω dans X : si $z \in \Omega$, posons $\delta = \delta_z > 0$, et $x_n = x_n(z) \in \mathrm{X}$; la série $\sum h^n x_n$ converge pour $h = \delta$, donc la suite $(\delta^n \|x_n\|)$ est bornée, disons par M. Alors

$$\|f(z+h) - f(z)\| \le \sum_{n=1}^{+\infty} |h|^n\, \|x_n\| \le \mathrm{M} \sum_{n=1}^{+\infty} \Bigl(\frac{|h|}{\delta}\Bigr)^n = \mathrm{M}\, \frac{|h|}{\delta - |h|}$$

tend vers 0 avec h, donc f est continue au point z.

Si f est analytique dans le disque ouvert $\mathrm{D}(0, \mathrm{R})$ et si ρ est un nombre $< \mathrm{R}$, la fonction $\theta \to \|f(\rho\, \mathrm{e}^{\mathrm{i}\theta})\|$ est continue sur le compact $[0, 2\pi]$, ce qui permet de considérer

$$\mathrm{M}_\rho(f) = \max\{\|f(\rho\, \mathrm{e}^{\mathrm{i}\theta})\| : \theta \in [0, 2\pi]\}.$$

Lemme 2.2.3. *Si f est une fonction analytique dans le disque ouvert $\Omega = \mathrm{D}(0, \mathrm{R})$ du plan complexe, où $0 < \mathrm{R} \le +\infty$, à valeurs dans un espace de Banach X, et si*

$$\sum_{n=0}^{+\infty} z^n\, x_n \tag{2}$$

est son développement au voisinage de 0, valable pour $|z|$ assez petit, on a pour tout $\rho < \mathrm{R}$

$$\sup_{n \ge 0} \rho^n\, \|x_n\| \le \mathrm{M}_\rho(f).$$

Il en résulte que le développement en 0 donné par (2) converge dans tout le disque Ω, et que sa somme représente la fonction f en tout point du disque Ω.

Preuve.— Pour toute forme linéaire ξ continue sur X, la fonction scalaire $f_\xi = \xi \circ f$ est holomorphe dans le disque Ω. Fixons $\rho < \mathrm{R}$; la fonction $z \to \|f(z)\|$ est bornée par le nombre $\mathrm{M}_\rho = \mathrm{M}_\rho(f)$ sur le cercle de rayon ρ. Sur ce même cercle, la fonction f_ξ est bornée par $\mathrm{M}_\rho\, \|\xi\|$, donc par les inégalités de Cauchy on a pour les dérivées en 0 de f_ξ, pour tout entier $n \ge 0$

$$\rho^n\, \frac{|f_\xi^{(n)}(0)|}{n!} \le \mathrm{M}_\rho\, \|\xi\|.$$

Le développement de Taylor de f_ξ au voisinage de 0 est donné par

$$f_\xi(h) = \xi\Bigl(\sum_{n=0}^{+\infty} h^n\, x_n\Bigr) = \sum_{n=0}^{+\infty} h^n\, \xi(x_n)$$

donc $n!\,\xi(x_n) = f_\xi^{(n)}(0)$ pour tout $n \ge 0$, et on a

$$\rho^n\, |\xi(x_n)| \le \mathrm{M}_\rho\, \|\xi\|.$$

Par le théorème de Hahn-Banach, il existe une forme linéaire ξ_n telle que $\|\xi_n\| \le 1$ et $\xi_n(x_n) = \|x_n\|$; on en déduit

$$(3) \qquad \rho^n\, \|x_n\| \le \mathrm{M}_\rho.$$

Ces majorations impliquent que la série (2) converge dans tout le disque ouvert $\mathrm{D}(0,\mathrm{R})$; posons $\mathrm{F}(z) = \sum_{n=0}^{+\infty} z^n x_n$; pour toute forme linéaire continue ξ sur X, la fonction scalaire $\mathrm{F}_\xi(z) = \xi(\mathrm{F}(z)) = \sum_{n=0}^{+\infty} z^n \xi(x_n)$ est holomorphe dans $\mathrm{D}(0,\mathrm{R})$, et elle coïncide avec f_ξ dans un voisinage de 0 ; ces deux fonctions sont donc égales en tout point de $\mathrm{D}(0,\mathrm{R})$ (principe des zéros isolés, par exemple). On a donc pour tout $z \in \mathrm{D}(0,\mathrm{R})$

$$\xi\bigl(\mathrm{F}(z) - f(z)\bigr) = 0$$

pour toute forme linéaire continue ξ, donc $\mathrm{F}(z) = f(z)$ par le théorème de Hahn-Banach.

///

Corollaire 2.2.4 (théorème de Liouville vectoriel). *Si $f : \mathbb{C} \to \mathrm{X}$ est analytique et bornée sur $\mathbb{C}$, alors f est constante ; si $f : \mathbb{C} \to \mathrm{X}$ est analytique et tend vers 0 à l'infini, alors f est identiquement nulle.*

Preuve. — On sait d'après le lemme précédent, appliqué avec $\mathrm{R} = +\infty$, que f est représentée sur $\mathbb{C}$ tout entier par sa série entière à l'origine $\sum_{n=0}^{+\infty} z^n x_n$. Si f reste bornée sur $\mathbb{C}$, il existe un nombre M qui majore tous les $\mathrm{M}_\rho(f)$, donc pour tout $n > 0$ et tout $\rho > 0$ on obtient par (3) que

$$\|x_n\| \le \frac{\mathrm{M}}{\rho^n},$$

ce qui montre, en faisant tendre ρ vers l'infini, que $x_n = 0$ pour tout $n > 0$, donc $f = x_0$ est constante. Si f tend vers 0 à l'infini, elle est d'abord bornée donc constante, et la constante $x_0 \in \mathrm{X}$ doit être nulle.

///

Formule du rayon spectral

Théorème. *Soit a un élément d'une algèbre de Banach unitaire complexe* A *; le spectre de a est non vide, et on a*

$$\mathrm{rs}(a) = \lim_n \|a^n\|^{1/n}.$$

Preuve. — Posons
$$r_* = \liminf_n \|a^n\|^{1/n}.$$
Si $z \in \mathbb{C}$ est tel que $|z| > r_*$, c'est-à-dire que
$$1 > |z|^{-1}\, r_* = \liminf_n \|z^{-n} a^n\|^{1/n},$$
il existe certainement des entiers m tels que $\|z^{-m} a^m\| < 1$; par conséquent, la série $\sum z^{-n} a^n$ converge dans A et sa somme est l'inverse de $1_{\mathrm{A}} - z^{-1}a$, d'après la proposition 2.1.2. On voit donc que $z 1_{\mathrm{A}} - a$ est inversible quand $|z| > r_*$: le spectre $\sigma(a)$ est contenu dans le disque fermé centré en 0 et de rayon r_*. Considérons maintenant la fonction résolvante
$$f(z) = (z 1_{\mathrm{A}} - a)^{-1}$$
qui est définie pour tout z de l'ensemble ouvert $\rho(a) \subset \mathbb{C}$. Si $z \in \rho(a)$, cherchons à calculer $f(z-h)$ pour $h \in \mathbb{C}$ assez petit ; posons $u = z 1_{\mathrm{A}} - a$, et considérons l'inverse $v = f(z)$ de u ; en écrivant $(z-h) 1_{\mathrm{A}} - a = u - h 1_{\mathrm{A}} = u(1_{\mathrm{A}} - h\, v)$, on voit que
$$f(z-h) = (1_{\mathrm{A}} - h\, v)^{-1} v = v + h\, v^2 + h^2\, v^3 + \cdots$$
pour $|h| < \|v\|^{-1}$, c'est-à-dire que $f(z-h)$ est représenté par une série de la forme $\sum h^n b_n(z)$, avec $(b_n(z)) \subset \mathrm{A}$, pour tout $z - h$ dans un voisinage de z. Il en résulte que la fonction f est analytique dans $\rho(a)$.

Montrons d'abord que le spectre est non vide. Dans le cas contraire, $\rho(a) = \mathbb{C}$ et f est analytique sur $\mathbb{C}$ tout entier ; de plus pour $|z|$ assez grand on a
$$f(z) = z^{-1}\, (1_{\mathrm{A}} - z^{-1} a)^{-1} = \sum_{n=0}^{+\infty} \frac{a^n}{z^{n+1}},$$
d'où résulte que
$$\|f(z)\| \le \sum_{n=0}^{+\infty} \frac{\|a\|^n}{|z|^{n+1}} = \frac{1}{|z| - \|a\|}$$
et
$$\lim_{|z| \to +\infty} f(z) = 0.$$
Si f était analytique sur $\mathbb{C}$, le théorème de Liouville 2.2.4 donnerait $f = 0$ partout : or ceci est impossible car les valeurs de f sont des inverses, donc des éléments non nuls.

Posons $\mathrm{R} = \mathrm{rs}(a)^{-1}$ (peut-être égal à $+\infty$) ; on a vu que $\mathrm{R} \ge 1/r_*$. Pour tout z non nul tel que $|z| < \mathrm{R}$, on est sûr que $z^{-1} \in \rho(a)$, ce qui permet de considérer
$$g(z) = (1_{\mathrm{A}} - z a)^{-1} = z^{-1} f(z^{-1}).$$
La première expression permet de prolonger la définition de g par $g(0) = 1_{\mathrm{A}}$. Puisque f est holomorphe dans $\rho(a)$, il est clair que g est holomorphe dans le disque ouvert de rayon R privé de 0 ; de plus, au voisinage de 0 on voit que
$$g(z) = 1_{\mathrm{A}} + z\, a + z^2\, a^2 + \cdots + z^n\, a^n + \cdots$$

qui est holomorphe dans ce voisinage, donc g est holomorphe dans $\mathrm{D}(0, \mathrm{R})$. D'après le lemme 2.2.3, on a

$$\sup_n \rho^n \|a^n\| < \infty$$

pour tout $\rho < \mathrm{R}$, donc en posant

$$r^* = \limsup_n \|a^n\|^{1/n}$$

on obtient $\rho\, r^* \le 1$, ce qui donne $\mathrm{R}\, r^* \le 1$ quand $\rho \to \mathrm{R}$. On a donc en regroupant nos informations

$$1/r_* \le \mathrm{R} \le 1/r^*.$$

Comme il est clair que $r_* \le r^*$, on obtient l'égalité $r_* = r^* = \lim_n \|a^n\|^{1/n} = 1/\mathrm{R}$. Ceci donne la formule du rayon spectral,

$$\mathrm{rs}(a) = \lim_n \|a^n\|^{1/n}.$$

///

Exemples.

1. Commençons par une application classique de la non vacuité du spectre : si dans une algèbre de Banach unitaire complexe A tout élément non nul est inversible, alors $\mathrm{A} = \mathbb{C}1_\mathrm{A}$ (l'ensemble des multiples scalaires de l'unité ; voir le théorème 7.5) ; si B est une algèbre de Banach unitaire complexe et I un idéal maximal, il résulte que le quotient $\mathrm{A} = \mathrm{B}/\mathrm{I}$ vérifie la conclusion précédente, ce qui signifie que I est un hyperplan de B.

2. Pour une isométrie $\mathrm{U} \in \mathcal{L}(\mathrm{X})$, on a $\|\mathrm{U}^n\| = 1$ pour tout n, donc $\mathrm{rs}(\mathrm{U}) = 1$. On a de façon générale, si u est inversible dans une algèbre de Banach

$$\sigma(u^{-1}) = \{1/\lambda : \lambda \in \sigma(u)\}.$$

Si $\mathrm{U} \in \mathcal{L}(\mathrm{X})$ est une isométrie surjective, donc inversible, on aura aussi $\mathrm{rs}(\mathrm{U}^{-1}) = 1$, ce qui entraîne que le spectre de U est contenu dans le cercle unité.

3. Si b est un élément hermitien d'une C^*-algèbre, on a $\|b^2\| = \|b^*b\| = \|b\|^2$. Ensuite b^2 est hermitien et on continue, $\|b^4\| = \|b\|^4$, etc... pour montrer que $\|b^{2^n}\| = \|b\|^{2^n}$ pour tout n, et conclure que $\mathrm{rs}(b) = \|b\|$. Si a est normal, on introduit $b = a^*a$ hermitien, et on écrit $\|a^2\|^2 = \|a^*a^*aa\| = \|a^*aa^*a\| = \|a^*a\|^2 = \|a\|^4$, etc... On a ainsi obtenu le résultat important qui suit.

Proposition 2.2.5. *Si a est un élément normal d'une C^*-algèbre, on a $\mathrm{rs}(a) = \|a\|$.*

Exercice. Estimer les normes des puissances P^n de l'opérateur de primitive (sur $\mathrm{L}_2(0,1)$) de l'exercice 1.2.5 et en déduire que $\sigma(\mathrm{P}) = \{0\}$. Retrouver le même résultat en calculant explicitement $\mathrm{R}_\lambda(\mathrm{P})$ pour tout nombre complexe $\lambda \neq 0$.

Intégrale de Riemann vectorielle

Soit X un espace de Banach ; on dit qu'une fonction $g : [a, b] \to \mathrm{X}$ est *en escalier* s'il existe une subdivision $\tau : t_0 = a < t_1 < \ldots < t_n = b$ de l'intervalle telle que g soit constante sur chaque intervalle ouvert (t_{i-1}, t_i), pour $i = 1, \ldots, n$. Si $x_i \in \mathrm{X}$ désigne la valeur (vectorielle) constante de g sur l'intervalle (t_{i-1}, t_i), on pose

$$\int_a^b g(t)\,\mathrm{d}t = \sum_{i=1}^n (t_i - t_{i-1}) g(x_i) \in \mathrm{X}.$$

On vérifie comme dans le cas scalaire que cette somme ne dépend que de la fonction g, et pas de la subdivision choisie. On a la majoration facile

$$\Big\| \int_a^b g(t)\,\mathrm{d}t \Big\| \le \int_a^b \|g(t)\|\,\mathrm{d}t.$$

Si f est continue de $[a, b]$ dans X, on vérifie que les sommes de Riemann

$$\sum_{i=1}^n (t_i - t_{i-1}) f(\xi_i) \in \mathrm{X},$$

où les ξ_i sont des points quelconques de $[t_{i-1}, t_i]$, convergent vers un vecteur $\mathrm{I} \in \mathrm{X}$ lorsque le pas $\delta(\tau) = \max\{|t_i - t_{i-1}| : i = 1, \ldots, n\}$ de la subdivision τ tend vers 0. Le vecteur limite I s'appelle l'intégrale de f,

$$\int_a^b f(t)\,\mathrm{d}t = \mathrm{I} = \lim_{\delta(\tau)\to 0} \sum_{i=1}^n (t_i - t_{i-1}) f(\xi_i) \in \mathrm{X}.$$

On obtient, en passant à la limite à partir des fonctions en escalier, la majoration simple mais importante

$$\Big\| \int_a^b f(t)\,\mathrm{d}t \Big\| \le \int_a^b \|f(t)\|\,\mathrm{d}t.$$

À partir de cette majoration on montre facilement le résultat qui suit.

Proposition. *Si une suite (f_n) de fonctions continues de $[a, b]$ dans un espace de Banach* X *converge uniformément sur $[a, b]$ vers une fonction f, on a*

$$\int_a^b f(t)\,\mathrm{d}t = \lim_n \int_a^b f_n(t)\,\mathrm{d}t.$$

Si une série $\sum u_n$ de fonctions continues de $[a, b]$ dans X *converge uniformément sur $[a, b]$, on a*

$$\int_a^b \Big(\sum_{n=0}^{+\infty} u_n(t) \Big)\,\mathrm{d}t = \sum_{n=0}^{+\infty} \int_a^b u_n(t)\,\mathrm{d}t.$$

2.3. Premiers rudiments de calcul fonctionnel holomorphe

Dans tout le paragraphe qui suit, a désigne un élément d'une algèbre de Banach unitaire complexe A, $\mathrm{U} = \mathrm{D}(0, \mathrm{R}) \subset \mathbb{C}$ est un disque ouvert centré en 0 et de rayon R qui contient le spectre de a, c'est-à-dire que $\mathrm{rs}(a) < \mathrm{R}$. Supposons que φ soit une fonction complexe, holomorphe dans U. On sait que φ est représentée dans le disque ouvert de rayon R par sa série de Taylor, qui est absolument convergente en tout point de $\mathrm{D}(0, \mathrm{R})$,

$$\forall z \in \mathrm{D}(0, \mathrm{R}), \quad \varphi(z) = \sum_{n=0}^{+\infty} c_n z^n$$

où les (c_n) sont des nombres complexes. Choisissons ρ tel que $\mathrm{rs}(a) < \rho < \mathrm{R}$. On sait que $\sum |c_n| \rho^n < +\infty$. Considérons la série vectorielle

$$\sum_{n=0}^{+\infty} c_n a^n. \tag{4}$$

Puisque $\mathrm{rs}(a) = \lim \|a^n\|^{1/n} < \rho$, on aura $\|a^n\| \leq \rho^n$ pour n assez grand, ce qui montre que la série (4) est normalement convergente, donc convergente dans l'espace complet A. Il est raisonnable de poser

$$\widetilde{\varphi}(a) = \sum_{n=0}^{+\infty} c_n a^n \in \mathrm{A}.$$

On a choisi d'écrire $\widetilde{\varphi}(a)$ plutôt que $\varphi(a)$, pour éviter des confusions lors du premier contact avec cette notion. Par la suite on écrira simplement $\varphi(a)$ (et on définira ainsi e^a pour tout élément $a \in \mathrm{A}$, sans utiliser de tilde de précaution, qu'on ne saurait pas où mettre de toute façon). On a aussi la formule de Cauchy, écrite de façon incorrecte mais suggestive,

$$\widetilde{\varphi}(a) = \frac{1}{2\pi \mathrm{i}} \int_{\gamma_\rho} \frac{\varphi(z)}{z - a} \, dz$$

où γ_ρ désigne le cercle de rayon ρ, centré en 0 et parcouru une fois dans le sens direct, ou de façon correcte comme

$$\widetilde{\varphi}(a) = \frac{1}{2\pi \mathrm{i}} \int_{\gamma_\rho} \varphi(z)(z1_\mathrm{A} - a)^{-1} \, dz = \frac{1}{2\pi \mathrm{i}} \int_{\gamma_\rho} \varphi(z) \, \mathrm{R}_z(a) \, dz.$$

En effet, lorsque $|z| = \rho$, $a - z1_\mathrm{A}$ est inversible et on peut écrire

$$(z1_\mathrm{A} - a)^{-1} = z^{-1}(1_\mathrm{A} - z^{-1}a)^{-1} = \sum_{n=0}^{+\infty} z^{-n-1} a^n,$$

où la série est normalement convergente sur le cercle γ_ρ. On a donc

$$\frac{1}{2\pi \mathrm{i}} \int_{\gamma_\rho} \varphi(z)(z1_\mathrm{A} - a)^{-1} \, dz = \sum_{n=0}^{+\infty} \Bigl(\frac{1}{2\pi \mathrm{i}} \int_{\gamma_\rho} z^{-n-1} \varphi(z) \, dz \Bigr) a^n = \sum_{n=0}^{+\infty} c_n a^n.$$

Si on paramétrise le cercle par $\theta \in [0, 2\pi] \to \rho\, \mathrm{e}^{\mathrm{i}\theta}$, on obtient l'expression

$$\widetilde{\varphi}(a) = \int_0^{2\pi} \varphi(\rho\, \mathrm{e}^{\mathrm{i}\theta})(\rho\, \mathrm{e}^{\mathrm{i}\theta}\, 1_{\mathrm{A}} - a)^{-1}\, \rho\, \mathrm{e}^{\mathrm{i}\theta}\ \frac{\mathrm{d}\theta}{2\pi}$$

qui permet d'obtenir la majoration

$$\|\widetilde{\varphi}(a)\| \le \rho\, \mathrm{M}_\rho(\varphi)\ \max\{\|\mathrm{R}_\lambda(a)\| : |\lambda| = \rho\},$$

où on a posé comme précédemment pour une fonction g continue dans le disque U

$$\mathrm{M}_\rho(g) = \max\{|g(\rho\, \mathrm{e}^{\mathrm{i}\theta})| : \theta \in \mathbb{R}\}.$$

Corollaire. *Soit* U *un disque ouvert de rayon* R *centré en* 0*, contenant le spectre de* a *; si une suite* (φ_n) *de fonctions holomorphes dans* U *tend uniformément vers* φ *sur tout compact de* U*, alors*

$$\widetilde{\varphi}(a) = \lim_n \widetilde{\varphi_n}(a).$$

Preuve.— Soit ρ tel que $\mathrm{rs}(a) < \rho < \mathrm{R}$; la fonction vectorielle $z \to (z1_{\mathrm{A}} - a)^{-1} \in \mathrm{A}$ est continue sur le cercle de rayon ρ, donc bornée en norme sur ce cercle par un certain nombre C. On peut majorer la norme de la différence des deux intégrales représentant $\varphi_n(a)$ et $\varphi(a)$, c'est-à-dire

$$\widetilde{\varphi_n}(a) - \widetilde{\varphi}(a) = \int_0^{2\pi} \Big(\varphi_n(\rho\, \mathrm{e}^{\mathrm{i}\theta}) - \varphi(\rho\, \mathrm{e}^{\mathrm{i}\theta})\Big)(\rho\, \mathrm{e}^{\mathrm{i}\theta}\, 1_{\mathrm{A}} - a)^{-1}\, \rho\, \mathrm{e}^{\mathrm{i}\theta}\ \frac{\mathrm{d}\theta}{2\pi},$$

par $\mathrm{C}\, \rho\, \mathrm{M}_\rho(\varphi - \varphi_n)$, qui tend vers 0 quand n tend vers l'infini.

///

Corollaire. *Si* φ *et* ψ *sont holomorphes dans un disque ouvert* $\mathrm{U} = \mathrm{D}(0, \mathrm{R})$ *contenant le spectre de* a*, on a*

$$\big(\widetilde{\varphi\psi}\big)\,(a) = \widetilde{\varphi}(a)\, \widetilde{\psi}(a).$$

Preuve.— On peut trouver deux suites de polynômes (P_n) et (Q_n) qui tendent respectivement vers φ et ψ, uniformément sur tout compact de U (il suffit de prendre les sommes partielles des séries de Taylor). Pour des polynômes il est facile de voir que $(\mathrm{P}_n\mathrm{Q}_n)\,(a) = \mathrm{P}_n(a)\,\mathrm{Q}_n(a)$. On obtient le résultat en passant à la limite, grâce à la continuité du produit dans A.

///

Exemples. Désormais on laisse tomber les tildes.

1. Pour tout entier $k \ge 0$, $(\varphi^k)\,(a) = (\varphi(a))^k$, et donc pour tout polynôme P on a $(\mathrm{P}(\varphi))\,(a) = \mathrm{P}(\varphi(a))$.

2. On a $\mathrm{e}^{sa}\, \mathrm{e}^{ta} = \mathrm{e}^{(s+t)\,a}$; en particulier, e^a est inversible, d'inverse e^{-a}. On note que si $ab = ba$, on a $\mathrm{e}^a\, b^k = b^k\, \mathrm{e}^a$ pour tout k, donc $\mathrm{e}^{a+b} = \mathrm{e}^a\, \mathrm{e}^b$, *mais ceci ne reste pas vrai en général*, sans hypothèse de commutation.

Dans le cas C^*-algèbre, pour tout a on voit (sur la série) que $\mathrm{e}^{a^*} = (\mathrm{e}^a)^*$; si a est hermitien, $\mathrm{e}^{\mathrm{i}a}$ est donc unitaire.

Corollaire. *Si φ est holomorphe dans un disque ouvert* $\mathrm{U} = \mathrm{D}(0, \mathrm{R})$ *contenant le spectre de a, on a l'égalité*

$$(\mathrm{e}^{\varphi})\,(a) = \mathrm{e}^{\varphi(a)}\,.$$

Preuve. — On trouve une suite de polynômes (P_n) qui tend vers la fonction exponentielle, uniformément sur tout compact de $\mathbb{C}$. Pour un polynôme P_n, il résulte du corollaire précédent (et de l'exemple 1 ci-dessus) que

$$(\mathrm{P}_n(\varphi))\,(a) = \mathrm{P}_n(\varphi(a)).$$

La suite $(\mathrm{P}_n(\varphi))$ tend uniformément vers e^{φ} sur tout compact de U, donc $(\mathrm{P}_n(\varphi))(a)$ tend vers $(\mathrm{e}^{\varphi})(a)$; de l'autre côté, $\mathrm{P}_n(\varphi(a))$ tend vers $\mathrm{e}^{\varphi(a)}$.

///

Remarque. La fonction exponentielle ne joue pas un rôle très particulier dans ce qui précède. Le même résultat vaudrait pour toute fonction entière ψ (une fonction ψ holomorphe sur $\mathbb{C}$), donnant

$$(\psi(\varphi))\,(a) = \psi(\varphi(a)).$$

Proposition. *Tout élément $b \in \mathrm{A}$ tel que $\|b - 1_{\mathrm{A}}\| < 1$ peut s'écrire e^c pour un certain $c \in \mathrm{A}$.*

Preuve. — On applique ce qui précède avec $\mathrm{U} = \mathrm{D}(0, 1)$, $\varphi(z) = \ln(1 - z)$. Posons $a = 1_{\mathrm{A}} - b$; si $\|a\| < 1$, on aura $\mathrm{rs}(a) < 1$, ce qui nous met dans le cadre voulu. On obtient

$$\mathrm{e}^{\varphi(a)} = (\mathrm{e}^{\varphi})(a) = 1_{\mathrm{A}} - a = b.$$

///

Notes du chapitre 2

([a]) L'endomorphisme continu M de $\mathrm{X} = \mathrm{C}([0, 1])$ défini par $(\mathrm{M}f)(t) = t\,f(t)$ pour toute $f \in \mathrm{X}$ (opération de multiplication par la fonction identité $t \to t$) n'a pas de valeur propre : si $(\mathrm{M} - \lambda\,\mathrm{Id}_{\mathrm{X}})\,f = 0$, on a $(t - \lambda)f(t) = 0$ pour tout t, donc $f(t) = 0$ pour tout $t \neq \lambda$, donc $f = 0$ par continuité.

([b]) Si $g \in \mathrm{A}(\mathrm{D})$ était l'inverse de $z \to z$ dans l'algèbre $\mathrm{A}(\mathrm{D})$ et $\widetilde{g}$ l'extension holomorphe de g dans le disque unité ouvert D, la fonction $z \in \mathrm{D} \to 1 - z\,\widetilde{g}(z)$ serait holomorphe dans le disque unité ouvert, continue sur le disque fermé et *nulle* sur le cercle unité. D'après le principe du maximum, elle devrait être nulle dans le disque entier, ce qui est impossible puisqu'elle vaut 1 à l'origine.

3. Calcul fonctionnel continu

L'objectif principal du chapitre est de construire un homomorphisme isométrique φ_{T} de l'algèbre des fonctions continues $\mathrm{C}(\sigma(\mathrm{T}))$ vers l'algèbre $\mathcal{L}(\mathrm{H})$, lorsque T est un opérateur normal (borné) sur un espace de Hilbert complexe H ; cet homomorphisme prolongera la correspondance $\mathrm{P} \rightarrow \mathrm{P}(\mathrm{T})$ définie sur les polynômes $\mathrm{P} \in \mathbb{C}[\mathrm{X}]$. En réalité, le travail pourra être effectué dans une C^*-algèbre unitaire A. Le gain de généralité est un peu illusoire, mais les notations sont plus légères dans ce cadre abstrait.

Un *homomorphisme d'algèbres* $\varphi : \mathrm{A} \rightarrow \mathrm{B}$ doit préserver les opérations des algèbres ; c'est donc une application linéaire, et telle que $\varphi(ab) = \varphi(a)\varphi(b)$ pour tous $a, b \in \mathrm{A}$. Si on ajoute le mot *unitaire*, on demandera aussi que $\varphi(1_{\mathrm{A}}) = 1_{\mathrm{B}}$. Si on a des algèbres *de Banach*, on ajoutera la condition que φ soit continu.

3.1. Diviseurs de zéro topologiques

Lemme. *Soient* A *une algèbre de Banach unitaire complexe,* $a \in \mathrm{A}$ *un élément non inversible et* $v \in \mathrm{A}$ *un élément inversible ; on a*

$$\|v^{-1}\|^{-1} \le \|a - v\| \,;$$

il en résulte, en désignant par u *l'élément de norme* 1 *donné par* $u = \|v^{-1}\|^{-1}\, v^{-1}$*, que*

$$\|au\| \le 2\,\|a - v\|, \quad \|ua\| \le 2\,\|a - v\|.$$

Preuve.— L'élément $v^{-1}a$ ne peut pas être inversible, sinon a le serait aussi ; il en résulte que $\|1_{\mathrm{A}} - v^{-1}a\| \ge 1$ par la proposition 2.1.2, donc

$$1 \le \|1_{\mathrm{A}} - v^{-1}a\| = \|v^{-1}(v - a)\| \le \|v^{-1}\|\, \|v - a\|,$$

ce qui donne le premier résultat. Ensuite, en posant $t = \|v^{-1}\|^{-1} \le \|a - v\|$, on a pour $u = tv^{-1}$

$$\begin{aligned}\|ua\| &= t\,\|v^{-1}a\| = t\,\|1_{\mathrm{A}} - 1_{\mathrm{A}} + v^{-1}a\| \le t + t\,\|1_{\mathrm{A}} - v^{-1}a\| \\ &\le t + t\,\|v^{-1}\|\,\|v - a\| = t + \|v - a\| \le 2\,\|a - v\|.\end{aligned}$$

Pour $\|au\|$ changer de côté : réécrire tout avec av^{-1}.

///

Remarque. Lorsque a est non inversible, on a obtenu la formule

$$\mathrm{dist}(a, \mathrm{G}(\mathrm{A})) \ge \frac{1}{2}\, \inf\{\|ax\| : \|x\| = 1\}.$$

Si S est un sous-ensemble d'un espace topologique X, le *bord* ou *frontière* de S est l'ensemble

$$\partial \mathrm{S} = \overline{\mathrm{S}} \cap \overline{\mathrm{S}^c}.$$

Si X est un espace métrique et si $s \in \partial\mathrm{S}$, il existe une suite $(x_n) \subset \mathrm{S}$ et une suite $(y_n) \subset \mathrm{S}^c$ telles que $\lim_n x_n = s$, $\lim_n y_n = s$: le point s peut être approché à la fois de l'intérieur et de l'extérieur.

Proposition. *Si a appartient au bord $\partial\,\mathrm{G}(\mathrm{A})$ (dans l'espace topologique* A*) de l'ensemble* G(A)*, il existe une suite $(u_n) \subset \mathrm{A}$ d'éléments inversibles de norme 1 telle que*

$$a\,u_n \to 0, \quad u_n\,a \to 0.$$

Preuve.— Puisque a est au bord de l'ouvert G(A), il n'est pas dedans : l'élément a est non inversible, mais il existe une suite (v_n) d'inversibles telle que $\|a - v_n\| \to 0$. Pour chaque indice n, l'élément $u_n = \|v_n^{-1}\|^{-1} v_n^{-1}$ est inversible, de norme un et par le lemme précédent,

$$\|au_n\| \le 2\,\|a - v_n\| \to 0,$$

et aussi $\|u_n a\| \to 0$.

///

Corollaire 3.1.1. *Si* A *est une algèbre de Banach unitaire complexe, si $a \in \mathrm{A}$ et si $\lambda \in \partial\,\sigma(a)$, il existe une suite $(u_n) \subset \mathrm{A}$ d'éléments de norme 1 telle que*

$$a\,u_n - \lambda u_n \to 0.$$

Preuve.— Si λ est au bord du spectre, il est clair que $a - \lambda 1_{\mathrm{A}}$ est au bord de G(A) et il suffit d'appliquer ce qui précède.

///

Définition. On dira que a est un *diviseur de zéro topologique dans* A (en abrégé : dzt) s'il existe une suite $(u_n) \subset \mathrm{A}$ d'éléments de norme 1 telle que au_n tende vers 0.

On a privilégié un côté ; on aurait pu utiliser l'autre côté, ou bien une définition bilatère, mais la précédente conviendra à nos besoins.

Remarque. *Tout dzt $a \in \mathrm{A}$ est non inversible dans* A.

Si a^{-1} existe et si au_n tend vers 0, alors $a^{-1}(au_n) = u_n$ tend vers 0, ce qui est contradictoire avec l'hypothèse $\|u_n\| = 1$ pour tout n.

Remarque 3.1.2. *L'image d'un dzt $a \in \mathrm{A}$ par un homomorphisme d'algèbres de Banach unitaires* isométrique $\varphi : \mathrm{A} \to \mathrm{B}$ *reste un dzt dans* B.

En effet $\varphi(au_n) = \varphi(a)\varphi(u_n)$ tend vers 0, et $\|\varphi(u_n)\| = 1$ pour tout n.

On a le résultat simple qui suit :

pour tout homomorphisme $\varphi : \mathrm{A} \to \mathrm{B}$ d'algèbres de Banach unitaires, l'image de G(A) *est contenue dans* G(B) *; si* A *est complexe et si $a \in \mathrm{A}$, on a donc*

$$\sigma(\varphi(a)) \subset \sigma(a),$$

qui se complète par le suivant.

Corollaire. *Pour tout homomorphisme* $\varphi : \mathrm{A} \to \mathrm{B}$ *d'algèbres de Banach unitaires* isométrique*, les éléments de* $\partial \mathrm{G}(\mathrm{A})$ *ont une image non inversible, et* $\varphi(\partial \mathrm{G}(\mathrm{A})) \subset \partial \mathrm{G}(\mathrm{B})$*. Si* A *est complexe et si* $a \in \mathrm{A}$*, le bord du spectre de* a *reste dans le spectre de l'image,*

$$\partial\, \sigma(a) \subset \partial\, \sigma(\varphi(a)).$$

Preuve. — Il suffit de montrer la première affirmation ; l'autre en découle car on voit que $\varphi(a - \lambda 1_{\mathrm{A}}) = \varphi(a) - \lambda 1_{\mathrm{B}}$. Si $a \in \partial \mathrm{G}(\mathrm{A})$, on a vu que a est un dzt, et on a dit que l'image d'un dzt par un homomorphisme *isométrique* reste un dzt, donc cette image $\varphi(a)$ reste non inversible dans B. Comme a est au bord de G(A), on peut trouver une suite (v_n) d'inversibles qui tendent vers a ; leurs images $\varphi(v_n)$ sont inversibles dans B et tendent vers $\varphi(a)$, qui est donc au bord de G(B).

///

Exemple. Désignons par A l'espace des fonctions continues dans le disque unité fermé $\overline{\mathrm{D}}$ du plan complexe, et holomorphes dans le disque ouvert. Munissons cet espace de la norme uniforme et du produit ponctuel des fonctions ; on obtient ainsi une algèbre de Banach complexe, qui est clairement [a] isomorphe (et même *isométrique*) à *l'algèbre du disque* de l'exemple 2.2.2. Il est clair que la fonction $z \to z$ définit un élément $a \in \mathrm{A}$ dont le spectre est égal au disque unité fermé (pour tout $\lambda \in \overline{\mathrm{D}}$, la fonction $z - \lambda$ ne peut pas avoir d'inverse dans A puisqu'elle s'annule au point $\lambda \in \overline{\mathrm{D}}$). Considérons un compact K contenant le cercle unité $\mathbb{T}$ et contenu dans $\overline{\mathrm{D}}$, et l'homomorphisme φ de restriction de A dans l'algèbre $\mathrm{B} = \mathrm{C}(\mathrm{K})$. Cet homomorphisme est isométrique par le principe du maximum (puisque $\mathbb{T} \subset \mathrm{K}$), et l'image $b = \varphi(a)$ a pour spectre dans C(K) l'ensemble des valeurs de la fonction sur K (exemples 2.2.1), donc $\sigma(b) = \mathrm{K}$. On voit bien que le bord $\mathbb{T}$ du spectre de a reste dans le (bord du) spectre de b, mais le bord du spectre de b peut être plus grand que $\partial\, \sigma(a)$.

Exercice. Si φ est un homomorphisme isométrique de A dans B, si $a \in \mathrm{A}$ et $b = \varphi(a)$, montrer que toute composante connexe V de l'ensemble résolvant $\rho(b)$ qui rencontre $\rho(a)$ est contenue dans $\rho(a)$ (cependant on a $\rho(b) \supset \rho(a)$). Ainsi, *le spectre* $\sigma(a)$ *dans la petite algèbre est la réunion de* $\sigma(\varphi(a))$ *(dans la grosse algèbre) et de certaines composantes connexes de* $\rho(\varphi(a))$. Montrer en particulier que la composante connexe de l'infini dans $\rho(a)$ est préservée dans $\rho(b)$. Par conséquent, la partie du (bord du) spectre de a qui « touche » la composante de l'infini de $\rho(a)$ est préservée par φ.

Corollaire. *Si* A *est complexe et si le spectre de* $a \in \mathrm{A}$ *est d'intérieur vide dans* $\mathbb{C}$*, le spectre de* a *est préservé par tout homomorphisme isométrique de* A *dans une autre algèbre de Banach unitaire* B.

Preuve. — On a ici $\sigma(a) = \partial\, \sigma(a)$, donc par le corollaire précédent

$$\sigma(a) = \partial\, \sigma(a) \subset \partial\, \sigma(\varphi(a)) \subset \sigma(\varphi(a)),$$

et on a toujours l'inclusion inverse $\sigma(a) \supset \sigma(\varphi(a))$.

///

Cas où $\mathrm{A} = \mathcal{L}(\mathrm{X})$

Si T est un dzt dans $\mathcal{L}(\mathrm{X})$, il existe une suite $(\mathrm{U}_n) \subset \mathcal{L}(\mathrm{X})$ d'opérateurs de norme 1 telle que TU_n tende vers 0 en norme d'opérateur. Puisque $\|\mathrm{U}_n\| = 1$, on peut trouver $y_n \in \mathrm{X}$ tel que $\|y_n\| \le 2$ et $\|\mathrm{U}_n y_n\| = 1$. La suite $x_n = \mathrm{U}_n y_n$ est alors une suite de vecteurs de norme 1 telle que $\mathrm{T}x_n \to 0$. On peut dire que 0 est une *valeur propre approchée* pour l'opérateur T.

Réciproquement, supposons qu'il existe une suite $(x_n) \subset \mathrm{X}$ de vecteurs de norme 1 telle que $\lim \mathrm{T}x_n = 0$; soit ξ_n une forme linéaire de norme 1 telle que $\xi_n(x_n) = 1$. L'opérateur V_n défini par $\mathrm{V}_n y = \xi_n(y) x_n$ est de norme un et TV_n tend vers 0.

Corollaire 3.1.3. *Pour tout endomorphisme* $\mathrm{T} \in \mathcal{L}(\mathrm{X})$ *d'un espace de Banach complexe* $\mathrm{X} \neq \{0\}$, *il existe* $\lambda \in \mathbb{C}$ *et une suite* $(x_n) \subset \mathrm{X}$ *de vecteurs de norme un telle que*

$$\mathrm{T}x_n - \lambda x_n \to 0.$$

Preuve. — On a montré que $\sigma(\mathrm{T})$ est un compact non vide ; on peut donc trouver un point λ au bord de $\sigma(\mathrm{T})$ (par exemple un point de module maximal), et on applique ce qui précède.

///

Exercice. Spectre des opérateurs de multiplication par $a \in \mathrm{A}$: si A est une algèbre de Banach unitaire, montrer que le spectre de a dans A est égal au spectre de l'opérateur $\mathrm{M}_a \in \mathcal{L}(\mathrm{A})$ de multiplication par a, défini par $\mathrm{M}_a(b) = ab$ pour tout $b \in \mathrm{A}$.

Exemple. Le nombre complexe 1 est valeur propre approchée pour le shift S sur $\ell_2(\mathbb{N})$, car le vecteur

$$x^{(n)} = \frac{1}{\sqrt{n}} \sum_{j=0}^{n-1} \mathbf{e}_j \in \ell_2(\mathbb{N})$$

vérifie $\|x^{(n)}\| = 1$ et $\|\mathrm{S}x^{(n)} - x^{(n)}\|^2 = 2/n$, qui tend vers 0. En modifiant cet exemple, on vérifie que tout complexe de module un est valeur propre approchée pour le shift S.

Spectre des éléments hermitiens d'une C^**-algèbre*

Si a est un élément hermitien d'une C^*-algèbre A, considérons un point frontière λ du spectre de a ; il existe des éléments $u \in \mathrm{A}$ de norme 1 tels que $\|(a - \lambda 1_\mathrm{A})u\| < \varepsilon/2$; en multipliant par u^* on obtient $\|u^*(a - \lambda 1_\mathrm{A})u\| < \varepsilon/2$; comme $\|v^*\| = \|v\|$ pour tout v, on a aussi $\|u^*(a - \overline{\lambda} 1_\mathrm{A})u\| < \varepsilon/2$ en prenant l'adjoint, puis par différence on voit que

$$|\lambda - \overline{\lambda}| = |\lambda - \overline{\lambda}| \, \|u^* u\| < \varepsilon,$$

ce qui prouve que λ est réel.

Le bord du spectre étant contenu dans la droite réelle, il en résulte que tout le spectre est réel. En effet, il existe $\lambda_0 = \sigma_0 + \mathrm{i}\tau_0 \in \sigma(a)$ tel que $|\operatorname{Im} \lambda_0| = |\tau_0|$ soit maximal parmi tous les éléments λ du spectre. Alors $\sigma_0 + \mathrm{i}(1+\varepsilon)\tau_0$ ne peut pas être dans $\sigma(a)$ pour $\varepsilon > 0$, donc λ_0 est dans le bord du spectre, donc réel, donc $\tau_0 = 0$; ceci montre que tout le spectre est réel.

Proposition. *Soit* A *une* C*-*algèbre unitaire ; si* $a \in \mathrm{A}$ *est hermitien on a*

$$\sigma(a) \subset \mathbb{R}.$$

Lemme 3.1.4. *Soit* A *une* C*-*algèbre ; si* $a \in \mathrm{A}$ *est normal on a*

$$\|ax\| = \|a^*x\|$$

pour tout $x \in \mathrm{A}$ *; en conséquence,* $\|ax - \lambda x\| = \|a^*x - \overline{\lambda}x\|$ *pour tout* $\lambda \in \mathbb{C}$.

Preuve.— Quand a est normal,

$$\|ax\|^2 = \|x^*a^*ax\| = \|x^*aa^*x\| = \|a^*x\|^2\,;$$

de plus, $(a - \lambda 1_\mathrm{A})^* = a^* - \overline{\lambda}1_\mathrm{A}$ et $a - \lambda 1_\mathrm{A}$ est normal.

///

Proposition 3.1.5. *Soit* A *une* C*-*algèbre ; si* $a \in \mathrm{A}$ *est normal et* $\lambda \in \sigma(a)$, *il existe une suite* $(u_n) \subset \mathrm{A}$ *d'éléments de norme* 1 *telle que*

$$au_n - \lambda u_n \to 0, \quad a^*u_n - \overline{\lambda}u_n \to 0.$$

Preuve.— L'élément $b = (a - \lambda 1_\mathrm{A})$ est normal et non inversible, donc $c = b^*b$ est non inversible (parce que $b^*b = b\,b^*$ et que b est non inversible : on montre —exercice facile[b]— que si $xy = yx$ est inversible, x est inversible). On a donc $0 \in \sigma(c)$; puisque c est hermitien, son spectre est contenu dans $\mathbb{R}$, tous les points du spectre de c sont au bord (pour la topologie de $\mathbb{C}$) et d'après le corollaire 3.1.1, il existe une suite $(u_n) \subset \mathrm{A}$ d'éléments de norme un telle que $cu_n \to 0$; par conséquent,

$$\lim_n \|bu_n\|^2 = \lim_n \|u_n^*b^*bu_n\| = \lim_n \|u_n^*cu_n\| = 0,$$

donc $au_n - \lambda u_n \to 0$, et l'autre propriété est automatique par le lemme 3.1.4.

///

Vecteurs propres approchés communs à deux endomorphismes qui commutent

Lemme 3.1.6. *On suppose que* $\mathrm{S}, \mathrm{T} \in \mathcal{L}(\mathrm{X})$, *avec* X *espace de Banach complexe, que* $\mathrm{TS} = \mathrm{ST}$, *et qu'il existe une suite* $(x_k) \subset \mathrm{X}$ *de vecteurs de norme* 1 *telle que* $\mathrm{S}\,x_k \to 0$ *; on peut trouver* $\lambda \in \mathbb{C}$ *et une suite* $(y_k) \subset \mathrm{X}$ *de vecteurs de norme* 1 *tels que*

$$\mathrm{S}y_k \to 0\,; \quad \mathrm{T}y_k - \lambda y_k \to 0.$$

Preuve. — Quand on a un vecteur $x \neq 0$ tel que $\mathrm{S}x = 0$, on peut raisonner sur le sous-espace non nul $\mathrm{Y} = \ker \mathrm{S}$, qui est stable par T, et utiliser le corollaire 3.1.3 appliqué à la restriction de T à ce sous-espace. L'idée ici est de considérer l'objet global (x_k) comme un vecteur ξ d'un espace normé Y à construire, vecteur qui vérifie « moralement » $\mathrm{S}\,\xi = 0$; mais il faudra aussi introduire les images morales par T de ce vecteur, à savoir $\mathrm{T}\xi$, $\mathrm{T}^2\xi$, etc... et plus généralement tous les $\mathrm{P}(\mathrm{T})\xi$ pour tous les polynômes $\mathrm{P} \in \mathbb{C}[\mathrm{X}]$. On va en fait assimiler ces « vecteurs moraux » $\mathrm{P}(\mathrm{T})\xi$ aux polynômes P eux-mêmes, et définir un espace Y qui sera le complété de l'espace des polynômes pour une norme qui reflétera le comportement des suites $(\mathrm{P}(\mathrm{T})\,x_k)$. Par un argument d'extraction de suite diagonale, on peut supposer que la suite (x_k) de l'énoncé du lemme est telle que la limite

$$\lim_k \|\mathrm{P}(\mathrm{T})\,x_k\|$$

existe pour tout P appartenant à l'ensemble dénombrable des polynômes à coefficients $c_j = a_j + \mathrm{i}b_j$ avec a_j, b_j rationnels. Par prolongement par continuité, la limite existera pour tous les polynômes $\mathrm{P} \in \mathbb{C}[\mathrm{X}]$. On définit une semi-norme sur l'espace $\mathrm{Y}_0 := \mathbb{C}[\mathrm{X}]$ des polynômes à coefficients complexes en posant

$$\forall \mathrm{P} \in \mathbb{C}[\mathrm{X}], \quad \|\mathrm{P}\| = \lim_k \|\mathrm{P}(\mathrm{T})\,x_k\|.$$

On notera $\mathbf{e}_j$ le monôme X^j, pour tout $j \geq 0$. Ainsi $\|\mathbf{e}_0\| = \lim_k \|\mathrm{T}^0 x_k\| = 1$. On définit ensuite une application linéaire τ sur Y_0, qui doit être le reflet de T, en posant $\tau(\mathrm{P}) = \mathrm{XP}$ (le produit de P par le monôme X) pour tout $\mathrm{P} \in \mathrm{Y}_0$. On voit que τ agit comme un opérateur de décalage sur la base $(\mathbf{e}_j)_{j \geq 0}$ de l'espace des polynômes. On a

$$\|\tau\,\mathrm{P}\| = \|\mathrm{XP}\| = \lim_k \|\mathrm{TP}(\mathrm{T})\,x_k\| \leq \|\mathrm{T}\| \lim_k \|\mathrm{P}(\mathrm{T})\,x_k\| = \|\mathrm{T}\|\,\|\mathrm{P}\|,$$

donc $\|\tau\| \leq \|\mathrm{T}\|$.

Supposons d'abord que la semi-norme sur Y_0 soit une norme. Désignons par Y le complété de l'espace normé Y_0 ; l'opérateur τ se prolonge par continuité en un endomorphisme de Y. Il existe par conséquent par le corollaire 3.1.3 une valeur propre approchée λ pour τ, donc pour tout n il existe un vecteur $y_n \in \mathrm{Y}$ de norme 1, que l'on peut choisir de la forme $y_n = \mathrm{P}_n \in \mathrm{Y}_0$ à cause de la densité de Y_0, tel que $\|\tau\,\mathrm{P}_n - \lambda \mathrm{P}_n\| < 2^{-n}$, c'est-à-dire que pour k assez grand

$$\|\mathrm{TP}_n(\mathrm{T})\,x_k - \lambda \mathrm{P}_n(\mathrm{T})\,x_k\| < 2^{-n}\,;$$

posons $u_{k,n} = \mathrm{P}_n(\mathrm{T})\,x_k$; on a $\lim_k \|u_{k,n}\| = \|\mathrm{P}_n\| = 1$. On voit que $\mathrm{S}\,u_{k,n} = \mathrm{SP}_n(\mathrm{T})\,x_k = \mathrm{P}_n(\mathrm{T})\mathrm{S}\,x_k$ tend vers 0 quand $k \to +\infty$. Pour chaque $n \geq 0$, posons $y_n = \|u_{k,n}\|^{-1} u_{k,n}$, où k est choisi assez grand pour que $\|\mathrm{S}\,u_{k,n}\| < 2^{-n}$, $\|\mathrm{T}\,u_{k,n} - \lambda u_{k,n}\| < 2^{-n}$ et $\|u_{k,n}\| > 1/2$, de sorte que $\|\mathrm{S}\,y_n\| < 2^{-n+1}$ et $\|\mathrm{T}\,y_n - \lambda y_n\| < 2^{-n+1}$. Ceci termine la preuve de ce premier cas.

Si la semi-norme n'est pas une norme, soit P un polynôme non nul de degré minimal d tel que $\|\mathrm{P}\| = 0$; on ne peut pas avoir $d = 0$ car $\|\mathbf{e}_0\| = 1$. Puisque $d \geq 1$, on peut factoriser P sous la forme $\mathrm{P} = (\mathrm{X} - \lambda)\mathrm{Q}$, $\lambda \in \mathbb{C}$, et on aura

$$0 = \|\mathrm{P}\| = \|(\mathrm{X} - \lambda)\mathrm{Q}\| = \lim_k \|(\mathrm{T} - \lambda\,\mathrm{Id})\mathrm{Q}(\mathrm{T})\,x_k\|$$

et $\lim_k \|\mathrm{Q}(\mathrm{T})\,x_k\| \neq 0$ par la minimalité de d ; la suite $y_k = \|\mathrm{Q}(\mathrm{T})\,x_k\|^{-1}\mathrm{Q}(\mathrm{T})\,x_k$ vérifie les conclusions voulues.

///

3.2. Calcul fonctionnel continu pour les éléments normaux

On considère une C^*-algèbre unitaire A. On notera $\mathbf{C}$ l'espace vectoriel de toutes les familles $\mathbf{c} = (c_{m,n})_{m,n \ge 0}$ de nombres complexes, dont un nombre fini seulement est non nul. Si $a \in \mathrm{A}$ est normal, on pose

$$\mathrm{F}_{\mathbf{c}}(a) = \sum_{m,n} c_{m,n}\, a^m (a^*)^n \in \mathrm{A}.$$

Si b commute avec a et avec a^*, il est clair que $b\, a^m (a^*)^n = a^m (a^*)^n b$ pour tous $m, n \ge 0$, par conséquent en passant aux combinaisons linéaires on obtient que

$$b\, \mathrm{F}_{\mathbf{c}}(a) = \mathrm{F}_{\mathbf{c}}(a)\, b\,;$$

en particulier, comme $b = a^k (a^*)^\ell$ commute avec a et avec a^* pour tous les entiers $k, \ell \ge 0$, on a

$$a^k (a^*)^\ell\, \mathrm{F}_{\mathbf{c}}(a) = \mathrm{F}_{\mathbf{c}}(a)\, a^k (a^*)^\ell,$$

et il en résulte en passant aux combinaisons linéaires que $\mathrm{F}_{\mathbf{c}}(a)$ commute avec tous les autres éléments de la forme $\mathrm{F}_{\mathbf{d}}(a)$, $\mathbf{d} \in \mathbf{C}$. Comme l'adjoint de $\mathrm{F}_{\mathbf{c}}(a)$ est égal à $\mathrm{F}_{\mathbf{c}^*}(a)$, où $\mathbf{c}^* = (\overline{c}_{n,m})$, on voit que tous les éléments de la forme $\mathrm{F}_{\mathbf{c}}(a)$ sont normaux.

À chaque $\mathbf{c} \in \mathbf{C}$ on associera la fonction $f_{\mathbf{c}}$ définie sur $\mathbb{C}$ par

$$f_{\mathbf{c}}(z) = \sum_{m,n} c_{m,n} z^m \overline{z}^n \in \mathbb{C}.$$

Le résultat essentiel pour la suite est le suivant.

Lemme 3.2.1. *Si* A *est une* C^**-algèbre unitaire et si* $a \in \mathrm{A}$ *est normal, on a*

$$\sigma(\mathrm{F}_{\mathbf{c}}(a)) = \{f_{\mathbf{c}}(\lambda) : \lambda \in \sigma(a)\}.$$

Il en résulte que

$$\|\mathrm{F}_{\mathbf{c}}(a)\| = \max\{|f_{\mathbf{c}}(\lambda)| : \lambda \in \sigma(a)\} = \|f_{\mathbf{c}}\|_{\mathrm{C}(\sigma(a))}$$

pour tout $\mathbf{c} \in \mathbf{C}$.

Preuve.— Soit $\mathrm{K} = \sigma(a)$ et posons $b = \mathrm{F}_{\mathbf{c}}(a)$; si $\lambda \in \mathrm{K}$, on peut trouver d'après la proposition 3.1.5 une suite d'éléments $(u_n) \subset \mathrm{A}$ de norme un telle que

$$au_n - \lambda u_n \to 0, \quad a^* u_n - \overline{\lambda} u_n \to 0.$$

Il en résulte par récurrence sur $k, \ell \ge 0$ que

$$a^k (a^*)^\ell u_n - \lambda^k \overline{\lambda}^\ell u_n \to 0\,; \tag{1}$$

montrons par exemple comment passer de ℓ à $\ell + 1$: on suppose la ligne précédente vérifiée ; par continuité du produit avec a^* et commutation, on obtient à partir de l'hypothèse de récurrence

$$a^* \bigl(a^k (a^*)^\ell u_n - \lambda^k \overline{\lambda}^\ell u_n\bigr) = a^k (a^*)^{\ell+1} u_n - \lambda^k \overline{\lambda}^\ell a^* u_n \to 0$$

et par la propriété de départ

$$\lambda^k \overline{\lambda}^\ell (a^* u_n - \overline{\lambda} u_n) = \lambda^k \overline{\lambda}^\ell a^* u_n - \lambda^k \overline{\lambda}^{\ell+1} u_n \to 0$$

d'où le résultat voulu en $k, \ell+1$, en additionnant les deux lignes ; par conséquent, en combinant un nombre fini des résultats obtenus par la propriété (1), on voit que

$$\mathrm{F}_{\mathbf{c}}(a)\, u_n - f_{\mathbf{c}}(\lambda)\, u_n = \sum_{k,\ell} c_{k,\ell} \left(a^k (a^*)^\ell u_n - \lambda^k \overline{\lambda}^\ell u_n \right) \to 0.$$

On en déduit que $\mathrm{F}_{\mathbf{c}}(a) - f_{\mathbf{c}}(\lambda) 1_{\mathrm{A}}$ est un dzt, donc est non inversible et on a montré que

$$\{ f_{\mathbf{c}}(\lambda) : \lambda \in \mathrm{K} \} \subset \sigma(\mathrm{F}_{\mathbf{c}}(a)).$$

La preuve de l'inclusion inverse s'appuie sur le lemme 3.1.6. Posons $b = \mathrm{F}_{\mathbf{c}}(a)$; si on suppose que $\mu \in \sigma(b)$, il existe puisque b est normal une suite (u_n) de norme un telle que $bu_n - \mu u_n \to 0$, d'après la proposition 3.1.5 ; l'opérateur $\mathrm{S} \in \mathcal{L}(\mathrm{A})$ défini par $\mathrm{S}x = bx - \mu x$ pour tout $x \in \mathrm{A}$ vérifie $\mathrm{S} u_n \to 0$, et l'opérateur $\mathrm{T} \in \mathcal{L}(\mathrm{A})$ de multiplication par a, défini par $\mathrm{T}x = ax$, commute avec S ; on peut donc trouver λ et une suite (v_n) de norme un telle que

$$\mathrm{S} v_n = b v_n - \mu v_n \to 0, \quad \mathrm{T} v_n - \lambda v_n = a v_n - \lambda v_n \to 0,$$

ce qui entraîne $a^* v_n - \overline{\lambda} v_n \to 0$. On obtient comme précédemment

$$\mathrm{F}_{\mathbf{c}}(a)\, v_n - f_{\mathbf{c}}(\lambda)\, v_n \to 0,$$

et il en résulte que $\mu = f_{\mathbf{c}}(\lambda)$.

On a ainsi identifié le spectre de $b = \mathrm{F}_{\mathbf{c}}(a)$; comme b est normal, on sait d'après la proposition 2.2.5 que

$$\|b\| = \mathrm{rs}(b) = \max\{ |\mu| : \mu \in \sigma(b) \} = \max\{ |f_{\mathbf{c}}(\lambda)| : \lambda \in \sigma(a) \}.$$

///

Remarque. Le résultat précédent est à peu près évident lorsque $\mathrm{T} \in \mathcal{L}(\mathrm{H})$ est normal et compact. Dans ce cas, il existe une base hilbertienne $(e_i)_{i \in \mathrm{I}}$ de H telle que $\mathrm{T} e_i = \lambda_i e_i$ pour tout i, et de plus pour tout $\varepsilon > 0$ il n'existe qu'un nombre fini d'indices $i \in \mathrm{I}$ tels que $|\lambda_i| \ge \varepsilon$. L'opérateur $\mathrm{S} = \mathrm{F}_{\mathbf{c}}(\mathrm{T})$ est l'opérateur diagonal dont les coefficients diagonaux sont les $f_{\mathbf{c}}(\lambda_i)$, la norme de S est donc le sup des $|f_{\mathbf{c}}(\lambda_i)|$, qui est majoré par le sup de $|f_{\mathbf{c}}|$ sur le spectre K de T puisque chaque λ_i est dans le spectre. Inversement, si λ est dans le spectre de T, ou bien λ est valeur propre de T, et λ est l'un des λ_i, donc $\|\mathrm{F}_{\mathbf{c}}(\mathrm{T})\| \ge \|\mathrm{S}(e_i)\| = |f_{\mathbf{c}}(\lambda_i)| = |f_{\mathbf{c}}(\lambda)|$, ou bien $\lambda = 0$ est limite d'une suite de λ_i, ce qui conduit au même résultat puisque $f_{\mathbf{c}}$ définit une fonction continue sur K. On a donc bien $\|\mathrm{F}_{\mathbf{c}}(\mathrm{T})\| = \|f_{\mathbf{c}}\|_{\mathrm{C(K)}}$.

Faisons quelques pas de plus ; si deux éléments $\mathbf{c}, \mathbf{d} \in \mathbf{C}$ définissent la même fonction f sur $\mathrm{K} = \sigma(a)$, c'est-à-dire si

$$\forall z \in \mathrm{K}, \quad f_{\mathbf{c}}(z) = f(z) = f_{\mathbf{d}}(z),$$

on aura $f_{\mathbf{c}-\mathbf{d}} = f_{\mathbf{c}} - f_{\mathbf{d}} = 0$ sur K et d'après le lemme 3.2.1 on aura

$$\|\mathrm{F}_{\mathbf{c}}(a) - \mathrm{F}_{\mathbf{d}}(a)\|_{\mathrm{A}} = \|\mathrm{F}_{\mathbf{c}-\mathbf{d}}(a)\|_{\mathrm{A}} = \|f_{\mathbf{c}-\mathbf{d}}\|_{\mathrm{C(K)}} = 0,$$

ce qui montre que $\mathrm{F}_{\mathbf{c}}(a) = \mathrm{F}_{\mathbf{d}}(a)$. Ceci prouve qu'il est possible d'associer à toute fonction $f \in \mathrm{C(K)}$ de la forme

$$\forall z \in \mathrm{K}, \quad f(z) = \sum_{m,n} c_{m,n}\, z^m \overline{z}^n \tag{2}$$

un élément $\psi(f) \in \mathrm{A}$ qui ne dépend que de la fonction f,

$$\psi(f) = \sum_{m,n} c_{m,n}\, a^m (a^*)^n \in \mathrm{A}.$$

De plus on a vu que

$$\|\psi(f)\|_{\mathrm{A}} = \|f\|_{\mathrm{C(K)}}.$$

Désignons par $\mathcal{P}$ le sous-ensemble de C(K) formé de toutes les fonctions de la forme (2), c'est-à-dire de la forme $z \in \mathrm{K} \to f_{\mathbf{c}}(z)$ pour un $\mathbf{c} \in \mathbf{C}$; il est clair que $\mathcal{P}$ est une sous-algèbre de C(K), stable par conjugaison complexe. Il est facile de vérifier que ψ est linéaire sur $\mathcal{P}$; si f est de la forme (2), alors la fonction conjuguée $\overline{f}$, qui a la forme

$$\forall z \in \mathrm{K}, \quad \overline{f(z)} = \sum_{m,n} \overline{c_{m,n}}\, z^n \overline{z}^m$$

a pour image

$$\psi(\overline{f}) = \sum_{m,n} \overline{c_{m,n}}\, a^n (a^*)^m = \psi(f)^*.$$

Pour tous $k, \ell \geq 0$, la fonction $z \in \mathrm{K} \to z^k \overline{z}^\ell f(z)$, représentée par

$$\sum_{m,n} c_{m,n}\, z^{m+k} \overline{z}^{n+\ell}$$

a pour image

$$\sum_{m,n} c_{m,n}\, a^{m+k} (a^*)^{n+\ell} = a^k (a^*)^\ell \psi(f),$$

ce qui montre, en passant aux combinaisons linéaires, que $\psi(gf) = \psi(g)\psi(f)$ quand g est une autre fonction de $\mathcal{P}$, de la forme

$$g(z) = \sum_{k,\ell} d_{k,\ell}\, z^k \overline{z}^\ell.$$

Lemme. *Soient* K *un compact non vide ; les éléments non inversibles de* C(K) *sont exactement les dzt.*

Preuve. — On pourrait se contenter de dire : la C*-algèbre C(K) est commutative, donc tous ses éléments sont normaux et le résultat provient de la proposition 3.1.5, appliquée à l'élément 0 du spectre. Il est cependant intéressant de décrypter ce résultat dans le cas particulier présent. Supposons f non inversible dans C(K) ; on sait alors qu'il existe $s_0 \in \mathrm{K}$ tel que $f(s_0) = 0$; pour tout $\varepsilon > 0$ petit, la fonction inversible complexe

$$g_\varepsilon(s) = |f(s)| + \mathrm{i}\varepsilon$$

est proche de l'élément hermitien non inversible $|f|$, et le minimum du module de g_ε est ε, atteint en s_0 ; la fonction

$$h_\varepsilon = \frac{\varepsilon}{g_\varepsilon(s)}$$

est de norme un, et

$$f(s)h_\varepsilon(s) = \varepsilon\,\frac{f(s)}{g_\varepsilon(s)}$$

est plus petite en module que ε partout. En prenant successivement $\varepsilon = 2^{-n}$ on construira une suite (h_n) de fonctions de norme 1 telle que $fh_n \to 0$.

///

Corollaire 3.2.2. *Pour tout homomorphisme isométrique φ de* C(K) *(complexe) dans une algèbre de Banach unitaire complexe* B, *on a*

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

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

Preuve. — On a vu que $\sigma(f) = f(\mathrm{K})$ dans les exemples 2.2.1 ; d'après le lemme précédent, tout élément de $\sigma(f)$ est un dzt, donc son image est non inversible d'après la remarque 3.1.2. On a donc $\sigma(f) \subset \sigma(\varphi(f))$, ce qui entraîne l'égalité.

///

Passons au théorème sur le calcul fonctionnel continu. Si K est un compact de $\mathbb{C}$, on notera ζ_K la fonction $z \in \mathrm{K} \to z \in \mathbb{C}$. Un $*$-*homomorphisme* entre deux C*-algèbres est un homomorphisme d'algèbres φ qui vérifie de plus

$$\varphi(x^*) = \varphi(x)^*$$

pour tout x dans la C*-algèbre de départ.

Théorème 3.2.3. *Soient* A *une* C**-algèbre unitaire, a un élément normal dans* A, *et posons* $\mathrm{K} = \sigma(a)$ *; il existe un et un seul $*$-homomorphisme de* C**-algèbres unitaires* $\varphi_a : \mathrm{C(K)} \to \mathrm{A}$ *tel que* $\varphi_a(\zeta_\mathrm{K}) = a$.

L'homomorphisme φ_a est isométrique ; pour toute fonction f continue sur K, *on note*

$$f(a) = \varphi_a(f)\,;$$

on a $f(a)^* = \overline{f}(a)$ et $f(a)$ est normal ; en fait, $f(a)$ commute avec tout élément $b \in \mathrm{A}$ qui commute avec a et avec a^*. On a de plus

$$\sigma(f(a)) = \sigma(f) = f(\sigma(a)).$$

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

Preuve.— Si φ est un tel homomorphisme, on a $\varphi(1) = 1_{\mathrm{A}}$ (unitaire), $\varphi(\zeta) = a$ (imposé), $\varphi(\overline{\zeta}) = a^*$ ($*$-homomorphisme), donc $\varphi(\zeta^k\overline{\zeta}^\ell) = a^k(a^*)^\ell$. Pour toute fonction f sur K de la forme

$$f = \sum_{m,n} c_{m,n}\, \zeta^m \overline{\zeta}^n,$$

on voit que l'image ne peut être que

$$\sum_{m,n} c_{m,n}\, a^m (a^*)^n = \mathrm{F}_{\mathbf{c}}(a).$$

Les fonctions telles que f forment l'algèbre $\mathcal{P}$ introduite précédemment ; cette algèbre $\mathcal{P}$ contient les constantes, est stable par conjugaison complexe et sépare les points de K (la seule fonction ζ suffit à séparer les points de K). D'après le théorème de Stone-Weierstrass, cette algèbre $\mathcal{P}$ est dense dans C(K). Si φ_1 et φ_2 sont deux solutions du problème, on a dit qu'on doit avoir $\varphi_1 - \varphi_2 = 0$ sur $\mathcal{P}$, donc partout sur C(K) par densité et continuité.

Inversement, on a défini avant l'énoncé du théorème un homomorphisme ψ de $\mathcal{P}$ dans A, qui associe à toute fonction $f \in \mathrm{C(K)}$ de la forme $f_{\mathbf{c}}$ l'élément bien défini $\psi(f) = \mathrm{F}_{\mathbf{c}}(a)$, et on a vu que

$$\|\psi(f)\|_{\mathrm{A}} = \|f\|_{\mathrm{C(K)}}.$$

On peut donc trouver un prolongement unique φ_a de ψ en application linéaire continue de C(K) dans A. Posons $f(a) = \varphi_a(f)$ pour toute $f \in \mathrm{C(K)}$. Pour toute suite (f_n) de fonctions de $\mathcal{P}$ qui converge uniformément sur K vers la fonction f, la suite $(\psi(f_n))$ tend en norme dans A vers $f(a)$, puisque φ_a est continu. Il en résulte par continuité de la norme que

$$\|f(a)\| = \lim_n \|\psi(f_n)\| = \lim_n \|f_n\|_{\mathrm{C(K)}} = \|f\|_{\mathrm{C(K)}},$$

ce qui montre que l'application $\varphi_a : f \in \mathrm{C(K)} \to f(a) \in \mathrm{A}$ est isométrique.

Par construction on a $\varphi_a(\zeta) = a$. Il reste à voir que φ_a est un homomorphisme. Si (f_n) converge uniformément vers f sur K et (g_n) converge uniformément vers g sur K, alors $f(a)g(a) = \lim_n \psi(f_n g_n) = (fg)(a)$ (utiliser la continuité du produit par rapport au couple de variables) ; on aura aussi $\overline{f}(a) = \lim_n \psi(\overline{f}_n) = \lim_n f_n(a)^* = f(a)^*$; donc φ_a est un $*$-homomorphisme de C*-algèbres de Banach unitaires complexes, isométrique. Il en résulte que $\sigma(f(a)) = f(\mathrm{K})$, d'après un principe général sur C(K) (corollaire 3.2.2). Les éléments $f(a)$ sont normaux puisque

$$f(a)^* f(a) = \varphi_a(\overline{f} f) = \varphi_a(f \overline{f}) = f(a) f(a)^*.$$

Plus généralement, si b commute avec a et avec a^*, on en déduit que $bf_n(a) = f_n(a)b$ pour tout n, donc $bf(a) = f(a)b$ par continuité du produit par b, à droite et à gauche. Ainsi $f(a)$ commute avec tout élément b qui commute avec a et avec a^*.

Supposons que f soit une fonction continue sur $\mathrm{K} = \sigma(a)$. Alors $f(a)$ est normal, ce qui permet d'appliquer à $f(a)$ le calcul fonctionnel défini précédemment. L'ensemble $\mathrm{L} = f(\mathrm{K}) \subset \mathbb{R}$ est compact, et c'est le spectre de $f(a)$. On voit alors que l'application $g \in \mathrm{C(L)} \to g \circ f \in \mathrm{C(K)}$ est un $*$-homomorphisme χ de C*-algèbres de C(L) dans C(K), et $\varphi_a \circ \chi$ est un $*$-homomorphisme qui transforme ζ_{L} en $f(a)$ et $\overline{\zeta}_{\mathrm{L}}$ en $f(a)^*$; d'après l'unicité, la composition $\varphi_a \circ \chi$ est égale à l'homomorphisme $\varphi_{f(a)}$ associé à l'opérateur hermitien $f(a)$. On a donc $(g \circ f)(a) = g(f(a))$ pour toute fonction continue g sur L.

///

Exemples évidents.

1. Supposons que T soit diagonal dans une base orthonormée, avec coefficients diagonaux (λ_n) ; l'opérateur T est alors normal. Pour toute fonction continue f définie sur $\mathbb{R}$, l'opérateur $f(\mathrm{T})$ est l'opérateur diagonal de coefficients $(f(\lambda_n))$; démonstration : passer à la limite à partir du cas $f = f_{\mathbf{c}}$.

2. Supposons que T soit l'opérateur $\mathrm{M}_\varphi : \mathrm{L}_2(0,1) \to \mathrm{L}_2(0,1)$ de multiplication par une fonction φ réelle continue. On voit que pour tout $\mathbf{c} \in \mathbf{C}$ l'opérateur $\mathrm{F}_{\mathbf{c}}(\mathrm{M}_\varphi)$ est l'opérateur de multiplication par la fonction $s \in [0,1] \to f_{\mathbf{c}}(\varphi(t))$, donc à la limite $f(\mathrm{M}_\varphi)$ est l'opérateur de multiplication par $s \to f(\varphi(s))$, c'est-à-dire que $f(\mathrm{M}_\varphi) = \mathrm{M}_{f\circ\varphi}$. On peut aussi raisonner en disant que $f \to \mathrm{M}_{f\circ\varphi}$ est bien l'unique homomorphisme décrit dans le théorème 3.2.3.

Corollaire. *Soient* A *une* C^**-algèbre,* $a \in \mathrm{A}$ *un élément normal et* f *une fonction complexe continue sur* $\sigma(a)$ *; alors, si* $f(\sigma(a)) \subset \mathbb{R}$*,* $f(a)$ *est hermitien ; si* $f(\sigma(a))$ *est contenu dans le cercle unité* $\mathbb{T}$ *du plan complexe, alors* $f(a)$ *est unitaire.*

Preuve.— Si la fonction f est réelle sur $\mathrm{K} = \sigma(a)$, on a $f = \overline{f}$ sur K et on peut écrire $f(a)^* = \overline{f}(a) = f(a)$, qui est donc hermitien. Supposons que $|f| = 1$ sur K et posons $u = f(a)$; on a $u^*u = \overline{f}(a)f(a) = (\overline{f}f)(a) = \varphi_a(\mathbf{1}) = 1_{\mathrm{A}}$, et le même calcul donne $uu^* = 1_{\mathrm{A}}$, donc u est unitaire.

///

Remarque. Les réciproques sont vraies par le théorème spectral : si $f(a)$ est hermitien, son spectre, égal à $f(\sigma(a))$, est réel, c'est-à-dire que f est réelle sur $\sigma(a)$, etc...

Exemples.

1. Supposons que u soit un élément unitaire de A, et supposons que $-1 \notin \sigma(u)$; la fonction $f(z) = \mathrm{i}(z-1)/(z+1)$ est alors définie et continue sur $\sigma(u)$, et à valeurs réelles. Il en résulte que $f(u)$ est hermitien. L'élément $f(u)$ est égal à $\mathrm{i}(u - 1_{\mathrm{A}})(u + 1_{\mathrm{A}})^{-1}$.

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

2. Pour tout $s \in \mathbb{R}$ considérons la fonction f_s définie sur $\mathbb{R}$ par $f_s(t) = \mathrm{e}^{\mathrm{i}\,st}$. Si $\mathrm{T} \in \mathcal{L}(\mathrm{H})$ est hermitien, on peut considérer pour tout s l'opérateur $\mathrm{U}_s = f_s(\mathrm{T}) = \mathrm{e}^{\mathrm{i}\,s\mathrm{T}}$. C'est un opérateur unitaire puisque f_s est à valeurs dans $\mathbb{T}$. De plus $f_{s_1}f_{s_2} = f_{s_1+s_2}$ pour tous s_1, s_2, donc $\mathrm{U}_{s_1}\mathrm{U}_{s_2} = \mathrm{U}_{s_1+s_2}$. On dit qu'on a un *groupe* d'opérateurs unitaires.

On peut montrer (théorème de Stone) une réciproque de ce fait, mais elle demande de considérer des opérateurs autoadjoints *non bornés*.

3. Soient T un opérateur normal sur H et F un sous-espace fermé de H, stable par T et par T^* ; alors la projection orthogonale P_{F} commute avec T et T^*, donc avec tout opérateur $f(\mathrm{T})$. Si S désigne la restriction de T à F, alors S est un opérateur normal sur F, et $\sigma(\mathrm{S}) \subset \sigma(\mathrm{T})$; pour toute fonction continue f sur le spectre de T l'opérateur $f(\mathrm{S})$ est la restriction de $f(\mathrm{T})$ à F.

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

T_1 sur H_1 dont le spectre est égal à K_1 : si $\lambda \notin \mathrm{K}_1$, la fonction g définie par $g(z) = z - \lambda$ en tout point $z \in \mathrm{K}_1$ et $g(z) = 1$ pour tout $z \in \mathrm{K}_2$ est une fonction de $\mathrm{C}(\sigma(\mathrm{T}))$, inversible dans cette algèbre, donc son image est un opérateur inversible dans $\mathcal{L}(\mathrm{H})$. L'image est l'opérateur égal à $\mathrm{T}_1 - \lambda\,\mathrm{Id}_{\mathrm{H}_1}$ sur H_1 et à $\mathrm{Id}_{\mathrm{H}_2}$ sur H_2. Il en résulte que $\mathrm{T}_1 - \lambda\,\mathrm{Id}_{\mathrm{H}_1}$ est inversible pour tout $\lambda \notin \mathrm{K}_1$. On a donc $\sigma(\mathrm{T}_1) \subset \mathrm{K}_1$ et de la même façon $\sigma(\mathrm{T}_2) \subset \mathrm{K}_2$. Inversement si $\lambda \in \mathrm{K}_1$, on sait que $\mathrm{T}_2 - \lambda\,\mathrm{Id}_{\mathrm{H}_2}$ est inversible ; alors $\mathrm{T}_1 - \lambda\,\mathrm{Id}_{\mathrm{H}_1}$ ne peut pas être inversible, sinon $\mathrm{T} - \lambda\,\mathrm{Id}_{\mathrm{H}}$ le serait aussi.

3.3. Application aux hermitiens positifs. La racine carrée

Lemme 3.3.1. *Soient* A *une* C^**-algèbre unitaire et* $a \in \mathrm{A}$ *un élément hermitien ; si le spectre de* a *est contenu dans* $[0, +\infty[$*, l'élément* a *peut s'écrire comme le carré d'un élément hermitien* b *à spectre positif*

$$a = b^2, \quad b^* = b, \quad \sigma(b) \subset [0, +\infty[.$$

Réciproquement, si $a = b^2$ *avec* b *hermitien, alors* a *est hermitien et* $\sigma(a) \subset [0, +\infty[$.

Preuve. — Supposons que a soit hermitien et que son spectre $\mathrm{K} = \sigma(a)$ soit contenu dans $[0, +\infty[$; notons $f \in \mathrm{C}(\mathrm{K})$ l'application $t \to \sqrt{t}$; par le théorème 3.2.3, on sait que $b = f(a) = f(a)^*$, donc b est hermitien ; de plus $f^2 = \zeta_{\mathrm{K}}$, donc $b^2 = f(a)^2 = a$. Inversement, on voit que b^2 est hermitien quand b est hermitien, et b^2 a un spectre positif (par calcul fonctionnel, ou calcul polynomial, qui se réduit ici à observer que

$$b^2 + s^2 1_{\mathrm{A}} = (b - \mathrm{i}s1_{\mathrm{A}})(b + \mathrm{i}s1_{\mathrm{A}})$$

est inversible quand s est réel).

///

Un élément hermitien de A dont le spectre est contenu dans $[0, +\infty[$ est appelé *positif*. Dans le cas particulier de la C^*-algèbre $\mathcal{L}(\mathrm{H})$ des opérateurs sur un espace de Hilbert complexe H, la notion d'élément hermitien positif peut se décrire de plusieurs autres façons, dont une seule fait appel aux vecteurs de H, l'équivalence des trois autres étant valable dans toute C^*-algèbre. Cependant, le passage du point (*ii*) ci-dessous aux autres points n'est pas évident dans le cas abstrait, et il est repoussé en appendice (lemme 7.1) ; c'est en fait le point clé pour pouvoir ramener le cas abstrait au cas de $\mathcal{L}(\mathrm{H})$ traité ici !

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

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

Preuve. — On a vu au lemme 3.3.1 que l'équivalence de (*iii*) et (*iv*) est un fait général. L'implication (*iii*) $\Rightarrow$ (*ii*) est évidente. Supposons (*ii*) vérifiée ; l'opérateur $\mathrm{T} = \mathrm{S}^*\mathrm{S}$ est hermitien et on a

$$\langle \mathrm{S}^*\mathrm{S}x, x\rangle = \langle \mathrm{S}x, \mathrm{S}x\rangle \geq 0$$

pour tout $x \in \mathrm{H}$, donc (*ii*) $\Rightarrow$ (*i*). Enfin, on voit que (*i*) implique que $\sigma(\mathrm{T}) \subset [0, +\infty[$: tout élément λ du spectre de T est dans le bord de ce spectre, donc il existe une suite $(x_n) \subset \mathrm{H}$ de vecteurs de norme 1 telle que $\mathrm{T}x_n - \lambda x_n \to 0$, ce qui donne par produit scalaire

$$\langle \mathrm{T}x_n, x_n\rangle - \lambda\langle x_n, x_n\rangle = \langle \mathrm{T}x_n, x_n\rangle - \lambda \to 0,$$

donc $\lambda = \lim_n \langle \mathrm{T}x_n, x_n\rangle$ est ≥ 0.

///

On note $\mathcal{L}(\mathrm{H})_+$ l'ensemble des éléments positifs de $\mathcal{L}(\mathrm{H})$. Pour $\mathrm{T} \in \mathcal{L}(\mathrm{H})_+$ et $\alpha > 0$, on pose $\mathrm{T}^\alpha = f(\mathrm{T})$, où $f \in \mathrm{C}(\sigma(\mathrm{T}))$ est l'application $t \to t^\alpha$. Pour $\alpha, \beta > 0$ on a $\mathrm{T}^{\alpha+\beta} = \mathrm{T}^\alpha\mathrm{T}^\beta$ et, par la dernière partie du théorème 3.2.3, on a $(\mathrm{T}^\alpha)^\beta = \mathrm{T}^{\alpha\beta}$. Le cas le plus important de cette construction est le cas $\alpha = 1/2$, mais l'unicité démontrée ci-dessous pour la racine carrée se généralise.

Proposition. *Pour $a \in \mathrm{A}$ hermitien et positif, il existe un et un seul $b \in \mathrm{A}$ hermitien et positif tel que $b^2 = a$.*

Preuve. — On a déjà vu au lemme 3.3.1 l'existence d'une racine hermitienne positive, passons maintenant à la démonstration de l'unicité. Soit b un élément hermitien positif tel que $b^2 = a$; considérons le spectre $\mathrm{K} = \sigma(b) \subset [0, +\infty[$, et considérons sur K la fonction $f : s \to s^2$, puis sur $\mathrm{L} = f(\mathrm{K}) \subset [0, +\infty[$ la fonction g définie par $g(t) = \sqrt{t}$. On notera que $\mathrm{L} = \sigma(f(b)) = \sigma(a)$ par le théorème spectral. Du fait que $\mathrm{K} \subset [0, +\infty[$, on vérifie que $g(f(s)) = \sqrt{s^2} = s$ pour tout $s \in \mathrm{K}$, donc le résultat de composition nous donne, puisque $g \circ f = \zeta_{\mathrm{K}}$

$$b = (g \circ f)(b) = g(f(b)) = g(b^2) = g(a) = \sqrt{a}.$$

///

Bien entendu il n'y a pas unicité si on ne demande pas que la racine soit positive : il suffit de considérer $-\sqrt{a}$ pour avoir une autre racine hermitienne.

Exemple. Reprenons l'opérateur P de primitive sur $[0, 1]$ (exercice 1.2.5). L'opérateur $\mathrm{P}^*\mathrm{P}$ est hermitien positif, et il associe à toute fonction f une fonction $g = \mathrm{P}^*\mathrm{P}f$ telle que $g'' = -g$; l'action de $\mathrm{P}^*\mathrm{P}$ est donc une double intégration (au signe près). Il existe un opérateur T sur $\mathrm{L}_2(0, 1)$ qui est la racine carrée de $\mathrm{P}^*\mathrm{P}$; on peut imaginer que son action ressemble à *une* intégration, mais ce n'est en tout cas pas l'action de primitive ordinaire. De plus cet opérateur n'est pas évident à expliciter, par exemple sous la forme d'un opérateur à noyau ; on sait pourtant que $\sqrt{\mathrm{P}^*\mathrm{P}}$ est de Hilbert-Schmidt, car ses valeurs propres, racines de celles de $\mathrm{P}^*\mathrm{P}$, sont de carré sommable (mais pas sommables, elles décroissent en $1/n$).

Certains exemples se décrivent avec la transformation de Fourier ; sur $\mathbb{R}$, l'opérateur qui associe à une fonction φ de classe C^∞ et à support compact la fonction $\mathrm{T}\varphi = -\varphi'' + \varphi$ se décrit du côté Fourier par

$$\widehat{(\mathrm{T}\varphi)}(t) = (1 + |t|^2)\, \widehat{\varphi}(t).$$

Cet opérateur n'est pas défini sur $L_2(\mathbb{R})$, mais il est facile de concevoir un inverse borné S pour T, agissant par

$$\widehat{(\mathrm{S}f)}(t) = (1+|t|^2)^{-1}\widehat{f}(t)$$

de $L_2(\mathbb{R})$ dans $L_2(\mathbb{R})$. Il est facile de voir que S est hermitien positif, et sa racine agit sur $L_2(\mathbb{R})$ par

$$(\widehat{\sqrt{\mathrm{S}}\,f})(t) = (1+|t|^2)^{-1/2}\widehat{f}(t).$$

Décomposition polaire

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

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

Preuve.— Pour tout vecteur $x \in \mathrm{H}$ on a

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

en particulier, $\ker \mathrm{T} = \ker|\mathrm{T}|$; l'adhérence G de l'image de $|\mathrm{T}|$ est l'orthogonal de $\ker \mathrm{T}$, et on a $\mathrm{H} = \mathrm{G} \oplus \ker \mathrm{T}$, somme directe orthogonale. On va expliquer la construction de u_0, restriction de u au morceau G. Tout d'abord, si $y = |\mathrm{T}|\,x \in \mathrm{im}(|\mathrm{T}|)$, nous devons nécessairement poser $u_0 y = \mathrm{T}x$ pour réaliser la factorisation voulue. Notons que si $y = |\mathrm{T}|\,x'$ est une autre représentation de y, on aura $x' - x \in \ker|\mathrm{T}| = \ker \mathrm{T}$, donc $\mathrm{T}x' = \mathrm{T}x$. Cela montre que l'on peut légitimement poser $u_0 y = \mathrm{T}x$, pour tout y dans l'image de $|\mathrm{T}|$, où x est n'importe quel vecteur tel que $|\mathrm{T}|\,x = y$.

On remarque ensuite que $\|u_0 y\| = \|\mathrm{T}x\| = \|\,|\mathrm{T}|\,x\,\| = \|y\|$, c'est-à-dire que u_0 est une isométrie de $\mathrm{im}(|\mathrm{T}|)$ dans H. Puisque H est complet, cette isométrie se prolonge en isométrie $\overline{u}_0$ de l'adhérence G, à valeurs dans H. Posons pour finir, si $x = y + z$, avec $y \in \mathrm{G}$ et $z \in \ker \mathrm{T}$

$$ux = \overline{u}_0 y$$

c'est-à-dire $u = \overline{u}_0 \circ \mathrm{P}_\mathrm{G}$. On vérifie pour finir que $u \circ |\mathrm{T}| = \mathrm{T}$, que u est nulle sur $\ker \mathrm{T}$ et que $\|u\| \leq 1$.

///

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

Remarque. *Si* T *est injectif,* u *est isométrique. Si* T *est injectif à image dense,* u *est unitaire.*

En effet, G est égal à H lorsque $\ker|\mathrm{T}| = \ker \mathrm{T} = \{0\}$, donc $u = \overline{u}_0$ dans ce cas. De plus, l'image de u contient l'image de T d'après la factorisation ; si T est injectif à image dense, u est une isométrie à image dense, donc surjective, donc unitaire.

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

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

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

Preuve. — Puisque u est nul sur $\ker \mathrm{T}$, on en déduit que $\mathrm{im}(u^*)$ est orthogonale à $\ker \mathrm{T}$, c'est-à-dire $\mathrm{im}(u^*) \subset \mathrm{G}$, où G est l'adhérence de $\mathrm{im}(|\mathrm{T}|)$ (qui est égale à $(\ker \mathrm{T})^\perp$). On a vu que u est isométrique sur G ; il en résulte que $\langle u^*uy, y\rangle = \|uy\|^2 = \langle y, y\rangle$ pour tout $y \in \mathrm{G}$; par polarisation, on obtient $\langle u^*uy_1, y_2\rangle = \langle y_1, y_2\rangle$ pour tous $y_1, y_2 \in \mathrm{G}$, et comme $u^*uy_1 \in \mathrm{G}$ on en déduit que $u^*u_{|\mathrm{G}} = \mathrm{Id}_\mathrm{G}$; puisque u^*u est nul sur l'orthogonal de G, on en déduit que $u^*u = \pi_\mathrm{G}$.

On a en particulier $u^*u\,|\mathrm{T}| = |\mathrm{T}|$, donc $u^*\mathrm{T} = |\mathrm{T}|$ et

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

L'opérateur $u\,|\mathrm{T}|\,u^*$ est hermitien positif et son carré est $\mathrm{T}\mathrm{T}^*$: il est donc égal à $|\mathrm{T}^*|$. De plus, la relation $u^*|\mathrm{T}^*| = u^*u\,|\mathrm{T}|\,u^* = |\mathrm{T}|\,u^* = \mathrm{T}^*$ montre que u^* est la phase de T^*. En appliquant à T^* ce qu'on a vu pour T, on voit que uu^* est la projection orthogonale sur l'orthogonal du noyau de T^*, c'est-à-dire la projection orthogonale sur l'adhérence de l'image de T.

///

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

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

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

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

3.4. Hermitiens en scalaires réels

La plupart des résultats de calcul fonctionnel restent vrais pour un opérateur hermitien T défini sur un espace de Hilbert *réel* H. Une première approche est de reprendre, dans le cas réel, l'extension de l'homomorphisme défini sur les fonctions polynomiales. Bien que la théorie du spectre demande normalement des scalaires complexes, il est possible de travailler avec l'ensemble $\mathrm{K} = \sigma_\mathbb{R}(\mathrm{T}) \subset \mathbb{R}$ formé des réels λ tels que $\mathrm{T} - \lambda\,\mathrm{Id}_\mathrm{H}$ ne soit pas inversible. Comme K est contenu dans $\mathbb{R}$, le théorème de Weierstrass nous indique que les fonctions polynomiales à coefficients réels sont denses dans l'espace $\mathrm{C}_\mathbb{R}(\mathrm{K})$ des fonctions réelles continues sur K. On obtient le résultat qui suit.

Théorème. *Soient* H *un espace de Hilbert réel,* $\mathrm{T} \in \mathcal{L}(\mathrm{H})$ *un opérateur hermitien, et posons* $\mathrm{K} = \sigma_\mathbb{R}(\mathrm{T})$ *; il existe un et un seul homomorphisme d'algèbres unitaires réelles* $\varphi_\mathrm{T} : \mathrm{C}_\mathbb{R}(\mathrm{K}) \to \mathcal{L}(\mathrm{H})$ *tel que* $\varphi_\mathrm{T}(\zeta_\mathrm{K}) = \mathrm{T}$.

L'homomorphisme φ_T *est isométrique. Si on note* $f(\mathrm{T}) = \varphi_\mathrm{T}(f)$ *pour toute fonction* f *réelle continue sur* K*, on a que* $f(\mathrm{T})$ *est hermitien ; en fait,* $f(\mathrm{T})$ *commute avec tout opérateur* $\mathrm{S} \in \mathcal{L}(\mathrm{H})$ *qui commute avec* T*. On a de plus*

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

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

Une deuxième approche (qu'on doit aussi utiliser en partie avec la première méthode) est de *complexifier* la situation : on introduit un espace de Hilbert complexe $\mathrm{H}_{\mathbb{C}}$ et une extension hermitienne $\mathrm{T}_{\mathbb{C}}$ de T à $\mathrm{H}_{\mathbb{C}}$. On voit que pour tout réel λ, l'opérateur $\mathrm{T}_{\mathbb{C}} - \lambda\,\mathrm{Id}_{\mathrm{H}_{\mathbb{C}}}$ est inversible dans $\mathcal{L}(\mathrm{H}_{\mathbb{C}})$ si et seulement si $\mathrm{T} - \lambda\,\mathrm{Id}_{\mathrm{H}}$ est inversible, ce qui ramène le bizarre spectre $\sigma_{\mathbb{R}}(\mathrm{T})$ au spectre complexe usuel $\sigma(\mathrm{T}_{\mathbb{C}})$.

On peut décrire $\mathrm{H}_{\mathbb{C}}$ comme l'ensemble des sommes formelles $x + \mathrm{i}y$, $x, y \in \mathrm{H}$ et on peut identifier H au $\mathbb{R}$-sous-espace vectoriel de $\mathrm{H}_{\mathbb{C}}$ formé des $x = x + \mathrm{i}0_{\mathrm{H}}$. L'extension complexe $\mathrm{T}_{\mathbb{C}}$ est définie naturellement en posant pour tout $x + \mathrm{i}y \in \mathrm{H}_{\mathbb{C}}$

$$\mathrm{T}_{\mathbb{C}}(x + \mathrm{i}y) = \mathrm{T}x + \mathrm{i}\,\mathrm{T}y.$$

On montre alors que pour toute fonction continue *réelle* f sur le spectre, le $\mathbb{R}$-sous-espace H est invariant par $f(\mathrm{T}_{\mathbb{C}})$, ce qui permet de considérer que $f(\mathrm{T})$ est la restriction de $f(\mathrm{T}_{\mathbb{C}})$ au $\mathbb{R}$-sous-espace H.

Notes du chapitre 3

([a]) Dans le second point de vue, on considère des fonctions définies sur le disque unité, alors que dans l'exemple 2.2.2 on a des fonctions sur le cercle unité : l'isomorphisme de la seconde algèbre vers la première est simplement l'application r de restriction au cercle unité. Elle est isométrique par le principe du maximum : si $f \in \mathrm{A}$, on a puisque f est holomorphe dans le disque ouvert et continue sur le disque fermé

$$\|f\|_{\mathrm{A}} = \max\{|f(z)| : z \in \overline{\mathrm{D}}\} = \max\{|f(z)| : |z| = 1\} = \|rf\|_{\mathrm{A}(\mathrm{D})}.$$

([b]) Supposons $xy = yx$, et xy inversible : il existe u tel que $u(xy) = (xy)u = 1_{\mathrm{A}}$; alors $(uy)x = u(xy) = 1_{\mathrm{A}}$ et

$$yu = (uy)x(yu) = uy$$

est l'inverse de x.

4. Opérateurs de Fredholm

4.1. Sous-espaces de dimension et de codimension finie

Lemme 4.1.1. Projection sur un sous-espace de dimension finie. *Si* E *est un sous-espace vectoriel de dimension finie d'un espace normé* X, *il existe un projecteur* P *continu sur* X *tel que* $\mathrm{P}(\mathrm{X}) = \mathrm{E}$.

Preuve. — On choisit une base $e_1, \ldots, e_n$ de E, et on désigne par e_j^*, $j = 1, \ldots, n$ les formes linéaires coordonnées sur E ; par le théorème de Hahn-Banach on peut prolonger ces formes linéaires sur E en formes linéaires continues $x_1^*, \ldots, x_n^*$ sur X. On pose

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

Il est clair que P est continue, et il est facile de vérifier que P est une projection de X sur E.

///

Corollaire 4.1.2. *Si* E *est un sous-espace vectoriel de dimension finie d'un espace normé* X, *il existe un sous-espace vectoriel fermé* $\mathrm{Y} \subset \mathrm{X}$ *tel que* $\mathrm{X} = \mathrm{E} \oplus \mathrm{Y}$.

Preuve. — Il suffit de prendre pour Y le noyau de la projection P de X sur E donnée par le lemme précédent.

///

On pourrait envisager la définition qui suit dans un cadre purement algébrique, mais nous ne l'utiliserons que pour un sous-espace vectoriel *fermé*. Nous avons donc ajouté ce mot « fermé » dans la définition.

Définition. Un sous-espace vectoriel fermé $\mathrm{Y} \subset \mathrm{X}$ est *de codimension finie dans* X si le quotient X/Y est de dimension finie. La *codimension* de Y dans X est la dimension de l'espace vectoriel quotient X/Y. On la notera $\operatorname{codim}_\mathrm{X} \mathrm{Y}$ ou simplement $\operatorname{codim} \mathrm{Y}$ si aucune confusion n'est à craindre.

Si X/Y est de dimension finie n et si on relève une base $(\hat{f}_1, \ldots, \hat{f}_n)$ du quotient en des vecteurs $(f_1, \ldots, f_n)$ de X, on obtient un système libre qui engendre un sous-espace vectoriel F de X, de dimension finie égale à $n = \operatorname{codim} \mathrm{Y}$, tel que $\mathrm{X} = \mathrm{Y} \oplus \mathrm{F}$. Il est clair qu'affirmer que $\mathrm{X} = \mathrm{Y} \oplus \mathrm{F}$, avec F de dimension finie d, équivaut à dire que Y est de codimension finie égale à d. Notons π_Y la projection canonique de X sur X/Y, qui associe à chaque vecteur $x \in \mathrm{X}$ sa classe $\pi_\mathrm{Y}(x)$ modulo Y, c'est-à-dire

$$\pi_\mathrm{Y}(x) = \{x + y : y \in \mathrm{Y}\}.$$

On rappelle que la norme de l'espace quotient X/Y est définie par

$$\forall x \in \mathrm{X}, \quad \|\pi_\mathrm{Y}(x)\| = \inf\{\|x + y\| : y \in \mathrm{Y}\},$$

ce qui signifie que la norme de la classe est l'inf des normes des éléments de la classe. Si on écrit la norme quotient sous la forme équivalente

$$\|\pi_\mathrm{Y}(x)\| = \inf\{\|x - y\| : y \in \mathrm{Y}\},$$

on voit que $\|\pi_{\mathrm{Y}}(x)\|$ est aussi la distance de x au sous-espace Y : si $x \notin \mathrm{Y}$, cette distance est > 0 puisque Y est fermé, c'est-à-dire que $\|\pi_{\mathrm{Y}}(x)\| > 0$ quand la classe n'est pas la classe nulle ; il s'agit donc bien d'une *norme* sur le quotient. On va faire quelques remarques simples mais utiles.

1. *Si* $\mathrm{X} = \mathrm{Y} + \mathrm{F}$ *avec* F *de dimension finie, alors* Y *est de codimension finie et* $\operatorname{codim} \mathrm{Y} \le \dim \mathrm{F}$. *La codimension de* Y *est* $\le k$ *si et seulement s'il existe* F *de dimension* $\le k$ *tel que* $\mathrm{X} = \mathrm{Y} + \mathrm{F}$.

Si $\mathrm{X} = \mathrm{Y} + \mathrm{F}$, alors $\mathrm{X}/\mathrm{Y} = \pi_{\mathrm{Y}}(\mathrm{F})$, qui est donc un espace de dimension finie $\le \dim \mathrm{F}$.

2. *Si* $\mathrm{Y} \cap \mathrm{F} = \{0\}$ *avec* F *de dimension finie, alors* $\operatorname{codim} \mathrm{Y} \ge \dim \mathrm{F}$. *La codimension de* Y *est* $\ge k$ *si et seulement s'il existe* F *de dimension* k *tel que* $\mathrm{Y} \cap \mathrm{F} = \{0\}$.

Si $\mathrm{Y} \cap \mathrm{F} = \{0\}$, la projection π_{Y} est injective sur F, donc l'image $\pi_{\mathrm{Y}}(\mathrm{F})$ a la même dimension que F et par conséquent $\dim \mathrm{X}/\mathrm{Y} \ge \dim \mathrm{F}$. Si $\dim \mathrm{X}/\mathrm{Y} \ge k$, on peut trouver k vecteurs indépendants $\hat{f}_1, \ldots, \hat{f}_k$ dans X/Y ; si on relève ces vecteurs en $f_1, \ldots, f_k \in \mathrm{X}$, l'espace $\mathrm{F} = \operatorname{Vect}(f_1, \ldots, f_k)$ est de dimension k et $\mathrm{Y} \cap \mathrm{F} = \{0\}$.

3. Codim dans codim : *si* Z *est de codimension* k *dans* Y *et* Y *de codimension* ℓ *dans* X, *alors* Z *est de codimension* $k + \ell$ *dans* X.

On peut écrire $\mathrm{X} = \mathrm{Y} \oplus \mathrm{E}$ et $\mathrm{Y} = \mathrm{Z} \oplus \mathrm{F}$, avec E et F de dimensions finies égales aux codimensions respectives ; alors la relation $\mathrm{X} = \mathrm{Z} \oplus (\mathrm{E} \oplus \mathrm{F})$ montre que

$$\operatorname{codim}_{\mathrm{X}} \mathrm{Z} = \operatorname{codim}_{\mathrm{Y}} \mathrm{Z} + \operatorname{codim}_{\mathrm{X}} \mathrm{Y}.$$

4. *L'intersection d'un sous-espace fermé* Z *avec un sous-espace de codimension finie* Y *est de codimension finie dans* Z.

L'image $\pi_{\mathrm{Y}}(\mathrm{Z})$ est de dimension finie, donc admet une base $(\hat{f}_1, \ldots, \hat{f}_d)$, qu'on peut relever en $(f_1, \ldots, f_d)$ dans Z ; on pose $\mathrm{F} = \operatorname{Vect}(f_1, \ldots f_d)$, et on constate alors que $\mathrm{Z} = (\mathrm{Z} \cap \mathrm{Y}) + \mathrm{F}$. En effet, si $z \in \mathrm{Z}$ on écrit $\pi_{\mathrm{Y}}(z) = \sum_{j=1}^{d} c_j \hat{f}_j$, et on voit que $z - \sum_{j=1}^{d} c_j f_j$ est dans $\mathrm{Z} \cap \mathrm{Y}$.

5. *L'intersection* $\mathrm{Z} \cap \mathrm{Y}$ *d'un sous-espace fermé de dimension infinie* $\mathrm{Z} \subset \mathrm{X}$ *avec un sous-espace* Y *de codimension finie dans* X *est de dimension infinie.*

On a vu que l'espace $\mathrm{Z} \cap \mathrm{Y}$ est de codimension finie dans Z, donc $\mathrm{Z} = (\mathrm{Z} \cap \mathrm{Y}) \oplus \mathrm{F}$, avec $\dim \mathrm{F} < +\infty$; il n'est pas possible que $\mathrm{Z} \cap \mathrm{Y}$ soit de dimension finie, puisque $\dim \mathrm{Z} = +\infty$.

Projection de $\mathrm{Z} + \mathrm{F}$ *sur* F

On suppose que Z est un sous-espace fermé d'un espace normé X, et F un sous-espace de dimension finie tel que $\mathrm{Z} \cap \mathrm{F} = \{0\}$. On a

$$\delta = \min\{\operatorname{dist}(f, \mathrm{Z}) : f \in \mathrm{S}_{\mathrm{F}}\} > 0$$

car la sphère unité S_{F} de F est compacte et disjointe du fermé Z. On en déduit par homogénéité

$$\|z + f\| \ge \delta \, \|f\|$$

pour tous $f \in \mathrm{F}$, $z \in \mathrm{Z}$. La projection naturelle P de Z+F sur F, définie par $\mathrm{P}(z+f) = f$, est donc de norme $\le \delta^{-1}$, et la projection $\mathrm{Q} = \mathrm{Id} - \mathrm{P}$ de $\mathrm{Z} + \mathrm{F}$ sur Z est alors de norme $\le 1 + \delta^{-1}$, c'est-à-dire que

$$\|z + f\| \ge \delta(1+\delta)^{-1}\|z\|$$

pour tous $f \in \mathrm{F}$, $z \in \mathrm{Z}$. Si π_F désigne la projection de X sur X/F, on a donc pour tout $z \in \mathrm{Z}$

$$\text{(Q)} \qquad \|\pi_\mathrm{F}(z)\|_{\mathrm{X}/\mathrm{F}} \ge \delta(1+\delta)^{-1}\|z\|.$$

Lemme 4.1.3. *Si* Z *est un sous-espace fermé d'un espace normé* X *et* F *un sous-espace de dimension finie, la somme* $\mathrm{Z} + \mathrm{F}$ *est fermée dans* X.

Preuve. — Supposons d'abord que $\mathrm{Z} \cap \mathrm{F} = \{0\}$; supposons qu'une suite de vecteurs $(z_n + f_n) \subset \mathrm{Z} + \mathrm{F}$ tende vers un vecteur $x \in \mathrm{X}$. Soit P la projection (continue) de $\mathrm{Z} + \mathrm{F}$ sur F discutée ci-dessus ; la suite $(z_n + f_n)$ est bornée puisque convergente vers x, donc son image (f_n) par P est bornée dans F. Par Bolzano, on trouve une sous-suite (f_{n_k}) qui tend vers un $f \in \mathrm{F}$, et par différence z_{n_k} converge aussi, vers un $z \in \mathrm{Z}$ puisque Z est fermé. Finalement, $x = z + f \in \mathrm{Z} + \mathrm{F}$.

Dans le cas général, on écrit $\mathrm{F} = (\mathrm{F} \cap \mathrm{Z}) \oplus \mathrm{F}_1$ (donc $\mathrm{F}_1 \subset \mathrm{F}$ est de dimension finie) et on constate que $\mathrm{Z} + \mathrm{F} = \mathrm{Z} + \mathrm{F}_1$ avec $\mathrm{Z} \cap \mathrm{F}_1 = \{0\}$, qui est fermé d'après la première partie de la preuve.

///

4.2. Plongements d'un espace de Banach dans un autre

À partir de maintenant il est important de supposer que les espaces considérés sont **complets**. On suppose que X et Y sont deux espaces de Banach, tous les deux réels ou tous les deux complexes.

Définition. On dira que $\mathrm{T} \in \mathcal{L}(\mathrm{X}, \mathrm{Y})$ est un *plongement* de X dans Y s'il existe une constante $c > 0$ telle que

$$\text{(P)} \qquad \forall x \in \mathrm{X}, \quad \|\mathrm{T}x\| \ge c\,\|x\|.$$

Il est clair que les opérateurs proches de T vérifient encore la propriété, mais pour un $c' < c$ voisin de c : par l'inégalité triangulaire, on aura pour tout $x \in \mathrm{X}$

$$\|(\mathrm{T}+\mathrm{S})x\| \ge \|\mathrm{T}x\| - \|\mathrm{S}x\| \ge (c - \|\mathrm{S}\|)\,\|x\|$$

ce qui montre que $\mathrm{T}' = \mathrm{T} + \mathrm{S}$ est encore un plongement si $\|\mathrm{S}\| < c$.

L'image par un plongement T de tout sous-espace fermé de X est fermée dans Y ; cette propriété sera montrée plus loin dans le cas plus général des *presque-plongements*, mais montrons tout de suite que l'image T(X) est fermée : si une suite $(\mathrm{T}x_n)$ converge vers un point $y \in \mathrm{Y}$, elle est de Cauchy ; comme $\|x_n - x_m\| \le c^{-1}\|\mathrm{T}x_n - \mathrm{T}x_m\|$, on voit que la suite (x_n) est de Cauchy dans X, donc converge vers un $x \in \mathrm{X}$ et $y = \mathrm{T}x$.

On peut remarquer que T est un plongement si et seulement s'il n'existe pas de suite $(x_n) \subset \mathrm{X}$ de vecteurs de norme 1 telle que $\mathrm{T}x_n \to 0$: si T n'est pas un plongement, il n'existe aucune constante $c > 0$ vérifiant la définition de plongement, par exemple $c = 2^{-n}$ ne convient pas, c'est-à-dire que pour tout n il existe un vecteur $x_n \in \mathrm{X}$ tel que $\|\mathrm{T}x_n\| < 2^{-n}\|x_n\|$; en multipliant x_n par un scalaire > 0, on peut se ramener à une suite $(x_n) \subset \mathrm{X}$ de vecteurs de norme 1, telle que $\mathrm{T}x_n \to 0$. L'implication inverse est évidente.

En particulier, un endomorphisme $\mathrm{T} \in \mathcal{L}(\mathrm{X})$ est un plongement si et seulement si 0 n'est pas valeur propre approchée de T.

Par ailleurs, il est clair que T est un plongement de X dans Y si et seulement si T induit un isomorphisme T_1 de X sur l'image T(X) : l'opérateur T_1 est l'élément de $\mathcal{L}(\mathrm{X}, \mathrm{T}(\mathrm{X}))$ défini par $\mathrm{T}_1 x = \mathrm{T}x$ pour tout $x \in \mathrm{X}$. Si T est un plongement, T est injectif et T_1 est bijectif, donc on peut définir l'inverse algébrique U de T_1, opérant de T(X) vers X ; l'inégalité (P) de la définition des plongements montre que $\|\mathrm{U}\| \le 1/c$, donc $\mathrm{U} \in \mathcal{L}(\mathrm{T}(\mathrm{X}), \mathrm{X})$ et T_1 est bien un isomorphisme d'espaces de Banach entre X et T(X).

Stabilité de la codimension de l'image

Lemme 4.2.1. *Si* T *est un plongement de* X *dans* Y *et si la codimension de* T(X) *dans* Y *(finie ou infinie) est* $\ge k$*, il en est de même pour les opérateurs* T' *voisins de* T*.*

Si T *est un plongement et si la codimension de* T(X) *dans* Y *est finie égale à* k*, il en est de même pour les opérateurs* T' *voisins de* T*.*

Si T *est un plongement et si la codimension de* T(X) *dans* Y *est infinie, il en est de même pour les opérateurs* T' *voisins de* T*.*

Preuve. — Si la codimension de T(X) dans Y est $\ge k$, on peut trouver un sous-espace $\mathrm{F} \subset \mathrm{Y}$ de dimension k tel que $\mathrm{T}(\mathrm{X}) \cap \mathrm{F} = \{0\}$; alors la projection π_F de Y sur Y/F induit un plongement de $\mathrm{Z} = \mathrm{T}(\mathrm{X})$ dans Y/F d'après l'inégalité (Q), et $\pi_\mathrm{F} \circ \mathrm{T}$ est aussi un plongement par composition évidente,

$$\forall x \in \mathrm{X}, \quad \|\pi_\mathrm{F}(\mathrm{T}x)\|_{\mathrm{Y}/\mathrm{F}} \ge \delta(1+\delta)^{-1}\|\mathrm{T}x\| \ge \delta(1+\delta)^{-1}\, c\, \|x\| \,;$$

il en est de même pour les voisins T' de T (car alors $\pi_\mathrm{F} \circ \mathrm{T}'$ est voisin de $\pi_\mathrm{F} \circ \mathrm{T}$) ; en particulier $\pi_\mathrm{F} \circ \mathrm{T}'$ est injectif quand T' est voisin de T, donc $\mathrm{F} \cap \mathrm{T}'(\mathrm{X}) = \{0\}$ et par conséquent $\mathrm{codim}_\mathrm{Y}\, \mathrm{T}'(\mathrm{X}) \ge k$.

Si la codimension de T(X) dans Y est égale à k, on a $\mathrm{Y} = \mathrm{T}(\mathrm{X}) \oplus \mathrm{F}$ pour un certain F de dimension k ; l'application $\pi_\mathrm{F} \circ \mathrm{T}$ reste un plongement comme avant, mais de plus $\pi_\mathrm{F}(\mathrm{T}(\mathrm{X})) = \mathrm{Y}/\mathrm{F}$, donc $\pi_\mathrm{F} \circ \mathrm{T}$ est surjective de X sur Y/F, et $\pi_\mathrm{F} \circ \mathrm{T}$ est un isomorphisme de X sur Y/F ; ceci passe aux voisins T' (stabilité des isomorphismes : résulte de la proposition 2.1.2) et donne le résultat voulu.

Soit T un plongement tel que $\mathrm{codim}_\mathrm{Y}\, \mathrm{T}(\mathrm{X}) = +\infty$, et soit B une boule ouverte dans $\mathcal{L}(\mathrm{X}, \mathrm{Y})$, centrée en T et de rayon assez petit pour que tous ses éléments soient des plongements ; pour chaque entier k, l'ensemble W_k des opérateurs $\mathrm{S} \in \mathrm{B}$ vérifiant l'inégalité $\mathrm{codim}_\mathrm{Y}\, \mathrm{S}(\mathrm{X}) \ge k$ est ouvert dans $\mathcal{L}(\mathrm{X}, \mathrm{Y})$ d'après ce qui précède, ainsi que l'ensemble V_k des S tel que $\mathrm{codim}_\mathrm{Y}\, \mathrm{S}(\mathrm{X}) = k$. Il en résulte que V_k est aussi fermé dans B (son complémentaire est formé de l'ouvert W_{k+1} et des ouverts V_j, $j < k$). Par la connexité de la boule B, chacun de ces ensembles V_k est vide ou égal à B. Si $\mathrm{codim}_\mathrm{Y}\, \mathrm{T}(\mathrm{X})$ est infinie, aucun de ces ensembles n'est égal à B : ils sont donc tous vides, et la boule B est entièrement formée d'opérateurs S avec $\mathrm{codim}_\mathrm{Y}\, \mathrm{S}(\mathrm{X}) = +\infty$.

///

Faute de mieux, on adoptera ici la terminologie suivante (qui m'est tout à fait personnelle).

Définition 4.2.2. On dira que $\mathrm{T} \in \mathcal{L}(\mathrm{X}, \mathrm{Y})$ est un *presque-plongement* de X dans Y s'il existe un sous-espace (fermé) de codimension finie $\mathrm{X}_1 \subset \mathrm{X}$ et une constante $c > 0$ tels que

$$\text{(PP)} \qquad \forall x \in \mathrm{X}_1, \quad \|\mathrm{T}x\| \geq c\,\|x\|.$$

Il est clair que les opérateurs proches de T vérifient encore la propriété, avec le même sous-espace X_1 mais pour un $c' < c$ voisin de c. Si S est de rang fini, son noyau $\ker \mathrm{S}$ est de codimension finie, donc $\mathrm{X}_2 = \mathrm{X}_1 \cap \ker \mathrm{S}$ est encore de codimension finie, et sur X_2 l'opérateur $\mathrm{T} + \mathrm{S}$ vérifie l'équation (PP), donc $\mathrm{T} + \mathrm{S}$ est un presque-plongement.

Si T est un presque-plongement, on voit que le noyau de T est de dimension finie, puisqu'il ne peut pas rencontrer le sous-espace X_1 de codimension finie, sur lequel T est injectif, ailleurs qu'en 0 : on a donc $\dim \ker \mathrm{T} \leq \operatorname{codim}_{\mathrm{X}} \mathrm{X}_1 < +\infty$; de plus, l'image par T de tout sous-espace fermé de X est fermée dans Y : si Z est un sous-espace fermé de l'espace X_1 de la définition, il est clair que $\mathrm{T}(\mathrm{Z})$ est complet, donc fermé ; en effet, toute suite de Cauchy (y_n) dans $\mathrm{T}(\mathrm{Z})$ peut s'écrire $y_n = \mathrm{T}x_n$ avec $(x_n) \subset \mathrm{Z} \subset \mathrm{X}_1$, donc d'après (PP)

$$\|x_m - x_n\| \leq c^{-1} \|\mathrm{T}x_m - \mathrm{T}x_n\|$$

tend vers 0 avec m, n ; la suite de Cauchy (x_n) converge vers un $x \in \mathrm{Z}$ puisque Z est complet, et $y_n = \mathrm{T}x_n$ converge vers $\mathrm{T}x \in \mathrm{T}(\mathrm{Z})$. Si Z est un sous-espace fermé quelconque, on écrit $\mathrm{Z} = (\mathrm{Z} \cap \mathrm{X}_1) \oplus \mathrm{F}$ avec F de dimension finie, puis $\mathrm{T}(\mathrm{Z}) = \mathrm{T}(\mathrm{Z} \cap \mathrm{X}_1) + \mathrm{T}(\mathrm{F})$ est fermé comme somme d'un sous-espace fermé et d'un sous-espace de dimension finie, par le lemme 4.1.3.

On a déjà montré la moitié du résultat qui suit.

Proposition 4.2.3. *Un opérateur* $\mathrm{T} \in \mathcal{L}(\mathrm{X}, \mathrm{Y})$ *est un presque-plongement si et seulement si son noyau est de dimension finie et son image fermée. Dans ce cas, la restriction* T_1 *de* T *à tout supplémentaire* X_1 *du noyau* $\ker \mathrm{T}$ *est un plongement de* X_1 *dans* Y.

Preuve. — Supposons $\ker \mathrm{T}$ de dimension finie et $\mathrm{T}(\mathrm{X})$ fermé. D'après le corollaire 4.1.2, on peut trouver X_1 fermé tel que $\mathrm{X} = \mathrm{X}_1 \oplus \ker \mathrm{T}$; alors $\mathrm{T}(\mathrm{X}_1) = \mathrm{T}(\mathrm{X})$ est un espace de Banach, et $\mathrm{T}_1 = \mathrm{T}_{|\mathrm{X}_1} \in \mathcal{L}(\mathrm{X}_1, \mathrm{T}(\mathrm{X}))$ est une bijection continue. D'après le théorème des isomorphismes de Banach (corollaire 9.3), la bijection inverse est continue, donc T_1 est un isomorphisme de X_1 sur $\mathrm{T}(\mathrm{X}_1)$. Si $\mathrm{U} \in \mathcal{L}(\mathrm{T}(\mathrm{X}), \mathrm{X}_1)$ est l'inverse de T_1, on voit que

$$\forall x \in \mathrm{X}_1, \quad \|x\| = \|\mathrm{UT}x\| \leq \|\mathrm{U}\|\,\|\mathrm{T}x\|$$

ce qui donne la propriété de plongement pour T_1, avec $c = \|\mathrm{U}\|^{-1}$.

///

4.3. Indice et opérateurs de Fredholm

Soient X, Y deux espaces de Banach et $\mathrm{T} \in \mathcal{L}(\mathrm{X}, \mathrm{Y})$ un presque-plongement ; il existe un sous-espace fermé $\mathrm{X}_1 \subset \mathrm{X}$, de codimension finie, tel que la restriction de T à X_1 soit un isomorphisme de X_1 sur le sous-espace fermé $\mathrm{Y}_1 = \mathrm{T}(\mathrm{X}_1)$. La quantité

$$\operatorname{codim}_{\mathrm{X}} \mathrm{X}_1 - \operatorname{codim}_{\mathrm{Y}} \mathrm{T}(\mathrm{X}_1),$$

qui est finie ou égale à $-\infty$, ne dépend pas du choix de X_1 ; on l'appelle *l'indice* de T, qui est noté $\operatorname{ind}(\mathrm{T})$. L'indice est $-\infty$ lorsque le sous-espace fermé $\mathrm{T}(\mathrm{X}_1)$ est de codimension *infinie* dans Y.

L'indépendance de l'indice par rapport au choix de X_1 est facile à prouver si X_2 est un autre choix avec $\mathrm{X}_2 \subset \mathrm{X}_1$; dans ce cas on peut écrire $\mathrm{X}_1 = \mathrm{X}_2 \oplus \mathrm{E}$ avec $\dim \mathrm{E} < +\infty$; comme T est injectif sur X_1 et que E est contenu dans X_1, on a $\mathrm{T}(\mathrm{X}_1) = \mathrm{T}(\mathrm{X}_2) \oplus \mathrm{T}(\mathrm{E})$ et de plus $\dim \mathrm{T}(\mathrm{E}) = \dim(\mathrm{E})$; alors

$$\operatorname{codim}_{\mathrm{X}} \mathrm{X}_2 - \operatorname{codim}_{\mathrm{Y}} \mathrm{T}(\mathrm{X}_2) = (\operatorname{codim}_{\mathrm{X}} \mathrm{X}_1 + \dim \mathrm{E}) - (\operatorname{codim}_{\mathrm{Y}} \mathrm{T}(\mathrm{X}_1) + \dim \mathrm{T}(\mathrm{E})).$$

Dans le cas général, pour comparer les valeurs obtenues pour deux choix X_1 et X_2, on passera par l'intermédiaire de l'espace $\mathrm{X}_3 = \mathrm{X}_1 \cap \mathrm{X}_2$, qui est de codimension finie.

Lemme. *L'indice est localement constant : si* $\mathrm{T} \in \mathcal{L}(\mathrm{X}, \mathrm{Y})$ *est un presque-plongement, les voisins de* T *ont le même indice que* T.

Preuve. — Il existe $c > 0$ et X_1 de codimension finie tels que $\|\mathrm{T}x\| \ge c\,\|x\|$ quand $x \in \mathrm{X}_1$. On aura, pour tout S tel que $\|\mathrm{S} - \mathrm{T}\| < c/2$, l'inégalité $\|\mathrm{S}x\| \ge (c/2)\,\|x\|$ quand $x \in \mathrm{X}_1$. On peut donc utiliser X_1 pour calculer l'indice de S ; mais on a vu que $\operatorname{codim}_{\mathrm{Y}} \mathrm{S}(\mathrm{X}_1)$ reste constant dans un voisinage de T, donc également

$$\operatorname{ind}(\mathrm{S}) = \operatorname{codim}_{\mathrm{X}} \mathrm{X}_1 - \operatorname{codim}_{\mathrm{Y}} \mathrm{S}(\mathrm{X}_1).$$

///

L'indice d'un presque-plongement T peut aussi se calculer par la formule

$$\operatorname{ind}(\mathrm{T}) = \dim \ker \mathrm{T} - \operatorname{codim}_{\mathrm{Y}} \mathrm{T}(\mathrm{X}).$$

En effet, d'après le corollaire 4.1.2, on peut trouver X_1 fermé tel que $\mathrm{X} = \mathrm{X}_1 \oplus \ker \mathrm{T}$ et on a vu à la proposition 4.2.3 que T induit un isomorphisme de X_1 sur $\mathrm{T}(\mathrm{X}_1) = \mathrm{T}(\mathrm{X})$. L'indice calculé avec ce choix de X_1 est égal à

$$\operatorname{codim}_{\mathrm{X}} \mathrm{X}_1 - \operatorname{codim}_{\mathrm{Y}} \mathrm{T}(\mathrm{X}_1) = \dim \ker \mathrm{T} - \operatorname{codim}_{\mathrm{Y}} \mathrm{T}(\mathrm{X}).$$

Définition 4.3.1. Soient X, Y deux espaces de Banach ; on dit que $\mathrm{T} \in \mathcal{L}(\mathrm{X}, \mathrm{Y})$ est un *opérateur de Fredholm* s'il existe un sous-espace fermé $\mathrm{X}_1 \subset \mathrm{X}$ de codimension finie, tel que la restriction de T à X_1 soit un isomorphisme de X_1 sur $\mathrm{Y}_1 = \mathrm{T}(\mathrm{X}_1)$, et que $\operatorname{codim}_{\mathrm{Y}} \mathrm{T}(\mathrm{X}_1) < +\infty$. Autrement dit : on a $\operatorname{codim}_{\mathrm{X}} \mathrm{X}_1 < +\infty$, $\operatorname{codim}_{\mathrm{Y}} \mathrm{T}(\mathrm{X}_1) < +\infty$ et il existe une constante $c > 0$ telle que

$$\forall x \in \mathrm{X}_1, \quad \|\mathrm{T}x\| \ge c\,\|x\|.$$

L'indice $\operatorname{ind}(\mathrm{T}) = \operatorname{codim}_{\mathrm{X}} \mathrm{X}_1 - \operatorname{codim}_{\mathrm{Y}} \mathrm{T}(\mathrm{X}_1)$ d'un opérateur de Fredholm est un entier fini, élément de $\mathbb{Z}$. Un opérateur de Fredholm est un presque-plongement dont l'image est de codimension finie. Les opérateurs de Fredholm ont donc les propriétés des presque-plongements : le noyau est de dimension finie, tout sous-espace fermé de X a une image T(X) fermée dans Y.

Proposition. *Un opérateur* $\mathrm{T} \in \mathcal{L}(\mathrm{X}, \mathrm{Y})$ *est de Fredholm si et seulement si son noyau est de dimension finie et son image fermée de codimension finie. On a dans ce cas*

$$\operatorname{ind}(\mathrm{T}) = \dim \ker \mathrm{T} - \operatorname{codim}_{\mathrm{Y}} \mathrm{T}(\mathrm{X}).$$

Preuve. — Supposons que T soit Fredholm au sens de la définition 4.3.1 ; puisque T est un presque-plongement, on sait que $\dim \ker \mathrm{T}$ est finie, et on sait que l'image $\mathrm{T}(\mathrm{X})$ est fermée ; de plus, $\mathrm{T}(\mathrm{X})$ est de codimension finie dans Y puisque $\mathrm{T}(\mathrm{X}_1)$ est déjà de codimension finie. La formule pour l'indice a déjà été vue dans le cas des presque-plongements.

///

Exemples.

1. Le shift $\mathrm{S} \in \mathcal{L}(\ell_2(\mathbb{N}))$ est une isométrie dont l'image est l'hyperplan fermé de $\ell_2(\mathbb{N})$ constitué des vecteurs $x = (x_n)_{n\in\mathbb{N}}$ tels que $x_0 = 0$. On voit que S est un opérateur de Fredholm, et $\operatorname{ind}(\mathrm{S}) = 0 - 1 = -1$.
L'adjoint S^* a un noyau de dimension 1, et il est surjectif ; son indice est $1 = -\operatorname{ind}(\mathrm{S})$. C'est un fait général : si $\mathrm{T} \in \mathcal{L}(\mathrm{H}, \mathrm{K})$ est de Fredholm, son adjoint est de Fredholm et $\operatorname{ind}(\mathrm{T}^*) = -\operatorname{ind}(\mathrm{T})$.

2. L'opérateur $\mathrm{T} : f \to f''$ est un opérateur de Fredholm de l'espace $\mathrm{X} = \mathrm{C}^2([0,1])$ dans $\mathrm{Y} = \mathrm{C}([0,1])$. Il est clairement surjectif et son noyau est de dimension 2 (fonctions affines). On a donc $\operatorname{ind}(\mathrm{T}) = 2$.

Exercice 4.3.2. Si T et U sont Fredholm, $\mathrm{U} \circ \mathrm{T}$ est Fredholm et

$$\operatorname{ind}(\mathrm{U} \circ \mathrm{T}) = \operatorname{ind}(\mathrm{T}) + \operatorname{ind}(\mathrm{U}).$$

Indication. Si $\mathrm{T} \in \mathcal{L}(\mathrm{X}, \mathrm{Y})$ est un plongement sur $\mathrm{X}_1 \subset \mathrm{X}$ et $\mathrm{U} \in \mathcal{L}(\mathrm{Y}, \mathrm{Z})$ un plongement sur $\mathrm{Y}_1 \subset \mathrm{Y}$, considérer le sous-espace $\mathrm{Y}_2 = \mathrm{T}(\mathrm{X}_1) \cap \mathrm{Y}_1$ de Y, ainsi que le sous-espace $\mathrm{X}_2 = \mathrm{X}_1 \cap \mathrm{T}^{-1}(\mathrm{Y}_2)$ de X ; compter les dimensions.

Les presque-plongements s'appellent plus habituellement des *semi-Fredholm* d'indice généralisé égal à $-\infty$. On a aussi des semi-Fredholm d'indice $+\infty$, dont l'image est fermée de codimension finie, mais le noyau de dimension infinie. Leur étude peut se ramener au cas de l'indice généralisé $-\infty$ en passant à l'application transposée (définie du dual Y^* dans le dual X^*) ; cette approche par passage au dual est très agréable dans le cas *réflexif*, par exemple pour les espaces de Hilbert, mais plus délicate dans le cas général. Nous allons plutôt procéder directement.

Lemme 4.3.3. *Si* $\mathrm{T} \in \mathcal{L}(\mathrm{X}, \mathrm{Y})$ *a une image fermée de codimension finie dans* Y *et un noyau de dimension infinie, il en est de même pour* $\mathrm{T} + \mathrm{S}$ *lorsque la norme* $\|\mathrm{S}\|$ *est assez petite.*

Preuve. — On suppose que l'opérateur $\mathrm{T} \in \mathcal{L}(\mathrm{X}, \mathrm{Y})$ a une image Y_0 fermée de codimension finie dans Y et un noyau de dimension infinie. Si on considère T comme une surjection de X sur l'espace de Banach Y_0, on peut appliquer le théorème de l'application ouverte (théorème 9.2), qui fournit un nombre K vérifiant : pour tout $y \in \mathrm{Y}_0$, il existe un vecteur $x = \rho(y)$ dans X tel que $\mathrm{T}x = y$ et $\|\rho(y)\| \le \mathrm{K}\,\|y\|$.

Il n'est pas nécessaire de nommer l'application ρ de Y_0 dans X qui associe à chaque $y \in \mathrm{Y}_0$ le vecteur $x \in \mathrm{X}$ (tel que nous l'avons dit, nous utilisons l'axiome du choix), mais cela permet de voir que nous sommes dans un schéma connu : supposons que $\mathrm{S}_0 \in \mathcal{L}(\mathrm{X}, \mathrm{Y})$ est à valeurs dans Y_0 et vérifie $\|\mathrm{S}_0\| \le 1/(2\mathrm{K})$; considérons la série d'applications (non linéaires) de Y_0 dans X,

$$\sigma = \rho + \rho \circ \mathrm{S}_0 \circ \rho + \rho \circ \mathrm{S}_0 \circ \rho \circ \mathrm{S}_0 \circ \rho + \rho \circ \mathrm{S}_0 \circ \rho \circ \mathrm{S}_0 \circ \rho \circ \mathrm{S}_0 \circ \rho + \cdots ;$$

grâce aux conditions de norme, il est facile de voir que la série σy converge dans X pour tout $y \in \mathrm{Y}_0$ et que $\|\sigma y\| \le 2\,\mathrm{K}\,\|y\|$; grâce au fait que $\mathrm{T}\rho y = y$ pour tout y, on voit que $\mathrm{T}\sigma y = y + \mathrm{S}_0 \sigma y$, donc $(\mathrm{T} - \mathrm{S}_0)(\sigma y) = y$. On a ainsi montré que $\mathrm{T}_1 = \mathrm{T} - \mathrm{S}_0$ est surjectif de X sur Y_0. Comme le noyau $\mathrm{N} = \ker \mathrm{T}_0$ est de dimension infinie, on peut trouver une suite (u_n) de vecteurs de norme 1 dans N, telle que $\|u_n - u_m\| \ge 1/2$ quand $m \ne n$ (corollaire 8.1). On a $\mathrm{T}_1 u_n = -\mathrm{S}_0 u_n$; si $\|\mathrm{S}_0\| < 1/(16\mathrm{K})$, on aura $\|\sigma(\mathrm{S}_0 u_n)\| < 1/8$. Les vecteurs $v_n = u_n + \sigma(\mathrm{S}_0 u_n)$ sont dans le noyau de $\mathrm{T}_1 = \mathrm{T} - \mathrm{S}_0$, de norme ≤ 2, et quand $m \ne n$ on a

$$\|v_m - v_n\| \ge \|u_m - u_n\| - \|\sigma(\mathrm{S}_0 u_m)\| - \|\sigma(\mathrm{S}_0 u_n)\| \ge 1/2 - 2/8 = 1/4 ;$$

le noyau de T_1 ne peut pas être de dimension finie puisque sa boule unité n'est pas compacte : la suite $(v_n/2) \subset \mathrm{B}_{\ker \mathrm{T}_1}$ n'a pas de sous-suite de Cauchy.

Introduisons une projection P linéaire continue de Y sur Y_0. Si $\mathrm{S} \in \mathcal{L}(\mathrm{X}, \mathrm{Y})$ est assez petit et $\mathrm{S}_0 = \mathrm{PS}$, on a vu que $\mathrm{T} - \mathrm{PS}$ a pour image Y_0, et a un noyau de dimension infinie. Pour passer à $\mathrm{T} - \mathrm{S}$, il suffit d'ajouter l'opérateur de rang fini $(\mathrm{P} - \mathrm{Id}_\mathrm{Y})\mathrm{S}$; ce sera un bon exercice([a]) pour le lecteur !

///

Perturbations des opérateurs de Fredholm

Proposition 4.3.4. *Si* $t \in [0,1] \to \mathrm{T}_t \in \mathcal{L}(\mathrm{X}, \mathrm{Y})$ *est un chemin continu de* $[0,1]$ *dans l'espace normé* $\mathcal{L}(\mathrm{X}, \mathrm{Y})$*, et si chaque* T_t *est semi-Fredholm, alors l'indice est constant.*

Preuve. — Par le lemme 4.3.3 et par les résultats précédents de ce chapitre, nous savons que l'indice est localement constant. Le résultat est donc clair par la connexité de $[0,1]$: pour chaque valeur $i \in \mathbb{Z} \cup \{-\infty\} \cup \{+\infty\}$, l'ensemble E_i des points t de $[0,1]$ où l'indice de T_t vaut i est un ouvert, et ces ouverts couvrent $[0,1]$ puisque l'indice de T_t est défini pour tout $t \in [0,1]$ d'après l'hypothèse. Chacun des ouverts est donc aussi fermé (comme complémentaire de la réunion des autres ouverts), donc chaque ensemble E_i est vide ou égal à $[0,1]$. Si i_0 est l'indice de T_0, l'ensemble E_{i_0} est non vide, donc égal à $[0,1]$, ce qui signifie que l'indice est constant égal à i_0.

///

Lemme. *Pour tout opérateur compact* $\mathrm{T} \in \mathcal{L}(\mathrm{X}, \mathrm{Y})$ *et pour tout* $\varepsilon > 0$, *il existe un sous-espace* X_1 *de codimension finie tel que* $\|\mathrm{T}_{|\mathrm{X}_1}\| \le \varepsilon$.

Preuve.— Considérons le compact $\mathrm{K} = \overline{\mathrm{T}(\mathrm{B}_\mathrm{X})}$. Pour chaque $y \in \mathrm{K}$, on trouve par Hahn-Banach une forme linéaire continue $\xi_y \in \mathrm{Y}^*$ telle que $\|y\| = |\xi_y(y)|$; l'ensemble des $y' \in \mathrm{K}$ tels que $\|y'\| < \varepsilon + |\xi_y(y')|$ est un voisinage ouvert de y. Par compacité, on peut recouvrir K par un nombre fini de tels ouverts. On peut sélectionner ainsi un ensemble fini de formes linéaires y_j^*, $j = 1, \ldots, \mathrm{N}$ tel que

$$\forall y \in \mathrm{K}, \quad \|y\| \le \varepsilon + \max_{1 \le j \le \mathrm{N}} |y_j^*(y)|.$$

On choisit pour X_1 l'intersection finie des noyaux des formes linéaires $y_j^* \circ \mathrm{T}$. Pour tout $x \in \mathrm{B}_{\mathrm{X}_1}$, on aura $\mathrm{T}x \in \mathrm{K}$ et $y_j^*(\mathrm{T}x) = 0$ pour $j = 1, \ldots, \mathrm{N}$, donc

$$\|\mathrm{T}x\| \le \varepsilon + \max_{1 \le j \le \mathrm{N}} |y_j^*(\mathrm{T}x)| = \varepsilon.$$

///

Lemme 4.3.5. *Si* $\mathrm{T} \in \mathcal{L}(\mathrm{X}, \mathrm{Y})$ *est un presque-plongement et si* $\mathrm{S} \in \mathcal{L}(\mathrm{X}, \mathrm{Y})$ *est compact, la somme* $\mathrm{T} + \mathrm{S}$ *est un presque-plongement.*

Preuve.— On a un sous-espace X_1 de codimension finie sur lequel T est un plongement, c'est-à-dire qu'il existe une constante $c > 0$ telle que

$$\forall x \in \mathrm{X}_1, \quad \|\mathrm{T}x\| \ge c\,\|x\|,$$

et on peut trouver un sous-espace X_2 de codimension finie sur lequel S est de norme $< c/2$; par l'inégalité triangulaire, $\mathrm{T}+\mathrm{S}$ est un plongement de constante $c/2$ sur $\mathrm{X}_1 \cap \mathrm{X}_2$, qui est de codimension finie dans X.

///

Théorème. *Si* $\mathrm{T} \in \mathcal{L}(\mathrm{X}, \mathrm{Y})$ *est Fredholm et* $\mathrm{S} \in \mathcal{L}(\mathrm{X}, \mathrm{Y})$ *compact, la somme* $\mathrm{T} + \mathrm{S}$ *est Fredholm, de même indice que* T.

Preuve.— On passe de T à $\mathrm{T} + \mathrm{S}$ en suivant le chemin continu $u \to \mathrm{T} + u\mathrm{S}$, $u \in [0, 1]$. Tous les éléments sont semi-Fredholm par le lemme précédent, et on conclut par la proposition 4.3.4.

///

Illustration.

1. Si φ est une fonction continue ≥ 0 sur $[0, 1]$, l'opérateur $\mathrm{V}_1 : f \to f'' - \varphi f$ est surjectif de l'espace X_1 formé des fonctions de classe C^2 sur $[0, 1]$, nulles en 0 et en 1, sur l'espace $\mathrm{Y} = \mathrm{C}([0, 1])$.

On voit que $\mathrm{T}_1 : f \to f''$ est un isomorphisme de X_1 sur Y, donc il est Fredholm d'indice 0 ; on montre que l'injection de X_1 dans Y est compacte par le théorème d'Ascoli (les éléments de la boule unité de X_1 vérifient $|f''| \le 1$ et $|f'| \le 1$ sur $[0, 1]$), et l'opérateur de multiplication par φ est continu de Y dans Y, donc $\mathrm{S}_1 : f \to \varphi f$ est compact de X_1 dans Y. L'opérateur $\mathrm{V}_1 = \mathrm{T}_1 - \mathrm{S}_1$ proposé est donc Fredholm d'indice 0. Mais V_1 est injectif car

$$\int_0^1 (\mathrm{V}_1 f)(t)\overline{f(t)}\,\mathrm{d}t = \int_0^1 \big(f''(t) - \varphi(t) f(t)\big)\overline{f(t)}\,\mathrm{d}t = -\int_0^1 \big(|f'(t)|^2 + \varphi(t)|f(t)|^2\big)\,\mathrm{d}t$$

ne peut être nul que si $f = 0$. Il en résulte que V_1 est surjectif, donc un isomorphisme : pour toute fonction continue g sur $[0,1]$, il existe une unique fonction f de classe C^2, nulle aux deux bords, telle que

$$f'' - \varphi\, f = g.$$

Si on définit V_2 sur l'espace X_2 plus grand formé des fonctions f de classe C^2 sur $[0,1]$ telles que $f(0) = f(1)$, par la même formule $\mathrm{V}_2 : f \in \mathrm{X}_2 \to f'' - \varphi f \in \mathrm{Y}$, il est évident que V_2 ne peut pas être injectif puisque V_1 couvrait déjà toute l'image Y. On a ainsi montré par une méthode un peu bizarre qu'il existe toujours des fonctions non nulles f telles que $f'' = \varphi f$.

Pour le point précédent, on pouvait aussi revenir à l'opérateur T, tel qu'il avait été défini sur $\mathrm{X} = \mathrm{C}^2([0,1])$ par $\mathrm{T} : f \in \mathrm{X} \to f'' \in \mathrm{Y}$; cet opérateur était Fredholm, d'indice égal à 2. L'opérateur $\mathrm{V} : f \in \mathrm{X} \to f'' - \varphi f \in \mathrm{Y}$ est aussi Fredholm d'indice 2 (perturbation compacte de T), pour toute fonction continue φ ; on a $\dim\ker \mathrm{T} \ge \mathrm{ind}(\mathrm{T}) = 2$, ce qui donne l'information bien raisonnable suivante : pour toute fonction continue φ, l'équation différentielle $y'' = \varphi\, y$ possède au moins deux solutions indépendantes.

2. On va généraliser en deux dimensions. Considérons l'espace H_2 des fonctions f de deux variables de la forme

$$f(x,y) \sim \sum_{m,n \in \mathbb{Z}} c_{m,n}\, \mathrm{e}^{\mathrm{i}mx}\, \mathrm{e}^{\mathrm{i}ny}$$

pour lesquelles on suppose que

$$\|f\|^2 = \sum_{m,n\in\mathbb{Z}} (1 + m^2 + n^2)^2\, |c_{m,n}|^2 < +\infty\,;$$

la fonction $f \to \|f\|$ définit une norme d'espace de Hilbert sur cet espace H_2. L'opérateur $\mathrm{L} : u \to -\Delta u + u$ agit sur les polynômes trigonométriques par

$$\sum_{m,n} c_{m,n} e_m \otimes e_n \to \sum_{m,n} (1 + m^2 + n^2)\, c_{m,n} e_m \otimes e_n.$$

D'après la définition de la norme de H_2, il est clair que cet opérateur est borné de H_2 à valeurs dans $\mathrm{H}_0 = \mathrm{L}_2([0,2\pi]^2)$; en fait il est facile de vérifier qu'il se prolonge en un isomorphisme T de H_2 sur H_0, donc T est Fredholm d'indice 0.

Considérons une fonction continue périodique $\varphi(x,y)$ telle que $\varphi(x,y) > 0$ en tout point, et posons

$$(\mathrm{S}u)(x,y) = (\varphi(x,y) - 1)\, u(x,y).$$

Cet opérateur est compact de H_2 dans H_0 : on voit d'abord que l'injection $u \to u$ est compacte de H_2 dans H_0 (utiliser les coefficients dans la base hilbertienne $(e_m \otimes e_n)_{m,n\in\mathbb{Z}}$), et on compose avec l'opérateur borné de H_0 dans H_0 donné par la multiplication par la fonction bornée $\varphi - 1$. On sait alors que $\mathrm{T} + \mathrm{S}$ est Fredholm d'indice 0 ; mais $\mathrm{T} + \mathrm{S}$ est injectif car

$$\langle (\mathrm{T}+\mathrm{S})u, u\rangle = \int_{[0,2\pi]^2} (-\Delta u + \varphi u)\, \overline{u}\, \mathrm{d}x\,\mathrm{d}y = \int_{[0,2\pi]^2} \bigl(|\nabla u|^2 + \varphi |u|^2\bigr)\, \mathrm{d}x\,\mathrm{d}y$$

ne peut être nul que si u est nulle. On en déduit que cet opérateur est surjectif : pour toute fonction $g \in \mathrm{H}_0$, il existe $u \in \mathrm{H}_2$ telle que

$$-\Delta u + \varphi\, u = g.$$

3. Considérons l'espace de Hilbert complexe $\mathrm{H} = \mathrm{L}_2(0, 2\pi)$, où $[0, 2\pi]$ est muni de la probabilité $d\theta/(2\pi)$. Posons $e_n(\theta) = \mathrm{e}^{\mathrm{i}n\theta}$, pour tout $n \in \mathbb{Z}$; la famille $(e_n)_{n\in\mathbb{Z}}$ est une base hilbertienne de H. Désignons par $\mathrm{H}_2(\mathbb{T})$ le sous-espace fermé de H engendré par les vecteurs $(e_n)_{n\geq 0}$, et désignons par P la projection orthogonale de H sur $\mathrm{H}_2(\mathbb{T})$: on a $\mathrm{P}e_n = e_n$ si $n \geq 0$ et $\mathrm{P}e_n = 0$ si $n < 0$. Pour chaque fonction complexe continue φ sur $[0, 2\pi]$, telle que $\varphi(0) = \varphi(2\pi)$, définissons l'endomorphisme T_φ de $\mathrm{H}_2(\mathbb{T})$ qui associe à $f \in \mathrm{H}_2(\mathbb{T})$ la projection sur $\mathrm{H}_2(\mathbb{T})$ du produit φf,

$$\forall f \in \mathrm{H}_2(\mathbb{T}), \quad \mathrm{T}_\varphi f = \mathrm{P}(\varphi f) \in \mathrm{H}_2(\mathbb{T}).$$

Si φ ne s'annule pas, l'opérateur T_φ est de Fredholm, et son indice se calcule à partir de l'indice (autour de 0) du chemin fermé $\theta \in [0, 2\pi] \to \varphi(\theta) \in \mathbb{C}$. Si par exemple

$$\varphi_0(\theta) = \mathrm{e}^{-\mathrm{i}\theta},$$

on voit que T_{φ_0} agit comme l'adjoint S^* de l'opérateur de shift sur $\ell_2(\mathbb{N})$,

$$\mathrm{T}_{\varphi_0} e_0 = 0, \quad \text{et} \quad \mathrm{T}_{\varphi_0} e_{n+1} = e_n \quad \text{pour} \quad n \geq 0.$$

On ramène le cas général à ces exemples simples par déformation continue du chemin : on commence par approcher φ uniformément par un polynôme trigonométrique

$$\mathrm{Q}(\theta) = \sum_{n=-\mathrm{N}}^{\mathrm{N}} c_n \, \mathrm{e}^{\mathrm{i}n\theta} ;$$

si l'approximation est suffisamment bonne, la déformation continue $t \to (1-t)\varphi + t\mathrm{Q}$, pour $t \in [0, 1]$, est formée de fonctions qui ne s'annulent pas sur $[0, 2\pi]$. On peut trouver un polynôme (algébrique) P de degré 2N tel que $\mathrm{Q}(\theta) = \mathrm{e}^{-\mathrm{i}\mathrm{N}\theta}\,\mathrm{P}(\mathrm{e}^{\mathrm{i}\theta})$. On factorise P sous la forme $a \prod_j (\mathrm{X} - \lambda_j)$, où aucune racine n'est sur le cercle unité de $\mathbb{C}$. Par une nouvelle déformation portant uniquement sur le choix d'un multiple scalaire du polynôme, on peut amener P à avoir la forme

$$\widetilde{\mathrm{P}} = \prod_{|\lambda_j|<1} (\mathrm{X} - \lambda_j) \prod_{|\lambda_j|>1} (1 - \mathrm{X}/\lambda_j).$$

Ensuite, on déplace vers 0 les ℓ racines de P intérieures au cercle, et les autres « vers l'infini » pour arriver finalement à $\mathrm{P}_1 = \mathrm{X}^\ell$ et $\mathrm{Q}_1(\theta) = \mathrm{e}^{\mathrm{i}k\theta}$, $k = \ell - \mathrm{N}$. L'opérateur associé à la fonction Q_1 est une puissance du shift S si $k \geq 0$, ou une puissance du shift inverse S^* si $k < 0$. L'indice de cet opérateur est égal à $-k$, et l'indice du chemin est k ; par ailleurs l'indice de l'opérateur et l'indice du chemin associé sont restés constants pendant la déformation. Ainsi,

si φ est une fonction complexe continue sur $[0, 2\pi]$ qui ne s'annule pas sur cet intervalle et telle que $\varphi(0) = \varphi(2\pi)$, l'opérateur T_φ est de Fredholm, et son indice est l'opposé de l'indice du chemin fermé $\gamma : \theta \in [0, 2\pi] \to \gamma(\theta) = \varphi(\theta)$ autour de 0,

$$\operatorname{ind} \mathrm{T}_\varphi = -\frac{1}{2\mathrm{i}\pi} \int_\gamma \frac{dz}{z} \cdot$$

Formulation classique de l'alternative de Fredholm

À l'époque de l'article de Fredholm (1903), il n'y avait pas plus d'espaces de Banach que de théorie des opérateurs compacts. Fredholm s'intéressait à la résolution d'équations intégrales de la forme suivante : étant donnés un noyau $k(x, y)$ et une fonction continue g sur $[0, 1]$, peut-on trouver f continue sur $[0, 1]$ qui vérifie l'équation intégrale

$$\forall x \in [0,1], \quad f(x) = \int_0^1 k(x,y) f(y) \, \mathrm{d}y + g(x).$$

Si le noyau k est continu sur le carré, il résulte du théorème d'Ascoli que l'opérateur intégral T_k de noyau k est compact de $\mathrm{C}([0, 1])$ dans lui-même, et on est en train d'essayer de résoudre une équation de la forme $(\mathrm{Id} - \mathrm{T}_k)f = g$. Fredholm découvre dans le langage de l'époque que l'indice de $\mathrm{Id} - \mathrm{T}_k$ est nul, ce qui le conduit à formuler ce qui est resté connu sous le nom *d'alternative de Fredholm*.

Quelques années après, sous l'influence de F. Riesz, on est arrivé à peu de chose près à la formulation « classique » suivante : soit S un opérateur compact sur X ; on sait que la transposée ${}^t\mathrm{S}$ est compacte de X^* dans X^*. On a *l'alternative* suivante :

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

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

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

4.4. Valeurs propres des opérateurs compacts

Lemme. *Si* $\mathrm{T} \in \mathcal{L}(\mathrm{X})$ *est un presque-plongement, l'intersection* Y *des images* $\mathrm{T}^n(\mathrm{X})$ *est fermée et* $\mathrm{T}(\mathrm{Y}) = \mathrm{Y}$.

Preuve. — Sous l'hypothèse de presque-plongement, on voit par récurrence que l'image $\mathrm{T}^{n+1}(\mathrm{X}) = \mathrm{T}(\mathrm{T}^n(\mathrm{X}))$ du sous-espace fermé $\mathrm{T}^n(\mathrm{X})$ est fermée, donc

$$\mathrm{Y} = \bigcap_{n \ge 0} \mathrm{T}^n(\mathrm{X})$$

est un sous-espace vectoriel fermé ; il est clair que $\mathrm{T}(\mathrm{Y}) \subset \mathrm{Y}$. De plus, le noyau N de T est de dimension finie, donc la suite décroissante de sous-espaces vectoriels $\mathrm{N} \cap \mathrm{T}^n(\mathrm{X})$ est stationnaire à partir d'un certain entier $k \ge 0$, d'où résulte que $\mathrm{N} \cap \mathrm{Y} = \mathrm{N} \cap \mathrm{T}^k(\mathrm{X})$. On en déduit que $\mathrm{N} \cap \mathrm{T}^k(\mathrm{X}) \subset \mathrm{Y}$. Si $y \in \mathrm{Y}$, on pourra écrire $y = \mathrm{T}^{k+n+1}x_n$ pour tout entier $n \ge 0$; posons $z_n = \mathrm{T}^{k+n}x_n$; alors $\mathrm{T}z_n = y$ pour tout $n \ge 0$, et $z_n \in \mathrm{T}^{k+n}(\mathrm{X})$; on a pour tout $p > 0$ que $z_0, z_p \in \mathrm{T}^k(\mathrm{X})$ et $z_0 - z_p \in \mathrm{N}$, donc $z_0 - z_p \in \mathrm{Y}$; on en déduit que $z_0 \in z_p + \mathrm{Y} \subset \mathrm{T}^{k+p}(\mathrm{X})$ pour tout p, donc $z_0 \in \mathrm{Y}$ et $y = \mathrm{T}z_0 \in \mathrm{T}(\mathrm{Y})$.

///

Lemme. *Soient* X *un espace de Banach complexe,* $\mathrm{X} \neq \{0\}$ *et* $\mathrm{T} \in \mathcal{L}(\mathrm{X})$ *un opérateur de Fredholm ; si* 0 *appartient au bord du spectre de* T*, les noyaux et les images de* T^n *se stabilisent : il existe un entier* $p \geq 1$ *tel que* $\ker \mathrm{T}^p = \ker \mathrm{T}^{p+1}$, $\mathrm{T}^p(\mathrm{X}) = \mathrm{T}^{p+1}(\mathrm{X})$ *; de plus,*

$$\mathrm{X} = \ker \mathrm{T}^p \oplus \mathrm{T}^p(\mathrm{X}),$$

ces deux sous-espaces étant stables par T*, et* $\dim \ker \mathrm{T}^p < +\infty$*. La valeur* 0 *est valeur propre de* T*, et* 0 *est un point isolé dans le spectre de* T*. La restriction de* T *à* $\mathrm{T}^p(\mathrm{X})$ *est un isomorphisme de l'espace de Banach* $\mathrm{T}^p(\mathrm{X})$*.*

Preuve. — Puisque 0 est dans le bord du spectre de T, l'opérateur T est limite d'opérateurs inversibles, qui sont Fredholm d'indice nul, et T lui-même est Fredholm, donc $\mathrm{ind}(\mathrm{T}) = 0$. Le noyau $\mathrm{N} = \ker \mathrm{T}$ de T n'est pas nul, sinon T serait un isomorphisme, autrement dit 0 est valeur propre de T ; il en résulte en passant que l'entier p de l'énoncé ne peut pas être nul, donc il sera ≥ 1. On sait que T est un presque-plongement ; par la preuve du lemme précédent, on peut trouver un entier k tel que $\mathrm{N} \cap \mathrm{T}^k(\mathrm{X}) = \mathrm{N} \cap \mathrm{Y}$, où $\mathrm{Y} = \bigcap_n \mathrm{T}^n(\mathrm{X})$. L'espace fermé Y est invariant par T ; considérons pour tout $\lambda \in \mathbb{C}$

$$\mathrm{V}_\lambda = \mathrm{T}_{|\mathrm{Y}} - \lambda\, \mathrm{Id}_\mathrm{Y} \in \mathcal{L}(\mathrm{Y})\,;$$

on voit d'abord que V_0 est Fredholm, d'indice ≥ 0 : en effet $\ker \mathrm{V}_0 = \mathrm{N} \cap \mathrm{Y} \subset \mathrm{N}$ est de dimension finie, et V_0 est surjectif par le lemme précédent ; pour λ proche de 0, l'opérateur V_λ est donc Fredholm de même indice ≥ 0 que V_0 ; mais comme l'indice de V_λ est ≤ 0 quand $\lambda \in \rho(\mathrm{T})$ (V_λ est alors injectif car $\mathrm{T} - \lambda\,\mathrm{Id}$ est inversible), et 0 étant adhérent à $\rho(\mathrm{T})$ par hypothèse, on déduit que V_0 est d'indice nul ; il en résulte que V_0 est injectif, c'est-à-dire que $\mathrm{N} \cap \mathrm{Y} = \mathrm{N} \cap \mathrm{T}^k(\mathrm{X}) = \{0\}$, et V_0 est un isomorphisme de Y. On en déduit que la suite des noyaux est stationnaire : si $\mathrm{T}^{k+1}x = 0$, $\mathrm{T}^k x$ est à la fois dans N et dans $\mathrm{T}^k(\mathrm{X})$, donc $\mathrm{T}^k x = 0$. Comme chaque T^n est Fredholm d'indice 0 par le résultat de l'exercice 4.3.2, les images se stabilisent au même moment, donc $\mathrm{Y} = \mathrm{T}^k(\mathrm{X})$.

L'espace X se décompose en somme directe de $\mathrm{Y} = \mathrm{T}^k(\mathrm{X})$ et de l'espace de dimension finie $\mathrm{N}_k = \ker \mathrm{T}^k$: en effet, l'intersection $\mathrm{N}_k \cap \mathrm{T}^k(\mathrm{X})$ est réduite à $\{0\}$ et

$$\dim \mathrm{N}_k = \mathrm{codim}\, \mathrm{T}^k(\mathrm{X})$$

puisque $\mathrm{ind}(\mathrm{T}^k) = 0$; vérifions l'affirmation sur l'intersection : si $y = \mathrm{T}^k x$ est en même temps dans N_k, on a $0 = \mathrm{T}^k y = \mathrm{T}^{2k} x$. Comme $2k \geq k$, on sait que $\ker \mathrm{T}^{2k} = \ker \mathrm{T}^k$, donc $x \in \ker \mathrm{T}^k$ et $y = \mathrm{T}^k x = 0$. Il est clair que $\mathrm{X}_0 = \ker \mathrm{T}^k$ et $\mathrm{X}_1 = \mathrm{T}^k(\mathrm{X}) = \mathrm{Y}$ sont stables par T ; on a ainsi montré l'essentiel du lemme, avec $p = k$. Il reste à montrer que 0 est isolé dans le spectre : si $\mathrm{T}_j \in \mathcal{L}(\mathrm{X}_j)$ désigne la restriction de T à X_j, pour $j = 0, 1$, on voit que T_0 est un endomorphisme nilpotent en dimension finie ($\mathrm{T}_0^k = 0$) donc 0 est sa seule valeur propre et $\mathrm{T}_0 - \lambda\,\mathrm{Id}_{\mathrm{X}_0}$ est inversible pour tout $\lambda \neq 0$; on a vu que $\mathrm{T}_1 = \mathrm{V}_0$ est un isomorphisme de X_1 sur X_1 ; il existe donc $\varepsilon_1 > 0$ tel que $\mathrm{T}_1 - \lambda\,\mathrm{Id}_{\mathrm{X}_1}$ reste un isomorphisme de X_1 quand $|\lambda| < \varepsilon_1$. Il en résulte que $\mathrm{T} - \lambda\,\mathrm{Id}_\mathrm{X}$ est inversible quand $0 < |\lambda| < \varepsilon_1$, ce qui montre que la valeur spectrale 0 est isolée dans $\sigma(\mathrm{T})$.

///

Théorème. *Soient* X *un espace de Banach complexe et* $\mathrm{S} \in \mathcal{L}(\mathrm{X})$ *un opérateur compact ; le spectre de* S *est fini, ou bien peut être rangé dans une suite* (λ_n) *tendant vers* 0*. Chaque valeur spectrale* $\lambda \neq 0$ *est une valeur propre de multiplicité finie, ce qui veut dire que l'opérateur* $\mathrm{T} = \lambda\,\mathrm{Id}_\mathrm{X} - \mathrm{S}$ *admet la décomposition du lemme précédent.*

Preuve. — Si $\sigma(\mathrm{S}) = \{0\}$ il n'y a rien à démontrer. Sinon, soit $\lambda \neq 0$ un point frontière du spectre de S ; alors $\mathrm{T} = \lambda\,\mathrm{Id_X} - \mathrm{S}$ est Fredholm, et $0 \in \partial\,\sigma(\mathrm{T})$. D'après le lemme précédent, λ est isolé dans le spectre de S. Puisque tous les points frontière non nuls sont isolés, il en résulte que tous les points non nuls du spectre de S sont isolés (voir plus loin). Le spectre de S n'a donc qu'un nombre fini de points dans toute couronne compacte $\{r \le |z| \le \mathrm{R}\}$ telle que $r > 0$, ce qui permet de les ranger s'il y a lieu dans une suite qui tend vers 0. Les propriétés de chaque valeur spectrale sont données par le lemme précédent, appliqué à $\mathrm{T} = \lambda\,\mathrm{Id_X} - \mathrm{S}$.

Justifions l'affirmation topologique précédente : les points non isolés du spectre forment un compact $\mathrm{K} \subset \sigma(\mathrm{S})$; si K est vide, il n'y a rien à montrer, sinon soit μ un point de K de module maximal ; le point μ ne peut pas être intérieur au spectre, sinon tout un voisinage de μ dans $\mathbb{C}$ serait formé de points du spectre, donc des points non isolés, et il existerait parmi eux des $\mu_1 \in \mathrm{K}$ de module plus grand que $|\mu|$. Ainsi $\mu \in \partial\,\sigma(\mathrm{S})$, et donc $\mu = 0$, sinon μ serait isolé d'après l'argument du paragraphe précédent. Puisque $\mu = 0$, on a bien montré que tous les $\lambda \neq 0$ du spectre de S sont isolés dans le spectre.

///

Remarque. Si $\mathrm{T} \in \mathcal{L}(\mathrm{X})$ est un opérateur de Fredholm, X complexe, on peut écrire $\mathrm{X} = \mathrm{X}_1 \oplus \mathrm{E} = \mathrm{T}(\mathrm{X}_1) \oplus \mathrm{F}$, où E, F sont de dimension finie et où T induit un isomorphisme $\mathrm{T}_1 \in \mathcal{L}(\mathrm{X}_1, \mathrm{T}(\mathrm{X}_1))$. Si $\mathrm{P} \in \mathcal{L}(\mathrm{X}, \mathrm{T}(\mathrm{X}_1))$ est la projection de X sur $\mathrm{T}(\mathrm{X}_1)$ parallèlement à F et si $\mathrm{S} = \mathrm{T}_1^{-1}\mathrm{P} \in \mathcal{L}(\mathrm{X})$, on constate que $\mathrm{ST} - \mathrm{Id_X}$ est nul sur X_1, donc de rang fini, donc compact, et de même $\mathrm{TS} - \mathrm{Id_X}$ est compact. Il en résulte que l'image $\widehat{\mathrm{T}}$ de T dans l'algèbre quotient $\mathrm{C} = \mathcal{L}(\mathrm{X})/\mathcal{K}(\mathrm{X})$ est inversible, d'inverse $\widehat{\mathrm{S}}$.

Inversement, si T est *inversible modulo les compacts*, c'est-à-dire si $\widehat{\mathrm{T}}$ est inversible, il existe S_1 et S_2, dans la même classe, tels que $\mathrm{S}_1\mathrm{T} - \mathrm{Id_X}$ et $\mathrm{T}\,\mathrm{S}_2 - \mathrm{Id_X}$ soient compacts ; alors $\mathrm{S}_1\mathrm{T}$ et $\mathrm{T}\,\mathrm{S}_2$ sont Fredholm comme perturbations compactes de l'identité. Il en résulte que le noyau de T est de dimension finie, et que l'image de T contient un sous-espace fermé de codimension finie, donc im(T) elle-même est fermée de codimension finie (petit exercice). Finalement T est Fredholm *si et seulement si* il est inversible modulo les compacts.

Si X est de dimension infinie, cette algèbre de Calkin C est une algèbre de Banach unitaire complexe : le spectre de $\widehat{\mathrm{T}}$ est donc non vide. Si $\lambda \in \sigma(\widehat{\mathrm{T}})$, non seulement $\mathrm{T} - \lambda\,\mathrm{Id_X}$ n'est pas inversible, mais il ne peut pas être Fredholm. Ce spectre $\sigma(\widehat{\mathrm{T}})$ est une partie fermée non vide de $\sigma(\mathrm{T})$, appelée *spectre essentiel*. La description donnée au théorème précédent est en fait valable pour tout $\lambda \in \sigma(\mathrm{T})$ qui appartient à la composante connexe de l'infini dans le complémentaire du spectre essentiel ; lorsque T est compact, son spectre essentiel est réduit à $\{0\}$.

Notes du chapitre 4

([a]) Soit $\mathrm{S} \in \mathcal{L}(\mathrm{X}, \mathrm{Y})$ de rang un, et posons $\mathrm{H} = \ker(\mathrm{S})$; c'est un hyperplan de X ; le noyau de $\mathrm{T} + \mathrm{S}$ est de dimension infinie, puisqu'il contient le sous-espace de dimension infinie $\ker(\mathrm{T}) \cap \mathrm{H}$. Montrons ensuite que l'image de $\mathrm{T} + \mathrm{S}$ est fermée.

Si $\ker(\mathrm{T}) \subset \mathrm{H}$, l'application S se factorise par le quotient $\widehat{\mathrm{X}} = \mathrm{X}/\ker(\mathrm{T})$: il existe un opérateur de rang un $\widehat{\mathrm{S}}$ tel que $\mathrm{S} = \widehat{\mathrm{S}} \circ \pi$, où π est la surjection canonique de X sur $\widehat{\mathrm{X}}$, et de même $\mathrm{T} = \widehat{\mathrm{T}} \circ \pi$. Mais $\widehat{\mathrm{T}}$ est un plongement de $\widehat{\mathrm{X}}$ dans Y, donc $\widehat{\mathrm{T}} + \widehat{\mathrm{S}}$ est un presque-plongement (remarque après la définition 4.2.2, ou bien cas très particulier du lemme 4.3.5), et il en résulte que l'image $\mathrm{im}(\widehat{\mathrm{T}}+\widehat{\mathrm{S}})$ est fermée. Comme π est surjective, on voit que $\mathrm{im}(\mathrm{T}+\mathrm{S}) = \mathrm{im}(\widehat{\mathrm{T}}+\widehat{\mathrm{S}})$ est fermée dans Y.

Si $\ker(\mathrm{T})$ n'est pas contenu dans $\mathrm{H} = \ker(\mathrm{S})$, on peut trouver un vecteur $x_0 \notin \mathrm{H}$ tel que $\mathrm{T}x_0 = 0$. On a alors $\mathrm{X} = \mathrm{H} \oplus \mathbb{K}x_0$, donc $\mathrm{T}(\mathrm{X}) = \mathrm{T}(\mathrm{H})$. L'image $(\mathrm{T}+\mathrm{S})(\mathrm{X})$ est égale à la somme de $(\mathrm{T}+\mathrm{S})(\mathrm{H}) = \mathrm{T}(\mathrm{H}) = \mathrm{T}(\mathrm{X})$ qui est fermé, et du sous-espace de dimension finie $(\mathrm{T}+\mathrm{S})(\mathbb{K}x_0) = \mathbb{K}\mathrm{S}x_0$. Par le lemme 4.1.3, on conclut que $(\mathrm{T}+\mathrm{S})(\mathrm{X})$ est fermé.

Pour finir ce premier cas, on note que l'image de $\mathrm{T}+\mathrm{S}$ contient $(\mathrm{T}+\mathrm{S})(\mathrm{H}) = \mathrm{T}(\mathrm{H})$, qu'il existe x tel que $\mathrm{X} = \mathrm{H} \oplus \mathbb{K}x$ et F de dimension finie tel que $\mathrm{Y} = \mathrm{T}(\mathrm{X}) + \mathrm{F}$. Alors

$$(\mathrm{T}+\mathrm{S})(\mathrm{X}) + (\mathbb{K}\mathrm{T}x + \mathrm{F}) \supset \mathrm{T}(\mathrm{H}) + \mathbb{K}\mathrm{T}x + \mathrm{F} = \mathrm{T}(\mathrm{X}) + \mathrm{F} = \mathrm{Y},$$

donc l'image de $\mathrm{T}+\mathrm{S}$ est de codimension finie. Si $\mathrm{S} \in \mathcal{L}(\mathrm{X},\mathrm{Y})$ est de rang fini, on procédera par récurrence en écrivant S comme une somme finie d'opérateurs de rang un.

5. Opérateurs autoadjoints non bornés

5.1. Opérateurs non bornés

Soient X et Y deux espaces vectoriels sur $\mathbb{K} = \mathbb{R}$ ou $\mathbb{K} = \mathbb{C}$; une *application linéaire partiellement définie* T de X dans Y (un peu plus loin, on dira un *opérateur*) est définie par la donnée d'un sous-espace vectoriel $\mathrm{dom}(\mathrm{T})$ de X appelé *domaine* de T et d'une application linéaire (usuelle) $\mathrm{L_T}$ de $\mathrm{dom}(\mathrm{T})$ dans Y. Autrement dit, la donnée T est celle de $(\mathrm{X}, \mathrm{Y}, \mathrm{dom}(\mathrm{T}), \mathrm{L_T})$. Dans la suite, pour tout vecteur $x \in \mathrm{dom}(\mathrm{T})$ on écrira simplement $\mathrm{T}x = \mathrm{L_T}\,x$. Si T est une application linéaire partiellement définie, le *graphe* de T est le sous-espace vectoriel du produit $\mathrm{X} \times \mathrm{Y}$ égal à

$$\mathrm{Gr}(\mathrm{T}) = \{(x, \mathrm{T}x) : x \in \mathrm{dom}(\mathrm{T})\}.$$

La restriction à $\mathrm{Gr}(\mathrm{T})$ de la première projection est injective. Réciproquement, appelons *graphe partiel* tout sous-espace vectoriel G de $\mathrm{X} \times \mathrm{Y}$ tel que la restriction de la première projection à G soit injective : si $(0, y) \in \mathrm{G}$, alors $y = 0$. On voit facilement que tout graphe partiel est le graphe d'une unique application linéaire partiellement définie T. La correspondance qui à T associe son graphe est une correspondance bijective entre applications linéaires partiellement définies et graphes partiels :

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

On appellera *image* de T le sous-espace vectoriel $\mathrm{im}(\mathrm{T})$ de Y formé de tous les vecteurs $\mathrm{T}x$, pour x variant dans $\mathrm{dom}(\mathrm{T})$; c'est l'image du graphe $\mathrm{Gr}(\mathrm{T})$ par la seconde projection du produit $\mathrm{X} \times \mathrm{Y}$.

Désormais on dira *opérateur* au lieu d'application linéaire partiellement définie. On appelle *extension* d'un opérateur T_0 tout opérateur T_1 tel que $\mathrm{Gr}(\mathrm{T}_0) \subset \mathrm{Gr}(\mathrm{T}_1)$. On écrit alors $\mathrm{T}_0 \subset \mathrm{T}_1$; cela signifie que le domaine de T_1 est plus grand que celui de T_0 et que T_1 *prolonge* T_0, en gardant les mêmes valeurs sur $\mathrm{dom}(\mathrm{T}_0)$.

Soient S et T deux opérateurs de X dans Y ; on définit l'opérateur $\mathrm{S} + \mathrm{T}$ en posant $\mathrm{dom}(\mathrm{S} + \mathrm{T}) = \mathrm{dom}(\mathrm{S}) \cap \mathrm{dom}(\mathrm{T})$ et en posant $(\mathrm{S} + \mathrm{T})(x) = \mathrm{S}x + \mathrm{T}x$ pour tout vecteur $x \in \mathrm{dom}(\mathrm{S} + \mathrm{T})$. Si S est une application linéaire usuelle de X dans Y, elle définit un opérateur de la façon la plus évidente : on pose $\mathrm{dom}(\mathrm{S}) = \mathrm{X}$ et $\mathrm{S}x \in \mathrm{Y}$ aura le sens habituel pour tout $x \in \mathrm{X}$; dans ce cas, si T est un opérateur (non borné) de X dans Y, le domaine de $\mathrm{S} + \mathrm{T}$ sera égal à celui de l'opérateur T. Cette remarque sera utilisée lorsque $\mathrm{X} = \mathrm{Y}$ et $\mathrm{S} = \lambda\,\mathrm{Id_X}$, pour introduire l'opérateur $\mathrm{T} - \lambda\,\mathrm{Id_X}$, de même domaine que T.

Soient X, Y et Z des espaces vectoriels, T un opérateur de X dans Y et S un opérateur de Y dans Z ; on définit la composition ST de ces deux opérateurs en posant d'abord $\mathrm{dom}(\mathrm{ST}) = \{x \in \mathrm{dom}(\mathrm{T}) : \mathrm{T}x \in \mathrm{dom}(\mathrm{S})\}$ et en posant $(\mathrm{ST})x = \mathrm{ST}x$ pour tout $x \in \mathrm{dom}(\mathrm{ST})$.

L'attitude habituelle quand on travaille avec les opérateurs bornés continus est d'essayer de les prolonger le plus vite possible à l'espace complet convenable : penser à la transformation de Fourier, qui est définie sur $L_1(\mathbb{R})$ par la formule intégrale usuelle ; on appelle aussi transformation de Fourier son extension par continuité depuis $L_1(\mathbb{R}) \cap L_2(\mathbb{R})$ jusqu'à l'espace complet $L_2(\mathbb{R})$. Pour comprendre les définitions de ce paragraphe, il faut se dire qu'on adopte l'attitude radicalement opposée : ici, on ne prend **aucune** initiative de prolongement ; si T_1 est défini sur D_1 et T_2 sur D_2, la seule chose que nous sommes obligés d'admettre est que les deux sont définis sur $D_1 \cap D_2$. On ne cherche surtout pas à aller plus loin.

Définition. Soient X et Y deux espaces de Banach ; un opérateur T de X dans Y est dit *densément défini* si son domaine dom(T) est dense dans X. Un opérateur de X dans Y est dit *fermé* si son graphe est un sous-espace fermé de $X \times Y$. Un opérateur de X dans Y est dit *fermable* s'il admet une extension fermée.

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

Exemples.

A. On prend $X = Y = L_2(\mathbb{R})$, et on définit un opérateur D_0 (D comme « dérivation ») de la façon suivante : $\mathrm{dom}(D_0)$ est l'espace $\mathcal{D}(\mathbb{R})$ des fonctions C^∞ à support compact et on pose $D_0 f = f'$ pour $f \in \mathrm{dom}(D_0)$. Cet opérateur est densément défini puisque $\mathcal{D}(\mathbb{R})$ est dense dans $L_2(\mathbb{R})$; il est assez facile de voir qu'il n'est pas fermé. En revanche il est fermable, et on déterminera plus loin sa fermeture.

B. Considérons $X = Y = L_2(0,1)$; désignons par P l'opérateur borné de « primitive nulle en 0 », défini par $(\mathrm{P}f)(t) = \int_0^t f(s)\,\mathrm{d}s$. On a vu que P est injectif, donc P définit une bijection de X sur $\mathrm{P(X)} = \mathrm{im(P)}$. On peut donc définir l'opérateur $\mathrm{T} = \mathrm{P}^{-1}$ de domaine im(P) en posant pour tout $g \in \mathrm{im(P)}$

$$(\mathrm{T}g = f) \Leftrightarrow (\mathrm{P}f = g).$$

Cet opérateur T est donc injectif lui-aussi, de im(P) dans $L_2(0,1)$ (en fait il est *bijectif* de im(P) *sur* $L_2(0,1)$). On voit facilement que im(P) est dense, donc $\mathrm{P}^{-1} = \mathrm{T}$ est densément défini. Puisque P est continu, son graphe est fermé, donc P^{-1} est fermé puisque son graphe s'obtient à partir de celui de P par l'homéomorphisme $(x,y) \to (y,x)$ de $L_2(0,1) \times L_2(0,1)$ sur lui-même.

Intégration par parties et dérivées généralisées, distributions

Supposons que deux fonctions intégrables f, g soient données sur $[a,b]$ et que l'on pose ensuite

$$\mathrm{F}(x) = \alpha + \int_a^x f(t)\,\mathrm{d}t, \quad \mathrm{G}(x) = \beta + \int_a^x g(t)\,\mathrm{d}t.$$

En utilisant le théorème de Fubini, on montre que

$$\text{(IPP)} \qquad \int_a^b \mathrm{F}(t)g(t)\,\mathrm{d}t = \Big[\mathrm{F}(t)\mathrm{G}(t)\Big]_a^b - \int_a^b f(t)\mathrm{G}(t)\,\mathrm{d}t.$$

Pour simplifier la preuve, on commence par remarquer que si on ajoute une constante c à F, les deux membres augmentent de $c\int_a^b g(t)\,\mathrm{d}t = c\,(\mathrm{G}(b) - \mathrm{G}(a))$. On peut donc se ramener au cas où $\mathrm{F}(a) = 0$, et de même on peut se ramener à $\mathrm{G}(b) = 0$. Sous ces deux hypothèses, on écrit

$$\int_a^b \mathrm{F}(t)g(t)\,\mathrm{d}t = \int_a^b \Bigl(\int_a^t f(s)\,\mathrm{d}s\Bigr) g(t)\,\mathrm{d}t = \int_a^b f(s)\Bigl(\int_s^b g(t)\,\mathrm{d}t\Bigr)\,\mathrm{d}s = -\int_a^b f(s)\mathrm{G}(s)\,\mathrm{d}s.$$

On dira qu'une fonction f continue sur $\mathbb{R}$ admet une *dérivée généralisée* g si g est une fonction mesurable, intégrable sur tout intervalle borné (on dit que f est *localement intégrable*), et si

$$\text{(DG)} \qquad f(t) - f(s) = \int_s^t g(u)\,\mathrm{d}u$$

pour tous réels $s < t$. La dérivée généralisée est unique (en tant que classe de fonctions définies Lebesgue-presque-partout) : si g_1 et g_2 étaient deux dérivées généralisées de f, on aurait $\int_s^t (g_2 - g_1) = 0$ pour tous $s < t$, ce qui entraîne que $d := g_2 - g_1$ est orthogonale à toutes les fonctions en escalier. Considérons un intervalle $[a, b]$ quelconque ; puisque les fonctions en escalier sont denses dans $\mathrm{L}_1(a, b)$, on peut trouver une fonction en escalier h telle que $\int_a^b |d - h| < \varepsilon$. Alors, puisque d est orthogonale à la fonction en escalier $\mathrm{sign}(h)$, on a

$$\int_a^b |h| = \int_a^b (h - d)\,\mathrm{sign}(h) \le \int_a^b |d - h| < \varepsilon$$

donc $\int_a^b |d| < 2\varepsilon$, pour tout $\varepsilon > 0$, et $d = 0$ presque partout, sur tout intervalle donc aussi sur $\mathbb{R}$, comme on le voulait.

Si φ est une fonction de l'espace $\mathcal{D}(\mathbb{R})$ des fonctions C^∞ à support compact, la formule (IPP), appliquée à un intervalle $[a, b]$ contenant le support de φ, nous donne si g est la dérivée généralisée de f

$$\int_{\mathbb{R}} f\varphi' = \int_a^b f(u)\varphi'(u)\,\mathrm{d}u = -\int_a^b g(u)\varphi(u)\,\mathrm{d}u = -\int_{\mathbb{R}} g\varphi\,;$$

le terme $\Bigl[f\varphi\Bigr]_a^b$ est nul car φ est nulle aux bornes de l'intervalle. Cette formule nous conduira un peu plus loin à une autre définition, celle de la *dérivée au sens des distributions*. Une *distribution* sur $\mathbb{R}$ est une forme linéaire sur $\mathcal{D}(\mathbb{R})$, possédant une certaine propriété de continuité([a]) que nous n'aurons pas à utiliser. Pour une fonction fixée f, localement intégrable sur $\mathbb{R}$, on peut considérer la forme linéaire T_f sur $\mathcal{D}(\mathbb{R})$ définie par

$$\forall\varphi \in \mathcal{D}(\mathbb{R}), \quad \mathrm{T}_f(\varphi) = \int_{\mathbb{R}} f\varphi.$$

Cette forme linéaire T_f est la distribution associée à la fonction f.

Lemme 5.1.1. *La correspondance $f \to \mathrm{T}_f$ est injective : si la forme linéaire T_f est nulle sur $\mathcal{D}(\mathbb{R})$, la fonction f est nulle presque-partout sur $\mathbb{R}$.*

Preuve. — Supposons que $\mathrm{T}_f = 0$, fixons une fonction $\theta \in \mathcal{D}(\mathbb{R})$; la fonction θf est dans $\mathrm{L}_1(\mathbb{R})$, donc il existe pour tout $\varepsilon > 0$ une fonction $\varphi \in \mathcal{D}(\mathbb{R})$ telle que

$$\int_{\mathbb{R}} |\varphi - \theta f| < \varepsilon\,;$$

pour tout $\alpha > 0$ la fonction $\psi_\alpha : x \to (\alpha^2 + \varphi(x)^2)^{-1/2}\varphi(x)$ est dans $\mathcal{D}(\mathbb{R})$; puisque $\mathrm{T}_f = 0$, on a donc

$$\int \frac{\varphi^2}{\sqrt{\alpha^2+\varphi^2}} = \int \frac{\varphi^2}{\sqrt{\alpha^2+\varphi^2}} - \mathrm{T}_f(\theta\psi_\alpha) = \int (\varphi - \theta f)\frac{\varphi}{\sqrt{\alpha^2+\varphi^2}} \le \int |\varphi - \theta f| < \varepsilon,$$

ce qui entraîne $\int |\varphi| \le \varepsilon$ en faisant tendre α vers 0, puis $\int |\theta f| < 2\varepsilon$, pour tout $\varepsilon > 0$. Il en résulte que $\theta f = 0$, pour toute θ, donc $f = 0$ (presque-partout).

///

Pour une fonction fixée f, localement intégrable sur $\mathbb{R}$, on peut considérer une autre forme linéaire ℓ sur $\mathcal{D}(\mathbb{R})$, définie par

$$\forall \varphi \in \mathcal{D}(\mathbb{R}), \quad \ell(\varphi) = -\int_{\mathbb{R}} f\varphi'.$$

C'est la *dérivée de f au sens des distributions* (plus généralement, on peut associer à toute distribution T une nouvelle distribution T′ définie par $\mathrm{T}'(\varphi) = -\mathrm{T}(\varphi')$; on vient de le faire pour $\mathrm{T} = \mathrm{T}_f$).

Lemme 5.1.2. *Si la dérivée distribution de f est représentée par une fonction localement intégrable g, la fonction f «est» continue et la fonction g est aussi la dérivée généralisée de f au sens de (DG).*

Preuve. — On a supposé qu'il existe une fonction g localement intégrable qui représente la dérivée distribution de f, c'est-à-dire telle que

$$\forall \varphi \in \mathcal{D}(\mathbb{R}), \quad \ell(\varphi) = -\int_{\mathbb{R}} f\varphi' = \int_{\mathbb{R}} g\varphi.$$

Définissons la fonction G par $\mathrm{G}(t) = \int_0^t g(u)\,\mathrm{d}u$. L'intégration par parties montre que

$$\forall \varphi \in \mathcal{D}(\mathbb{R}), \quad \int_{\mathbb{R}} f\varphi' = \int_{\mathbb{R}} \mathrm{G}\varphi'.$$

Autrement dit, la dérivée distribution de la fonction $f - \mathrm{G}$ est nulle. Si nous montrons que $f - \mathrm{G}$ est constante égale à c (presque partout), on aura bien que f admet un représentant continu, à savoir $\mathrm{G} + c$, et on écrira pour ce représentant

$$f(t) - f(s) = \mathrm{G}(t) - \mathrm{G}(s) = \int_s^t g(u)\,\mathrm{d}u$$

pour tous $s < t$, comme voulu.

Supposons donc que la dérivée distribution de h localement intégrable soit nulle, c'est-à-dire que T_h soit nulle sur le sous-espace de $\mathcal{D}$ formé de toutes les φ'. Il est facile de voir que ce sous-espace est égal à l'hyperplan de $\mathcal{D}$ formé des fonctions d'intégrale nulle([b]) ; si on fixe $\theta_0 \in \mathcal{D}(\mathbb{R})$, telle que $\int \theta_0 = 1$, on voit que $\varphi - (\int_{\mathbb{R}} \varphi)\, \theta_0$ est la dérivée d'une fonction $\psi \in \mathcal{D}(\mathbb{R})$, pour toute $\varphi \in \mathcal{D}(\mathbb{R})$, donc

$$\int_{\mathbb{R}} h\Bigl(\varphi - \Bigl(\int_{\mathbb{R}} \varphi\Bigr)\, \theta_0\Bigr) = \mathrm{T}_h(\psi') = 0,$$

c'est-à-dire que

$$\forall \varphi \in \mathcal{D}(\mathbb{R}), \quad \mathrm{T}_h(\varphi) = \int_{\mathbb{R}} h\varphi = \Bigl(\int_{\mathbb{R}} h\theta_0\Bigr) \int_{\mathbb{R}} \varphi.$$

On voit que T_h est représentée par la fonction constante égale à $\int_{\mathbb{R}} h\theta_0$; par le lemme d'unicité 5.1.1, on en déduit que h est égale à cette constante.

///

Remarque 5.1.3. Supposons que la forme linéaire ℓ, dérivée distribution de f, soit L_2-continue, c'est-à-dire continue quand on munit l'espace $\mathcal{D}(\mathbb{R})$ de la norme induite par $\mathrm{L}_2(\mathbb{R})$; comme $\mathcal{D}(\mathbb{R})$ est dense dans $\mathrm{L}_2(\mathbb{R})$, cette forme linéaire se prolonge alors en forme linéaire continue sur $\mathrm{L}_2(\mathbb{R})$. Il existe donc une fonction $g \in \mathrm{L}_2(\mathbb{R})$ qui représente cette forme linéaire continue, c'est-à-dire telle que

$$\forall \varphi \in \mathcal{D}(\mathbb{R}), \quad \ell(\varphi) = -\int_{\mathbb{R}} f\varphi' = \int_{\mathbb{R}} g\varphi.$$

D'après le lemme précédent, f « est » continue et g vérifie la relation (DG).

Un exemple de fermeture

Exemple. On va étudier en détail la fermeture de l'opérateur D_0 de l'exemple **A**, défini sur $\mathrm{L}_2(\mathbb{R})$. On va montrer que D_0 est fermable et identifier sa fermeture.

Supposons que (f, g) soit dans l'adhérence de $\mathrm{Gr}(\mathrm{D}_0)$; il existe une suite $(f_n) \subset \mathcal{D}(\mathbb{R})$ telle que $(f_n, f_n') \to (f, g)$ dans $\mathrm{L}_2 \times \mathrm{L}_2$, c'est-à-dire que $f_n \to f$ dans L_2 et $f_n' \to g$ dans L_2. Puisque la suite (f_n') converge, elle est bornée dans L_2 par une constante C, donc en appliquant Cauchy-Schwarz, on obtient

$$\forall t, u \in \mathbb{R}, \quad \bigl|f_n(u) - f_n(t)\bigr| = \Bigl|\int_t^u f_n'(s)\, \mathrm{d}s\Bigr| \le |u-t|^{1/2}\, \|f_n'\|_2 \le \mathrm{C}\, |u-t|^{1/2}.$$

La suite des fonctions $(f_n - f_n(0))$ est donc équicontinue, et uniformément bornée sur tout compact de $\mathbb{R}$; par Ascoli, on peut supposer, quitte à passer à une sous-suite, que la suite de fonctions $(f_n - f_n(0))$ converge uniformément sur tout compact vers une fonction F continue sur $\mathbb{R}$. En intégrant sur $(0, 1)$ on déduit de la convergence uniforme que

$$\int_0^1 f_n(s)\, \mathrm{d}s - f_n(0) \to \int_0^1 \mathrm{F}(s)\, \mathrm{d}s,$$

mais on déduit de la convergence L_2 que

$$\int_0^1 f_n(s)\,\mathrm{d}s \to \int_0^1 f(s)\,\mathrm{d}s,$$

ce qui implique que $\lim f_n(0) = c$ existe ; la suite (f_n) tend donc uniformément vers $\mathrm{F}+c$ sur tout compact, et vers f dans $\mathrm{L}_2(\mathbb{R})$. Il en résulte que $f = \mathrm{F} + c$ (comme élément de l'espace $\mathrm{L}_2(\mathbb{R})$). On peut donc choisir la fonction continue $\mathrm{F} + c$ comme représentant de f : on dira que f «est» continue. Par les convergences établies ci-dessus, on a

$$\forall t, u \in \mathbb{R}, \quad f(u) - f(t) = \lim_n \int_t^u f_n'(s)\,\mathrm{d}s = \int_t^u g(s)\,\mathrm{d}s.$$

On voit donc que g est la dérivée généralisée de f, et $g \in \mathrm{L}_2(\mathbb{R})$. On introduit l'ensemble

$$\mathrm{G}_1 = \bigl\{(f, g) \in \mathrm{L}_2(\mathbb{R}) \times \mathrm{L}_2(\mathbb{R}) : \forall t < u, \quad f(u) = f(t) + \int_t^u g(s)\,\mathrm{d}s\bigr\}$$

(pour être vraiment correct, on devrait dire : l'ensemble des couples (f, g) tels que la classe f admette un représentant $\tilde{f}$ pour lequel, pour tous $t < u$, on ait $\tilde{f}(u) =\ldots$). On vient de montrer que l'adhérence de $\mathrm{Gr}(\mathrm{D}_0)$ est contenue dans G_1 ; pour savoir que D_0 est fermable, il suffit de voir que G_1 est un graphe : c'est clairement un espace vectoriel, et si $(0, g) \in \mathrm{G}_1$, g est la dérivée généralisée de 0, donc $g = 0$ par l'unicité de la dérivée généralisée.

On va montrer que l'adhérence du graphe de D_0 est *égale* à G_1. Si $(f, g) \in \mathrm{G}_1$, la fonction g est la dérivée généralisée de f et on a vu que pour toute $\varphi \in \mathcal{D}(\mathbb{R})$

$$\text{(DG1)} \qquad \int_{\mathbb{R}} f\varphi' = -\int_{\mathbb{R}} g\varphi.$$

Désignons par G_2 l'ensemble des couples $(f, g) \in \mathrm{L}_2(\mathbb{R}) \times \mathrm{L}_2(\mathbb{R})$ tels que l'égalité ci-dessus soit vraie pour toute fonction $\varphi \in \mathcal{D}(\mathbb{R})$. On vient d'indiquer que $\mathrm{G}_1 \subset \mathrm{G}_2$; il est clair que G_2 est fermé dans $\mathrm{L}_2(\mathbb{R}) \times \mathrm{L}_2(\mathbb{R})$, comme intersection de fermés. On va voir que G_2 est contenu dans l'adhérence de $\mathrm{Gr}(\mathrm{D}_0)$, par la méthode usuelle de régularisation-troncature. Soit (θ_n) une suite de fonctions de $\mathcal{D}(\mathbb{R})$, positives et d'intégrale 1, dont les supports tendent vers $\{0\}$; on peut obtenir une telle suite en posant

$$\forall t \in \mathbb{R}, \quad \theta_n(t) = 2^n \theta_0(2^n t)\,;$$

on sait, par la théorie de la convolution, que $\theta_n * f$ tend vers f dans $\mathrm{L}_2(\mathbb{R})$, pour toute $f \in \mathrm{L}_2(\mathbb{R})$; de plus $\theta_n * f$ est C^∞ et sa dérivée est égale à $\theta_n' * f$. Si $(f, g) \in \mathrm{G}_2$, on va vérifier que $\theta_n * g$ est aussi la dérivée de $\theta_n * f$. Par intégration par parties et Fubini, on voit que

$$-\int_{\mathbb{R}} (\theta_n * f)'\varphi = \int_{\mathbb{R}} (\theta_n * f)\varphi' = -\int_{\mathbb{R}} (\theta_n * g)\varphi$$

pour toute $\varphi \in \mathcal{D}(\mathbb{R})$. En effet,

$$\int_{\mathbb{R}} \Bigl(\int_{\mathbb{R}} \theta_n(s) f(t-s)\,\mathrm{d}s\Bigr)\varphi'(t)\,\mathrm{d}t = \int_{\mathbb{R}} \Bigl(\int_{\mathbb{R}} f(t-s)\varphi'(t)\,\mathrm{d}t\Bigr)\theta_n(s)\,\mathrm{d}s$$

$$= \int_{\mathbb{R}} \Big(\int_{\mathbb{R}} f(u)\varphi'(u+s)\,\mathrm{d}u \Big) \theta_n(s)\,\mathrm{d}s = -\int_{\mathbb{R}} \Big(\int_{\mathbb{R}} g(u)\varphi(u+s)\,\mathrm{d}u \Big) \theta_n(s)\,\mathrm{d}s$$

et on termine en refaisant le même chemin à l'envers. Par la densité de $\mathcal{D}(\mathbb{R})$ dans $\mathrm{L}_2(\mathbb{R})$, il en résulte que $(\theta_n * f)' = \theta_n * g$. Le couple $(\theta_n * f, \theta_n * g)$ tend vers (f, g) dans $\mathrm{L}_2 \times \mathrm{L}_2$. On a donc pu trouver un couple de fonctions C^∞, de la forme (f_1, f_1') avec $f_1 = \theta_n * f$ pour un n grand, qui est proche de (f, g). Posons maintenant $\chi_n(x) = \chi(x/n)$, où $\chi \in \mathcal{D}(\mathbb{R})$ est égale à 1 dans un voisinage de 0. La fonction $h_n = \chi_n(x) f_1(x)$ est C^∞ à support compact, donc dans le domaine de D_0, et on va voir que le couple (h_n, h_n') tend vers (f_1, f_1') en norme L_2. C'est facile pour $h_n \to f_1$ (convergence dominée). Pour les dérivées, on a

$$h_n' = \frac{1}{n}\,\chi'\Big(\frac{x}{n}\Big) f_1(x) + \chi_n(x) f_1'(x),$$

et ça marche car $\chi_n f_1'$ tend vers f_1' dans $\mathrm{L}_2(\mathbb{R})$ pour la même raison de convergence dominée, tandis que

$$\int_{\mathbb{R}} \Big| \frac{1}{n}\,\chi'\Big(\frac{x}{n}\Big) f_1(x) \Big|^2 \mathrm{d}x \le \frac{1}{n^2}\, \|\chi'\|_\infty^2 \, \|f_1\|_2^2 \to 0.$$

On a ainsi montré en deux temps (régularisation, troncature) que le couple $(f, g) \in \mathrm{G}_2$ est dans l'adhérence du graphe de D_0 ; finalement, on a établi que

$$\overline{\mathrm{Gr}(\mathrm{D}_0)} \subset \mathrm{G}_1 \subset \mathrm{G}_2 \subset \overline{\mathrm{Gr}(\mathrm{D}_0)},$$

donc les trois ensembles sont égaux.

On appelle $\mathrm{H}^1(\mathbb{R})$ (espace de Sobolev) l'espace des fonctions $f \in \mathrm{L}_2(\mathbb{R})$ qui admettent une dérivée généralisée $g \in \mathrm{L}_2(\mathbb{R})$; on a vu au lemme 5.1.2 qu'il revient au même de savoir que la dérivée distribution est représentée par une fonction de $\mathrm{L}_2(\mathbb{R})$. Désormais on notera simplement la *dérivée généralisée* de f par $g = f'$. La fermeture de l'opérateur D_0 de l'exemple **A** est donc l'opérateur $\mathrm{D} = \overline{\mathrm{D}}_0$ de $\mathrm{L}_2(\mathbb{R})$ dans lui-même dont le domaine est $\mathrm{H}^1(\mathbb{R})$ et qui est défini par $\mathrm{D}f = f'$ pour $f \in \mathrm{H}^1(\mathbb{R})$.

On définit aussi l'espace $\mathrm{H}^1([0,1])$ des fonctions $f \in \mathrm{L}_2([0,1])$ (en fait f sera continue) pour lesquelles existe une fonction $g \in \mathrm{L}_2([0,1])$ telle que

$$\text{(DG2)} \qquad f(t) = f(0) + \int_0^t g(s)\,\mathrm{d}s,$$

pour tout $t \in [0,1]$. Si on se rappelle l'opérateur-exemple P de $\mathrm{L}_2([0,1])$ dans lui-même qui associe à chaque $g \in \mathrm{L}_2([0,1])$ sa « primitive » nulle en zéro, on voit que $\mathrm{H}^1([0,1])$ est égal à $\mathrm{im}(\mathrm{P}) + \mathbb{K}\mathbf{1}$.

Si f est dans $\mathrm{H}^1([0,1])$, la fonction g vérifiant (DG2) est unique (par la densité dans L_2 des fonctions en escalier). Il est commode de noter f' cette unique fonction g, en l'appelant dérivée *généralisée* de f, pour ne pas oublier que ce n'est pas une dérivée au sens ordinaire.

Continuons sur la notion de dérivée généralisée. C'est la propriété (DG1) qui permet d'étendre la définition de H^1 au cas de plusieurs dimensions. Par exemple, on dit que

$f \in \mathrm{H}^1(\mathbb{R}^2)$ si $f \in \mathrm{L}_2(\mathbb{R}^2)$ et s'il existe deux fonctions $g_1, g_2 \in \mathrm{L}_2(\mathbb{R}^2)$ qui seront les *dérivées partielles faibles de f au sens des distributions*, ce qui signifie que

$$\int_{\mathbb{R}^2} f(x) \frac{\partial \varphi}{\partial x_j}(x) \,\mathrm{d}x = -\int_{\mathbb{R}^2} g_j(x) \varphi(x) \,\mathrm{d}x$$

pour $j = 1, 2$ et pour toute fonction $\varphi \in \mathcal{D}(\mathbb{R}^2)$. Les fonctions de cet espace $\mathrm{H}^1(\mathbb{R}^2)$ ne sont plus nécessairement continues, ni même bornées sur les compacts de $\mathbb{R}^2$. Si Ω est un ouvert de $\mathbb{R}^2$, on définit de la même façon un espace $\mathrm{H}^1(\Omega)$, où f et les g_j sont dans $\mathrm{L}_2(\Omega)$ et où les fonctions test φ sont maintenant limitées à l'espace $\mathcal{D}(\Omega)$ des fonctions C^∞ à support compact dans Ω.

On est conduit naturellement à ces espaces de Sobolev si on généralise l'exemple D_0 : définissons un opérateur ∇_0 de $\mathrm{X} = \mathrm{L}_2(\mathbb{R}^2)$ dans l'espace de Hilbert $\mathrm{Y} = \mathrm{L}_2(\mathbb{R}^2, d\lambda, \mathbb{K}^2)$ des fonctions définies sur $\mathbb{R}^2$, à valeurs dans l'espace vectoriel $\mathbb{K}^2$ (en général $\mathbb{R}^2$, mais ça peut être $\mathbb{C}^2$) de la façon suivante. Le domaine de ∇_0 est $\mathcal{D}(\mathbb{R}^2)$ (fonctions à valeurs réelles ou complexes). Si $\varphi \in \mathrm{dom}(\nabla_0)$, on définit $\nabla_0 \varphi \in \mathrm{Y}$ comme la fonction vectorielle

$$t \in \mathbb{R}^2 \to (\nabla \varphi)(t) \in \mathbb{K}^2.$$

Si on détermine la fermeture de ∇_0 comme on l'a fait ci-dessus pour D_0, on trouve que le domaine de la fermeture $\nabla = \overline{\nabla}_0$ est $\mathrm{H}^1(\mathbb{R}^2)$, l'opérateur ∇ associant à chaque fonction $f \in \mathrm{H}^1(\mathbb{R}^2)$ son gradient généralisé, dont les deux composantes sont les dérivées partielles généralisées. Si on effectue le même travail sur un ouvert Ω, on est conduit à l'espace $\mathrm{H}^1_0(\Omega)$, sous-espace de $\mathrm{H}^1(\Omega)$ dont on reparlera plus loin.

5.2. Spectre des opérateurs fermés

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

Répétons pour enfoncer le clou : la valeur λ est régulière pour T s'il existe un opérateur borné $\mathrm{S} \in \mathcal{L}(\mathrm{X})$ qui vérifie les deux propriétés suivantes :

– pour tout $x \in \mathrm{dom}(\mathrm{T})$, on a $x = \mathrm{S}(\mathrm{T}x - \lambda x)$;
– on a $\mathrm{S}(\mathrm{X}) \subset \mathrm{dom}(\mathrm{T})$, et pour tout $y \in \mathrm{X}$, on a $y = (\mathrm{T} - \lambda \,\mathrm{Id}_\mathrm{X})\mathrm{S}y$.

Soit T un opérateur sur un espace de Banach complexe X ; désignons par Ω_T l'ensemble des $\lambda \in \mathbb{C}$ qui sont valeur régulière de T ; pour $\lambda \in \Omega_\mathrm{T}$, on pose

$$\mathrm{R}_\lambda(\mathrm{T}) = (\lambda \,\mathrm{Id}_\mathrm{X} - \mathrm{T})^{-1} \in \mathcal{L}(\mathrm{X})$$

et on appelle $\mathrm{R}_\lambda(\mathrm{T})$ la *résolvante* de T.

Seuls les opérateurs fermés sont intéressants pour la théorie spectrale : en effet, si T admet une valeur régulière λ, l'opérateur $(\lambda \,\mathrm{Id}_\mathrm{X} - \mathrm{T})^{-1}$ est continu donc à graphe fermé ; on en déduit que son inverse $\lambda \,\mathrm{Id}_\mathrm{X} - \mathrm{T}$ est fermé, et il en résulte facilement que T lui-même est fermé. Autrement dit : si T n'est pas fermé, T n'admet aucune valeur régulière, donc on a toujours $\sigma(\mathrm{T}) = \mathbb{C}$.

Soit T un opérateur fermé d'un espace de Banach X dans lui-même ; remarquons que pour tout $\lambda \in \mathbb{C}$, l'opérateur $\lambda\,\mathrm{Id}_\mathrm{X} - \mathrm{T}$ est fermé. Si $\lambda\,\mathrm{Id}_\mathrm{X} - \mathrm{T}$ est bijectif de dom(T) sur X, alors λ est une valeur régulière car $(\lambda\,\mathrm{Id}_\mathrm{X} - \mathrm{T})^{-1}$ est fermé, défini sur un espace de Banach, donc continu par le théorème du graphe fermé 9.4.

Remarque 5.2.1. Dans le cas d'un opérateur borné $\mathrm{T} \in \mathcal{L}(\mathrm{X})$, on sait que T commute avec $\mathrm{R}_\lambda(\mathrm{T})$ pour tout $\lambda \notin \sigma(\mathrm{T})$; ici, la situation est un peu plus délicate, car $\mathrm{R}_\lambda(\mathrm{T})\mathrm{T}$ est *a priori* un opérateur non borné dont le domaine est dom(T) alors que $\mathrm{T}\mathrm{R}_\lambda(\mathrm{T})$ est défini sur tout l'espace X. En fait le deuxième opérateur est une extension du premier,

$$\mathrm{R}_\lambda(\mathrm{T})\mathrm{T} \subset \mathrm{T}\mathrm{R}_\lambda(\mathrm{T})\,;$$

en effet, et en posant $\mathrm{R} = \mathrm{R}_\lambda(\mathrm{T})$ pour raccourcir, pour tout vecteur $x \in \mathrm{dom(T)}$ on a $x = \mathrm{R}(\lambda x - \mathrm{T}x)$ et pour tout $y \in \mathrm{X}$, on a $y = (\lambda\,\mathrm{Id}_\mathrm{X} - \mathrm{T})\mathrm{R}y$, donc pour $x = y \in \mathrm{dom(T)}$,

$$\lambda\mathrm{R}x - \mathrm{RT}x = x = \lambda\mathrm{R}x - \mathrm{TR}x\,;$$

par conséquent, on a bien $\mathrm{RT}x = \mathrm{TR}x$. De plus, comme $y = \lambda\mathrm{R}y - \mathrm{TR}y$ pour tout $y \in \mathrm{X}$, on voit que l'opérateur $\mathrm{T}\mathrm{R}_\lambda(\mathrm{T}) = \mathrm{TR} = \lambda\mathrm{R}_\lambda(\mathrm{T}) - \mathrm{Id}_\mathrm{X}$ est borné sur X (comme on pouvait le suspecter à partir du moment où cet opérateur $\mathrm{T}\mathrm{R}_\lambda(\mathrm{T})$ était défini sur l'espace de Banach X tout entier).

Exemples.

1. Considérons l'opérateur M de multiplication par la fonction $t \to t$ sur $\mathrm{L}_2(\mathbb{R})$, défini sur le domaine

$$\mathrm{dom(M)} = \{f \in \mathrm{L}_2(\mathbb{R}) : \int_\mathbb{R} |t|^2\, |f(t)|^2 \,\mathrm{d}t < +\infty\},$$

qui agit sur $f \in \mathrm{dom(M)}$ par $(\mathrm{M}f)(t) = tf(t)$; on a bien alors $\mathrm{M}f \in \mathrm{L}_2(\mathbb{R})$. On peut décrire l'appartenance de f au domaine dom(M) en une seule formule,

$$\int_\mathbb{R} (1 + |t|^2)\, |f(t)|^2 \,\mathrm{d}t < +\infty.$$

On montre assez facilement que M est fermé en utilisant les outils de la théorie de l'intégration. Cherchons le spectre de M ; on suppose d'abord que $\lambda \in \mathbb{R}$; soit $\mathrm{B} = \mathrm{B}(\lambda, \varepsilon)$, $\varepsilon > 0$; on peut considérer la fonction $f = \mathbf{1}_\mathrm{B}$, qui est dans dom(M), et qui est non nulle dans $\mathrm{L}_2(\mathbb{R})$. On a

$$|\lambda f - \mathrm{M}f| = |\lambda - t|\,\mathbf{1}_\mathrm{B} \le \varepsilon\,\mathbf{1}_\mathrm{B}$$

car $\mathbf{1}_\mathrm{B}$ est nulle là où $|t - \lambda| > \varepsilon$. Ceci montre que $\|\lambda f - \mathrm{M}f\|_2 \le \varepsilon\,\|f\|_2$; si l'inverse $\mathrm{R}_\lambda(\mathrm{M})$ de $\lambda\,\mathrm{Id} - \mathrm{M}$ existait, il devrait vérifier $\|\mathrm{R}_\lambda(\mathrm{M})\| \ge 1/\varepsilon$, pour tout $\varepsilon > 0$, ce qui est impossible. Il en résulte que $\lambda \in \sigma(\mathrm{M})$.

On suppose inversement que $\lambda \notin \mathbb{R}$. Considérons la fonction continue g définie sur $\mathbb{R}$ par $g(t) = (\lambda - t)^{-1}$; elle est bornée sur $\mathbb{R}$ par $|\operatorname{Im}\lambda|^{-1}$. La multiplication M_g est bornée sur $\mathrm{L}_2(\mathbb{R})$ puisque g est bornée, et on va voir que $\mathrm{M}_g = \mathrm{R}_\lambda(\mathrm{M})$. Si $f \in \mathrm{dom(M)}$, on voit que $\mathrm{M}_g(\lambda f - \mathrm{M}f) = g(\lambda - t)f$ est égale à f ; on a bien $\mathrm{M}_g(\lambda f - \mathrm{M}f) = f$ en tant que classe. Inversement, si $h \in \mathrm{L}_2(\mu)$, on vérifie que $\mathrm{M}_g(h) \in \mathrm{dom(M)}$; en effet,

$$\int_\mathbb{R} |t|^2\, |(\mathrm{M}_g h)(t)|^2 \,\mathrm{d}t = \int_\mathbb{R} |t|^2\, |g(t)h(t)|^2 \,\mathrm{d}t = \int_\mathbb{R} |tg(t)|^2\, |h(t)|^2 \,\mathrm{d}t < +\infty$$

parce que $tg(t)$ est bornée sur $\mathbb{R}$; ensuite on a bien $(\lambda\,\mathrm{Id} - \mathrm{M})(\mathrm{M}_g(h)) = h$. On a ainsi montré que $\mathrm{M}_g = \mathrm{R}_\lambda(\mathrm{M})$.

En bref, le spectre de M est exactement l'ensemble des $\lambda \in \mathbb{R}$: on a $\sigma(\mathrm{M}) = \mathbb{R}$.

2. Nous allons montrer maintenant que le spectre de l'opérateur $\mathrm{T} = \mathrm{P}^{-1}$ de l'exemple **B** est vide : évidemment, 0 est valeur régulière de T et $\mathrm{R}_0(\mathrm{T}) = -\mathrm{P}$. Pour $\lambda \neq 0$, cherchons à résoudre l'équation $\lambda x - \mathrm{T}x = y$, pour $y \in \mathrm{X}$ donné (on cherche $x \in \mathrm{dom}(\mathrm{T}) = \mathrm{im}(\mathrm{P})$). Puisque l'opérateur T est surjectif, on peut écrire $y = \mathrm{T}z$, avec $z = \mathrm{P}y \in \mathrm{dom}(\mathrm{T})$. En appliquant P on trouve $\lambda \mathrm{P}x - x = z$, soit $\lambda^{-1}x - \mathrm{P}x = -\lambda^{-1}z$. On sait que λ^{-1} n'est pas dans le spectre de P (qui est réduit à $\{0\}$) donc on peut résoudre,

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

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

3. Opérateur diagonal. Pour toute suite scalaire $(\mu_n)_{n\geq 0}$, on définit un opérateur (en général non borné) sur $\ell_2(\mathbb{N})$ dont le domaine est l'espace vectoriel

$$\mathrm{E} = \{x \in \ell_2 : \sum |\mu_n x_n|^2 < +\infty\}$$

et qui est défini pour $x \in \mathrm{E}$ par $(\mathrm{T}x)_n = \mu_n x_n$. On montre facilement que le spectre de T est l'adhérence dans $\mathbb{C}$ de l'ensemble des valeurs $(\mu_n)_{n\geq 0}$. Comme toute partie fermée non vide F de $\mathbb{C}$ admet une suite dense, on trouve que pour toute partie fermée non vide F de $\mathbb{C}$, on peut construire un opérateur T d'un espace de Hilbert H dont le spectre $\sigma(\mathrm{T})$ soit égal à F. L'opérateur $\mathrm{T} = \mathrm{P}^{-1}$ fournit un cas où $\sigma(\mathrm{T}) = \emptyset$.

4. Considérons l'opérateur D de dérivation sur $\mathrm{L}_2(\mathbb{R})$, défini sur le domaine $\mathrm{H}^1(\mathbb{R})$ par $\mathrm{D}f = f'$ (dérivée généralisée) ; on voit d'abord que si ω est réel, la valeur imaginaire pure $\lambda = \mathrm{i}\omega$ n'est pas régulière, car on va trouver une suite $(f_n) \subset \mathrm{dom}(\mathrm{D})$ telle que $\|f_n\| = 1$ et

$$(\lambda\,\mathrm{Id} - \mathrm{D})f_n = \lambda f_n - f_n' \to 0.$$

À cet effet choisissons $\theta \in \mathcal{D}(\mathbb{R})$ telle que $\|\theta\|_2 = 1$, puis posons

$$f_n(t) = \frac{1}{\sqrt{n}}\,\theta\Big(\frac{t}{n}\Big)\,\mathrm{e}^{\mathrm{i}\omega t}.$$

On vérifie que $\|f_n\|_2 = \|\theta\|_2 = 1$, et

$$\lambda f_n(t) - f_n'(t) = -\frac{1}{n}\,\frac{1}{\sqrt{n}}\,\theta'\Big(\frac{t}{n}\Big)\,\mathrm{e}^{\mathrm{i}\omega t}$$

dont la norme est $n^{-1}\|\theta'\|_2$, qui tend vers 0. Il ne peut donc pas exister d'inverse borné pour $\lambda\,\mathrm{Id} - \mathrm{D}$.

Pour montrer que les valeurs $\lambda \notin \mathrm{i}\mathbb{R}$ sont régulières, on divise en deux cas : si $\lambda = a + \mathrm{i}b$ avec a, b réels et $a > 0$, on montre([c]) que

$$(\mathrm{R}_\lambda g)(t) = \mathrm{e}^{\lambda t}\int_t^{+\infty} \mathrm{e}^{-\lambda s}\,g(s)\,\mathrm{d}s.$$

Dans le cas $a < 0$, on intègre depuis $-\infty$. On conclut que $\sigma(\mathrm{D}) = \mathrm{i}\mathbb{R}$.

Le raisonnement utilisé dans l'exemple 2 précédent montre que

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

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

Preuve.— On veut résoudre pour tout $y \in \mathrm{X}$ l'équation

$$(\lambda^{-1}\, \mathrm{Id}_\mathrm{X} - \mathrm{T}^{-1})x = y. \tag{1}$$

Il est très facile d'identifier x si on suppose que $y \in \mathrm{dom}(\mathrm{T})$: dans ce cas on obtient $\mathrm{T}(x - \lambda\mathrm{T}^{-1}x) = \lambda\mathrm{T}y$ en appliquant l'opérateur $\lambda\mathrm{T}$ aux deux membres de l'équation, ce qui fournit tout de suite $\mathrm{T}x - \lambda x = \lambda\mathrm{T}y$, donc $x = -\lambda\mathrm{R}_\lambda(\mathrm{T})\mathrm{T}y$.

Le cas général où on suppose seulement $y \in \mathrm{X}$ est un peu plus délicat (voir la remarque 5.2.1) ; on cherche $x \in \mathrm{dom}(\mathrm{T}^{-1}) = \mathrm{im}(\mathrm{T})$. Puisque λ est une valeur régulière pour l'opérateur T, on peut écrire $y = (\lambda\, \mathrm{Id}_\mathrm{X} - \mathrm{T})z$, avec $z = \mathrm{R}_\lambda(\mathrm{T})y$. On sait alors que $z \in \mathrm{dom}(\mathrm{T}) = \mathrm{im}(\mathrm{T}^{-1})$, donc il existe $u \in \mathrm{dom}(\mathrm{T}^{-1})$ tel que $z = \mathrm{T}^{-1}u$. L'équation proposée est donc

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

Il en résulte que $x_0 = -\lambda u$ convient. Par ailleurs, la solution x_0 de l'équation (1) est unique, car $\lambda^{-1}\, \mathrm{Id}_\mathrm{X} - \mathrm{T}^{-1}$ est injectif : si $v \in \mathrm{dom}(\mathrm{T}^{-1})$ et $\lambda^{-1}v - \mathrm{T}^{-1}v = 0$, alors $v = \lambda\mathrm{T}^{-1}v \in \mathrm{im}(\mathrm{T}^{-1}) = \mathrm{dom}(\mathrm{T})$; en appliquant T, on obtient $\mathrm{T}v = \lambda v$, qui implique $v = 0$ puisque λ est régulière pour T. On a donc trouvé l'unique solution de l'équation (1),

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

qui est bien dans $\mathrm{dom}(\mathrm{T}^{-1})$ (car u y est). On a ainsi montré que $\lambda^{-1}\, \mathrm{Id}_\mathrm{X} - \mathrm{T}^{-1}$ est une bijection de $\mathrm{dom}(\mathrm{T}^{-1})$ sur X, dont l'inverse est $-\lambda\mathrm{T}\, \mathrm{R}_\lambda(\mathrm{T})$; pour finir, on a vu à la remarque 5.2.1 que $\mathrm{T}\, \mathrm{R}_\lambda(\mathrm{T})$ est borné, car égal à $\lambda\mathrm{R}_\lambda(\mathrm{T}) - \mathrm{Id}_\mathrm{X}$.

///

Proposition. *Le spectre d'un opérateur fermé* T *d'un espace de Banach complexe* X *dans lui-même est une partie fermée de* $\mathbb{C}$*, et l'application* $\lambda \to \mathrm{R}_\lambda(\mathrm{T})$ *est analytique sur le complémentaire du spectre de* T*, à valeurs dans* $\mathcal{L}(\mathrm{X})$*.*

Preuve.— Soit λ_0 une valeur régulière pour T ; en introduisant $\mathrm{T}_0 = \mathrm{T} - \lambda_0\, \mathrm{Id}_\mathrm{X}$ on se ramène à étudier la situation au voisinage de $\lambda = 0$. L'opérateur T_0 est une bijection de $\mathrm{dom}(\mathrm{T})$ sur X, d'inverse $\mathrm{S} = -\mathrm{R}_{\lambda_0}(\mathrm{T}) \in \mathcal{L}(\mathrm{X})$; on veut montrer qu'il existe $\varepsilon_0 > 0$ tel que ε soit valeur régulière de T_0 dès que $|\varepsilon| < \varepsilon_0$. L'opérateur borné S vérifie les deux propriétés suivantes :

– pour tout $y \in \mathrm{X}$, on a $\mathrm{S}y \in \mathrm{dom}(\mathrm{T}_0) = \mathrm{dom}(\mathrm{T})$ et $\mathrm{T}_0\mathrm{S}y = y$;

– pour tout $x \in \mathrm{dom}(\mathrm{T})$, on a $\mathrm{S}\mathrm{T}_0x = x$.

Considérons l'opérateur borné $\mathrm{V}_\varepsilon = (\mathrm{Id}_\mathrm{X} - \varepsilon\mathrm{S})^{-1} \in \mathcal{L}(\mathrm{X})$, qui est certainement défini quand $|\varepsilon| < \varepsilon_0 = \|\mathrm{S}\|^{-1}$. On va voir que l'opérateur $\mathrm{S}_\varepsilon = \mathrm{S}\mathrm{V}_\varepsilon = \mathrm{V}_\varepsilon\mathrm{S}$ convient comme inverse borné de $\mathrm{T}_0 - \varepsilon\, \mathrm{Id}_\mathrm{X}$. Si $y \in \mathrm{X}$, le vecteur $\mathrm{S}_\varepsilon y = \mathrm{S}\mathrm{V}_\varepsilon y$ est dans l'image de S, donc dans $\mathrm{dom}(\mathrm{T})$, et

$$(\mathrm{T}_0 - \varepsilon\, \mathrm{Id}_\mathrm{X})\mathrm{S}\mathrm{V}_\varepsilon y = \mathrm{T}_0\mathrm{S}\mathrm{V}_\varepsilon y - \varepsilon\mathrm{S}\mathrm{V}_\varepsilon = \mathrm{V}_\varepsilon y - \varepsilon\mathrm{S}\mathrm{V}_\varepsilon y = (\mathrm{Id}_\mathrm{X} - \varepsilon\mathrm{S})\mathrm{V}_\varepsilon y = y\,;$$

si $x \in \mathrm{dom}(\mathrm{T})$, on a

$$\mathrm{S}_\varepsilon(\mathrm{T}_0 - \varepsilon\,\mathrm{Id}_\mathrm{X})x = \mathrm{V}_\varepsilon \mathrm{S}(\mathrm{T}_0 - \varepsilon\,\mathrm{Id}_\mathrm{X})x = \mathrm{V}_\varepsilon x - \varepsilon \mathrm{V}_\varepsilon \mathrm{S} x = \mathrm{V}_\varepsilon(\mathrm{Id}_\mathrm{X} - \varepsilon \mathrm{S})x = x$$

ce qui montre que $\mathrm{T}_0 - \varepsilon\,\mathrm{Id}_\mathrm{X}$ est bijective de $\mathrm{dom}(\mathrm{T})$ sur X, d'inverse $\mathrm{S}_\varepsilon = -\mathrm{R}_\varepsilon(\mathrm{T}_0)$. En développant V_ε au moyen de la série géométrique usuelle, on voit que

$$-\mathrm{R}_\varepsilon(\mathrm{T}_0) = \mathrm{S} + \varepsilon \mathrm{S}^2 + \varepsilon^2 \mathrm{S}^3 + \cdots \tag{2}$$

L'écriture (2) montre que $\varepsilon \to \mathrm{R}_\varepsilon(\mathrm{T}_0)$ est développable en série au voisinage de 0, c'est-à-dire que $\lambda \to \mathrm{R}_\lambda(\mathrm{T})$ est analytique dans l'ouvert Ω_T.

///

5.3. Adjoint hilbertien

Soient X et Y deux espaces de Banach et T un opérateur densément défini de X dans Y ; on définit *le transposé* de T, qui est un opérateur de Y^* dans X^*, de la façon suivante : le domaine de ${}^t\mathrm{T}$ est défini comme étant l'ensemble des $y^* \in \mathrm{Y}^*$ telles que la forme linéaire $y^* \circ \mathrm{T} : x \in \mathrm{dom}(\mathrm{T}) \to y^*(\mathrm{T}x)$ soit continue (en ayant muni l'espace vectoriel $\mathrm{dom}(\mathrm{T})$ de la norme induite par celle de X) : ceci signifie qu'il existe une constante C (dépendant de y^*) telle que

$$\forall x \in \mathrm{dom}(\mathrm{T}), \quad |y^*(\mathrm{T}x)| \le \mathrm{C}\,\|x\|.$$

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

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

pour tous $x \in \mathrm{dom}(\mathrm{T})$ et $y^* \in \mathrm{dom}({}^t\mathrm{T})$.

Lorsque X et Y sont deux espaces de Hilbert et T un opérateur densément défini de X dans Y, on définit un opérateur T^* de Y dans X de la façon suivante : on posera $\mathrm{T}^*y = x$ si la forme linéaire ℓ_y associée à $y \in \mathrm{H}$ est dans $\mathrm{dom}({}^t\mathrm{T})$, et si la forme linéaire $\ell_x = x^*$ est égale à ${}^t\mathrm{T}(\ell_y)$. Le vecteur y est donc dans le domaine de T^* si et seulement si la forme linéaire $\ell : u \in \mathrm{dom}(\mathrm{T}) \to \langle \mathrm{T}u, y\rangle$ est continue sur $\mathrm{dom}(\mathrm{T})$ (muni de la norme de X), c'est-à-dire qu'il existe une constante C (dépendant de y) telle que

$$\forall u \in \mathrm{dom}(\mathrm{T}), \quad |\langle \mathrm{T}u, y\rangle| \le \mathrm{C}\,\|u\|\,;$$

le couple $(y, x) \in \mathrm{Y} \times \mathrm{X}$ est dans le graphe de T^* si et seulement si

$$\langle \mathrm{T}u, y\rangle = \langle u, x\rangle \tag{3}$$

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

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

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

(3) définit un ensemble fermé de couples (y, x), ce qui montre que T^* est toujours un opérateur fermé.

Redisons les choses d'une autre façon, qui sera très utile plus loin. Sur l'espace $\mathrm{X} \times \mathrm{Y}$ on introduit le produit scalaire

$$\langle (x, y), (x', y') \rangle_{\mathrm{X} \times \mathrm{Y}} = \langle x, x' \rangle + \langle y, y' \rangle$$

et on procède de même sur $\mathrm{Y} \times \mathrm{X}$. Ces deux espaces produit deviennent ainsi deux espaces de Hilbert. Soit $\mathrm{U}_{\mathrm{X},\mathrm{Y}} \in \mathcal{L}(\mathrm{X} \times \mathrm{Y}, \mathrm{Y} \times \mathrm{X})$ l'opérateur unitaire qui à $(x, y) \in \mathrm{X} \times \mathrm{Y}$ associe $(-y, x)$; l'adjoint $\mathrm{U}^*_{\mathrm{X},\mathrm{Y}} = \mathrm{U}^{-1}_{\mathrm{X},\mathrm{Y}}$ vérifie $\mathrm{U}^*_{\mathrm{X},\mathrm{Y}}(y, x) = (x, -y)$, c'est-à-dire que $\mathrm{U}^*_{\mathrm{X},\mathrm{Y}} = -\mathrm{U}_{\mathrm{Y},\mathrm{X}}$. Le couple (y, x) est dans le graphe de T^* si et seulement si on a pour tout $z \in \mathrm{dom}(\mathrm{T})$

$$0 = \langle y, -\mathrm{T}z \rangle + \langle x, z \rangle = \langle (y, x), (-\mathrm{T}z, z) \rangle_{\mathrm{Y} \times \mathrm{X}},$$

ce qui signifie que (y, x) est orthogonal à toutes les images par $\mathrm{U}_{\mathrm{X},\mathrm{Y}}$ des points $(z, \mathrm{T}z)$ du graphe de T, c'est-à-dire que (y, x) est orthogonal au sous-espace $\mathrm{U}_{\mathrm{X},\mathrm{Y}}(\mathrm{Gr}(\mathrm{T}))$.

Le graphe $\mathrm{Gr}(\mathrm{T}^*)$ *de l'opérateur adjoint* T^* *est l'orthogonal de* $\mathrm{U}_{\mathrm{X},\mathrm{Y}}(\mathrm{Gr}(\mathrm{T}))$ *dans l'espace de Hilbert* $\mathrm{Y} \times \mathrm{X}$.

On voit à nouveau que $\mathrm{Gr}(\mathrm{T}^*)$ est fermé. Bien entendu, en modifiant notre interprétation,

$$0 = \langle -x, z \rangle + \langle y, \mathrm{T}z \rangle = \langle (-x, y), (z, \mathrm{T}z) \rangle_{\mathrm{X} \times \mathrm{Y}},$$

pour tout $z \in \mathrm{dom}(\mathrm{T})$, et il revient au même de dire que

$\mathrm{U}_{\mathrm{Y},\mathrm{X}}(\mathrm{Gr}(\mathrm{T}^*))$ *est l'orthogonal de* $\mathrm{Gr}(\mathrm{T})$ *dans l'espace de Hilbert* $\mathrm{X} \times \mathrm{Y}$.

Voici une première occasion d'utiliser cette approche.

Proposition. *Soient* X *et* Y *deux espaces de Hilbert et* T *un opérateur densément défini de* X *dans* Y *; si* T *est fermé, alors* T^* *est densément défini,* $(\mathrm{T}^*)^* = \mathrm{T}$ *et on a la décomposition orthogonale*

$$\mathrm{X} \times \mathrm{Y} = \mathrm{Gr}(\mathrm{T}) \otimes \mathrm{U}_{\mathrm{Y},\mathrm{X}}(\mathrm{Gr}(\mathrm{T}^*)).$$

Preuve. — Supposons T densément défini et fermé. Pour montrer que T^* est densément défini, on va montrer que $y = 0_{\mathrm{Y}}$ est le seul vecteur de Y orthogonal à $\mathrm{dom}(\mathrm{T}^*)$. Si y est orthogonal à $\mathrm{dom}(\mathrm{T}^*)$, le couple $(y, 0_{\mathrm{X}})$ est orthogonal à $\mathrm{Gr}(\mathrm{T}^*)$, donc $(0_{\mathrm{X}}, y)$ est orthogonal à $\mathrm{U}_{\mathrm{Y},\mathrm{X}}(\mathrm{Gr}(\mathrm{T}^*)) = \mathrm{Gr}(\mathrm{T})^\perp$,

$$(0_{\mathrm{X}}, y) \in \mathrm{Gr}(\mathrm{T})^{\perp\perp} ;$$

puisque T est fermé, $\mathrm{Gr}(\mathrm{T})$ est un sous-espace fermé, donc $\mathrm{Gr}(\mathrm{T}))^{\perp\perp} = \mathrm{Gr}(\mathrm{T})$; on obtient ainsi $(0_{\mathrm{X}}, y) \in \mathrm{Gr}(\mathrm{T})$, d'où $y = \mathrm{T}\,0_{\mathrm{X}} = 0_{\mathrm{Y}}$.

On sait qu'en général $\mathrm{U}_{\mathrm{Y},\mathrm{X}}(\mathrm{Gr}(\mathrm{T}^*))$ est l'orthogonal de $\mathrm{Gr}(\mathrm{T})$; si $\mathrm{Gr}(\mathrm{T})$ est fermé, on en déduit que $\mathrm{Gr}(\mathrm{T})$ est l'orthogonal de $\mathrm{U}_{\mathrm{Y},\mathrm{X}}(\mathrm{Gr}(\mathrm{T}^*))$; mais on a dit que le graphe de $(\mathrm{T}^*)^*$ est l'orthogonal de $\mathrm{U}_{\mathrm{Y},\mathrm{X}}(\mathrm{Gr}(\mathrm{T}^*))$. Il en résulte que $\mathrm{Gr}((\mathrm{T}^*)^*) = \mathrm{Gr}(\mathrm{T})$, ce qui contient toute l'information voulue pour conclure que $(\mathrm{T}^*)^* = \mathrm{T}$. On a de plus la décomposition orthogonale $\mathrm{H} = \mathrm{Z} \oplus \mathrm{Z}^\perp$, vraie pour tout sous-espace fermé Z d'un espace de Hilbert H.

///

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

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

///

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

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

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

Exemples.

1. On va vérifier que l'opérateur M est autoadjoint. On voit facilement que dom(M) est dense dans $\mathrm{L}_2(\mathbb{R})$ (parce que dom(M) contient toutes les fonctions de $\mathrm{L}_2(\mathbb{R})$ à support borné). Il est à peu près évident que M est symétrique,

$$\langle \mathrm{M}f, g\rangle = \int_{\mathbb{R}} \bigl(tf(t)\bigr)\,\overline{g(t)}\,\mathrm{d}t = \int_{\mathbb{R}} f(t)\,\overline{tg(t)}\,\mathrm{d}t = \langle f, \mathrm{M}g\rangle.$$

On en déduit $\mathrm{dom}(\mathrm{M}) \subset \mathrm{dom}(\mathrm{M}^*)$. Inversement, supposons que $g \in \mathrm{dom}(\mathrm{M}^*) \subset \mathrm{L}_2(\mathbb{R})$, et considérons pour tout $n \ge 0$ la fonction $f_n \in \mathrm{dom}(\mathrm{M})$ définie en posant pour presque tout $t \in \mathbb{R}$

$$f_n(t) = t\,\mathbf{1}_{[-n,n]}(t)g(t)\,;$$

puisque $g \in \mathrm{dom}(\mathrm{M}^*)$, il existe une constante C telle que pour tout $n \ge 0$, on ait $|\langle \mathrm{M}f_n, g\rangle| \le \mathrm{C}\,\|f_n\|_2$, ce qui donne

$$\int_{-n}^{n} t^2|g(t)|^2\,\mathrm{d}t \le \mathrm{C}\,\Bigl(\int_{-n}^{n} t^2|g(t)|^2\,\mathrm{d}t\Bigr)^{1/2}$$

d'où résulte que $\int_{\mathbb{R}} t^2|g(t)|^2\,\mathrm{d}t \le \mathrm{C}^2 < +\infty$, soit $g \in \mathrm{dom}(\mathrm{M})$. La vérification est finie.

2. On définit un opérateur $\mathrm{D}_{\mathbf{A}}$ sur $\mathrm{X} = \mathrm{L}_2(0,1)$ (D comme dérivation, **A** parce qu'il y aura des variations **B**, **C** de cet exemple), par son domaine

$$\mathrm{dom}(\mathrm{D}_{\mathbf{A}}) = \mathrm{H}^1_0 = \{f \in \mathrm{H}^1([0,1]) : f(0) = f(1) = 0\}$$

et on définit ensuite $\mathrm{D}_{\mathbf{A}} f = f'$ (la dérivée généralisée) pour toute $f \in \mathrm{dom}(\mathrm{D}_{\mathbf{A}})$. On vérifie d'abord que $\mathrm{H}^1_0 = \mathrm{dom}(\mathrm{D}_{\mathbf{A}})$ est dense dans X : c'est clair puisque H^1_0 contient $\mathcal{D}(]0,1[)$, qui est dense dans $\mathrm{L}_2(0,1)$; on note que si $g \in \mathrm{H}^1([0,1])$, $f \in \mathrm{H}^1_0$, on a en utilisant l'intégration par parties dans $\mathrm{H}^1([0,1])$

$$\langle g, \mathrm{D}_{\mathbf{A}} f\rangle = \int_0^1 g\,\overline{f'} = \bigl[g\,\overline{f}\bigr]_0^1 - \int_0^1 g'\,\overline{f} = -\int_0^1 g'\,\overline{f} = -\langle g', f\rangle$$

(le terme $\bigl[\,.\,\bigr]_0^1$ est nul parce que f est nulle aux bornes par définition de H^1_0). On a ainsi montré que le domaine de $\mathrm{D}^*_{\mathbf{A}}$ contient $\mathrm{H}^1([0,1])$, et que $\mathrm{D}^*_{\mathbf{A}} g = -g'$. Inversement, supposons que $g \in \mathrm{dom}(\mathrm{D}_{\mathbf{A}}^*)$. Dire que (g,h) est dans le graphe de $\mathrm{D}_{\mathbf{A}}^*$ signifie que $h = \mathrm{D}_{\mathbf{A}}^* g \in \mathrm{X}$ vérifie

$$\langle f, h\rangle = \langle \mathrm{D}_{\mathbf{A}} f, g\rangle$$

pour toute fonction $f \in \mathrm{H}_0^1$. On a donc si $(g,h) \in \mathrm{Gr}(\mathrm{D}_{\mathbf{A}}^*)$

$$\int_0^1 f\,\overline{h} = -\int_0^1 f'\,\overline{g} \tag{4}$$

pour toute $f \in \mathrm{H}_0^1$. Posons $\mathrm{H}(t) = \int_0^t h(s)\,\mathrm{d}s$. On obtient par intégration par parties

$$\int_0^1 f\,\overline{h} = \bigl[f\,\overline{\mathrm{H}}\bigr]_0^1 - \int_0^1 f'\,\overline{\mathrm{H}} = -\int_0^1 f'\,\overline{\mathrm{H}},$$

ce qui donne

$$\int_0^1 f'\,\overline{g} = \int_0^1 f'\,\overline{\mathrm{H}},$$

pour toute $f \in \mathrm{H}_0^1$. On remarque ([d]) que l'ensemble des f', lorsque $f \in \mathrm{H}_0^1$, est exactement l'ensemble de toutes les fonctions k de $\mathrm{X} = \mathrm{L}_2(0,1)$ qui sont d'intégrale nulle sur $[0,1]$. Cet ensemble des fonctions d'intégrale nulle est égal à $(\mathbb{C}\mathbf{1})^\perp$, et l'équation précédente indique que $\mathrm{H} - g$ est orthogonale à $(\mathbb{C}\mathbf{1})^\perp$, donc $\mathrm{H} - g \in (\mathbb{C}\mathbf{1})^{\perp\perp} = \mathbb{C}\mathbf{1}$. On obtient que $\mathrm{H} - g$ est une fonction constante, donc $g = \mathrm{H} + \mathrm{C}te$; comme H est une fonction de $\mathrm{H}^1([0,1])$, il en résulte que $g \in \mathrm{H}^1$. On a déjà vu que $\mathrm{H}^1 \subset \mathrm{dom}(\mathrm{D}_{\mathbf{A}}^*)$, et on a maintenant $\mathrm{dom}(\mathrm{D}_{\mathbf{A}}^*) \subset \mathrm{H}^1$, donc $\mathrm{dom}(\mathrm{D}_{\mathbf{A}}^*) = \mathrm{H}^1$ et pour $g \in \mathrm{dom}(\mathrm{D}_{\mathbf{A}}^*)$ on a $\mathrm{D}_{\mathbf{A}}^* g = -g'$.

On voit ainsi apparaître une variante $\mathrm{D}_{\mathbf{B}}$ de l'opérateur de dérivation, dont le domaine est

$$\mathrm{dom}(\mathrm{D}_{\mathbf{B}}) = \mathrm{H}^1([0,1])$$

et on définit ensuite $\mathrm{D}_{\mathbf{B}} f = f'$; on a obtenu que $\mathrm{D}_{\mathbf{A}}^* = -\mathrm{D}_{\mathbf{B}}$. D'après le résultat sur les double-adjoints, on a aussi $\mathrm{D}_{\mathbf{B}}^* = -\mathrm{D}_{\mathbf{A}}$. Si on travaille sur les complexes, on aura envie de récrire les formules sous la forme $(\mathrm{i}\mathrm{D}_{\mathbf{A}})^* = \mathrm{i}\mathrm{D}_{\mathbf{B}}$, $(\mathrm{i}\mathrm{D}_{\mathbf{B}})^* = \mathrm{i}\mathrm{D}_{\mathbf{A}}$, mais ceci ne donne pas encore un exemple de dérivation autoadjointe sur $(0,1)$. C'est l'objet du prochain exercice. L'exemple $\mathrm{i}\mathrm{D}_{\mathbf{A}}$ est un opérateur symétrique, mais non auto-adjoint. L'exemple $\mathrm{i}\mathrm{D}_{\mathbf{B}}$ n'est pas symétrique, puisque son adjoint $\mathrm{i}\mathrm{D}_{\mathbf{A}}$ a un domaine plus petit que le domaine de $\mathrm{i}\mathrm{D}_{\mathbf{B}}$.

Exercice.

a. On définit un opérateur $\mathrm{D}_{\mathbf{C}}$ sur $\mathrm{X} = \mathrm{L}_2(0,1)$ par son domaine

$$\mathrm{dom}(\mathrm{D}_{\mathbf{C}}) = \{f \in \mathrm{H}^1([0,1]) : f(0) = f(1)\}$$

et on définit ensuite $\mathrm{D}_{\mathbf{C}} f = f'$ (la dérivée généralisée) pour toute $f \in \mathrm{dom}(\mathrm{D}_{\mathbf{C}})$. Montrer que $\mathrm{i}\mathrm{D}_{\mathbf{C}}$ est auto-adjoint.

b. Soit D l'opérateur sur $\mathrm{L}_2(\mathbb{R})$, de domaine $\mathrm{H}^1(\mathbb{R})$, qui agit par $\mathrm{D}f = f'$ (dérivée généralisée). Montrer que $\mathrm{i}\mathrm{D}$ est un opérateur autoadjoint sur $\mathrm{L}_2(\mathbb{R})$.

Proposition. *Soient* H *un espace de Hilbert et* B *un opérateur borné, hermitien et injectif ; alors* B^{-1} *est autoadjoint.*

Preuve. — Le domaine de $\mathrm{T} = \mathrm{B}^{-1}$ est l'image $\mathrm{im}(\mathrm{B})$ de B ; comme B est hermitien injectif, cette image est dense ; en effet, si z est orthogonal à $\mathrm{im}(\mathrm{B})$, on a

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

pour tout $y \in \mathrm{H}$ donc $\mathrm{B}z = 0$, donc $z = 0$ puisque B est injectif. Il est évident que T est symétrique : si $x_1, x_2 \in \mathrm{im}(\mathrm{B})$, on peut écrire $x_j = \mathrm{B}y_j$, $j = 1, 2$ et

$$\langle \mathrm{T}x_1, x_2\rangle = \langle y_1, \mathrm{B}y_2\rangle = \langle \mathrm{B}y_1, y_2\rangle = \langle x_1, \mathrm{T}x_2\rangle.$$

Il reste à montrer que le domaine de T^* n'est pas plus grand que $\mathrm{im}(\mathrm{B})$; si u est dans le domaine de T^*, il existe un vecteur $v \in \mathrm{H}$ tel que

$$\langle \mathrm{T}x, u\rangle = \langle x, v\rangle$$

pour tout $x = \mathrm{B}y$ dans $\mathrm{im}(\mathrm{B})$, donc puisque $\mathrm{T}x = y$

$$\langle y, u\rangle = \langle \mathrm{B}y, v\rangle = \langle y, \mathrm{B}v\rangle$$

pour tout $y \in \mathrm{H}$, ce qui donne $u = \mathrm{B}v$, donc $u \in \mathrm{im}(\mathrm{B}) = \mathrm{dom}(\mathrm{T})$.

///

Proposition 5.3.1. *Soient* H *et* K *deux espaces de Hilbert et* T *un opérateur fermé densément défini de* H *dans* K *; l'opérateur* $(\mathrm{Id}_\mathrm{H} + \mathrm{T}^*\mathrm{T})$ *est injectif, son image est égale à* H *et* $(\mathrm{Id}_\mathrm{H} + \mathrm{T}^*\mathrm{T})^{-1}$ *est un élément positif de* $\mathcal{L}(\mathrm{H})$*. L'opérateur* $\mathrm{T}^*\mathrm{T}$ *est autoadjoint et son spectre est contenu dans* $[0, +\infty[$*.*

Preuve. — Soit $x \in \mathrm{H}$; comme T est densément défini et fermé on a

$$\mathrm{H} \times \mathrm{K} = \mathrm{Gr}(\mathrm{T}) \oplus \mathrm{U}_{\mathrm{K},\mathrm{H}}(\mathrm{Gr}(\mathrm{T}^*)),$$

donc il existe deux vecteurs $\xi \in \mathrm{Gr}(\mathrm{T})$ et $\eta \in \mathrm{Gr}(\mathrm{T}^*)$ tels que $(x, 0) = \xi - \mathrm{U}_{\mathrm{K},\mathrm{H}}(\eta)$; en d'autres termes, il existe $z \in \mathrm{dom}(\mathrm{T})$ et $y \in \mathrm{dom}(\mathrm{T}^*)$ tels que

$$(x, 0) = (z, \mathrm{T}z) + (\mathrm{T}^*y, -y).$$

Alors $y = \mathrm{T}z$, donc $z \in \mathrm{dom}(\mathrm{T}^*\mathrm{T})$ et $x = (\mathrm{Id}_\mathrm{H} + \mathrm{T}^*\mathrm{T})\, z$, ce qui montre que $(\mathrm{Id}_\mathrm{H} + \mathrm{T}^*\mathrm{T})$ est surjectif. Soit $z \in \mathrm{dom}(\mathrm{T}^*\mathrm{T})$; comme $z \in \mathrm{dom}(\mathrm{T})$ et $\mathrm{T}z \in \mathrm{dom}(\mathrm{T}^*)$, on a

$$\langle (\mathrm{Id}_\mathrm{H} + \mathrm{T}^*\mathrm{T})\, z, z\rangle = \langle z + \mathrm{T}^*\mathrm{T}\, z, z\rangle = \langle z, z\rangle + \langle \mathrm{T}^*\mathrm{T}\, z, z\rangle = \|z\|^2 + \|\mathrm{T}z\|^2.$$

Alors

$$\|z\|^2 \le \|z\|^2 + \|\mathrm{T}z\|^2 = \langle (\mathrm{Id}_\mathrm{H} + \mathrm{T}^*\mathrm{T})\, z, z\rangle \le \|z\|\, \|(\mathrm{Id}_\mathrm{H} + \mathrm{T}^*\mathrm{T})\, z\|,$$

donc $\|z\| \le \|(\mathrm{Id}_\mathrm{H} + \mathrm{T}^*\mathrm{T})\, z\|$; il en résulte que $\mathrm{Id}_\mathrm{H} + \mathrm{T}^*\mathrm{T}$ est injectif, que l'opérateur inverse $\mathrm{B} = (\mathrm{Id}_\mathrm{H} + \mathrm{T}^*\mathrm{T})^{-1}$ est continu et que $\|\mathrm{B}\| \le 1$. Enfin, considérons $x_1 = \mathrm{B}y_1$, $x_2 = \mathrm{B}y_2$, avec $y_1, y_2 \in \mathrm{H}$; on a $x_1, x_2 \in \mathrm{dom}(\mathrm{T}^*\mathrm{T})$ et

$$\langle y_1, \mathrm{B}y_2\rangle = \langle (\mathrm{Id}_\mathrm{H} + \mathrm{T}^*\mathrm{T})\, x_1, x_2\rangle = \langle x_1, x_2\rangle + \langle \mathrm{T}x_1, \mathrm{T}x_2\rangle = \langle \mathrm{B}y_1, y_2\rangle,$$

donc B est hermitien ; de plus les égalités précédentes montrent que

$$\langle y_1, \mathrm{B}y_1\rangle = \langle x_1, x_1\rangle + \langle \mathrm{T}x_1, \mathrm{T}x_1\rangle \ge 0$$

donc B est un élément positif de $\mathcal{L}(\mathrm{H})$. Par la proposition précédente, l'opérateur inverse $\mathrm{B}^{-1} = \mathrm{Id}_\mathrm{H} + \mathrm{T}^*\mathrm{T}$ est autoadjoint, donc $\mathrm{T}^*\mathrm{T}$ est autoadjoint. Comme l'opérateur B est positif de norme ≤ 1, on a $\sigma(\mathrm{B}) \subset [0, 1]$; il en résulte que $\sigma(\mathrm{Id}_\mathrm{H} + \mathrm{T}^*\mathrm{T}) \subset [1, +\infty[$ et $\sigma(\mathrm{T}^*\mathrm{T}) \subset [0, +\infty[$.

///

5.4. Transformations et représentations

On peut appliquer la transformation de Cayley, qui est fondée sur l'utilisation de la fonction $t \in \mathbb{R} \to (1 + \mathrm{i}t)/(1 - \mathrm{i}t)$ à valeurs dans le cercle unité, pour associer à tout autoadjoint non borné T sur H un opérateur unitaire U sur H,

$$\mathrm{U} = (\mathrm{Id_H} + \mathrm{i\,T})(\mathrm{Id_H} - \mathrm{i\,T})^{-1}.$$

Sous certaines hypothèses, on peut aussi faire le chemin inverse, et associer à certains unitaires un autoadjoint, en général non borné.

Le calcul fonctionnel continu conduit au théorème de représentation suivant, résultat à partir duquel le calcul fonctionnel devient essentiellement trivial.

Théorème 5.4.1. *Soit* $\mathrm{T} \in \mathcal{L}(\mathrm{H})$ *un opérateur normal borné, défini sur un espace de Hilbert complexe* H *; il existe un espace mesuré* (Ω, μ)*, un opérateur unitaire* U *de* H *sur* $\mathrm{L}_2(\Omega, \mu)$ *et une fonction mesurable bornée* $g \in \mathrm{L}_\infty(\Omega, \mu)$ *tels que*

$$\mathrm{T} = \mathrm{U}^* \circ \mathrm{M}_g \circ \mathrm{U},$$

où $\mathrm{M}_g \in \mathcal{L}(\mathrm{L}_2(\Omega, \mu))$ *est l'opérateur de multiplication par la fonction* g*.*

Esquisse de preuve. — Donnons une idée de démonstration dans un cas plus simple, qui sert de brique fondamentale pour la preuve générale : on supposera qu'il existe $x_0 \in \mathrm{H}$, disons de norme 1, tel que l'espace vectoriel engendré par toutes les images $\mathrm{T}^m(\mathrm{T}^*)^n x_0$, $m, n \geq 0$, soit dense dans H. On note $\mathrm{K} = \sigma(\mathrm{T})$ comme d'habitude et on note $\zeta = \zeta_\mathrm{K}$ la fonction $z \in \mathrm{K} \to z \in \mathbb{C}$. On introduit une application linéaire u_0 de C(K) dans H en posant

$$\forall f \in \mathrm{C(K)}, \quad u_0 f = f(\mathrm{T})\, x_0.$$

L'image de la fonction constante $\mathbf{1}$ est le vecteur x_0 donné ; l'image de u_0 est dense dans H d'après notre hypothèse supplémentaire, car cette image contient tous les vecteurs $\mathrm{T}^m(\mathrm{T}^*)^n x_0$, images des fonctions $\zeta^m \overline{\zeta}^n$. Introduisons une forme linéaire ξ sur C(K) par

$$\forall f \in \mathrm{C(K)}, \quad \xi(f) = \langle u_0 f, x_0 \rangle = \langle f(\mathrm{T})\, x_0, x_0 \rangle\,;$$

la valeur $\xi(f)$ est réelle quand f est réelle, car $f(\mathrm{T})$ est alors hermitien, et positive si f est réelle ≥ 0, car $f(\mathrm{T})$ est alors hermitien positif. On a $\xi(\mathbf{1}) = \langle x_0, x_0 \rangle = 1$. D'après le théorème de représentation des formes linéaires positives sur C(K), on sait qu'il existe une probabilité μ sur K telle que

$$\forall f \in \mathrm{C(K)}, \quad \xi(f) = \int_\mathrm{K} f(t)\, \mathrm{d}\mu(t).$$

On voit que u_0 est isométrique de C(K), muni de la norme de $\mathrm{L}_2(\mathrm{K}, \mu)$, à valeurs dans l'espace de Hilbert H : en effet,

$$\|u_0 f\|_\mathrm{H}^2 = \langle f(\mathrm{T})\, x_0, f(\mathrm{T})\, x_0 \rangle = \langle \overline{f}(\mathrm{T}) f(\mathrm{T})\, x_0, x_0 \rangle = \langle u_0(\overline{f} f), x_0 \rangle = \int_\mathrm{K} |f|^2\, \mathrm{d}\mu.$$

On peut donc étendre u_0 en une isométrie U de $\mathrm{L}_2(\mathrm{K}, \mu)$ dans H, surjective car l'image de u_0 est dense. On voit que pour f continue,

$$\mathrm{TU}f = \mathrm{T}f(\mathrm{T})x_0 = u_0(\zeta f) = \mathrm{U}(\zeta f).$$

On voit donc que l'action de T dans H correspond, par l'équivalence unitaire donnée par U, à la multiplication par la fonction $g = \zeta_\mathrm{K}$ dans $\mathrm{L}_2(\mathrm{K}, \mu)$.

///

Corollaire. *Soit* T *un opérateur autoadjoint (non borné) sur un espace de Hilbert complexe* H *; il existe un espace mesuré* (Ω, μ)*, un opérateur unitaire* U *de* H *sur* $\mathrm{L}_2(\Omega, \mu)$ *et une fonction mesurable réelle* $g \in \mathrm{L}_\infty(\Omega, \mu)$ *tels que*

$$\mathrm{T} = \mathrm{U}^* \circ \mathrm{M}_g \circ \mathrm{U},$$

où M_g *est l'opérateur (non borné) de multiplication par la fonction réelle* g.

Pour que l'énoncé soit complet, il faut préciser les correspondances entre domaines : on pose

$$\mathrm{dom}(\mathrm{M}_g) = \{f \in \mathrm{L}_2(\Omega, \mu) : gf \in \mathrm{L}_2(\Omega, \mu)\}$$

et on doit avoir $\mathrm{dom}(\mathrm{M}_g) = \mathrm{U}(\mathrm{dom}(\mathrm{T}))$.

Si on utilise comme unitaire U la transformation de type Fourier qui est définie sur $\mathrm{L}_1(\mathbb{R})$ par

$$\forall t \in \mathbb{R}, \quad \widehat{f}(t) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} f(x)\, \mathrm{e}^{-\mathrm{i}tx}\, \mathrm{d}t,$$

et qui est ensuite prolongée en isométrie de $\mathrm{L}_2(\mathbb{R})$, on voit que l'autoadjoint $\mathrm{iD} : f \to \mathrm{i}f'$, de domaine $\mathrm{H}^1(\mathbb{R})$, correspond par Fourier à l'opérateur M de multiplication par $t \to t$.

5.5. L'espace $\mathrm{H}_0^1(\Omega)$

On peut généraliser l'exemple **A** dans un ouvert Ω de $\mathbb{R}^d$; on considère l'opérateur ∇_0 sur $\mathrm{H} = \mathrm{L}_2(\Omega, \lambda)$ (λ la mesure de Lebesgue), défini sur le domaine $\mathrm{dom}(\nabla_0) = \mathcal{D}(\Omega)$ par $\varphi \in \mathcal{D}(\Omega) \to \nabla_0\varphi$, où

$$\forall t \in \Omega, \quad (\nabla_0\varphi)(t) = (\nabla\varphi)(t) = (\mathrm{D}_1\varphi(t), \ldots, \mathrm{D}_d\varphi(t)) \in \mathbb{K}^d$$

où D_j désigne la $j$ème dérivée partielle de φ. D'un point de vue plus global, on peut considérer que $t \in \Omega \to (\nabla\varphi)(t)$ est une fonction vectorielle, élément de l'espace de Hilbert $\mathrm{K} = \mathrm{L}_2(\Omega, \lambda, \mathbb{K}^d)$, et ∇_0 sera vu comme opérateur de H dans K. La fermeture de ∇_0 est l'opérateur ∇ dont le domaine est l'espace vectoriel $\mathrm{H}_0^1(\Omega)$ formé des fonctions $f \in \mathrm{L}_2(\Omega)$ qui sont limite dans L_2 d'une suite $(\varphi_n) \subset \mathcal{D}(\Omega)$ telle que $\mathrm{D}_j\varphi_n$ converge dans L_2 vers une fonction g_j pour $j = 1, \ldots, d$. Ces fonctions g_j auront la propriété

$$\int_\Omega f\, \mathrm{D}_j\psi = -\int_\Omega g_j\psi$$

pour toute $\psi \in \mathcal{D}(\Omega)$; ce sont les *dérivées partielles au sens des distributions* de f, qui seront encore notées $\mathrm{D}_j f$; l'opérateur fermé ∇ défini[e] sur $\mathrm{H}_0^1(\Omega)$ agit par

$$(\mathrm{D}_1 f, \ldots, \mathrm{D}_d f) = \nabla f \in \mathrm{K}.$$

On munit $\mathrm{H}_0^1(\Omega)$ de la norme du graphe,

$$\|f\|_{\mathrm{H}^1}^2 = \int_\Omega |f|^2 + \int_\Omega \sum_{j=1}^d |\mathrm{D}_j f|^2 = \int_\Omega \bigl(|f|^2 + |\nabla f|^2\bigr).$$

Muni de cette norme, l'espace H_0^1 est complet, et c'est un espace de Hilbert pour le produit scalaire

$$\langle f, g\rangle_{H^1} = \int_\Omega \big(f\overline{g} + \nabla f \cdot \overline{\nabla g}\big).$$

Si on est sur un ouvert Ω d'une variété riemannienne M, on a une notion de gradient pour les fonctions différentiables définies sur Ω, et on peut généraliser la discussion précédente à ce cadre. Il faut noter une différence : les valeurs du gradient $\nabla_m f$, $m \in M$, sont dans des espaces vectoriels différents quand m varie, puisque $\nabla_m f$ appartient à l'espace tangent $T_m M$ à la variété au point m. Cela n'empêche pas de définir un espace de Hilbert K analogue au précédent.

Injection de $H_0^1(\Omega)$ *dans* $L_2(\mathbb{R}^d)$

Si φ est une fonction de $\mathcal{D}(\Omega)$, on la prolongera en une fonction $E\varphi$ définie sur $\mathbb{R}^d$ en posant simplement $(E\varphi)(x) = \varphi(x)$ si $x \in \Omega$ et $(E\varphi)(x) = 0$ si $x \notin \Omega$. Il est clair que $E\varphi$ est C^∞ sur $\mathbb{R}^d$. On a

$$\int_{\mathbb{R}^d} |(E\varphi)(x)|^2 \,dx = \int_\Omega |\varphi(x)|^2 \,dx \le \|\varphi\|_{H^1}^2 \tag{5}$$

pour toute $\varphi \in \mathcal{D}(\Omega)$. Il est clair en particulier que l'application E définie sur $\mathcal{D}(\Omega)$ est continue de la norme $H^1(\Omega)$ vers $L_2(\mathbb{R}^d)$; comme l'espace $H_0^1(\Omega)$ est précisément défini comme adhérence de $\mathcal{D}(\Omega)$ dans $H^1(\Omega)$, il en résulte que l'application E se prolonge en application continue de $H_0^1(\Omega)$ dans $L_2(\mathbb{R}^d)$. Il s'agit moralement de l'injection canonique du premier espace dans le second. On a alors

Théorème. *Soit* $\Omega \subset \mathbb{R}^d$ *un ouvert* **borné** *; l'injection* E *est un opérateur compact de* $H_0^1(\Omega)$ *dans* $L_2(\mathbb{R}^d)$.

Preuve. — On va appliquer le théorème de Riesz-Fréchet-Kolmogorov sur la relative compacité dans $L_p(\mathbb{R}^d)$ d'un ensemble de fonctions A, ici avec $p = 2$, et

$$A = \{E\varphi : \varphi \in \mathcal{D}(\Omega),\ \|\varphi\|_{H^1} \le 1\}.$$

Ce théorème comprend trois clauses dont deux sont évidentes ici. Tout d'abord, A est borné dans $L_2(\mathbb{R}^d)$ d'après (5) ; ensuite, toutes les fonctions $E\varphi$ sont à support dans le compact fixé $\overline{\Omega}$ (c'est ici qu'intervient l'hypothèse Ω *borné*). Il reste à voir la clause sur l'équicontinuité des translations.

On montre que $\|f_v - f\|_2 \le |v|$ pour toute fonction $f \in A$ et tout vecteur $v \in \mathbb{R}^d$, où $f_v(x) = f(x - v)$. Pour cela on écrit

$$f(x+v) - f(x) = \int_0^1 (\nabla f)(x + sv) \cdot v \,ds,$$

puis avec Cauchy-Schwarz suivi de Jensen

$$\big(f(x+v) - f(x)\big)^2 = \Big(\int_0^1 (\nabla f)(x+sv) \cdot v \,ds\Big)^2 \le$$

$$\Big(\int_0^1 |(\nabla f)(x+sv)|\,|v|\,\mathrm{d}s \Big)^2 \le |v|^2 \int_0^1 |(\nabla f)(x+sv)|^2\,\mathrm{d}s.$$

Par Fubini on obtiendra

$$\|f_{-v} - f\|_2^2 = \int_{\mathbb{R}^d} \big(f(x+v) - f(x)\big)^2\,\mathrm{d}x \le |v|^2 \int_0^1 \int_{\mathbb{R}^d} |(\nabla f)(x+sv)|^2\,\mathrm{d}x\,\mathrm{d}s =$$

$$|v|^2 \int_0^1 \big\|\, |\nabla f|\, \big\|_{\mathrm{L}_2}^2\,\mathrm{d}s = |v|^2 \,\big\|\, |\nabla f|\, \big\|_2^2\,;$$

cette relation montrée pour des fonctions régulières se prolonge par densité à toutes les fonctions de $\mathrm{H}_0^1(\Omega)$; pour toute fonction f dans l'image $\mathrm{E}(\mathrm{H}_0^1(\Omega)) \subset \mathrm{H}^1(\mathbb{R}^d)$, elle se résume en

$$\forall v \in \mathbb{R}^d, \quad \|f_v - f\|_2 \le |v|\, \|\nabla f\|_2\,; \tag{6}$$

elle donne $\|f_v - f\|_2 \le |v|$ pour tout vecteur $v \in \mathbb{R}^d$ et toute fonction $f \in \mathrm{A}$, d'après la définition de A. On a ainsi montré l'équicontinuité des translations sur l'ensemble A, ce qui permet de conclure.

///

Inégalité de Poincaré

On désigne encore par Ω un ouvert borné de $\mathbb{R}^d$. Reprenons le contenu de l'inégalité (6), en prenant pour $v_0 = v$ un vecteur assez grand pour que $\Omega - v_0$ soit disjoint de Ω ; pour $x \in \Omega$, on a alors $f(x - v_0) = 0$. On aura donc $f_{v_0} - f = -f$ sur Ω, donc

$$\|f\|_{\mathrm{L}_2(\Omega)} = \|f_{v_0} - f\|_{\mathrm{L}_2(\Omega)} \le \|f_{v_0} - f\|_{\mathrm{L}_2(\mathbb{R}^d)} \le |v_0|\, \|\nabla f\|_{\mathrm{L}_2(\Omega)},$$

ce qui constitue *l'inégalité de Poincaré* : si Ω est un ouvert borné, il existe une constante C (dépendant de Ω) telle que

$$\int_\Omega |f(x)|^2\,\mathrm{d}x \le \mathrm{C}^2 \int_\Omega |(\nabla f)(x)|^2\,\mathrm{d}x,$$

pour toute fonction f de $\mathrm{H}_0^1(\Omega)$. On peut prendre pour C la norme du vecteur v_0.

///

La démonstration précédente montre que la condition Ω borné n'est pas nécessaire, puisque nous avons simplement utilisé l'existence d'un vecteur v_0 tel que $\Omega - v_0$ soit disjoint de Ω ; cela serait encore vrai pour une bande infinie mais de largeur finie. L'inégalité de Poincaré montre que dans le cas d'un ouvert borné, la quantité

$$\Big(\int_\Omega |(\nabla f)(x)|^2\,\mathrm{d}x \Big)^{1/2}$$

définit une norme équivalente sur l'espace $\mathrm{H}_0^1(\Omega)$.

Opérateur laplacien sur un ouvert borné

Gardons notre ouvert Ω borné dans $\mathbb{R}^d$. On remarque pour commencer que l'image de l'opérateur ∇ de $\mathrm{H} = \mathrm{L}_2(\Omega)$ dans $\mathrm{K} = \mathrm{L}_2(\Omega, \lambda, \mathbb{K}^d)$, de domaine $\mathrm{H}_0^1(\Omega)$, est fermée dans l'espace de Hilbert K : si une suite (∇f_n) converge vers une fonction vectorielle $k = (k_1, \ldots, k_d)$, l'inégalité de Poincaré montre que la suite (f_n) est de Cauchy dans L_2, donc converge vers $f \in \mathrm{L}_2(\Omega)$, ce qui montre que (f, k) est dans l'adhérence du graphe de l'opérateur ∇, qui est fermé, donc $k = \nabla f$ est bien dans $\mathrm{F} = \mathrm{im}(\nabla)$, qui est donc un sous-espace vectoriel fermé de l'espace de Hilbert K.

Cherchons le domaine de ∇^* ; une fonction vectorielle $k \in \mathrm{K}$ est dans $\mathrm{dom}(\nabla^*)$ si et seulement si la forme linéaire

$$f \in \mathrm{H}_0^1(\Omega) \to \int_\Omega \nabla f \cdot \overline{k}$$

est L_2-continue ; mais puisque l'espace image F est fermé, on peut écrire $k = \nabla g + k_1$ pour une certaine fonction $g \in \mathrm{H}_0^1(\Omega)$ et pour une $k_1 \perp \mathrm{F}$; en nous restreignant au cas où $f \in \mathcal{D}(\Omega)$ est une fonction test de la théorie des distributions, on a donc

$$\int_\Omega \nabla f \cdot \overline{k} = \int_\Omega \nabla f \cdot \overline{\nabla g} = -(\Delta(\bar{g}), f)$$

où Δ désigne le laplacien au sens des distributions ; dire que cette forme linéaire est L_2-bornée signifie que la distribution $\Delta(g)$ « est » une fonction de $\mathrm{L}_2(\Omega)$, donc

$$\mathrm{dom}(\nabla^*) = \{\nabla g + k_1 : g \in \mathrm{H}_0^1(\Omega),\ \Delta g \in \mathrm{L}_2(\Omega),\quad k_1 \perp \mathrm{F}\}.$$

On a donc

$$\mathrm{dom}(\nabla^* \nabla) = \{f \in \mathrm{H}_0^1(\Omega),\ \Delta f \in \mathrm{L}_2(\Omega)\},$$

et l'opérateur $\nabla^* \nabla$ agit sur les fonctions du domaine par

$$\nabla^* \nabla f = -\Delta f,$$

ce qui montre que $-\Delta$, défini de cette façon, est un opérateur autoadjoint positif (d'après le théorème 5.3.1). Il faut bien faire attention qu'il n'y a pas un unique « opérateur », au sens de ce chapitre, qui mérite de s'appeler Laplacien. Nous avons privilégié un cas de figure, qui correspond au problème de résoudre $\Delta u = f$, avec f donnée dans l'ouvert Ω, et u inconnue, mais qu'on veut nulle au bord (condition au bord dite *de Dirichlet*). Sur une variété riemannienne, la même approche à partir d'un opérateur gradient fournira un opérateur Laplacien.

La définition du domaine de Δ, qui a été obtenue par la théorie des pages précédentes et qui ne parle que de Δf sans isoler les dérivées partielles secondes individuellement, peut sembler un peu curieuse. En fait, si Ω est un ouvert régulier, on peut montrer que toutes les dérivées partielles secondes de f sont dans $\mathrm{L}_2(\Omega)$ quand $\Delta f \in \mathrm{L}_2(\Omega)$ (voir le livre de Brézis par exemple ; cette affirmation est facile par Fourier dans le cas $\Omega = \mathbb{R}^d$).

D'après la théorie générale, $\mathrm{Id} + \nabla^* \nabla$ est une bijection de $\mathrm{dom}(\nabla^* \nabla)$ sur $\mathrm{L}_2(\Omega)$, d'inverse B borné et hermitien. On va voir que B est de plus compact. Supposons que $h \in \mathrm{L}_2(\Omega)$, $f = \mathrm{B}h \in \mathrm{dom}(\nabla^* \nabla)$; on a

$$\langle h, \mathrm{B}h \rangle = \langle h, f \rangle = \langle (\mathrm{Id} + \nabla^* \nabla) f, f \rangle = \|f\|_2^2 + \|\nabla f\|_2^2.$$

On peut faire parler cette relation, de façon plus ou moins précise. Tout d'abord,

$$\|f\|_{\mathrm{H}^1}^2 = \|f\|_2^2 + \|\nabla f\|_2^2 = \langle h, f\rangle \le \|h\|_2 \|f\|_2 \le \|h\|_2 \|f\|_{\mathrm{H}^1}$$

montre que $\|\mathrm{B}h\|_{\mathrm{H}^1} = \|f\|_{\mathrm{H}^1} \le \|h\|_2$, ce qui montre que B envoie la boule unité de $\mathrm{L}_2(\Omega)$ dans celle de $\mathrm{H}_1 = \mathrm{H}_0^1(\Omega)$, qui est compacte dans $\mathrm{L}_2(\Omega)$; l'opérateur injectif B est donc un endomorphisme compact de $\mathrm{L}_2(\Omega)$. Comme tel, il admet une diagonalisation dans une base orthonormée. Remarquons que 1 n'est pas valeur propre de B : si $\mathrm{B}h = h$, les relations précédentes donnent

$$\|h\|_2^2 + \|\nabla h\|_2^2 = \langle h, h\rangle = \|h\|_2^2,$$

donc $\nabla h = 0$ et $h = 0$ par Poincaré, puisque $h = \mathrm{B}h \in \mathrm{H}_0^1(\Omega)$.

L'opérateur B est hermitien positif injectif, de norme 1 et 1 n'est pas valeur propre : on peut ranger ses valeurs propres dans une suite $\mu_1 \ge \mu_2 \ldots$ qui tend vers 0 par valeurs positives, et $\mu_1 < 1$. Les fonctions propres φ_n vérifient

$$\varphi_n - \Delta\varphi_n = \frac{1}{\mu_n}\,\varphi_n,$$

soit

$$-\Delta\varphi_n = \frac{1-\mu_n}{\mu_n}\,\varphi_n.$$

On obtient pour $-\Delta$ une suite croissante (λ_n) de valeurs propres > 0, qui tend vers $+\infty$.

Exemples

L'exemple le plus élémentaire du Laplacien-Dirichlet (on s'intéresse aux fonctions nulles au bord) est le cas de l'ouvert $\Omega = (0, \pi)$ de la droite réelle. Dans ce cas, on voit que

$$\varphi_n(x) = \sin(nx), \quad n = 1, 2, \ldots$$

donne une base orthogonale de $\mathrm{L}_2(0, \pi)$ formée de fonctions propres du « Laplacien », réduit ici à $f \to f''$; on trouve les valeurs propres de $-\Delta$,

$$\lambda_n = n^2, \quad n = 1, 2, \ldots$$

Un autre exemple simple est celui de la plaque carrée $(0, \pi)^2$; on trouve une suite double de fonctions propres,

$$\varphi_{m,n}(x, y) = \sin(mx)\sin(ny), \quad m, n = 1, 2, \ldots$$

et les valeurs propres de $-\Delta$,

$$\lambda_{m,n} = m^2 + n^2, \quad m, n = 1, 2, \ldots$$

Un autre cas calculable, bien que beaucoup plus difficile, est celui du disque unité de $\mathbb{R}^2$; on trouve encore une double infinité de fonctions propres,

$$\varphi_{n,k}(x, y) = \mathrm{J}_n(\mu_{n,k} r)\cos(n\theta), \quad n = 0, 1, \ldots$$

$$\psi_{n,k}(x, y) = \mathrm{J}_n(\mu_{n,k} r)\sin(n\theta), \quad n = 1, 2, \ldots$$

où $x = r\cos\theta$, $y = r\sin\theta$, où J_n est la fonction de Bessel d'indice n, pour $n = 0, 1, \ldots,$ et $(\mu_{n,k})_{k\ge1}$ est la liste des zéros > 0 de la fonction de Bessel J_n.

Quand on a une base de fonctions propres du Laplacien, on peut se faire une idée de la résolution de l'équation des ondes dans le domaine Ω,

$$\frac{\partial^2}{\partial t^2} = \Delta_x$$

où x représente le collectif des variables d'espace et t le temps. La théorie montre que ces fonctions propres sont nécessairement C^∞ dans l'ouvert Ω. Si les fonctions propres sont les $(\varphi_n(x))$, avec

$$-\Delta\varphi_n = \lambda_n\varphi_n,$$

on posera (sous réserve de justifications qu'on ne trouvera pas ici)

$$u(x,t) = \sum_n c_n \cos\bigl(\sqrt{\lambda_n}t\bigr)\,\varphi_n(x).$$

La fonction u reste nulle pour x au bord de l'ouvert. L'état initial, à l'instant $t = 0$, correspond au développement de la fonction « valeur initiale » dans la base de fonctions propres,

$$u(x,0) = \sum_n c_n\,\varphi_n(x).$$

On peut tenir des considérations analogues pour l'équation de la chaleur,

$$\frac{\partial}{\partial t} = \Delta_x$$

avec cette fois

$$u(x,t) = \sum_n c_n\,\mathrm{e}^{-\lambda_n t}\;\varphi_n(x).$$

Notes du chapitre 5

($^{\mathbf{a}}$) Disons simplement ceci : une suite $(\varphi_n) \subset \mathcal{D}(\mathbb{R})$ tend vers 0 dans l'espace $\mathcal{D}(\mathbb{R})$ si toutes ces fonctions sont à support dans un même compact, et si pour tout $k \in \mathbb{N}$ les dérivées k-ièmes $(\varphi_n^{(k)})$ convergent uniformément vers 0. Si T est une distribution, on doit alors avoir $\mathrm{T}\varphi_n \to 0$.

($^{\mathbf{b}}$) Soit $\psi \in \mathcal{D}(\mathbb{R})$ d'intégrale nulle, à support dans un intervalle compact $[a,b]$; la fonction Ψ définie par $\Psi(x) = \int_{-\infty}^x \psi(t)\,\mathrm{d}t$ est de classe C^∞, mais de plus elle est à support compact : on a bien sûr $\Psi(x) = 0$ quand $x \le a$, mais aussi pour $x \ge b$

$$\Psi(x) = \int_{-\infty}^x \psi(t)\,\mathrm{d}t = \int_{-\infty}^b \psi(t)\,\mathrm{d}t = \int_{\mathbb{R}} \psi(t)\,\mathrm{d}t = 0\,;$$

donc $\Psi \in \mathcal{D}(\mathbb{R})$ et ψ est bien une dérivée. Inversement, il est clair que les dérivées des fonctions de $\mathcal{D}(\mathbb{R})$ sont d'intégrale nulle.

($^{\mathbf{c}}$) On doit montrer que la fonction f définie sur $\mathbb{R}$ par

$$\forall t \in \mathbb{R}, \quad f(t) = \mathrm{e}^{\lambda t}\int_t^{+\infty} \mathrm{e}^{-\lambda s}\,g(s)\,\mathrm{d}s$$

est dans $\mathrm{L}_2(\mathbb{R})$ quand $g \in \mathrm{L}_2(\mathbb{R})$, et que sa dérivée généralisée vérifie $\lambda f - f' = g$. Pour le premier point, on peut observer que f est la convolution de la fonction $x \to \mathbf{1}_{\{x<0\}}\, \mathrm{e}^{\lambda x}$, qui est intégrable sur $\mathbb{R}$ parce que $\mathrm{Re}\,\lambda > 0$, avec la fonction $g \in \mathrm{L}_2(\mathbb{R})$, et se rappeler que $\mathrm{L}_1(\mathbb{R}) * \mathrm{L}_2(\mathbb{R}) \subset \mathrm{L}_2(\mathbb{R})$. Pour le second point, on utilisera la formule (IPP).

(**d**) Si $k \in \mathrm{L}_2(0,1)$ est d'intégrale nulle, on pose

$$\mathrm{K}(x) = \int_0^x k(t)\,\mathrm{d}t\,;$$

on voit que $\mathrm{K} \in \mathrm{H}^1([0,1])$ vérifie $\mathrm{K}(0) = \mathrm{K}(1) = 0$. Inversement, si K est dans H^1_0, on a

$$0 = \mathrm{K}(1) - \mathrm{K}(0) = \int_0^1 \mathrm{K}'(t)\,\mathrm{d}t.$$

On peut montrer que H^1_0 est en fait l'adhérence de $\mathcal{D}(]0,1[)$ dans $\mathrm{H}^1([0,1])$: soit $\mathrm{K} \in \mathrm{H}^1_0$, de dérivée généralisée $k \in \mathrm{L}_2(0,1)$, k d'intégrale nulle ; on peut approcher k (dans L_2) par une fonction k_0 d'intégrale nulle et de plus, nulle en dehors de $[\varepsilon, 1-\varepsilon]$, pour un $\varepsilon > 0$ suffisamment petit. En convolant k_0 avec une fonction de $\mathcal{D}(\mathbb{R})$ à petit support, on obtient une nouvelle fonction k_1, proche de k_0, d'intégrale nulle, et de plus $k_1 \in \mathcal{D}(]0,1[)$. La primitive K_1 de k_1 est encore dans $\mathcal{D}(]0,1[)$, et elle est proche en norme H^1 de la primitive K de k.

(**e**) On a défini $\mathrm{H}^1_0(\Omega)$ en prenant l'adhérence dans $\mathrm{L}_2(\mathbb{R})^{d+1}$ de l'ensemble des $(\varphi, \nabla\varphi)$ pour $\varphi \in \mathcal{D}(\Omega)$. Cette adhérence est exactement le graphe de l'opérateur ∇ de domaine $\mathrm{H}^1_0(\Omega)$: pour être complet, il faut voir que cette adhérence est bien un graphe, c'est-à-dire que si $(0, g)$ est la limite d'une suite $(\varphi_n, \nabla\varphi_n)$ pour $(\varphi_n) \subset \mathcal{D}(\Omega)$, alors $g = 0$. Si $\psi \in \mathcal{D}(\Omega)$ est donnée, on a pour tout n et $j = 1, \ldots, d$

$$\int_\Omega \bigl(\mathrm{D}_j\varphi_n\bigr)\psi\,\mathrm{d}x = -\int_\Omega \varphi_n\bigl(\mathrm{D}_j\psi\bigr)\,\mathrm{d}x\,;$$

on obtient à la limite, en notant g_j la $j$ième composante de la fonction vectorielle g

$$\int_\Omega g_j\psi\,\mathrm{d}x = 0.$$

Comme $\mathcal{D}(\Omega)$ est dense dans $\mathrm{L}_2(\Omega)$, il en résulte que $g_j = 0$ pour tout j, donc $g = 0$.

6. Appendice A. Calcul fonctionnel holomorphe

Un *cycle* Γ dans un ouvert Ω du plan complexe est une somme formelle $\Gamma = \sum_{n=1}^{\mathrm{N}} \gamma_n$, où chaque γ_n est un chemin fermé (continu) dans Ω. On pose quand f est une fonction complexe continue dans Ω

$$\int_\Gamma f(z)\,dz = \sum_{n=1}^{\mathrm{N}} \int_{\gamma_n} f(z)\,dz.$$

On considérera un segment (orienté) $\mathrm{I} = [\mathrm{A}, \mathrm{B}] \subset \mathbb{C}$ comme un chemin de A à B, et on calculera

$$\int_{\mathrm{I}} f(z)\,dz = \int_0^1 f\big((1-t)\mathrm{A} + t\mathrm{B}\big)\,(\mathrm{B} - \mathrm{A})\,dt$$

Si $\mathrm{I} = [\mathrm{A}, \mathrm{B}]$, on notera $[\mathrm{B}, \mathrm{A}] = -\mathrm{I}$. On voit que

$$\int_{-\mathrm{I}} f(z)\,dz = -\int_{\mathrm{I}} f(z)\,dz.$$

On note $\mathrm{H}(\Omega)$ l'espace vectoriel des fonctions complexes f qui sont holomorphes dans l'ouvert Ω. On désignera par Γ^* l'ensemble (compact) formé par la réunion des images des chemins fermés γ_n, $n = 1, \ldots, \mathrm{N}$.

Lemme 6.1. *Si Ω est un ouvert de $\mathbb{C}$ et Γ un cycle dans Ω, tel que*

$$\mathrm{ind}_\Gamma(z) = \frac{1}{2\pi \mathrm{i}} \int_\Gamma \frac{1}{t - z}\,dt$$

soit égal à 0 *pour tout $z \notin \Omega$, on a*

$$f(z)\,\mathrm{ind}_\Gamma(z) = \frac{1}{2\pi \mathrm{i}} \int_\Gamma \frac{f(t)}{t - z}\,dt$$

pour tout $z \in \Omega \setminus \Gamma^*$ et toute fonction $f \in \mathrm{H}(\Omega)$. Il en résulte que

$$\int_\Gamma f(t)\,dt = 0$$

pour toute fonction $f \in \mathrm{H}(\Omega)$.

Esquisse de preuve. — Soit $f \in \mathrm{H}(\Omega)$; on introduit la fonction h définie sur Ω^2 par

$$h(t, w) = \frac{f(t) - f(w)}{t - w}$$

quand $t \neq w$ et $h(t, t) = f'(t)$. On montre que la fonction g définie sur Ω par

$$g(w) = \int_\Gamma h(t, w)\,dt$$

est holomorphe. Par ailleurs, l'ensemble

$$\Omega_1 = \{w \in \mathbb{C} \setminus \Gamma^* : \mathrm{ind}_\Gamma(w) = 0\}$$

est un ouvert, et $\mathbb{C} = \Omega \cup \Omega_1$. On peut définir une fonction holomorphe g_1 par

$$\forall w \in \Omega_1, \quad g_1(w) = \int_\Gamma \frac{f(t)}{t - w}\, dt$$

et les fonctions g et g_1 coïncident sur $\Omega \cap \Omega_1$, ce qui permet de définir une fonction holomorphe G sur $\mathbb{C}$ en les recollant. Mais G tend vers 0 à l'infini, donc elle est nulle par le théorème de Liouville. La nullité de $\mathrm{G} = g$ sur Ω fournit le premier résultat. Pour le second résultat, appliquer le premier à la fonction $g(z) = (z - z_0) f(z)$, où z_0 est un point fixé de $\Omega \setminus \Gamma^*$.

Lemme 6.2. *Si Ω est un ouvert de $\mathbb{C}$ et* K *un compact contenu dans Ω, il existe un cycle Γ dans Ω, qui ne rencontre pas* K *et tel que*

$$\mathrm{ind}_\Gamma(z) = \frac{1}{2\pi \mathrm{i}} \int_\Gamma \frac{1}{t - z}\, dt \tag{1}$$

soit égal à 1 *pour tout $z \in \mathrm{K}$ et égal à* 0 *pour tout $z \notin \Omega$. Il en résulte que*

$$f(z) = \frac{1}{2\pi \mathrm{i}} \int_\Gamma \frac{f(t)}{t - z}\, dt$$

pour tout $z \in \mathrm{K}$ et toute fonction $f \in \mathrm{H}(\Omega)$.

Preuve. — La preuve est très simple dans son principe, mais un peu longue à raconter, surtout sans l'aide d'un dessin ; le lecteur aura tout intérêt à faire ce dessin lui-même. On choisit $\delta > 0$ tel que $\sqrt{2}\delta < \mathrm{dist}(\mathrm{K}, \Omega^c)$, on considère la grille formée de toutes les verticales $x = j\delta$ et toutes les horizontales $y = k\delta$, $j, k \in \mathbb{Z}$. Introduisons une terminologie locale qui va nous aider à décrire la situation : pour chaque couple $(j, k) \in \mathbb{Z}^2$, le *point de grille* $\mathrm{P}_{j,k}$ est le point $(j\delta, k\delta) \simeq \delta(j + \mathrm{i}k) \in \mathbb{C}$, le *carré de grille* $q = q_{j,k}$ est l'enveloppe convexe des quatre points de grille $\mathrm{P}_{j,k}, \mathrm{P}_{j+1,k}, \mathrm{P}_{j+1,k+1}, \mathrm{P}_{j,k+1}$. Le *bord orienté* (dans le sens direct) du carré de grille $q_{j,k}$ est le chemin fermé $\gamma_{j,k} = \gamma(q_{j,k})$ comprenant les quatre *segments de grille* $\mathrm{I}_{j,k} = [\mathrm{P}_{j,k}, \mathrm{P}_{j+1,k}]$, $\mathrm{J}_{j+1,k} = [\mathrm{P}_{j+1,k}, \mathrm{P}_{j+1,k+1}]$, suivi de $[\mathrm{P}_{j+1,k+1}, \mathrm{P}_{j,k+1}]$ et $[\mathrm{P}_{j,k+1}, \mathrm{P}_{j,k}]$ (qui sont égaux à $-\mathrm{I}_{j,k+1}$ et $-\mathrm{J}_{j,k}$; les segments de grille sont orientés, I est considéré différent de $-\mathrm{I}$). Chaque segment de grille I apparaît dans le bord orienté d'un unique carré de grille q, et le carré q' correspondant à $-\mathrm{I}$ est différent du carré q correspondant à I. Par exemple, $\mathrm{I}_{j,k}$ n'apparaît que dans $\gamma_{j,k}$, bord de $q = q_{j,k}$ et $-\mathrm{I}_{j,k}$ n'apparaît que dans $\gamma_{j,k-1}$, bord de $q' = q_{j,k-1}$.

On considère la famille (finie) $\mathcal{Q}$ de tous les carrés de grille q qui rencontrent K ; il résulte de la condition sur δ que ces carrés sont contenus dans Ω, et il est évident que la réunion de ces carrés q couvre K. On considère le cycle Γ_0 qui est la somme des parcours des bords orientés $\gamma(q)$ des carrés $q \in \mathcal{Q}$. Le cycle Γ_0 possède la propriété suivante : si un segment de grille I rencontre K, alors I et $-\mathrm{I}$ apparaissent exactement une fois dans Γ_0 ; en effet, si I rencontre K et fait partie du bord du carré de grille q, alors $-\mathrm{I}$ fait partie du bord d'un carré $q' \neq q$; évidemment q et q' rencontrent K, donc font partie de la famille $\mathcal{Q}$, par conséquent $\gamma(q)$ et $\gamma(q')$ sont deux termes de la somme Γ_0 ; l'un fait intervenir I et l'autre $-$I. Aucun autre carré de grille ne fait intervenir I ou $-$I.

Si z appartient à l'intérieur d'un carré de grille q dont le bord orienté est $\gamma(q)$, il est clair que l'indice de $\gamma(q)$ par rapport à z est égal à 1. Si de plus le carré q est dans la famille $\mathcal{Q}$, ce chemin fermé $\gamma(q)$ est l'un des chemins qui forment Γ_0 ; le point z est extérieur aux autres carrés de grille $q' \neq q$ de la famille $\mathcal{Q}$, donc l'indice de $\gamma(q')$ par rapport à z est égal à 0, et au total l'indice de Γ_0 par rapport à z est égal à 1 ; en revanche, si z' n'est pas dans Ω, il est extérieur à tous les carrés $q' \in \mathcal{Q}$ qui sont, eux, contenus dans Ω, donc l'indice de Γ_0 par rapport à z' est égal à 0. Appelons $(**)$ cette propriété d'un cycle Γ : pour tout point z qui est intérieur à un carré de grille de la famille $\mathcal{Q}$, l'indice de Γ par rapport à z est égal à 1, et pour tout point z qui n'est pas dans Ω, l'indice de Γ par rapport à z est égal à 0. Le cycle initial Γ_0 vérifie $(**)$.

On simplifie le cycle Γ_0 de proche en proche de la façon suivante : supposons Γ_n atteint, formé d'un nombre fini de chemins fermés γ_ℓ constitués par des segments de grille, et vérifiant la condition suivante : le cycle Γ_n vérifie la condition $(**)$, et si un segment de grille I rencontre K et apparaît dans Γ_n, alors I et $-$I apparaissent exactement une fois dans Γ_n.

On a vu que cette condition est vraie pour Γ_0. Expliquons le passage à Γ_{n+1} : si un des segments I de Γ_n rencontre K, alors I et $-$I apparaissent chacun une fois dans Γ_n ; si c'est dans deux chemins γ_ℓ, γ_m différents, on recolle ces deux chemins en un seul, en supprimant I et $-$I : si $\mathrm{I} = [\mathrm{A}, \mathrm{B}]$ apparaît dans γ_ℓ, faisons partir le chemin fermé γ_ℓ du point B ; le chemin fermé γ_ℓ doit conduire de B à A, pour se fermer avec I ; de même γ_m conduit de A à B, pour se fermer avec $-$I ; on obtient un chemin fermé sans I ni $-$I en suivant d'abord γ_ℓ de B à A puis γ_m de A à B. Si I et $-$I apparaissent dans le même chemin fermé γ_ℓ, on coupe ce chemin en deux chemins fermés en supprimant I et $-$I : en effet, le parcours γ' de B à A obtenu ci-dessus en supprimant I dans γ_ℓ doit traverser $-\mathrm{I} = [\mathrm{B}, \mathrm{A}]$; on repasse donc au point B, ce qui donne un premier chemin fermé γ_1, puis on repart de A jusqu'à A pour un deuxième chemin fermé γ_2. Pour toute fonction f holomorphe au voisinage du cycle on a

$$\int_{\Gamma_n} f(t)\,dt = \int_{\Gamma_{n+1}} f(t)\,dt$$

car on a supprimé deux parcours opposés. La condition $(**)$ reste donc vraie pour Γ_{n+1} ; comme on supprime en même temps I et $-$I, il est clair que l'autre condition sur Γ_n est préservée.

On arrive finalement à un cycle Γ qui ne rencontre plus K et vérifie $(**)$. Soit z un point de K ; il est contenu dans un carré $q \in \mathcal{Q}$; si z est intérieur à q, on sait que l'indice de Γ par rapport à z est égal à 1 ; si z est sur le bord de q, il est limite de points de l'intérieur de q, et par continuité, l'indice est égal à 1 aussi. Pour les points $z' \notin \Omega$ on sait que l'indice est nul : on a ainsi obtenu la condition (1) pour le cycle Γ.

///

Corollaire 6.3. *Soient $\Omega \subset \mathbb{C}$ un ouvert et $\mathrm{K} \subset \Omega$ un compact ; il existe un ouvert ω tel que $\mathrm{K} \subset \omega \subset \Omega$ et tel que toute $f \in \mathrm{H}(\Omega)$ soit limite uniforme sur ω de fonctions rationnelles de la forme*

$$z \to \sum_{j=1}^{n} \frac{c_j}{z - \lambda_j},$$

où les λ_j sont des points de $\Omega \setminus \mathrm{K}$ (et même de $\Omega \setminus \omega$).

Preuve. — À partir de K et Ω on construit le cycle Γ du lemme précédent. On désigne par $\varepsilon > 0$ la distance de K au fermé $\Gamma^* \cup \Omega^c$ (Γ^* est l'image de Γ) et on pose

$$\omega = \{z : d(z, \mathrm{K}) < \varepsilon/2\}.$$

On a bien $\mathrm{K} \subset \omega \subset \Omega$. Il suffit pour finir de discrétiser l'intégrale de Cauchy sur le cycle Γ du lemme précédent ; les points λ_j sont pris dans le parcours de Γ : ils sont donc dans l'ensemble $\Omega \setminus \omega$, qui est contenu dans $\Omega \setminus \mathrm{K}$.

///

Calcul fonctionnel holomorphe

Si a est un élément d'une algèbre de Banach unitaire complexe A et si f est une fonction holomorphe dans un voisinage du spectre de a, introduisons un ouvert Ω contenant $\mathrm{K} = \sigma(a)$ et tel que $f \in \mathrm{H}(\Omega)$, et un cycle Γ qui vérifie la condition du lemme 6.2. Considérons l'intégrale de Riemann vectorielle

$$f_\Gamma(a) = \frac{1}{2\pi \mathrm{i}} \int_\Gamma f(z)\,(z1_\mathrm{A} - a)^{-1}\,dz \in \mathrm{A}.$$

On montre que cette intégrale ne dépend pas du cycle Γ choisi, en appliquant une forme linéaire continue quelconque sur A et en utilisant le résultat connu dans le cas scalaire, donné par le lemme 6.1 ; ceci permet de nommer $\widetilde{f}(a)$ l'élément $f_\Gamma(a)$ de A ainsi obtenu. Expliquons cette indépendance : si Γ_1 et Γ_2 sont deux cycles dans $\Omega_1 \supset \mathrm{K}$ et $\Omega_2 \supset \mathrm{K}$, qui vérifient les conditions du lemme 6.2, et si g est holomorphe (scalaire) dans $\Omega_1 \setminus \mathrm{K}$ et $\Omega_2 \setminus \mathrm{K}$, considérons l'ouvert $\omega = (\Omega_1 \cup \Omega_2) \setminus \mathrm{K}$; la fonction g est holomorphe dans ω, le cycle $\Gamma_1 - \Gamma_2$ est un cycle dans ω qui vérifie les conditions du lemme 6.1, donc

$$\int_{\Gamma_1} g(t)\,dt = \int_{\Gamma_2} g(t)\,dt.$$

La fonction g est de la forme

$$g(z) = \xi\bigl(f(z)\,(z1_\mathrm{A} - a)^{-1}\bigr),$$

où ξ est une forme linéaire continue sur A.

Après avoir défini $\widetilde{f}(a)$, on peut développer la théorie du calcul fonctionnel holomorphe et étudier les propriétés de l'application $\mathcal{H}_a$ qui associe l'élément $\widetilde{f}(a) \in \mathrm{A}$ à chaque fonction f holomorphe au voisinage du spectre de $a \in \mathrm{A}$. Les premiers résultats consistent à montrer que $\widetilde{f}(a)$ est bien ce qu'on attend quand f est un polynôme ou une fraction rationnelle. On consultera Rudin, *Functional Analysis*, par exemple. Pour ne pas laisser le lecteur complètement sur sa faim, jetons lui quelques miettes. Lorsque $f = 1$ (la fonction constante) on peut prendre $\Omega = \mathbb{C}$ et pour Γ un cercle de très grand rayon r. En développant $(z1_\mathrm{A} - a)^{-1}$, lorsque $|z| = r$, en série de puissances de a/z on voit facilement (revenir à la section 2.3) que

$$\widetilde{1}(a) = \frac{1}{2\pi \mathrm{i}} \sum_{n=0}^{+\infty} \int_\Gamma \frac{a^n}{z^{n+1}}\,dz = a^0 = 1_\mathrm{A}.$$

On montre de la même façon que si $f_k(z) = z^k$, $k \geq 0$, on obtient $\mathcal{H}_a(f_k) = a^k$. Ainsi $\mathcal{H}_a$ coïncide avec le calcul polynomial usuel, dans le cas des fonctions polynomiales.

Posons $r_\lambda(z) = (z - \lambda)^{-1}$ et aussi $r_\lambda(a) = (a - \lambda 1_{\mathrm{A}})^{-1}$ si $a \in \mathrm{A}$ et si $\lambda \notin \sigma(a)$. Un premier objectif est de vérifier que $r_\lambda(a) = \widetilde{r_\lambda}(a)$. On note d'abord que

$$(a - \lambda 1_{\mathrm{A}}) \widetilde{f}(a) = \frac{1}{2\pi \mathrm{i}} \int_\Gamma (a - \lambda 1_{\mathrm{A}}) f(z) (z 1_{\mathrm{A}} - a)^{-1} \, dz$$

en utilisant des limites de sommes de Riemann et le fait que le produit par $a - \lambda 1_{\mathrm{A}}$ est une application linéaire continue de A dans A. On vérifie ensuite que si $f = r_\lambda$,

$$(a - \lambda 1_{\mathrm{A}}) \, r_\lambda(z) (z 1_{\mathrm{A}} - a)^{-1} = -r_\lambda(z) 1_{\mathrm{A}} + (z 1_{\mathrm{A}} - a)^{-1} \,;$$

il en résulte que $\widetilde{r_\lambda}(a) = r_\lambda(a)$ quand $\lambda \notin \sigma(a)$: on peut choisir un ouvert Ω contenant le spectre $\mathrm{K} = \sigma(a)$ et tel que $\lambda \notin \Omega$, puis le cycle Γ dans Ω ; comme $\lambda \notin \Omega$, l'intégrale de r_λ sur Γ est nulle, et on a dit que celle de $(z 1_{\mathrm{A}} - a)^{-1}$ vaut $2\pi \mathrm{i} \, 1_{\mathrm{A}}$. On a donc

$$(a - \lambda 1_{\mathrm{A}}) \, \widetilde{r_\lambda}(a) = 1_{\mathrm{A}}.$$

Ensuite la relation $r_\lambda(a) \, r_\mu(a) = (r_\lambda(a) - r_\mu(a))/(\lambda - \mu)$ permet de montrer que l'image par $\mathcal{H}_a$ de la fonction produit $z \to r_\lambda(z) \, r_\mu(z)$ est égale au produit $r_\lambda(a) \, r_\mu(a)$. En utilisant ensuite le résultat d'approximation donné par le corollaire 6.3, on montre que le calcul fonctionnel holomorphe donne un homomorphisme : si f, g sont holomorphes dans un ouvert Ω qui contient $\mathrm{K} = \sigma(a)$, on introduit l'ouvert ω du corollaire 6.3, puis un cycle Γ dans ω vérifiant les conditions du lemme 6.2 par rapport à ω. On peut approcher f et g uniformément sur $\Gamma^* \subset \omega$ par des combinaisons linéaires de fonctions r_λ, $\lambda \notin \omega$, et on en déduit que l'image par $\mathcal{H}_a$ de la fonction produit fg est égale à $\widetilde{f}(a) \, \widetilde{g}(a)$.

Quand les choses sont suffisamment avancées, on abandonne la notation provisoire $\widetilde{f}(a)$, et on écrit simplement $f(a)$, pour toute fonction f holomorphe dans un voisinage du spectre de a. On montre le *théorème spectral*,

$$\sigma\big(f(a)\big) = f\big(\sigma(a)\big).$$

Remarque. Il est évident que si f est nulle dans un voisinage du spectre de a, alors $f(a) = 0$; mais le fait que f soit nulle *sur le seul spectre* de a n'implique pas que $f(a) = 0$: l'exemple le plus simple est fourni par $f(z) = z$ et un élément non nul a tel que $\sigma(a) = \{0\}$, par exemple une matrice nilpotente dans $\mathrm{A} = \mathrm{M}_n(\mathbb{C})$. On a vu que dans ce cas, $f(a) = a \neq 0$, bien que f soit nulle sur le spectre.

À titre d'exercice très instructif, on pourra calculer $f(\mathrm{J})$ pour une matrice de Jordan J de taille $n \times n$, telle que $\mathrm{J}_{i,i+1} = 1$ si $1 \leq i < n$ et $\mathrm{J}_{i,j} = 0$ si $j - i \neq 1$. On pourra passer ensuite à $f(\mathrm{S}^*)$, où S^* est l'adjoint hilbertien du shift S de $\ell_2(\mathbb{N})$.

7. Appendice B. Représentations des C*-algèbres

7.1. Hermitiens positifs dans le cas abstrait

Soit A une C*-algèbre unitaire ; désignons par $\mathcal{P} = \mathcal{P}_\mathrm{A}$ la classe des éléments hermitiens positifs dans A, c'est-à-dire les hermitiens $b \in \mathrm{A}$ tels que $\sigma(b) \subset [0, +\infty[$, ou bien, ce qui revient au même par calcul fonctionnel, les hermitiens b qui sont le carré a^2 d'un autre hermitien a (Lemme 3.3.1) ; il est clair que l'ensemble $\mathcal{P}$ est stable par multiplication par un réel $s \geq 0$.

On note qu'un élément hermitien a est dans $\mathcal{P}$ si et seulement si pour $s > 0$ assez grand, on a $\|a - s1_\mathrm{A}\| \leq s$. En effet, d'après la proposition 2.2.5, on a

$$\|a - s1_\mathrm{A}\| = \mathrm{rs}(a - s1_\mathrm{A}) = \max\{|t - s| : t \in \sigma(a)\}$$

et puisque $\sigma(a) \subset \mathbb{R}$, on a

$$\max\{|t - s| : t \in \sigma(a)\} \leq s$$

si et seulement si tous les éléments t du spectre de a sont dans l'intervalle $[0, 2s]$, ce qui se produit si et seulement si a est positif et $\|a\| \leq 2s$.

Lemme. *L'ensemble $\mathcal{P}$ est un cône convexe : si a_1, a_2 sont dans $\mathcal{P}$ et si λ_1, λ_2 sont des scalaires ≥ 0, alors $\lambda_1 a_1 + \lambda_2 a_2 \in \mathcal{P}$.*

Preuve.— Si a_1 et a_2 sont dans $\mathcal{P}$, alors les inégalités vraies pour $s_j > \|a_j\|$, $j = 1, 2$

$$\|(a_1 + a_2) - (s_1 + s_2)1_\mathrm{A}\| \leq \|a_1 - s_1 1_\mathrm{A}\| + \|a_2 - s_2 1_\mathrm{A}\| \leq s_1 + s_2$$

impliquent que $a_1 + a_2 \in \mathcal{P}$.

///

Soit $x \in \mathrm{A}$ un élément quelconque, et posons $a = x^* x$; il est clair que a est hermitien, mais on veut montrer que $a \in \mathcal{P}$. Ce fait est évident lorsque $\mathrm{A} = \mathcal{L}(\mathrm{H})$, mais pas dans le cadre abstrait. Ce qui est facile à partir de la remarque précédente, c'est que $x^* x + x\,x^*$ est hermitien positif : on peut écrire $x = c + \mathrm{i} d$ avec c, d hermitiens ; on voit que

$$x^* x + x\,x^* = 2\,(c^2 + d^2),$$

donc $x^* x + x\,x^* \in \mathcal{P}$ d'après le lemme précédent.

Lemme 7.1. *Pour tout $x \in \mathrm{A}$, l'élément hermitien $x^* x$ est positif.*

Preuve.— Supposons que $a = x^* x$ ne soit pas positif, c'est-à-dire que $\sigma(a)$ contienne un point $\lambda_0 < 0$. On va commencer par voir qu'on pourrait alors construire un autre élément y tel que $y^* y$ soit *négatif*. L'idée est de « localiser » x, grâce au calcul fonctionnel, là où a serait négatif. Soit f une fonction continue sur $\mathbb{R}$ telle que $f(\lambda_0) = |\lambda_0|^{-1/2}$, et f nulle sur $[0, +\infty[$; posons $y = x f(a) \in \mathrm{A}$; on a

$$y^* y = \overline{f}(a)\, x^* x\, f(a) = \overline{f}(a)\, a\, f(a),$$

qui est l'image de la fonction $g : t \in \sigma(a) \to t\,|f(t)|^2$ par l'homomorphisme de calcul fonctionnel ; la fonction g est ≤ 0 sur $\sigma(a)$, et $g(\lambda_0) = -1$, donc $y^* y = g(a)$ est négatif, et par le théorème spectral, $-1 = g(\lambda_0)$ est dans le spectre de $y^* y = g(a)$.

Comme $y^* y + y\,y^*$ est positif, il en résulte que $y\,y^* = (y^* y + y\,y^*) + (-y^* y)$ est positif. Mais un exercice classique nous dit que $\sigma(y\,y^*) \setminus \{0\} = \sigma(y^* y) \setminus \{0\}$, donc on doit avoir $-1 \in \sigma(y\,y^*)$, contradiction.

///

Complément. Résolvons ce petit exercice : si 1 n'est pas dans le spectre d'un produit uv, c'est que $1_{\rm A} - uv$ est inversible. Mais alors on vérifie sans peine que

$$1_{\rm A} + v\,(1_{\rm A} - uv)^{-1}u$$

est l'inverse de $1_{\rm A} - vu$. On conclut que $1 \in \sigma(uv)$ si et seulement si $1 \in \sigma(vu)$. Dans la preuve du lemme précédent, on aurait aussi pu dire : on sait qu'il existe u de norme 1 tel que $y^*yu \sim -u$, donc $y\,y^*yu \sim -yu$, ce qui montre que $-1 \in \sigma(y\,y^*)$ car les relations $1 \sim \|y^*yu\| \le \|y^*\|\,\|yu\|$ montrent que $v = yu$ n'est pas nul.

Proposition 7.2. *Soient* A *une* C*-*algèbre unitaire et* $a \in {\rm A}$ *; les conditions suivantes sont équivalentes :*

(*i*) *il existe* $b \in {\rm A}$ *tel que* $a = b^*b$ *;*
(*ii*) *il existe* $b \in {\rm A}$ *tel que* $b = b^*$ *et* $a = b^2$ *;*
(*iii*) *l'élément* a *est hermitien et* $\sigma(a) \subset [0, +\infty[$.

Preuve.— Supposons (*i*) vérifiée ; l'élément $a = b^*b$ est hermitien et son spectre est dans $[0, +\infty[$ d'après le lemme 7.1. L'implication (*ii*) ⇒ (*i*) est évidente. L'équivalence entre (*ii*) et (*iii*) est contenue dans le lemme 3.3.1.

///

7.2. Théorème de représentation dans une algèbre $\mathcal{L}({\rm H})$

Soit A une C*-algèbre unitaire ; si $b \in {\rm A}$ est hermitien positif et $\|b\| \le 1$, le spectre de b est contenu dans $[0, 1]$, donc

$$\sigma(1_{\rm A} - b) = \{1 - \lambda : \lambda \in \sigma(b)\} \subset [0, 1]\,;$$

par conséquent, on a $\|1_{\rm A} - b\| = {\rm rs}(1_{\rm A} - b) \le 1$ par la proposition 2.2.5. Pour les mêmes raisons, pour tout hermitien positif b on a

$$\|1_{\rm A} + b\| = {\rm rs}(1_{\rm A} + b) = 1 + {\rm rs}(b) = 1 + \|b\|.$$

Si a est hermitien, $b = a^2$ est hermitien positif et pour tout t réel

$$\|1_{\rm A} + {\rm i}ta\|^2 = \|(1_{\rm A} - {\rm i}ta)(1_{\rm A} + {\rm i}ta)\| = \|1_{\rm A} + t^2a^2\| = 1 + t^2\|a^2\| = 1 + t^2\|a\|^2.$$

Lemme. *Soient* A *une* C*-*algèbre unitaire, et* f *une forme linéaire continue sur* A *telle que* $\|f\| \le 1$ *et* $f(1_{\rm A}) = 1$ *; alors,* f *est réelle sur les éléments hermitiens,* ≥ 0 *sur les hermitiens positifs et*

$$\forall x \in {\rm A}, \quad f(x^*) = \overline{f(x)}.$$

Preuve.— Supposons que f soit une forme linéaire (complexe) continue sur A, telle que $\|f\| \le 1$ et $f(1_{\rm A}) = 1$. On aura pour tout $a \in {\rm A}$ hermitien et tout t réel

$$|1 + {\rm i}tf(a)| = |f(1_{\rm A} + {\rm i}ta)| \le \|f\|\,\|1_{\rm A} + {\rm i}ta\| \le \|1_{\rm A} + {\rm i}ta\| \le 1 + t^2\|a\|^2.$$

En choisissant t petit tendant vers 0, on conclut que $f(a)$ doit nécessairement être un nombre réel. Si x est quelconque, écrivons $x = c + {\rm i}d$, c, d hermitiens ; alors $x^* = c - {\rm i}d$ et puisque $f(c)$, $f(d)$ sont réels, on a $f(x^*) = f(c) - {\rm i}f(d) = \overline{f(x)}$. Si b est hermitien positif et $\|b\| \le 1$,

$$1 - f(b) = f(1_{\rm A} - b) \le \|1_{\rm A} - b\| \le 1$$

montre que $f(b) \ge 0$.

///

Définition. Une forme linéaire f sur A, de norme 1 et telle que $f(1_{\mathrm{A}}) = 1$, est appelée un *état* de la C*-algèbre.

Lemme 7.3. *Pour tout hermitien positif* $a \in \mathrm{A}$, *il existe un état* f *tel que* $f(a) = \|a\|$.

Preuve.— Si $a \in \mathrm{A}$ est hermitien positif, on a vu que $\|1_{\mathrm{A}} + a\| = 1 + \|a\|$. Par le théorème de Hahn-Banach, il existe une forme linéaire f sur A, de norme 1, telle que

$$f(1_{\mathrm{A}} + a) = \|1_{\mathrm{A}} + a\| = 1 + \|a\|\,;$$

les deux nombres (complexes) $f(1_{\mathrm{A}})$ et $f(a)$ vérifient $|f(1_{\mathrm{A}})| \le \|1_{\mathrm{A}}\| = 1$ et $|f(a)| \le \|a\|$; la seule possibilité pour que leur somme soit $1 + \|a\|$ est que

$$f(1_{\mathrm{A}}) = 1\,; \quad f(a) = \|a\|.$$

La forme linéaire f est donc un état qui répond à la question.

///

Si f est un état sur A, on pose

$$\forall x, y \in \mathrm{A}, \quad \langle x, y \rangle_f = f(y^* x).$$

On a $\langle y, x \rangle_f = f(x^* y) = f((y^* x)^*) = \overline{f(y^* x)} = \overline{\langle x, y \rangle}_f$ par les propriétés des états, et d'autre part $\langle x, x \rangle_f = f(x^* x) \ge 0$ par le lemme 7.1. On obtient donc un produit scalaire sur A, peut-être dégénéré. On note que

$$\|x\|_f^2 := \langle x, x \rangle_f = f(x^* x) \le \|f\|\,\|x^* x\| = \|x\|^2\,;$$

si $a \in \mathrm{A}$ et si $x \in \mathrm{A}$ est tel que $\|x\|_f \le 1$, on obtient par Cauchy-Schwarz

$$\|ax\|_f^2 = f(x^* a^* a x) = \langle a^* a x, x \rangle_f \le \|a^* a x\|_f\,\|x\|_f \le \|a^* a x\|_f.$$

Si a est hermitien, on obtient ainsi de proche en proche pour tout entier $n \ge 1$

$$\|ax\|_f \le \|a^2 x\|_f^{1/2} \le \|a^4 x\|_f^{1/4} \le \ldots \le \|a^{2^n} x\|_f^{2^{-n}} \le \|a^{2^n} x\|^{2^{-n}} \le \|a\|\,\|x\|^{2^{-n}}$$

ce qui entraîne que $\|ax\|_f \le \|a\|$ à la limite. Si a est quelconque,

$$\|ax\|_f^2 \le \|a^* a x\|_f \le \|a^* a\| = \|a\|^2.$$

Il résulte de cette discussion que chaque $a \in \mathrm{A}$ définit un opérateur de multiplication sur l'espace A muni de la semi-norme donnée par f, et que

(1) $$\forall a, x \in \mathrm{A}, \quad \|ax\|_f \le \|a\|\,\|x\|_f.$$

On ne procède pas comme dans les lignes qui suivent quand on n'a pas peur des ensembles d'indices quelconques (non dénombrables) et des passages au quotient. Si on a peur, on suppose A séparable ; on peut alors trouver dans $\mathcal{P}$ une suite dense (a_n) et pour tout $n \geq 1$, un état f_n tel que $f_n(a_n) = \|a_n\|$, d'après le lemme 7.3. Il est clair par densité que pour tout hermitien positif a, on a

$$\sup_n f_n(a) = \|a\|.$$

En particulier l'état

$$\mathrm{F} = \sum_{n=1}^{+\infty} 2^{-n} f_n$$

est non nul sur tout hermitien positif a non nul. Si on remplace l'état f_n par l'état

$$g_n = (1 - 2^{-n}) f_n + 2^{-n}\mathrm{F},$$

on a encore

$$\sup_n g_n(a) = \|a\|$$

mais tous les produits scalaires $\langle\, , \rangle_{g_n}$ sont non dégénérés. Notons H_n le complété de A pour le produit scalaire donné par g_n. En prolongeant, grâce à la continuité donnée par l'inégalité (1), l'opération de multiplication $x \in \mathrm{A} \to ax$, on définit un opérateur de multiplication sur H_n, dont la norme est majorée par $\|a\|$. On introduit l'espace de Hilbert H des suites $\mathbf{x} = (x_n)$ telles que $x_n \in \mathrm{H}_n$ pour tout $n \geq 1$ et $\sum \|x_n\|^2 < +\infty$. À chaque $a \in \mathrm{A}$ on associe un opérateur M_a sur H,

$$\forall \mathbf{x} = (x_n) \in \mathrm{H}, \quad \mathrm{M}_a \mathbf{x} = (a x_n)_{n \geq 1}.$$

Il est clair que $\|\mathrm{M}_a\|_{\mathcal{L}(\mathrm{H})} \leq \|a\|$. Inversement, pour tout $n \geq 1$ considérons l'élément $\mathbf{x}^{(n)} = (x_k^{(n)})_k \in \mathrm{H}$ dont toutes les coordonnées $x_k^{(n)}$ sont nulles sauf $x_n^{(n)} = 1_\mathrm{A}$; on a $\|\mathbf{x}^{(n)}\|_\mathrm{H} = \|1_\mathrm{A}\|_n = g_n(1_\mathrm{A}) = 1$, donc

$$\|\mathrm{M}_a\|^2 \geq \sup_n \|\mathrm{M}_a \mathbf{x}^{(n)}\|^2 = \sup_n \|a\|_n^2 = \sup_n\, g_n(a^*a) = \|a^*a\| = \|a\|^2.$$

L'application $a \to \mathrm{M}_a$ est un $*$-homomorphisme de A dans $\mathcal{L}(\mathrm{H})$, et il est isométrique d'après les lignes qui précèdent. On a donc démontré, mais seulement dans le cas séparable, le théorème de représentation qui suit. La démonstration du cas général est identique pour le fond, mais demande quelques outils abstraits supplémentaires, tels que les sommes directes quelconques d'espaces de Hilbert. Nous ne l'écrirons pas.

Théorème 7.4. *Pour toute* C^**-algèbre unitaire* A*, il existe un espace de Hilbert complexe* H *et un* $*$*-homomorphisme isométrique de* C^**-algèbres unitaires de* A *dans* $\mathcal{L}(\mathrm{H})$.

Théorème 7.5 (Gelfand-Mazur). *Une algèbre de Banach unitaire complexe* A *dans laquelle tout élément non nul est inversible est formée des multiples scalaires de* 1_{A} *(elle est donc de dimension* 1 *sur* $\mathbb{C}$*).*

Preuve. — Soient $a \in \mathrm{A}$ et $\lambda \in \sigma(a)$; puisque $a - \lambda 1_{\mathrm{A}}$ n'est pas inversible, on doit avoir $a - \lambda 1_{\mathrm{A}} = 0$, donc tout élément de A est de la forme $\lambda 1_{\mathrm{A}}$ pour un $\lambda \in \mathbb{C}$.

Soit A une algèbre de Banach unitaire complexe *commutative* ; si I est un idéal maximal de A, il est fermé : en effet, I est contenu dans le fermé complémentaire de l'ouvert G(A) des inversibles ; son adhérence $\overline{\mathrm{I}}$ est encore disjointe de G(A), et c'est encore un idéal, par les propriétés de continuité des opérations. C'est donc un idéal propre plus grand que I, qui doit être égal à I par maximalité, donc $\mathrm{I} = \overline{\mathrm{I}}$ est fermé. De plus, I est un hyperplan : en effet, le quotient B = A/I est une algèbre de Banach dans laquelle tout élément non nul est inversible, donc B est de dimension un et I de codimension 1. Considérons la forme linéaire χ dont le noyau est I et qui vérifie la condition $\chi(1_{\mathrm{A}}) = 1$; cette forme linéaire est continue parce que son noyau I est fermé, et elle est de plus multiplicative : si $a \in \mathrm{A}$ et si $\chi(a) \neq 0$, définissons g par $g(x) = \chi(ax)/\chi(a)$. Alors g est une forme linéaire qui s'annule sur I, donc elle est proportionnelle à χ, mais $\chi(1_{\mathrm{A}}) = g(1_{\mathrm{A}}) = 1$, donc $\chi = g$ et $\chi(ax) = \chi(a)\chi(x)$ pour tout $x \in \mathrm{A}$. Une forme linéaire χ avec ces propriétés est appelée *caractère* de l'algèbre A.

Définition. Un *caractère* χ sur A est une forme linéaire continue non nulle sur A qui est aussi multiplicative : $\chi(ab) = \chi(a)\chi(b)$, pour tous $a, b \in \mathrm{A}$.

Comme χ n'est pas nulle et que $\chi(1_{\mathrm{A}} b) = \chi(1_{\mathrm{A}})\,\chi(b)$ pour tout $b \in \mathrm{A}$, il en résulte que $\chi(1_{\mathrm{A}}) = 1$. L'ensemble des caractères de A s'appelle le *spectre de* A, et il est noté Sp(A) ; c'est un sous-ensemble de la sphère unité du dual topologique A^* de A ; pour voir ceci, on note que la suite $(a^n)_{n \geq 0}$ est bornée quand $\|a\| \leq 1$, ce qui entraîne que $(\chi(a^n)) = (\chi(a)^n)$ est bornée, donc $|\chi(a)| \leq 1$, et $\|\chi\| = 1$ car $\chi(1_{\mathrm{A}}) = 1$. En particulier, χ est un état et on sait que $\chi(x^*) = \overline{\chi(x)}$ pour tout $x \in \mathrm{A}$.

Proposition 7.6. *Soit* A *une algèbre de Banach unitaire commutative ; un élément* $a \in \mathrm{A}$ *est inversible dans* A *si et seulement si* $\chi(a) \neq 0$ *pour tout caractère* χ *de* A *; de plus*

$$\sigma(a) = \{\chi(a) : \chi \in \mathrm{Sp(A)}\}.$$

Preuve. — Il est clair que $\chi(a) \neq 0$ quand a est inversible, car $\chi(a)\,\chi(a^{-1}) = \chi(1_{\mathrm{A}}) = 1$; si a n'est pas inversible, l'ensemble $a\mathrm{A}$ est un idéal propre dans A, qui est contenu dans un idéal maximal I. Si χ est le caractère dont le noyau est I, on aura $\chi(a) = 0$. Il en résulte que pour tout $a \in \mathrm{A}$,

$$\sigma(a) = \{\chi(a) : \chi \in \mathrm{Sp(A)}\}\,;$$

en effet, puisque $\chi(a - \chi(a)1_{\mathrm{A}}) = \chi(a) - \chi(a) = 0$, on voit que $a - \chi(a)1_{\mathrm{A}}$ n'est pas inversible, donc $\chi(a) \in \sigma(a)$ pour tout $\chi \in \mathrm{Sp(A)}$; inversement, si $a - \lambda 1_{\mathrm{A}}$ n'est pas inversible, il existe un caractère χ tel que $\chi(a - \lambda 1_{\mathrm{A}}) = 0$, donc $\lambda = \chi(a)$.

///

Exemple d'application : l'algèbre de Wiener

L'algèbre de Wiener W est l'algèbre des fonctions continues sur $\mathbb{T}$ ayant des coefficients de Fourier absolument sommables ; si on voit $\mathbb{T}$ comme le quotient $\mathbb{R}/2\pi\mathbb{Z}$, une fonction $f \in \mathrm{W}$ est de la forme

$$\forall \theta \in \mathbb{R}, \quad f(\theta) = \sum_{n \in \mathbb{Z}} a_n \,\mathrm{e}^{\mathrm{i} n\theta}$$

avec

$$\|f\|_{\mathrm{W}} = \sum_{n \in \mathbb{Z}} |a_n| < \infty.$$

Le produit de l'algèbre W est le produit ponctuel des fonctions sur $\mathbb{T}$ (ou sur $\mathbb{R}$, dans le modèle ci-dessus), mais la norme de W est la norme $\ell_1(\mathbb{Z})$ des coefficients de Fourier. Cette algèbre est clairement isomorphe à l'algèbre $\ell_1(\mathbb{Z})$ munie du produit de convolution des suites.

Une fonction f de W *est inversible dans* W *si et seulement si la fonction f ne s'annule pas sur $\mathbb{T}$.*

Cela provient du fait que les caractères de cette algèbre de Banach commutative sont les évaluations aux points de $\mathbb{T}$: soit χ un caractère sur W et soit $\lambda = \chi(\mathrm{e}^{\mathrm{i}\theta})$, où $\mathrm{e}^{\mathrm{i}\theta}$ désigne (incorrectement) la fonction $\theta \to \mathrm{e}^{\mathrm{i}\theta} \in \mathrm{W}$; puisque la famille $(\mathrm{e}^{\mathrm{i}n\theta})_{n\in\mathbb{Z}}$ est bornée dans W, il en résulte que la famille $(\lambda^n = \chi(\mathrm{e}^{\mathrm{i}n\theta}))_{n\in\mathbb{Z}}$ est bornée, donc $|\lambda| = 1$ et on peut écrire $\lambda = \mathrm{e}^{\mathrm{i}x}$ pour un certain réel x. Pour toute fonction $f(\theta) = \sum_{n\in\mathbb{Z}} a_n \,\mathrm{e}^{\mathrm{i}n\theta}$ dans W,

$$\chi(f) = \sum_{n \in \mathbb{Z}} a_n \lambda^n = \sum_{n \in \mathbb{Z}} a_n \,\mathrm{e}^{\mathrm{i}nx} = f(x).$$

Si f ne s'annule pas sur $\mathbb{T}$, on voit que $\chi(f) \neq 0$ pour tout caractère χ, donc f est inversible dans W d'après la proposition 7.6.

///

Le spectre Sp(A) est fermé pour la topologie de la convergence simple sur A (c'est la topologie w^* sur le dual) ; en effet, la propriété de multiplicativité des caractères χ et le fait que $\chi(1_{\mathrm{A}}) = 1$ passent à la limite simple. Puisque la boule unité du dual de A est compacte pour la convergence simple (Banach-Alaoglu), le spectre Sp(A) est donc compact pour cette topologie. Définissons une application de A dans l'espace C(Sp(A)) des fonctions complexes continues sur le compact Sp(A) ; pour tout $a \in \mathrm{A}$, soit $j(a)$ la fonction continue sur Sp(A) définie par

$$\forall \chi \in \mathrm{Sp(A)}, \quad j(a)(\chi) = \chi(a)\,;$$

cette application j n'est pas injective en général. Mais dans le cas où A est une C*-algèbre, cette application j est isométrique, et conduit à l'important théorème qui suit.

Théorème 7.7 (Gelfand). *Toute C*-algèbre unitaire commutative est isométriquement isomorphe à l'algèbre des fonctions continues sur* $\mathrm{K} = \mathrm{Sp(A)}$.

Preuve. — Si a est hermitien, on a $\mathrm{rs}(a) = \|a\|$, donc $\|a\|$ ou bien $-\|a\|$ est dans le spectre de a et il existe $\chi \in \mathrm{Sp(A)}$ tel que $|\chi(a)| = \|a\|$. On a vu que pour tout $x \in \mathrm{A}$, on a $\chi(x^*) = \overline{\chi(x)}$ par la propriété d'état. Ceci implique que pour tout $x \in \mathrm{A}$, il existe χ tel que

$$|\chi(x)|^2 = \chi(x^*x) = \|x^*x\| = \|x\|^2,$$

montrant ainsi que l'application $j : \mathrm{A} \to \mathrm{C(Sp(A))}$ est isométrique. L'image de A dans C(Sp(A)) est une sous-algèbre fermée par conjugaison complexe, et qui sépare évidemment les points de Sp(A), ce qui implique que j est surjective par le théorème de Stone-Weierstrass.

///

Remarque. Le lecteur attentif aura noté qu'on vient de redémontrer d'une façon un peu différente l'existence du calcul fonctionnel continu.

Proposition. *Soit* B *une* C^**-algèbre unitaire ; pour toute* C^**-algèbre unitaire* C *et pour tout* $*$*-homomorphisme* $\varphi : \mathrm{B} \to \mathrm{C}$, *on a* $\|\varphi\| \le 1$. *Si* φ *est injectif, alors* φ *est isométrique.*

Preuve. — La première partie est assez facile et routinière ; on note d'abord que l'image par φ d'un hermitien $a \in \mathrm{B}$ est un hermitien : comme φ est un $*$-homomorphisme, $\varphi(a)^* = \varphi(a^*) = \varphi(a)$ est hermitien ; ensuite, l'image par φ d'un hermitien positif dans B est un hermitien positif dans C : si b est hermitien positif, on peut écrire $b = a^2$, avec a hermitien, et $\varphi(b) = \varphi(a)^2$ est hermitien positif. Soit enfin $x \in \mathrm{B}$ tel que $\|x\| \le 1$; alors $b = x^*x$ est hermitien positif de norme ≤ 1, son spectre est dans $[0, 1]$ et il en est de même pour $1_\mathrm{B} - b$, qui est donc aussi hermitien positif ; les deux images $\varphi(b)$ et $1_\mathrm{C} - \varphi(b) = \varphi(1_\mathrm{B} - b)$ sont hermitiennes positives, ce qui implique que le spectre de $\varphi(b)$ est dans $[0, 1]$, donc

$$\|\varphi(x)\|^2 = \|\varphi(x)^*\varphi(x)\| = \|\varphi(b)\| = \mathrm{rs}(\varphi(b)) \le 1$$

d'après la proposition 2.2.5. On a bien $\|\varphi\| \le 1$.

Pour la deuxième partie, on suppose de plus que φ est injective, et on note qu'il suffirait de pouvoir remplacer la fin de la ligne précédente par

$$\|\varphi(x)\|^2 = \|\varphi(x)^*\varphi(x)\| = \|\varphi(b)\| = \|b\| = \|x\|^2.$$

Il suffit donc de montrer que φ préserve la norme des hermitiens. Étant donné un élément hermitien $a \in \mathrm{B}$, on peut considérer la sous-algèbre unitaire A de B engendrée par a ; c'est une C^*-algèbre commutative, et son image dans C est une sous-algèbre commutative. La question se ramène donc complètement au cas commutatif. Si B et C sont commutatives, on voit que $\chi \circ \varphi$ est un caractère de B, pour tout caractère χ de C. Désignons par g cette application (continue) du compact $\mathrm{K_C} = \mathrm{Sp(C)}$ dans le compact $\mathrm{K_B} = \mathrm{Sp(B)}$, définie par

$$\forall \chi \in \mathrm{Sp(C)}, \quad g(\chi) = \chi \circ \varphi.$$

Si g est surjective, on obtient le résultat facilement : pour tout $a \in \mathrm{B}$, il existe un caractère ψ sur B tel que $\|a\| = |\psi(a)|$, d'après la proposition 7.6 ; comme il existe $\chi \in \mathrm{Sp(C)}$ tel que $\psi = g(\chi)$, on a

$$\|\varphi(a)\| \ge |\chi(\varphi(a))| = |\psi(a)| = \|a\|.$$

Dans le cas contraire, $g(\mathrm{K_C})$ est un fermé de $\mathrm{K_B}$, différent de $\mathrm{K_B}$. On peut trouver une fonction continue f sur $\mathrm{K_B}$, nulle sur le fermé $g(\mathrm{K_C})$ mais qui n'est pas identiquement nulle. D'après le théorème de représentation, cette fonction provient d'un élément $a \in \mathrm{B}$, par la formule $f = j(a)$ qui se traduit en $f(\psi) = \psi(a)$ pour tout $\psi \in \mathrm{K_B}$; pour tout caractère χ de C, on aura alors

$$\chi(\varphi(a)) = g(\chi)(a) = f(g(\chi)) = 0,$$

ce qui montre que $\varphi(a) = 0$, et contredit l'injectivité de φ.

///

Remarque. Pour définir le calcul fonctionnel d'un élément normal, on a imposé la relation $\varphi_a(\overline{\zeta_\mathrm{K}}) = a^*$; c'est vrai qu'on s'est servi de cette propriété pour construire φ_a. Mais supposons donné un homomorphisme d'algèbres de Banach unitaires complexes de C(K) dans $\mathcal{L}(\mathrm{H})$, et dont l'image soit contenue dans la sous-algèbre engendrée par T normal et son adjoint ; on sait maintenant que cette sous-algèbre, commutative, est isomorphe à un espace C(L), donc on considérera que φ est à valeurs dans C(L) ; pour toute fonction continue f de module 1 sur K, on aura que $\mathrm{U} = \varphi(f)$ est une fonction sur le compact L dont toutes les puissances entières positives ou négatives sont uniformément bornées (par la norme de φ) ; il en résulte que U est unitaire et $\varphi(\overline{f}) = \varphi(f)^*$. Ensuite les combinaisons linéaires de fonctions de module 1 forment une algèbre dense, par Stone-Weierstrass, ce qui prouve que $\varphi(\overline{\zeta_\mathrm{K}})$ n'a pas le choix (les fonctions de module un séparent les points de $\mathrm{K} \subset \mathbb{C}$: il suffit de considérer toutes les fonctions sur $\mathbb{R}^2 \simeq \mathbb{C}$ de la forme

$$\varphi_{u,v}(x, y) = \mathrm{e}^{\mathrm{i}(ux+vy)},$$

u, v variant dans $\mathbb{R}$).

8. Appendice C. Perturbations strictement singulières

8.1. Suites basiques

Définition. On dit qu'une suite $(e_n)_{n\geq 0}$ de vecteurs d'un espace de Banach X est une *base de Schauder normalisée* de X si tous les vecteurs sont de norme 1 et si tout vecteur $x \in \mathrm{X}$ admet une représentation unique sous la forme d'une série convergente

$$x = \sum_{n=0}^{+\infty} a_n e_n,$$

où les (a_n) sont des scalaires. On peut associer à chaque vecteur x le sup des normes des sommes partielles de la série précédente ; on obtient ainsi une norme sur X, plus grande que la norme initiale, pour laquelle l'espace est encore complet. Il en résulte par les grands principes que cette norme est équivalente à la norme initiale, ce qui signifie que les projections P_n définies pour tout entier $n \geq 0$ par

$$\mathrm{P}_n x = \sum_{j=0}^{n} a_j e_j$$

sont uniformément bornées par une constante M, qu'on appelle la constante de basicité de la base (e_n).

On dit qu'une suite (e_n) est une *suite basique normalisée* si elle est une base de Schauder normalisée pour l'espace vectoriel fermé qu'elle engendre.

Exercice : Système de Schauder. On considère la fonction continue φ définie sur $\mathbb{R}$ par $\varphi(t) = 0$ hors de $[0, 1]$ et $\varphi(t) = 1 - |2t - 1|$ dans l'intervalle $[0, 1]$ (fonction «triangle»). Pour chaque entier $n \geq 0$ on considère les 2^n fonctions $\varphi_{n,j}(t) = \varphi(2^n t - j)$, où $j = 0, \ldots, 2^n - 1$. Montrer que cette famille de fonctions, ordonnée suivant l'ordre lexicographique des (n, j), forme une base de Schauder de l'espace des fonctions continues sur $[0, 1]$ nulles en 0 et en 1.

Algorithme pour la construction d'une suite basique

On suppose que X est un espace de Banach de dimension infinie. On se donne une suite bornée de nombres $c_n > 1$, par exemple $c_n = 2$ pour tout n, ou bien une suite (c_n) telle que $\lim_n c_n = 1$.

En commençant avec $n = 0$ et $\mathrm{A}_0 = \{0_{\mathrm{X}^*}\}$, on effectue l'algorithme suivant.

- A_n est un ensemble fini de formes linéaires sur X, de norme ≤ 1.
- Le sous-espace

$$\mathrm{X}_{n+1} = \bigcap_{\xi \in \mathrm{A}_n} \ker \xi = \{x \in \mathrm{X} : \forall x^* \in \mathrm{A}_n, \quad x^*(x) = 0\}$$

est un sous-espace de codimension finie de X (comme intersection finie d'hyperplans).

- Le vecteur e_{n+1} est un vecteur de norme un choisi dans X_{n+1} (le choix est quelconque ; cette souplesse sera utilisée plus loin).
- L'espace de dimension finie $E_{n+1} = \mathrm{Vect}(e_1, \ldots, e_{n+1})$ peut être approximativement normé par un nouvel ensemble fini A_{n+1} de formes linéaires de norme ≤ 1, de façon que $A_n \subset A_{n+1}$ et que

$$\|x\| \leq c_{n+1} \max\{|x^*(x)| : x^* \in A_{n+1}\}$$

pour tout $x \in E_{n+1}$.
- Retour au départ avec $n+1$ au lieu de n.

On initialise avec $A_0 = \{0\}$ de sorte que $X_1 = X$.

Lemme. *Tout vecteur y de l'adhérence Y de la réunion $Y_0 = \bigcup_{n=0}^{+\infty} E_n$ peut s'écrire de façon unique comme une série convergente de la forme*

$$y = \sum_{n=1}^{+\infty} a_n e_n.$$

De plus, $|a_n| \leq (c_n + c_{n-1}) \|y\|$ pour tout $n \geq 2$, et $|a_1| \leq c_1 \|y\|$.

Preuve. — Fixons un entier $m \geq 1$. Soit $y \in Y_0$; alors $y \in E_n$ pour un certain n, qu'on peut supposer $\geq m$; écrivons $y = \sum_{j \leq n} a_j e_j$ et posons $z = P_m y = \sum_{j \leq m} a_j e_j$; on voit que $x^*(y) = x^*(z)$ pour tout $x^* \in A_m$, puisque $e_{m+1}, \ldots, e_n$ ont été choisis dans le noyau des éléments de A_m. Par conséquent

$$\|y\| \geq \max\{|x^*(y)| : x^* \in A_m\} = \max\{|x^*(z)| : x^* \in A_m\} \geq c_m^{-1} \|z\|,$$

donc $\|z\| \leq c_m \|y\|$. Ceci montre que la projection naturelle P_m de Y_0 sur E_m est de norme $\leq c_m$, ce qui permet de la prolonger à Y par continuité. La suite (P_m) de ces projections est donc bornée, et converge vers l'identité sur le sous-espace dense Y_0, donc aussi sur Y. On a $P_m P_n = P_m$ si $m < n$; si $y \in Y$, les développements successifs des $P_m y$ font intervenir des coefficients $a_1, \ldots, a_n, \ldots$ tels que

$$P_m y = \sum_{j=1}^{m} a_j e_j.$$

La convergence de $P_m y$ vers y signifie que la série $\sum_j a_j e_j$ converge, et $y = \sum_{j=1}^{+\infty} a_j e_j$. De plus le coefficient a_n est obtenu par différence de P_n et P_{n-1},

$$a_n e_n = P_n y - P_{n-1} y,$$

donc $|a_n| \leq (c_n + c_{n-1}) \|y\|$ quand $n \geq 2$. On en déduit en passant l'indépendance linéaire des vecteurs : si $v = \sum_{j=1}^{m} b_j e_j = 0$, on aura $|b_j| \leq 0$.

Si $y = \sum_{j=1}^{+\infty} b_j e_j$ est une autre représentation du même vecteur, on verra que $P_m y = \sum_{j=1}^{m} b_j e_j$, ce qui donne $b_j = a_j$ par l'indépendance des vecteurs.

///

Corollaire 8.1. *Dans tout espace normé de dimension infinie, et pour tout $\varepsilon > 0$, on peut trouver une suite de vecteurs (e_n) telle que pour tous $m \neq n$ on ait*

$$\|e_m\| = \|e_n\| = 1, \quad \|e_m - e_n\| > 1 - \varepsilon.$$

Preuve.— On construit une suite basique (e_n) en prenant $c_n = (1-\varepsilon)^{-1} > 1$ pour tout n. On a vu que les projections P_n ont une norme majorée par c_n, donc pour $m < n$

$$1 = \|e_m\| = \|\mathrm{P}_m(e_m - e_n)\| \leq c_m \, \|e_m - e_n\|.$$

///

Remarque. Dans tout Hilbert de dimension infinie on peut trouver une suite infinie (e_n) telle que $\|e_m - e_n\| = \sqrt{2}$ quand $n \neq m$. Exercice facile : on ne peut pas aller plus loin dans un Hilbert.

Dans ℓ_1 on peut trouver une distance 2, le maximum théorique.

Corollaire. *La restriction d'un opérateur compact à un sous-espace de dimension infinie n'est jamais un plongement.*

Preuve.— Soient T un opérateur compact de X dans Y, et X_0 un sous-espace de dimension infinie de X ; on choisit une suite (e_n) comme ci-dessus dans le sous-espace. Si $\mathrm{T}_{|\mathrm{X}_0}$ était un plongement, il existerait une constante $c > 0$ telle que pour $m \neq n$

$$\|\mathrm{T}(e_m - e_n)\| \geq c \, \|e_m - e_n\| \geq c(1 - \varepsilon).$$

On aurait ainsi une suite *infinie* $(\mathrm{T}e_n)$ de points écartés, contenue dans le compact $\mathrm{K} = \overline{\mathrm{T}(\mathrm{B}_\mathrm{X})}$, ce qui est impossible.

///

8.2. Perturbations strictement singulières

Définition. Soient X, Y deux espaces de Banach ; on dira que $\mathrm{T} \in \mathcal{L}(\mathrm{X}, \mathrm{Y})$ est *strictement singulier* si la restriction de T à un sous-espace de dimension infinie $\mathrm{X}_0 \subset \mathrm{X}$ n'est jamais un plongement. Autrement dit, pour tout X_0 de dimension infinie et tout $\varepsilon > 0$, il existe $x_0 \in \mathrm{X}_0$ tel que

$$\|\mathrm{T}x_0\| < \varepsilon \, \|x_0\|.$$

On vient de voir que les opérateurs compacts sont strictement singuliers.

Lemme 8.2. *Pour que* $\mathrm{T} \in \mathcal{L}(\mathrm{X}, \mathrm{Y})$ *ne soit pas un presque-plongement, il faut et il suffit que pour tout* $\varepsilon > 0$, *il existe un sous-espace* $\mathrm{Z} \subset \mathrm{X}$ *de dimension infinie tel que la restriction de* T *vérifie* $\|\mathrm{T}_{|\mathrm{Z}}\| \leq \varepsilon$.

Preuve. — Si T n'est pas un presque-plongement, on sait que pour toute constante $c > 0$ et tout sous-espace V de codimension finie dans X, il existe un vecteur $v \in \mathrm{V}$ avec $\|\mathrm{T}v\| < c\|v\|$. On peut inclure cette information dans l'algorithme de sélection de suite basique : au moment de choisir e_n (de norme 1) dans l'espace de codimension finie X_n, on peut imposer que $\|\mathrm{T}e_n\| < 2^{-n-2}\varepsilon$.

Si Z désigne l'espace vectoriel fermé engendré par la suite (e_n) et si $x \in \mathrm{Z}$, on aura

$$x = \sum_{n=1}^{+\infty} a_n e_n, \quad \mathrm{T}x = \sum_{n=1}^{+\infty} a_n \mathrm{T}e_n$$

avec $|a_n| \le 4\,\|x\|$ (si on a pris $c_n = 2$ dans l'algorithme de construction de suite basique),

$$\|\mathrm{T}x\| \le \sum_{n=1}^{+\infty} |a_n|\,\|\mathrm{T}e_n\| \le 4\,\varepsilon\Bigl(\sum_{n=1}^{+\infty} 2^{-n-2}\Bigr)\|x\| = \varepsilon\,\|x\|.$$

L'implication inverse est évidente.

///

Lemme. *Si* $\mathrm{T} \in \mathcal{L}(\mathrm{X}, \mathrm{Y})$ *est un opérateur de Fredholm et si* $\mathrm{S} \in \mathcal{L}(\mathrm{X}, \mathrm{Y})$ *est strictement singulier, la somme* $\mathrm{T} + \mathrm{S}$ *est un presque-plongement de* X *dans* Y*.*

Preuve. — Il existe une constante $c > 0$ et un sous-espace $\mathrm{X}_1 \subset \mathrm{X}$ de codimension finie tels que $\|\mathrm{T}x\| \ge c\,\|x\|$, pour tout $x \in \mathrm{X}_1$; choisissons ε tel que $0 < \varepsilon < c$; si $\mathrm{T}+\mathrm{S}$ n'était pas un presque-plongement, il existerait un sous-espace Z_1 de dimension infinie tel que $\|(\mathrm{T}+\mathrm{S})z\| < \varepsilon\|z\|$, pour tout $z \in \mathrm{Z}_1$. Alors $\mathrm{Z} = \mathrm{Z}_1 \cap \mathrm{X}_1$ est de dimension infinie, et $\mathrm{S}_{|\mathrm{Z}}$ serait un plongement d'après l'inégalité triangulaire, puisque $c - \varepsilon > 0$ et que

$$\forall z \in \mathrm{Z}, \quad \|\mathrm{S}z\| \ge \|\mathrm{T}z\| - \|(\mathrm{T}+\mathrm{S})z\| \ge (c - \varepsilon)\,\|z\|\,;$$

ceci contredit le caractère strictement singulier de S.

///

Théorème. *Si* $\mathrm{T} \in \mathcal{L}(\mathrm{X}, \mathrm{Y})$ *est un opérateur de Fredholm et si* $\mathrm{S} \in \mathcal{L}(\mathrm{X}, \mathrm{Y})$ *est strictement singulier, la somme* $\mathrm{T} + \mathrm{S}$ *est Fredholm, de même indice que* T.

Preuve. — Le principe de la démonstration est identique à celui qu'on a donné pour les perturbations compactes : on passe de T à $\mathrm{T}+\mathrm{S}$ en suivant le chemin continu $u \to \mathrm{T}+u\mathrm{S}$, $u \in [0, 1]$, chemin continu de $[0, 1]$ dans $\mathcal{L}(\mathrm{X}, \mathrm{Y})$. Tous les éléments sont semi-Fredholm, l'indice est constant par la proposition 4.3.4.

///

9. Appendice D. Théorème des isomorphismes

On dit qu'un sous-ensemble non-vide C d'un espace vectoriel est *symétrique* si $\mathrm{C} = -\mathrm{C}$.

Lemme 9.1. *Si* C *est un convexe symétrique d'intérieur non vide dans un espace vectoriel normé, il existe* $r > 0$ *tel que la boule ouverte* $\mathrm{B}(0, r)$, *de rayon* r *et centrée à l'origine, soit contenue dans* C.

Preuve.— Par hypothèse, il existe $x \in \mathrm{C}$ et $r > 0$ tels que $\mathrm{B}(x, r) \subset \mathrm{C}$; pour tout vecteur u tel que $\|u\| < r$, on sait alors que $x + u, x - u \in \mathrm{C}$; comme C est symétrique, on a aussi $-x + u \in \mathrm{C}$ et

$$u = \frac{1}{2}\Big((x + u) + (-x + u)\Big) \in \mathrm{C}.$$

Ainsi, C contient tous les éléments u de $\mathrm{B}(0, r)$.

///

Théorème 9.2 de l'application ouverte. *Soient* X, Y *deux espaces de Banach et* T *une application linéaire continue* surjective *de* X *sur* Y *; il existe un nombre* K *vérifiant la propriété suivante : pour tout* $y \in \mathrm{Y}$, *il existe* $x \in \mathrm{X}$ *tel que* $\mathrm{T}x = y$ *et* $\|x\| \le \mathrm{K}\,\|y\|$. *L'application* T *est une* application ouverte *: pour tout ouvert* U *de* X, *l'image* $\mathrm{T(U)}$ *est un ouvert de* Y.

Preuve.— L'espace X est réunion des multiples entiers $n\,\mathrm{B_X}$ de la boule unité fermée $\mathrm{B_X}$, $n \ge 1$; puisque l'application T est surjective, l'espace Y est réunion des images $\mathrm{T}(n\,\mathrm{B_X}) = n\,\mathrm{T(B_X)}$, et *a fortiori* réunion des adhérences,

$$\mathrm{Y} = \bigcup_{n=1}^{+\infty} n\,\overline{\mathrm{T(B_X)}}.$$

D'après le théorème de Baire, l'un au moins de ces ensembles fermés $n\,\overline{\mathrm{T(B_X)}}$ a un intérieur non vide, et comme ils sont tous des homothétiques de $\overline{\mathrm{T(B_X)}}$, on déduit que $\overline{\mathrm{T(B_X)}}$ est d'intérieur non vide ; d'après le lemme 9.1, il existe $r > 0$ tel que le convexe symétrique $\mathrm{C} = \overline{\mathrm{T(B_X)}}$ contienne la boule ouverte $\mathrm{B}(0, r)$ dans Y, et comme ce convexe C est fermé on a aussi

$$r\,\mathrm{B_Y} = \overline{\mathrm{B}(0, r)} \subset \overline{\mathrm{T(B_X)}}\,; \tag{1}$$

si y est un vecteur non nul de Y, le vecteur $y_1 = r\,\|y\|^{-1} y$ est dans $r\,\mathrm{B_Y}$, donc pour tout $\beta > 0$ il existe d'après (1) un vecteur $x_1 \in \mathrm{X}$ tel que $\|x_1\| \le 1$ et $\|y_1 - \mathrm{T}x_1\| < \beta$; si on pose $\mathrm{M} = 1/r$, il en résulte par homogénéité que pour tout $y \in \mathrm{Y}$ et tout $\varepsilon > 0$, il existe $x \in \mathrm{X}$ tel que

$$\|y - \mathrm{T}x\| < \varepsilon \quad \text{et} \quad \|x\| \le \mathrm{M}\,\|y\|. \tag{2}$$

On va montrer qu'en fait : pour tout $y \in \mathrm{Y}$ et tout $\varepsilon > 0$, il existe $x \in \mathrm{X}$ tel que

$$y = \mathrm{T}x \quad \text{et} \quad \|x\| \le (\mathrm{M} + \varepsilon)\,\|y\|. \tag{3}$$

Si $y = 0_{\mathrm{Y}}$, il suffit de prendre $x = 0_{\mathrm{X}}$, donc on va supposer y non nul. La stratégie consiste à itérer l'opération indiquée dans (2). Donnons nous une suite (ε_n) de réels > 0 telle que $\sum_{n=0}^{+\infty} \varepsilon_n < \varepsilon \, \|y\|$. En appliquant la propriété (2) à y et ε_0, on trouve un vecteur $x_0 \in \mathrm{X}$ tel qu'en posant $w_0 = y - \mathrm{T}x_0$, on ait

$$y = \mathrm{T}x_0 + w_0, \quad \|w_0\| < \varepsilon_0 \ \text{ et } \ \|x_0\| \le \mathrm{M} \, \|y\|.$$

En appliquant (2) à w_0 et ε_1, on peut écrire $w_0 = \mathrm{T}v_1 + w_1$, avec $\|w_1\| < \varepsilon_1$ et avec $\|v_1\| \le \mathrm{M} \, \|w_0\| \le \mathrm{M} \, \varepsilon_0$. En remplaçant w_0 par sa nouvelle expression, on obtient

$$y = \mathrm{T}x_0 + w_0 = \mathrm{T}(x_0 + v_1) + w_1.$$

On continue de remplacer l'erreur w_n par $\mathrm{T}v_{n+1} + w_{n+1}$, de façon que $\|w_{n+1}\| < \varepsilon_{n+1}$ et $\|v_{n+1}\| \le \mathrm{M} \, \|w_n\| \le \mathrm{M} \, \varepsilon_n$ pour tout $n \ge 0$. On obtient de proche en proche

$$y = \mathrm{T}(x_0 + v_1 + \cdots + v_{n+1}) + w_{n+1}. \tag{4}$$

Puisque $\|v_{n+1}\| \le \mathrm{M} \, \varepsilon_n$ pour tout n, on voit que la série $x_0 + \sum v_n$ converge dans l'espace de Banach X ; si on pose

$$x = x_0 + v_1 + \cdots + v_n + \cdots \in \mathrm{X}$$

et si on tient compte du fait que $\lim w_{n+1} = 0_{\mathrm{Y}}$, on déduit de (4) que $y = \mathrm{T}x$. Enfin,

$$\|x\| \le \|x_0\| + \sum_{n=0}^{+\infty} \|v_{n+1}\| \le \mathrm{M} \, \|y\| + \sum_{n=0}^{+\infty} \varepsilon_n \le (\mathrm{M} + \varepsilon) \, \|y\|.$$

On a obtenu le premier résultat annoncé dans le théorème, avec $\mathrm{K} = \mathrm{M} + \varepsilon$.

Soit U un ouvert de X ; si y est un point de T(U), il existe $x \in \mathrm{U}$ tel que $y = \mathrm{T}x$, et il existe $s > 0$ tels que $\mathrm{B}(x, s) \subset \mathrm{U}$. On constate alors que la boule $\mathrm{B}(y, s/\mathrm{K})$ est contenue dans T(U) : si $\|w\| < s/\mathrm{K}$, il existe un vecteur $v \in \mathrm{X}$ tel que $\mathrm{T}v = w$ et $\|v\| < s$; le vecteur $x + v$ est dans U, et $\mathrm{T}(x + v) = y + w \in \mathrm{T(U)}$. On a montré que pour tout $y \in \mathrm{T(U)}$, il existe une boule ouverte contenant y et contenue dans T(U), donc T(U) est ouvert dans Y, et l'application T est ouverte.

///

Remarque. En général, il est *impossible* de trouver une application *linéaire* σ de relèvement de l'application surjective T, c'est-à-dire une application linéaire de Y dans X telle que $\mathrm{T} \circ \sigma = \mathrm{Id}_{\mathrm{Y}}$. Avec pas mal de travail, on peut trouver un relèvement *continu* (on utilise le fait que les espaces métriques admettent des partitions de l'unité continues).

Corollaire 9.3 (théorème des isomorphismes). *Soient* X, Y *deux espaces de Banach et* T *une application linéaire continue* bijective *de* X *sur* Y *; alors* T *est un isomorphisme, c'est-à-dire que l'application linéaire* T^{-1} *est continue de* Y *dans* X.

Preuve. — Soit K la constante donnée par le théorème précédent : si $y \in \mathrm{Y}$, il existe un et un seul $x \in \mathrm{X}$ tel que $y = \mathrm{T}x$, donc ce vecteur $x = \mathrm{T}^{-1}y$ vérifie $\|x\| \le \mathrm{K} \, \|y\|$; on a bien montré que

$$\forall y \in \mathrm{X}, \quad \|\mathrm{T}^{-1}y\| \le \mathrm{K} \, \|y\|.$$

///

Corollaire 9.4 (théorème du graphe fermé). *Soient* X, Y *deux espaces de Banach et* T *une application linéaire de* X *dans* Y *dont le graphe est fermé dans* $X \times Y$ *; alors* T *est continue de* X *dans* Y.

Preuve. — Si on munit le produit $X \times Y$ de la norme

$$\|(x, y)\| = \|x\|_X + \|y\|_Y,$$

la topologie normée correspondante est la topologie produit, et $X \times Y$ devient un espace de Banach. Le graphe de T est un sous-espace vectoriel fermé de $X \times Y$, donc Gr(T) lui-même est un espace de Banach. L'application π de Gr(T) dans X définie pour tout $(x, T(x)) \in \mathrm{Gr}(T)$ par

$$\pi(x, T(x)) = x$$

est continue (de norme ≤ 1) et bijective. D'après le corollaire précédent, son inverse $x \to (x, T(x))$ est continu, ce qui implique que $x \to T(x)$ est continue de X dans Y.

///

Index alphabétique

Index des notations