ANALYSE FONCTIONNELLE ET THÉORIE SPECTRALE

MT404

Année 2001-2002

Sommaire

0. Introduction

L'Analyse Fonctionnelle est née au début du 20ème siècle pour fournir un cadre abstrait et général à un certain nombre de problèmes, dont beaucoup sont issus de la physique, et où la question posée est la recherche d'une *fonction* vérifiant certaines propriétés, par exemple une équation aux dérivées partielles. La théorie moderne de l'intégration (Lebesgue, un peu après 1900) et la théorie des espaces de Hilbert se sont rejointes pour créer l'un des objets les plus importants, l'espace L_2 des fonctions de carré sommable, qui a permis en particulier de placer la théorie des séries de Fourier dans un cadre conceptuellement beaucoup plus clair et plus simple que celui qui était en vigueur à la fin du 19ème siècle.

La première partie du poly contient les éléments de base de l'Analyse Fonctionnelle : espaces normés et espaces de Banach, espaces de Hilbert, applications linéaires continues, dualité, topologies faibles. La seconde partie concerne la théorie spectrale. En très gros, il s'agit de la généralisation au cadre infini-dimensionnel de la théorie de la diagonalisation. La notion fondamentale est la notion de *spectre* d'une application linéaire continue d'un espace de Banach dans lui-même. Le *calcul fonctionnel* sera l'occasion de mettre en action nombre d'objets vus en licence.

Ce polycopié provient en grande partie du polycopié de Georges Skandalis pour l'édition 1998-1999 du même enseignement. Je le remercie vivement de m'avoir transmis ses fichiers, ce qui m'a considérablement allégé la tâche. J'encourage très vivement les étudiants à lire de vrais et bons *livres* d'Analyse Fonctionnelle, par exemple ceux de Brézis, Reed et Simon, Rudin, qui sont indiqués dans la brochure de la maîtrise.

La plupart des points du cours sont numérotés, par chapitre-section-type, par exemple le "théorème 2.3.4" serait le quatrième énoncé (théorème, proposition, lemme, corollaire) de la section 3 du chapitre 2 ; à l'intérieur de cette section, le théorème sera appelé "théorème 4", ailleurs dans le chapitre 2 il sera désigné par "théorème 3.4" et dans un autre chapitre "théorème 2.3.4". Les passages écrits en petits caractères contiennent des informations qui peuvent être omises en première lecture. Le poly se termine par un index terminologique et un index des notations.

Il existe une version longue de ce poly, disponible sur le web (voir plus loin). J'ai gardé ici la numérotation de ce poly *long*, pour des raisons évidentes de compatibilité. Certaines sections, et surtout certains théorèmes, remarques, etc... ont été supprimés dans la présente version. Il en résulte que la numérotation de cette version pourra sembler bizarre et pleine de trous par endroits.

Depuis deux ans, j'ai fait une tentative pour utiliser les "nouvelles technologies" en plaçant un certain nombre d'informations sur un site Web,

`http://www.math.jussieu.fr/~maurey/ths.html`

en particulier le texte de ce poly, des résumés de cours, etc... Je suis toujours content de recevoir des suggestions constructives pour améliorer l'efficacité de cet outil.

B. Maurey

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. Si f est une fonction réelle définie sur l'ensemble X, il est parfois rapide et agréable d'utiliser la notation des probabilistes $\{f > t\} = \{x \in \mathrm{X} : f(x) > t\}$. De temps en temps on notera 0_{X} le vecteur nul d'un espace vectoriel X, quand il semblera que cette notation lourde lève toute ambiguïté.

Chapitre 1. Espaces normés et applications linéaires continues

1.1. Normes, semi-normes ; espaces de Banach

On note $\mathbb{K}$ le corps $\mathbb{R}$ ou $\mathbb{C}$. Les espaces vectoriels considérés dans ce cours seront toujours des espaces vectoriels réels ou complexes.

Définition 1.1.1. Soit X un espace vectoriel sur $\mathbb{K}$; on appelle *semi-norme* sur X une application $p : \mathrm{X} \to \mathbb{R}_+$ vérifiant les propriétés suivantes :

(*i*) pour tout $x \in \mathrm{X}$ et tout $\lambda \in \mathbb{K}$, on a $p(\lambda x) = |\lambda|\, p(x)$;
(*ii*) pour tous $x, y \in \mathrm{X}$, on a $p(x + y) \le p(x) + p(y)$.

Si pour tout vecteur x non nul de X on a $p(x) > 0$, on dit que p est une *norme* sur X.

La propriété $p(x + y) \le p(x) + p(y)$ s'appelle *l'inégalité triangulaire* pour la semi-norme p. De l'inégalité triangulaire ci-dessus, on déduit :

Lemme 1.1.1. *Si p est une semi-norme sur* X*, on a $|p(x) - p(y)| \le p(x - y)$ pour tous vecteurs $x, y \in \mathrm{X}$.*

Rappelons qu'un sous-ensemble C d'un espace vectoriel X est dit *convexe* si pour tout couple (x, y) d'éléments de C, le *segment* $[x, y]$ est tout entier contenu dans C ; le segment $[x, y]$ est formé des *combinaisons convexes* des deux points x et y, c'est à dire tous les points de la forme $z = (1 - t)x + ty$, où t varie dans $[0, 1]$. Une fonction réelle f définie sur un sous-ensemble convexe C de X est dite *fonction convexe sur* C si

$$f((1 - t)x + ty) \le (1 - t)\, f(x) + t\, f(y)$$

pour tous $x, y \in \mathrm{C}$ et tout $t \in [0, 1]$. On dira qu'une fonction réelle q sur X est *positivement homogène* si elle vérifie que $q(\mu x) = \mu\, q(x)$ pour tout $x \in \mathrm{X}$ et tout nombre réel $\mu \ge 0$. Si q est positivement homogène et *sous-additive*, c'est à dire que $q(x + y) \le q(x) + q(y)$ pour tous $x, y \in \mathrm{X}$, alors q est une fonction convexe sur X, puisqu'on aura alors

$$q((1 - t)x + ty) \le q((1 - t)x) + q(ty) = (1 - t)q(x) + tq(y).$$

En particulier, les semi-normes sur X sont des fonctions convexes. Lorsque f est une fonction convexe définie sur un ensemble convexe $\mathrm{C} \subset \mathrm{X}$, les ensembles de la forme $\mathrm{C}_t = \{x \in \mathrm{C} : f(x) \le t\}$ sont des ensembles convexes, pour tout t réel (la réciproque n'est pas vraie).

Corollaire 1.1.3. *Pour que la fonction $p \ge 0$ soit une semi-norme sur l'espace vectoriel* X*, il faut et il suffit que $p(\lambda x) = |\lambda|\, p(x)$ pour tout scalaire $\lambda \in \mathbb{K}$ et pour tout vecteur $x \in \mathrm{X}$ et que l'ensemble $\{x \in \mathrm{X} : p(x) \le 1\}$ soit convexe.*

Démonstration. Supposons que $\mathrm{C}_p = \{x \in \mathrm{X} : p(x) \le 1\}$ soit convexe, et déduisons la sous-additivité de p ; soient x et y deux vecteurs de X, $a > p(x)$ et $b > p(y)$;

considérons les deux vecteurs $x_1 = a^{-1}x$ et $y_1 = b^{-1}y$; l'homogénéité de p implique que $p(x_1) < 1$ et $p(y_1) < 1$; formons ensuite la combinaison convexe

$$z = \frac{a}{a+b}\,x_1 + \frac{b}{a+b}\,y_1,$$

qui est dans C_p d'après l'hypothèse de convexité, c'est à dire que $p(z) \le 1$. Mais on vérifie immédiatement que $z = (a+b)^{-1}(x+y)$, et l'homogénéité de p transforme alors l'inégalité $p(z) \le 1$ en $p(x+y) \le a+b$. En faisant tendre a vers $p(x)$ et b vers $p(y)$ on obtient $p(x+y) \le p(x) + p(y)$.

//

Exemple 1.1.2. Pour $1 \le r < +\infty$, soit $\mathcal{L}_r = \mathcal{L}_r([0,1])$ l'espace vectoriel des fonctions f complexes définies sur $[0,1]$ telles que f soit mesurable et $\int_0^1 |f(s)|^r\,ds < +\infty$; la quantité

$$p(f) = \Bigl(\int_0^1 |f(s)|^r\,ds\Bigr)^{1/r}$$

est une semi-norme sur $\mathcal{L}_r$;

pour le vérifier, on voit d'abord que $p(\lambda f) = |\lambda|\,p(f)$ (facile), puis on montre que l'ensemble $\{f \in \mathcal{L}_r : p(f) \le 1\}$ est convexe. Cela provient de la convexité sur $[0,+\infty[$ de la fonction $u \to u^r$; on a alors si f, g sont deux éléments de $\mathcal{L}_r$ tels que $p(f) \le 1$, $p(g) \le 1$ et si $0 \le t \le 1$,

$$|(1-t)f(s) + tg(s)|^r \le \bigl((1-t)|f(s)| + t|g(s)|\bigr)^r \le (1-t)|f(s)|^r + t|g(s)|^r$$

pour tout $s \in [0,1]$, donc

$$\int_0^1 \bigl|(1-t)f(s) + tg(s)\bigr|^r\,ds \le (1-t)\int_0^1 |f(s)|^r\,ds + t\int_0^1 |g(s)|^r\,ds \le (1-t) + t = 1.$$

On appelle *espace normé* un espace vectoriel X muni d'une norme p. Si (X, p) est un espace normé, nous en ferons un *espace métrique* en définissant la distance d sur X par $d(x,y) = p(x-y)$, et nous munirons X de la topologie associée à cette métrique, que nous appellerons *topologie de la norme*. Soient $x \in \mathrm{X}$ et $r > 0$; on appelle *boule ouverte* de centre x et de rayon r le sous-ensemble $\mathrm{B}_p(x,r) = \{y \in \mathrm{X} : p(y-x) < r\}$ de X. Rappelons que dans la topologie de la norme sur X, les parties ouvertes sont les réunions de boules ouvertes ; une partie U de X est un voisinage de $x \in \mathrm{X}$ si et seulement s'il existe $r > 0$ tel que $\mathrm{B}(x,r) \subset \mathrm{U}$. La *boule fermée* de centre x et de rayon $r > 0$ est l'ensemble $\{y \in \mathrm{X} : p(y-x) \le r\}$. Par convention, la *boule unité* d'un espace normé X sera la boule *fermée* de centre 0_{X} et de rayon 1 ; on la notera B_{X}.

Proposition 1.1.4. *Soit* (X, p) *un espace normé ; l'application* $p : \mathrm{X} \to \mathbb{R}_+$ *est continue pour la topologie de la norme.*

Cela résulte immédiatement du lemme 1. En effet, si la suite $(x_n) \subset \mathrm{X}$ tend vers y, on aura

$$|p(x_n) - p(y)| \le p(x_n - y) = d(x_n, y) \to 0.$$

En général, nous noterons $\|x\|$ la norme d'un vecteur x d'un espace normé X, ou bien $\|x\|_{\mathrm{X}}$ s'il y a un risque de confusion ; parfois il sera encore commode de désigner la fonction norme par un symbole littéral, comme dans la notation (X, p) utilisée jusqu'ici.

La topologie et la structure d'espace vectoriel d'un espace normé sont compatibles, autrement dit, un espace normé est un espace vectoriel topologique au sens suivant :

Définition 1.1.3. Un *espace vectoriel topologique* est un espace vectoriel X sur $\mathbb{K}$ muni d'une topologie pour laquelle les deux applications $(x, y) \to x + y$ de $\mathrm{X} \times \mathrm{X}$ dans X et $(\lambda, x) \to \lambda x$ de $\mathbb{K} \times \mathrm{X}$ dans X sont continues.

Proposition 1.1.5. *Un espace normé, muni de la topologie de la norme, est un espace vectoriel topologique.*

Démonstration. Soit X un espace normé ; démontrons la continuité de l'application $(x, y) \to x + y$. Puisque la topologie provient d'une métrique, nous pouvons utiliser des suites convergentes. Soient donc (x_n) une suite qui converge vers x et (y_n) une suite qui converge vers y ; on aura

$$d(x_n + y_n, x + y) = \|x_n + y_n - x - y\| \le \|x_n - x\| + \|y_n - y\|$$

qui tend bien vers 0. La continuité de l'application $(\lambda, x) \to \lambda\, x$ se démontre de façon analogue : si (λ_n) converge vers $\lambda \in \mathbb{K}$ et si (x_n) converge vers $x \in \mathrm{X}$, on écrira

$$d(\lambda_n x_n, \lambda\, x) \le d(\lambda_n x_n, \lambda_n x) + d(\lambda_n x, \lambda\, x) = |\lambda_n|\, \|x_n - x\| + |\lambda_n - \lambda|\, \|x\|$$

qui tend vers 0 (noter que la suite (λ_n) est bornée puisqu'elle est convergente).

//

Exercice 1.1.4. Si C est un sous-ensemble convexe d'un espace normé X, montrer que son adhérence est convexe. Montrer que l'adhérence d'un sous-espace vectoriel est un sous-espace vectoriel.

Définition 1.1.5. Un *espace de Banach* est un espace vectoriel normé, complet pour la distance associée à la norme.

Si F est un sous-espace vectoriel fermé d'un espace de Banach E, il est lui aussi complet pour la norme induite par celle de E, donc F est un espace de Banach.

Exemples 1.1.6.

1. L'espace $\mathrm{C}[0, 1]$ (réel ou complexe) des fonctions scalaires continues sur $[0, 1]$, muni de la *norme uniforme*,

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

est un espace de Banach. Le fait qu'il soit complet est une traduction du théorème selon lequel une limite uniforme d'une suite de fonctions continues est une fonction continue.

2. Pour $1 \le p < +\infty$, l'espace $\mathrm{L}_p = \mathrm{L}_p([0, 1])$ des *classes de fonctions* f complexes sur $[0, 1]$ telles que f soit mesurable et $\int_0^1 |f(s)|^p\, ds < +\infty$ est normé par

$$\|f\|_p = \Bigl(\int_0^1 |f(s)|^p\, ds\Bigr)^{1/p}.$$

On a déjà vu que $f \to \|f\|_p$ est une semi-norme. Si $\|f\|_p = 0$ et si $f_1 \in \mathcal{L}_p$ est un représentant quelconque de f, on a $\int_0^1 |f_1(s)|^p\, ds = 0$. Comme la fonction $|f_1|^p$ est

≥ 0, cela entraîne que $f_1 = 0$ presque partout, donc f est la classe nulle, c'est à dire que $f = 0_{\mathrm{L}_p}$. On a ainsi montré que $f \to \|f\|_p$ est une norme sur L_p.

Cet espace L_p est de plus complet (voir le chapitre 3).

3. De façon analogue, on désigne par $\ell_p = \ell_p(\mathbb{N})$ l'espace vectoriel des suites scalaires $x = (x_n)_{n\in\mathbb{N}}$ telles que $\sum |x_n|^p < +\infty$. C'est un espace de Banach pour la norme

$$\|x\|_p = \Bigl(\sum_{n=0}^{+\infty} |x_n|^p\Bigr)^{1/p}.$$

L'espace ℓ_∞ est l'espace des suites scalaires $x = (x_n)$ bornées, normé par

$$\|x\|_\infty = \sup_n |x_n|.$$

L'espace ℓ_∞ est complet pour cette norme. L'espace c_0 est l'espace des suites scalaires (x_n) telles que $\lim x_n = 0$. C'est un sous-espace fermé de ℓ_∞, donc un espace de Banach.

Séries de vecteurs

Une série de vecteurs $\sum u_k$ dans un espace normé X est dite *convergente dans* X si la suite des sommes partielles (U_n) est convergente dans X, où la somme partielle U_n est définie pour tout $n \geq 0$ par

$$\mathrm{U}_n = \sum_{k=0}^{n} u_k \in \mathrm{X}.$$

Si la série converge dans X, la *somme de la série* est un vecteur de X, qui est la limite de la suite (U_n), et on note

$$\sum_{k=0}^{+\infty} u_k = \lim_n \mathrm{U}_n \in \mathrm{X}.$$

Il faut bien comprendre que la notion de somme de la série n'a aucun sens si on ne mentionne pas la topologie qui a été utilisée pour définir la notion de limite.

Un cas particulier est celui des séries $\sum u_k$ telles que $\sum \|u_k\| < +\infty$, que l'on peut appeler *absolument convergentes* ou bien *normalement convergentes*.

Sous la condition $\sum \|u_k\| < +\infty$, le reste de la série des normes

$$r_n = \sum_{k>n} \|u_k\|$$

est une suite numérique qui tend vers 0 quand $n \to +\infty$, et on peut écrire pour tous $\ell, m \geq n$, en supposant $\ell < m$ pour fixer les idées

$$\mathrm{U}_m - \mathrm{U}_\ell = u_{\ell+1} + \cdots + u_m,$$

$$\|\mathrm{U}_m - \mathrm{U}_\ell\| \leq \|u_{\ell+1}\| + \cdots + \|u_m\| \leq \sum_{k>n} \|u_k\| = r_n,$$

ce qui montre que la suite (U_n) est alors de Cauchy.

Quand X est complet, la condition $\sum \|u_k\| < +\infty$ garantit donc la convergence dans X de la série $\sum u_k$. En fait, on a

Proposition 1.1.6. *Soit* X *un espace normé ; pour que* X *soit complet, il faut et il suffit que pour toute série* $\sum u_k$ *de vecteurs de* X*, la condition* $\sum \|u_k\| < +\infty$ *entraîne que la série* $\sum u_k$ *est convergente dans* X.

Démonstration. On a déjà vu que si X est complet, les séries absolument convergentes sont convergentes dans X ; montrons la réciproque : soit (x_n) une suite de Cauchy de vecteurs de X ; pour tout entier $k \geq 0$, on peut trouver un entier N_k tel que $\|x_m - x_n\| < 2^{-k}$ pour tous entiers $m, n \geq \mathrm{N}_k$, et on peut supposer que $\mathrm{N}_{k+1} > \mathrm{N}_k$. Posons alors $u_0 = x_{\mathrm{N}_0}$ et

$$u_{k+1} = x_{\mathrm{N}_{k+1}} - x_{\mathrm{N}_k}$$

pour tout $k \geq 0$. Par construction, on a $\|u_{k+1}\| < 2^{-k}$, donc la série $\sum u_k$ converge dans X d'après l'hypothèse. Mais les sommes partielles (U_k) de cette série sont égales aux vecteurs (x_{N_k}), donc la sous-suite (x_{N_k}) converge vers le vecteur $\mathrm{U} \in \mathrm{X}$ somme de la série $\sum u_k$. Puisque la suite (x_n) est de Cauchy, on en déduit facilement que la suite entière (x_n) converge vers U, donc X est complet.

//

Notons que lorsque la série $\sum u_k$ converge dans X, on a l'inégalité

$$\Bigl\|\sum_{k=0}^{+\infty} u_k\Bigr\| \leq \sum_{k=0}^{+\infty} \|u_k\|,$$

en convenant que la somme de la série des normes vaut $+\infty$ lorsqu'elle est divergente. Cette inégalité est obtenue en passant à la limite dans la suite des inégalités triangulaires $\|\sum_{k=0}^{n} u_k\| \leq \sum_{k=0}^{n} \|u_k\|$.

Exemple 1.1.7. Pour tout $n \geq 0$, désignons par $\mathbf{e}_n$ le vecteur de ℓ_p, ou de c_0, dont les coordonnées sont $\mathbf{e}_{n,i} = 0$ si $i \neq n$ et $\mathbf{e}_{n,n} = 1$. On appellera $(\mathbf{e}_n)_{n\geq 0}$ la *suite canonique.* Soit $x = (x_n)$ un élément de c_0 ; le vecteur x est la somme (dans l'espace normé c_0) de la série $\sum_{k=0}^{+\infty} x_k \mathbf{e}_k$ (petit exercice pour le lecteur ; le même résultat vaut pour tous les espaces ℓ_p avec $p < +\infty$).

1.2. Applications linéaires continues

Théorème 1.2.1. *Soient* X *et* Y *deux espaces normés et* $\mathrm{T} : \mathrm{X} \to \mathrm{Y}$ *une application linéaire ; les propriétés suivantes sont équivalentes :*

(*i*) *l'application* T *est continue sur* X *;*

(*ii*) *l'application* T *est continue au point* 0_X *;*

(*iii*) *il existe un nombre* $\mathrm{M} \geq 0$ *tel que, pour tout* $x \in \mathrm{X}$ *on ait*

$$\|\mathrm{T}(x)\|_\mathrm{Y} \leq \mathrm{M}\, \|x\|_\mathrm{X}.$$

Démonstration. Il est clair que $(i) \Rightarrow (ii)$. Si T est continue en 0, il existe un nombre $\delta > 0$ tel que pour tout $u \in \mathrm{X}$, la condition $d_\mathrm{X}(u, 0) \leq \delta$ implique $d_\mathrm{Y}(\mathrm{T}(u), \mathrm{T}(0)) \leq 1$; autrement dit, $\|u\|_\mathrm{X} \leq \delta$ implique $\|\mathrm{T}(u)\|_\mathrm{Y} \leq 1$. Etant donné un vecteur x non nul quelconque dans X, le vecteur $u = \delta\, \|x\|_\mathrm{X}^{-1} x$ vérifie $\|u\|_\mathrm{X} \leq \delta$, donc $\|\mathrm{T}(u)\|_\mathrm{Y} \leq 1$, ce qui revient à dire que $\|\mathrm{T}(x)\|_\mathrm{Y} \leq \delta^{-1}\, \|x\|_\mathrm{X}$. On a ainsi montré que (iii) est vraie, avec $\mathrm{M} = \delta^{-1}$. Enfin, supposons (iii) vérifiée ; si une suite (x_n) de X tend vers un vecteur $x \in \mathrm{X}$, on aura

$$d(\mathrm{T}(x_n), \mathrm{T}(x)) = \|\mathrm{T}(x_n) - \mathrm{T}(x)\|_{\mathrm{Y}} = \|\mathrm{T}(x_n - x)\|_{\mathrm{Y}} \le \mathrm{M}\, \|x_n - x\|_{\mathrm{X}} \to 0,$$

ce qui montre que T est continue au point x, et ceci pour tout $x \in \mathrm{X}$.

//

Soient p et q deux semi-normes sur un espace vectoriel X ; on dit que p et q sont *équivalentes* s'il existe deux nombres réels $m > 0$ et $\mathrm{M} \ge 0$ tels que $m\,p \le q \le \mathrm{M}\,p$.

Corollaire 1.2.2. *Deux normes p et q sur un espace vectoriel X définissent la même topologie si et seulement si elles sont équivalentes.*

Soient X et Y deux espaces normés et S, T deux applications linéaires continues de X dans Y ; on sait que l'application $\mathrm{S}+\mathrm{T}$, qui associe à tout $x \in \mathrm{X}$ l'image $\mathrm{S}(x)+\mathrm{T}(x) \in \mathrm{Y}$, est linéaire. Vérifions rapidement sa continuité en 0 : si (x_n) tend vers 0 dans X, alors $(\mathrm{S}(x_n))$ et $(\mathrm{T}(x_n))$ tendent vers 0 dans Y donc $(\mathrm{S}+\mathrm{T})(x_n) = \mathrm{S}(x_n) + \mathrm{T}(x_n)$ tend vers 0 par la propriété d'espace vectoriel topologique.

L'ensemble des applications linéaires continues de X dans Y est donc un sous-espace vectoriel noté $\mathcal{L}(\mathrm{X}, \mathrm{Y})$ de l'ensemble des applications linéaires de X dans Y. On appelle aussi *opérateur borné* une application linéaire continue entre deux espaces normés. Dans le cas où $\mathrm{Y} = \mathrm{X}$, on note simplement $\mathcal{L}(\mathrm{X})$ l'espace des endomorphismes continus de X.

Soit $\mathrm{T} : \mathrm{X} \to \mathrm{Y}$ une application linéaire continue ; d'après le théorème 1, il existe une constante M telle que $\|\mathrm{T}(x)\|_{\mathrm{Y}} \le \mathrm{M}$ pour tout vecteur x de X tel que $\|x\|_{\mathrm{X}} \le 1$. On peut donc considérer la quantité (finie)

$$\|\mathrm{T}\| = \|\mathrm{T}\|_{\mathcal{L}(\mathrm{X},\mathrm{Y})} = \sup\{\|\mathrm{T}(x)\|_{\mathrm{Y}} : \|x\|_{\mathrm{X}} \le 1\}$$

qui s'appelle la *norme de l'application linéaire* T.

Proposition 1.2.3. *Soient* X *et* Y *deux espaces normés et* $\mathrm{T} : \mathrm{X} \to \mathrm{Y}$ *une application linéaire continue ; on pose*

$$\|\mathrm{T}\|_{\mathcal{L}(\mathrm{X},\mathrm{Y})} = \sup\{\|\mathrm{T}(x)\|_{\mathrm{Y}} : \|x\|_{\mathrm{X}} \le 1\}.$$

Pour tout $x \in \mathrm{X}$, on a

$$\|\mathrm{T}(x)\|_{\mathrm{Y}} \le \|\mathrm{T}\|_{\mathcal{L}(\mathrm{X},\mathrm{Y})}\, \|x\|_{\mathrm{X}}.$$

La constante $\|\mathrm{T}\|_{\mathcal{L}(\mathrm{X},\mathrm{Y})}$ est le plus petit nombre M *tel que l'inégalité $\|\mathrm{T}(x)\|_{\mathrm{Y}} \le \mathrm{M}\, \|x\|_{\mathrm{X}}$ soit vraie pour tout $x \in \mathrm{X}$. L'application $\mathrm{T} \to \|\mathrm{T}\|_{\mathcal{L}(\mathrm{X},\mathrm{Y})}$ est une norme sur $\mathcal{L}(\mathrm{X}, \mathrm{Y})$.*

Démonstration. Vérifions que $\mathrm{T} \to \|\mathrm{T}\|$ est une norme. Il est d'abord évident que $\|\mathrm{T}\| = 0$ implique que $\|\mathrm{T}(x)\| = 0$ pour tout $x \in \mathrm{X}$, c'est à dire $\mathrm{T}(x) = 0_{\mathrm{Y}}$ pour tout $x \in \mathrm{X}$ puisque Y est normé, donc T est l'application nulle. Montrons ensuite que $\mathrm{T} \to \|\mathrm{T}\|$ est une semi-norme ; il est facile de vérifier que $\|\lambda \mathrm{T}\| = |\lambda|\, \|\mathrm{T}\|$ pour tout $\lambda \in \mathbb{K}$; ensuite, pour tout x tel que $\|x\| \le 1$,

$$\|(\mathrm{S}+\mathrm{T})(x)\| = \|\mathrm{S}(x) + \mathrm{T}(x)\| \le \|\mathrm{S}(x)\| + \|\mathrm{T}(x)\| \le \|\mathrm{S}\| + \|\mathrm{T}\|,$$

d'où l'inégalité $\|\mathrm{S}+\mathrm{T}\| \le \|\mathrm{S}\| + \|\mathrm{T}\|$, obtenue en passant au sup sur x dans la boule unité de X.

//

Exemples 1.2.1.

1. Si X est un espace normé non nul, on a toujours $\| \mathrm{Id}_{\mathrm{X}} \| = 1$.

2. Soit $f \in \mathrm{C}([0,1])$ fixée ; on définit un endomorphisme M_f de $\mathrm{C}([0,1])$, l'application de multiplication par f, en posant $\mathrm{M}_f(g) = fg$ pour toute $g \in \mathrm{C}([0,1])$. On montre que $\|\mathrm{M}_f\| = \|f\|_\infty$.

La proposition suivante est facile mais importante.

Proposition 1.2.4. *Soient* X, Y *et* Z *des espaces normés,* $\mathrm{S} : \mathrm{X} \to \mathrm{Y}$ *et* $\mathrm{T} : \mathrm{Y} \to \mathrm{Z}$ *des applications linéaires continues ; on a*

$$\|\mathrm{T} \circ \mathrm{S}\| \le \|\mathrm{S}\| \, \|\mathrm{T}\|.$$

Démonstration. Soit x un vecteur de X ; on peut écrire

$$\|(\mathrm{T} \circ \mathrm{S})(x)\|_{\mathrm{Z}} = \|\mathrm{T}(\mathrm{S}(x))\|_{\mathrm{Z}} \le \|\mathrm{T}\| \, \|\mathrm{S}(x)\|_{\mathrm{Y}} \le \|\mathrm{T}\| \, \|\mathrm{S}\| \, \|x\|_{\mathrm{X}},$$

ce qui entraîne l'inégalité voulue.

//

Proposition 1.2.5. *Soient* X *et* Y *deux espaces normés ; si* Y *est un espace de Banach, l'espace* $\mathcal{L}(\mathrm{X}, \mathrm{Y})$ *est un espace de Banach.*

Démonstration. Supposons que Y soit un espace de Banach. Soit $\sum u_k$ une série normalement convergente dans $\mathcal{L}(\mathrm{X}, \mathrm{Y})$; pour tout vecteur $x \in \mathrm{X}$, on a $\|u_k(x)\| \le \|u_k\| \, \|x\|$, donc la série $\sum u_k(x)$ est normalement convergente dans Y. Puisque Y est complet, cette série converge dans Y et on peut poser pour tout $x \in \mathrm{X}$

$$\mathrm{U}(x) = \sum_{k=0}^{+\infty} u_k(x) \in \mathrm{Y}.$$

Il est facile de vérifier que l'application U ainsi définie de X dans Y est linéaire, et de plus pour tout $x \in \mathrm{X}$ on a $\|\mathrm{U}(x)\| \le \sum_{k=0}^{+\infty} \|u_k(x)\| \le \left(\sum_{k=0}^{+\infty} \|u_k\|\right) \|x\|$, ce qui montre que U est continue et

$$(*) \qquad \|\mathrm{U}\| \le \sum_{k=0}^{+\infty} \|u_k\|.$$

Il reste à voir que U est la limite dans $\mathcal{L}(\mathrm{X}, \mathrm{Y})$ de la suite (U_n) des sommes partielles. On a

$$(\mathrm{U} - \mathrm{U}_n)(x) = \sum_{j>n} u_j(x) = \sum_{k=0}^{+\infty} v_k(x)$$

où on a posé $v_k = u_{n+k+1}$ pour tout $k \ge 0$; en appliquant l'inégalité $(*)$ à la série $\sum v_k$ on obtient $\|\mathrm{U} - \mathrm{U}_n\| \le \sum_{k=0}^{+\infty} \|v_k\| = \sum_{k>n} \|u_k\|$, et cette quantité tend vers 0 lorsque $n \to +\infty$.

//

Image d'une série convergente. Soit $\sum u_k$ une série convergente de vecteurs dans l'espace normé X et soit $\mathrm{T} : \mathrm{X} \to \mathrm{Y}$ une application linéaire continue. Alors la série $\sum \mathrm{T}(u_k)$ converge dans Y et

$$\mathrm{T}\Big(\sum_{k=0}^{+\infty} u_k\Big) = \sum_{k=0}^{+\infty} \mathrm{T}(u_k).$$

Démonstration. La suite des sommes partielles $\mathrm{U}_n = \sum_{k=0}^{n} u_k$ converge dans X vers la somme U de la série, on a $\mathrm{T}(\mathrm{U}_n) = \sum_{k=0}^{n} \mathrm{T}(u_k)$ par linéarité de T, et $\mathrm{T}(\mathrm{U}_n)$ tend vers l'image T(U) de U, par la continuité de T.

//

1.3. Produits et quotients

Proposition 1.3.1. *Soient* X *et* Y *deux espaces normés ; il existe une norme sur* $\mathrm{X} \times \mathrm{Y}$ *qui définit la topologie produit.*

Remarque 1.3.1. On vérifie sans peine que $\mathrm{X} \times \mathrm{Y}$ est un espace de Banach si et seulement si X et Y sont des espaces de Banach.

Soient X un espace vectoriel et Y un sous-espace de X ; rappelons que X/Y est le quotient de X pour la relation d'équivalence R_Y telle que $x \,\mathrm{R}_\mathrm{Y}\, y \iff y - x \in \mathrm{Y}$. Le quotient X/Y est muni de l'unique structure d'espace vectoriel pour laquelle l'application quotient $\mathrm{X} \to \mathrm{X}/\mathrm{Y}$ est linéaire. La classe de 0_X est égale à Y, et c'est le vecteur nul de l'espace quotient X/Y ; les autres classes sont les translatés de Y (ce sont les sous-espaces affines $\mathrm{Y} + x$, parallèles à Y).

Proposition 1.3.2. *Soient* X *un espace normé et* Y *un sous-espace vectoriel fermé de* X *; notons* $\pi : \mathrm{X} \to \mathrm{X}/\mathrm{Y}$ *l'application quotient. La fonction* $q : \mathrm{X}/\mathrm{Y} \to \mathbb{R}_+$ *définie par*

$$q(\xi) = \inf\{\|x\| : x \in \mathrm{X},\ \pi(x) = \xi\}$$

est une norme sur X/Y.

Démonstration. Supposons que $q(\xi) = 0$ et montrons que ξ est la classe nulle dans X/Y, c'est à dire la classe d'équivalence égale au sous-espace Y ; c'est ici que l'hypothèse Y fermé est cruciale : dire que $q(\xi) = 0$ signifie qu'il existe des vecteurs x_n tels que $\pi(x_n) = \xi$ et tels que $\|x_n\| \to 0$. Si $y \in \xi$, la suite $(y - x_n)$ est dans la classe de 0, c'est à dire dans Y, et converge vers y ; il en résulte que $y \in \mathrm{Y}$ puisque Y est fermé, donc $\xi \subset \mathrm{Y}$ ce qui implique en fait $\xi = \mathrm{Y} = 0_{\mathrm{X}/\mathrm{Y}}$.

Montrons que q est une semi-norme. Il est clair que $q(\lambda\xi) = |\lambda|\, q(\xi)$ pour tout $\lambda \in \mathbb{K}$ et tout $\xi \in \mathrm{X}/\mathrm{Y}$. Soient $\xi, \xi' \in \mathrm{X}/\mathrm{Y}$ et $\varepsilon > 0$; on peut trouver $x, x' \in \mathrm{X}$ tels que $\pi(x) = \xi$, $\pi(x') = \xi'$ et $\|x\| \le q(\xi) + \varepsilon$, $\|x'\| \le q(\xi') + \varepsilon$; on a

$$q(\xi + \xi') \le \|x + x'\| \le \|x\| + \|x'\| \le q(\xi) + q(\xi') + 2\varepsilon,$$

d'où $q(\xi + \xi') \le q(\xi) + q(\xi')$ en faisant tendre ε vers 0.

//

La projection π vérifie $\|\pi\| \le 1$. Notons encore que l'image par π de la boule unité ouverte $\mathrm{B}_\mathrm{X}(0, 1)$ de X est exactement la boule unité ouverte du quotient X/Y.

Proposition 1.3.3. *Soient* X, Z *deux espaces normés,* Y *un sous-espace fermé de* X *et* $g \in \mathcal{L}(\mathrm{X}, \mathrm{Z})$ *nulle sur* Y *; il existe une unique* $h \in \mathcal{L}(\mathrm{X}/\mathrm{Y}, \mathrm{Z})$ *telle que* $g = h \circ \pi$ *(où* $\pi : \mathrm{X} \to \mathrm{X}/\mathrm{Y}$ *est l'application quotient) ; on a* $\|h\| = \|g\|$.

Proposition 1.3.4. *Soient* X *un espace de Banach et* Y *un sous-espace fermé ; alors* X/Y *est un espace de Banach.*

Démonstration. On va utiliser le critère de la proposition 1.6. Soit $\sum \xi_k$ une série normalement convergente dans le quotient. Pour tout entier $k \ge 0$ on peut trouver un représentant $u_k \in \xi_k$ tel que $\|u_k\| \le 2\|\xi_k\|$; la série $\sum u_k$ est elle aussi normalement convergente, donc convergente dans X puisque X est complet. Finalement, la série $\sum \xi_k$, image par l'application linéaire continue π de la série convergente $\sum u_k$, est convergente dans X/Y, ce qui termine la démonstration.

//

1.4. Principe de prolongement. Complété d'un espace normé

Lemme 1.4.1. *Soient* X *un espace normé,* X_0 *un sous-espace vectoriel de* X*, dense dans* X *et* F *un espace de Banach ; toute application linéaire continue* $\mathrm{T} : \mathrm{X}_0 \to \mathrm{F}$ *se prolonge de façon unique en application linéaire continue* $\widetilde{\mathrm{T}} : \mathrm{X} \to \mathrm{F}$, *et* $\|\widetilde{\mathrm{T}}\| = \|\mathrm{T}\|$.

C'est par ce procédé que l'on définit par exemple la transformée de Fourier sur $\mathrm{X} = \mathrm{F} = \mathrm{L}_2(\mathbb{R})$, à partir de sa définition intégrale sur le sous-espace dense $\mathrm{X}_0 = \mathrm{L}_1 \cap \mathrm{L}_2$.

Démonstration. Soient $x \in \mathrm{X}$ et $n \ge 0$; d'après la densité de X_0 dans X, l'ensemble

$$\mathrm{A}_n = \{y \in \mathrm{X}_0 : \|y - x\| < 2^{-n}\}$$

est non vide ; si $y, y' \in \mathrm{A}_n$, on a $\|y'-y\| \le \|y'-x\|+\|y-x\| \le 2^{-n+1}$, donc le diamètre de A_n tend vers 0 (l'ensemble A_n devrait s'appeler $\mathrm{A}_n(x)$, mais ce serait vraiment trop lourd). Puisque T est linéaire bornée, la suite $(\mathrm{T}(\mathrm{A}_n))$ est une suite décroissante de sous-ensembles non vides de F, de diamètres tendant vers 0. Précisément, on déduit de ce qui précède que si $v, w \in \overline{\mathrm{T}(\mathrm{A}_n)}$, on a $\|v - w\| \le \|\mathrm{T}\|\, 2^{-n+1}$. Puisque F est complet, on sait que $\bigcap_n \overline{\mathrm{T}(\mathrm{A}_n)}$ contient exactement un point. Appelons $\mathrm{S}(x)$ cet unique point.

Soit $(y_k)_k \subset \mathrm{X}_0$ une suite quelconque telle que $y_k \to x$, et soit n_0 quelconque ; on aura $y_k \in \mathrm{A}_{n_0}$ pour tout $k \ge k_0$, donc $\|\mathrm{T}(y_k) - \mathrm{S}(x)\| \le 2^{-n_0+1}\, \|\mathrm{T}\|$ pour $k \ge k_0$; ceci montre que $\mathrm{S}(x) = \lim_k \mathrm{T}(y_k)$ pour toute suite $(y_k) \subset \mathrm{X}_0$ telle que $x = \lim_k y_k$. Il est facile de vérifier que $x \to \mathrm{S}(x)$ est linéaire de X dans F, à partir de cette remarque (prendre $y_k \to x$ et $y'_k \to x'$). On a aussi

$$\|\mathrm{S}(x)\| = \lim_k \|\mathrm{T}(y_k)\| \le \|\mathrm{T}\| \lim_k \|y_k\| = \|\mathrm{T}\|\, \|x\|,$$

ce qui montre que S est continue et $\|\mathrm{S}\| \le \|\mathrm{T}\|$. Si $x \in \mathrm{X}_0$, il est clair que $\mathrm{T}(x)$ est l'unique point commun aux ensembles $\mathrm{T}(\mathrm{A}_n)$, donc $\mathrm{S}(x) = \mathrm{T}(x)$ dans ce cas, ce qui montre que S prolonge T ; il en résulte que $\|\mathrm{T}\| \le \|\mathrm{S}\|$, donc $\|\mathrm{T}\| = \|\mathrm{S}\|$. Si S_1 est une autre application continue qui prolonge T, on aura $\mathrm{S}_1(x) = \lim_k \mathrm{S}_1(y_k)$ par continuité de S_1, mais $\mathrm{S}_1(y_k) = \mathrm{T}(y_k)$ par hypothèse, donc $\mathrm{S}_1(x) = \lim_k \mathrm{T}(y_k) = \mathrm{S}(x)$ pour tout $x \in \mathrm{X}$, ce qui montre l'unicité de S. Il nous suffit pour finir de prendre $\widetilde{\mathrm{T}} = \mathrm{S}$.

//

Lemme 1.4.2. *Soient* X *un espace normé,* X_0 *un sous-espace vectoriel de* X*, dense dans* X *et* F *un espace de Banach ; si* (T_n) *est une suite d'applications linéaires de* X *dans* F*, telle que* $M = \sup_n \|T_n\| < +\infty$ *et telle que* $\lim_n T_n(x)$ *existe dans* F *pour tout* $x \in X_0$*, alors* $T(x) = \lim_n T_n(x)$ *existe dans* F *pour tout* $x \in X$*, et l'application* T *est linéaire continue de* X *dans* F*, avec* $\|T\| \le M$.

1.5. Complexifié d'un espace normé réel

Voir le poly long.

1.6. Dual d'un espace normé, application transposée

Rappelons que $\mathbb{K}$ désigne le corps $\mathbb{R}$ ou $\mathbb{C}$. Soit X un espace normé sur $\mathbb{K}$; on appelle *dual* (topologique) de X et on note X^* l'espace de Banach $X^* = \mathcal{L}(X, \mathbb{K})$. Cet espace est complet par la proposition 2.5.

Exemple 1.6.1. Si $x = (x_n)$ est un élément de ℓ_1, on lui associe une forme linéaire continue f_x sur c_0 en posant

$$\forall y = (y_n) \in c_0, \quad f_x(y) = \sum_{k=0}^{+\infty} x_k y_k.$$

De plus, la norme de f_x dans le dual de c_0 est égale à la norme de x dans ℓ_1.

L'application $x \to f_x$ est donc une application linéaire isométrique de ℓ_1 dans $(c_0)^*$. Elle est aussi surjective, donc bijective. D'une certaine façon, le dual de c_0 "*est*" égal à ℓ_1.

Soit X un espace normé complexe ; c'est, en particulier, un espace normé réel. Il y a deux notions distinctes de dual pour X : le dual en tant qu'espace réel $X^*_{\mathbb{R}} = \mathcal{L}_{\mathbb{R}}(X, \mathbb{R})$ et le dual en tant qu'espace complexe $X^*_{\mathbb{C}} = \mathcal{L}_{\mathbb{C}}(X, \mathbb{C})$. En fait, on peut identifier ces deux espaces. Notons $\mathrm{Re} : \mathbb{C} \to \mathbb{R}$ l'application $\mathbb{R}$-linéaire qui à un nombre complexe $a + ib$ associe sa partie réelle a (pour $a, b \in \mathbb{R}$).

Proposition 1.6.1. *L'application* $g \to \mathrm{Re} \circ g$ *est une bijection isométrique de* $X^*_{\mathbb{C}}$ *sur l'espace* $X^*_{\mathbb{R}}$.

Démonstration. Soit $g \in X^*_{\mathbb{C}}$; alors $x \to \mathrm{Re}\, g(x)$ est $\mathbb{R}$-linéaire, et $|\mathrm{Re}\, g(x)| \le |g(x)|$ pour tout $x \in X$, donc $\|\mathrm{Re} \circ g\| \le \|g\|$. Par ailleurs, pour tout x dans la boule unité de X, il existe $\lambda \in \mathbb{C}$ tel que $|\lambda| = 1$ et $\lambda g(x) = |g(x)|$, donc

$$|g(x)| = \lambda g(x) = g(\lambda x) = (\mathrm{Re} \circ g)(\lambda x) \le \|\mathrm{Re} \circ g\|,$$

par conséquent $\|g\| = \|\mathrm{Re} \circ g\|$. Par ailleurs, soit $\ell \in X^*_{\mathbb{R}}$, notons $g : x \to \ell(x) - i\ell(ix)$; on vérifie sans peine que g est $\mathbb{C}$-linéaire, et $\mathrm{Re}\, g = \ell$. On a donc prouvé que $g \to \mathrm{Re} \circ g$ est surjective ; comme elle est isométrique, elle est injective donc bijective. //

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

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

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

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

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

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

Vérifions que $\|{}^t\mathrm{T}\| \leq \|\mathrm{T}\|$. Soit $y^* \in \mathrm{Y}^*$ tel que $\|y^*\| \leq 1$. Pour tout vecteur $x \in \mathrm{X}$ tel que $\|x\| \leq 1$ on a

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

d'où il résulte que $\|{}^t\mathrm{T}(y^*)\| \leq \|\mathrm{T}\|$ en prenant le sup sur x dans la boule unité de X, puis $\|{}^t\mathrm{T}\| \leq \|\mathrm{T}\|$ en prenant le sup sur y^* dans la boule unité de Y^*. La démonstration des autres points est laissée en exercice.

1.7. Parties totales. Séparabilité

Définition 1.7.1. On dit qu'un espace topologique Z est *séparable* s'il existe une partie *dénombrable* $\mathrm{D} \subset \mathrm{Z}$ qui soit dense dans Z.

– 1. Les espaces $\mathbb{R}$ et $\mathbb{C}$ sont séparables (par exemple, $\mathbb{Q}$ est un sous-ensemble dénombrable dense dans $\mathbb{R}$). Plus généralement, tout fermé de $\mathbb{C}$ est séparable (exercice).

– 2. Tout espace normé de dimension finie est séparable : si F est un espace vectoriel de dimension finie sur $\mathbb{R}$ et si $(x_1, \ldots, x_n)$ est une base de F, l'ensemble dénombrable $\mathrm{D} = \{\sum_{i=1}^n \lambda_i x_i : \lambda_i \in \mathbb{Q}\}$ est dense dans F.

Proposition 1.7.1. *Pour que* X *normé soit séparable, il faut et il suffit qu'il existe une suite croissante* (F_n) *de sous-espaces de dimension finie de* X *telle que* $\bigcup_n \mathrm{F}_n$ *soit dense dans* X*. Si* X *est un espace normé séparable de dimension infinie, on peut trouver une suite croissante* (F_n) *de sous-espaces vectoriels de* X *telle que* $\dim \mathrm{F}_n = n$ *pour tout* $n \geq 0$ *et telle que la réunion* $\mathrm{F} = \bigcup_n \mathrm{F}_n$ *soit dense dans* X*.*

En effet, si $\mathrm{D} = \{d_0, d_1, \ldots, d_n, \ldots\}$ est dense et si $\mathrm{F}_n = \mathrm{Vect}(d_0, \ldots, d_{n-1})$, il est évident que $\bigcup_n \mathrm{F}_n$ est dense dans X puisque cet ensemble contient D. Inversement si $\bigcup \mathrm{F}_n$ est dense, on choisit D_n dénombrable dense dans F_n, et $\mathrm{D} = \bigcup_n \mathrm{D}_n$ sera dénombrable et dense dans X.

Pour établir la deuxième partie il suffit de modifier légèrement l'argument ci-dessus, en ne prenant le vecteur d_{n+1} *que s'il* n'est pas déjà dans $\mathrm{Vect}(d_0, \ldots, d_n)$: on pose $\mathrm{F}_0 = \{0\}$ et pour tout $n \geq 0$, en supposant F_n déjà défini, de dimension n, on désigne par k_n le plus petit indice m tel que $d_m \notin \mathrm{F}_n$ (s'il n'y avait pas de tel indice m, tous les vecteurs (d_m) seraient dans F_n, donc on aurait $\mathrm{X} = \mathrm{F}_n$ de dimension finie, contradiction). On pose $\mathrm{F}_{n+1} = \mathrm{Vect}(\mathrm{F}_n, d_{k_n})$. On vérifie que F_{n+1} contient $d_0, \ldots, d_{k_n}$, donc à la fin $\bigcup_n \mathrm{F}_n$ contient l'ensemble dense D.

Exemples 1.7.2.

1. Les espaces ℓ_p et les espaces $\mathrm{L}_p([0,1])$ sont séparables pour $1 \le p < \infty$.

   Pour tout $n \ge 0$, désignons par $\mathbf{e}_n$ le $n$ième vecteur de la *suite canonique* définie par $\mathbf{e}_n(k) = \delta_{n,k}$ pour tous $n, k \ge 0$. Le sous-espace $\mathrm{F}_n = \mathrm{Vect}(\mathbf{e}_0, \ldots, \mathbf{e}_{n-1})$ est formé des vecteurs y de ℓ_p dont les coordonnées y_j, $j \ge n$ sont nulles. Il est facile de voir que tout $x \in \ell_p$ est limite d'une suite de tels vecteurs y.

2. En revanche, ℓ_∞ et $\mathrm{L}_\infty([0,1])$ ne sont pas séparables.

   Pour $t \in [0,1]$, soit f_t la fonction indicatrice de $[0,t]$; alors $\|f_s - f_t\|_\infty = 1$ si $s \ne t$, et la famille (f_t) est non-dénombrable ; si (x_n) était dense dans L_∞, il existerait pour tout $t \in [0,1]$ un indice *unique* n tel que $\|f_t - x_n\|_\infty < 1/4$, ce qui donnerait une injection de $[0,1]$ dans $\mathbb{N}$.

Soient X un espace normé et D un sous-ensemble de X ; on dit que D est *total* dans X si le sous-espace vectoriel L (algébrique) engendré par D est dense dans X (ce sous-espace L est l'ensemble des combinaisons linéaires de vecteurs de D).

Par exemple, la suite canonique $(\mathbf{e}_n)_{n\ge0}$ est totale dans ℓ_p pour tout $p < \infty$. Elle n'est pas totale dans ℓ_∞, mais elle est totale dans c_0.

Proposition 1.7.2. *Pour qu'un espace normé* X *soit séparable, il faut et il suffit qu'il admette une partie dénombrable totale.*

Démonstration. Soit X un espace normé tel qu'il existe une suite (x_n) d'éléments totale dans X ; l'espace vectoriel L engendré par la suite est dense dans X, et il est égal à la réunion croissante des sous-espaces L_n de dimension finie définis par $\mathrm{L}_n = \mathrm{Vect}(x_0, \ldots, x_{n-1})$. On sait alors que X est séparable puisque $\bigcup_{n\ge0} \mathrm{L}_n$ est dense dans X (appliquer la proposition 1). Dans l'autre direction c'est trivial.

//

2. Espaces de Hilbert

2.1. Produits scalaires

Définition 2.1.1. Soient X et Y deux espaces vectoriels complexes ; une application $f : \mathrm{X} \to \mathrm{Y}$ est dite *antilinéaire* si, pour tous $x, y \in \mathrm{X}$ et tout $\lambda \in \mathbb{C}$ on a $f(x+y) = f(x) + f(y)$ et $f(\lambda x) = \overline{\lambda} f(x)$.

On notera que la composition de deux (ou d'un nombre pair) d'applications antilinéaires est une application linéaire ; de plus, la composition d'une linéaire et d'une antilinéaire est antilinéaire.

Définition 2.1.2. Soit X un espace vectoriel complexe ; on appelle *forme sesquilinéaire* sur X une application $\mathrm{B} : \mathrm{X} \times \mathrm{X} \to \mathbb{C}$ telle que, pour tout $y \in \mathrm{X}$, l'application $x \to \mathrm{B}(x, y)$ soit linéaire et telle que pour tout $x \in \mathrm{X}$, l'application $y \to \mathrm{B}(x, y)$ soit antilinéaire (de X dans $\mathbb{C}$).

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

Proposition 2.1.1 : identité de polarisation.

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

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

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

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

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

Corollaire 2.1.2. *Soient* X *un espace vectoriel complexe et* B *une forme sesquilinéaire sur* X *; les conditions suivantes sont équivalentes :*

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

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

//

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

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

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

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

Proposition 2.1.3 : inégalité de Cauchy-Schwarz. *Soit* X *un espace vectoriel muni d'un produit scalaire ; pour tous* $x, y \in \mathrm{X}$ *on a*

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

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

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

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

//

Corollaire 2.1.4. *Soit* X *un espace vectoriel muni d'un produit scalaire ; l'application* $x \to \langle x, x \rangle^{1/2}$ *est une semi-norme sur* X.

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

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

par la proposition 3.

//

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

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

2.2. Espaces de Hilbert, orthogonalité, bases

Un *produit scalaire* sur un espace vectoriel réel ou complexe X est donc une application $(x,y) \to \langle x,y\rangle$ de $\mathrm{X}\times\mathrm{X}$ dans $\mathbb{K}$ telle que

$x \in \mathrm{X} \to \langle x,y\rangle$ est $\mathbb{K}$-linéaire pour tout $y \in \mathrm{X}$ fixé,
$\langle y,x\rangle = \overline{\langle x,y\rangle}$ pour tous $x, y \in \mathrm{X}$,
$\langle x,x\rangle \geq 0$ pour tout $x \in \mathrm{X}$.

Certains auteurs exigent qu'un produit scalaire vérifie $\langle x,x\rangle > 0$ pour tout $x \neq 0_{\mathrm{X}}$. Dans ce cas la semi-norme $x \to \sqrt{\langle x,x\rangle}$ du corollaire précédent est une norme sur l'espace X.

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

Proposition 2.2.1. *Soit* X *un espace préhilbertien ; pour tout vecteur* $y \in \mathrm{X}$ *la forme linéaire* $\ell_y : x \to \langle x,y\rangle$ *est continue de* X *dans* $\mathbb{K}$. *L'application* $y \to \ell_y$ *est antilinéaire et isométrique de* X *dans* X^*.

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

//

Définition 2.2.1. On appelle *espace de Hilbert* un espace vectoriel H (réel ou complexe) muni d'un produit scalaire $(x,y) \to \langle x,y\rangle$ tel que la semi-norme $x \to \sqrt{\langle x,x\rangle}$ soit une norme sur H, qui rende cet espace **complet**.

Si H est un espace de Hilbert, on notera $\|x\| = \sqrt{\langle x,x\rangle}$ pour tout $x \in \mathrm{H}$. L'inégalité de Cauchy-Schwarz s'écrit alors $|\langle x,y\rangle| \leq \|x\|\,\|y\|$.

Exemple 2.2.2. L'espace $\mathrm{L}_2(\Omega,\mu)$ est un espace de Hilbert pour le produit scalaire

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

L'espace ℓ_2 est un cas particulier, obtenu lorsque $\Omega = \mathbb{N}$ est muni de la mesure de comptage (définie par $\mu(\{n\}) = 1$ pour tout $n \in \mathbb{N}$).

Définition 2.2.3. Soit H un espace de Hilbert ; on dit que les vecteurs x et y de H sont *orthogonaux* si $\langle x,y\rangle = 0$. Soit $(x_n)_{n\geq 0}$ une suite infinie de vecteurs de H ou bien $(x_1,\ldots,x_{\mathrm{N}})$ une suite finie ; on dit que la suite est *orthogonale* si les x_n sont deux à deux orthogonaux, c'est à dire si $\langle x_m,x_n\rangle = 0$ lorsque $m \neq n$; on dit que c'est une *suite orthonormée* si de plus, pour tout n, on a $\|x_n\| = 1$.

Si x est orthogonal à $y_1,\ldots,y_n$, alors x est orthogonal à toutes les combinaisons linéaires de $y_1,\ldots,y_n$ d'après la linéarité du produit scalaire par rapport à sa première variable. Le vecteur x est donc orthogonal au sous-espace vectoriel engendré

$$\mathrm{F} = \mathrm{Vect}(y_1,\ldots,y_n).$$

Si x est orthogonal à tous les vecteurs d'un ensemble A, alors x est aussi orthogonal à l'adhérence de A (parce que l'application $a \to \langle a,x\rangle$ est continue).

Lemme 2.2.3. *Soient* $(u_1,\ldots,u_n)$ *des vecteurs deux à deux orthogonaux d'un espace de Hilbert* H *; on a*

$$\Big\|\sum_{k=1}^{n} u_k\Big\|^2 = \sum_{k=1}^{n} \|u_k\|^2.$$

En particulier, des vecteurs orthogonaux non nuls *sont linéairement indépendants.*

Démonstration. Facile, en développant le carré scalaire $\langle \sum_{k=1}^n u_k, \sum_{k=1}^n u_k\rangle$.

//

Lemme 2.2.4. *Soit* $(e_1,\ldots,e_n)$ *une suite orthonormée finie dans un espace de Hilbert* H *; posons* $\mathrm{F} = \mathrm{Vect}(e_1,\ldots,e_n)$ *; pour tout vecteur* $x \in \mathrm{H}$*, le vecteur*

$$y = \sum_{i=1}^{n} \langle x, e_i\rangle\, e_i$$

est la projection orthogonale de x *sur* F*, c'est à dire que* $y \in \mathrm{F}$ *et que le vecteur* $x-y$ *est orthogonal* à F*.*

Démonstration. Il est évident que $y \in \mathrm{F}$, et il est clair que $\langle y, e_j\rangle = \langle x, e_j\rangle$ pour tout $j = 1,\ldots,n$, donc $x-y$ est orthogonal à tous les (e_j), ce qui implique que $x-y$ est orthogonal à F.

//

Lemme 2.2.5 : inégalité de Bessel. *Soient* H *un espace de Hilbert et* $(e_n)_{n\ge 0}$ *une suite orthonormée dans* H *; pour tout* $x \in \mathrm{H}$ *la série numérique* $\sum_k |\langle x, e_k\rangle|^2$ *est convergente et*

$$\sum_{k\ge 0} |\langle x, e_k\rangle|^2 \le \|x\|^2.$$

Démonstration. Il suffit de montrer le résultat pour une suite finie $e_1,\ldots,e_n$. On a vu que si on pose $y = \sum_{i=1}^n \langle x, e_i\rangle\, e_i$, le vecteur $x-y$ est orthogonal au sous-espace $\mathrm{F} = \mathrm{Vect}(e_1,\ldots,e_n)$, donc $x-y$ est orthogonal à $y \in \mathrm{F}$. On aura puisque $x = y + (x-y)$

$$\|x\|^2 = \|y\|^2 + \|x-y\|^2 \ge \|y\|^2 = \sum_{i=1}^{n} |\langle x, e_i\rangle|^2$$

d'où le résultat.

//

Lemme 2.2.6. *Soit* $(u_n)_{n\ge 0}$ *une suite orthogonale dans un espace de Hilbert* H *; la série de vecteurs* $\sum_k u_k$ *converge dans* H *si et seulement si* $\sum \|u_k\|^2 < +\infty$, *et dans ce cas*

$$\Bigl\|\sum_{k=0}^{+\infty} u_k\Bigr\|^2 = \sum_{k=0}^{+\infty} \|u_k\|^2.$$

Si $(e_n)_{n\ge 0}$ *est une suite orthonormée, la série de vecteurs* $\sum_k c_k e_k$ *converge si et seulement si* $\sum |c_k|^2 < +\infty$, *et dans ce cas on a* $\bigl\|\sum_{k=0}^{+\infty} c_k e_k\bigr\|^2 = \sum_{k=0}^{+\infty} |c_k|^2$.

Démonstration. Posons $\mathrm{U}_n = \sum_{i=0}^{n} u_i$. Si $m < n$ on a par orthogonalité

$$\|\mathrm{U}_n - \mathrm{U}_m\|^2 = \sum_{k=m+1}^{n} \|u_k\|^2.$$

A partir de là, il est clair que la suite (U_n) est de Cauchy dans H si et seulement si la série numérique $\sum_k \|u_k\|^2$ vérifie le critère de convergence de Cauchy. La norme de la somme de la série s'obtient en passant à la limite dans l'égalité du lemme 3.

//

Lemme 2.2.7. *Soit* $(e_n)_{n\ge 0}$ *une suite orthonormée dans* H *et soit* F *le sous-espace vectoriel fermé engendré par la suite* $(e_n)_{n\ge 0}$ *; pour tout vecteur* $y \in \mathrm{F}$, *on a*

$$y = \sum_{k=0}^{+\infty} \langle y, e_k\rangle\, e_k.$$

Démonstration. Posons $c_j = \langle y, e_j\rangle$ pour tout $j \ge 0$, et $z = \sum_{k=0}^{+\infty} c_k\, e_k$. Cette série converge d'après le lemme 5 et le lemme précédent, et $z \in \mathrm{F}$. Pour tout $j \ge 0$, on voit en passant à la limite grâce à la continuité de l'application $x \to \langle x, e_j\rangle$

$$\langle z, e_j\rangle = \lim_n \,\langle \sum_{i=0}^{n} c_i e_i, e_j\rangle = c_j = \langle y, e_j\rangle$$

ce qui montre que $y - z$ est orthogonal à chacun des vecteurs e_j, donc $y - z$ est orthogonal à F. Puisque $y - z \in \mathrm{F}$, il en résulte que $y - z = 0_{\mathrm{H}}$, d'où le résultat.

//

Définition 2.2.4. On appelle *base hilbertienne* d'un espace de Hilbert séparable H de dimension infinie une suite orthonormée $(e_n)_{n\ge 0}$ qui est de plus *totale* dans H. On dit aussi *base orthonormée* de H.

Certaines bases hilbertiennes sont naturellement indexées par un ensemble dénombrable spécifique, par exemple $\mathrm{I} = \mathbb{Z}$, plutôt que par l'ensemble $\mathbb{N}$. Du point de vue théorique, il n'y a pas de différence et nous écrirons les preuves avec $\mathrm{I} = \mathbb{N}$.

Proposition 2.2.8. *Supposons que* $(e_n)_{n\ge 0}$ *soit une base orthonormée de l'espace de Hilbert séparable* H *de dimension infinie. Pour tout vecteur* x *de* H, *on a*

$$x = \sum_{k=0}^{+\infty} \langle x, e_k\rangle\, e_k \quad \text{et} \quad \|x\|^2 = \sum_{k=0}^{+\infty} |\langle x, e_k\rangle|^2.$$

On voit qu'une base hilbertienne $(e_n)_{n\ge 0}$ de H est une suite orthonormée qui vérifie pour tout $x \in \mathrm{H}$ la première propriété indiquée dans la proposition précédente. En effet, cette propriété implique clairement que la suite $(e_n)_{n\ge 0}$ doit être totale dans H.

Démonstration. Par définition d'une base orthonormée, la suite $(e_n)_{n\ge 0}$ est totale dans H, ce qui signifie que le sous-espace vectoriel fermé F engendré par cette suite est égal à H. Il suffit d'appliquer le lemme 7 pour obtenir la première partie de la conclusion, et le lemme 6 pour la seconde.

//

Théorème 2.2.9. *Pour tout espace de Hilbert séparable* H *de dimension infinie, il existe une base orthonormée* $(e_n)_{n\ge 0}$.

Si H est de dimension finie, l'existence de base orthonormée (bien entendu *finie*) a été vue en DEUG. Le cas des espaces de Hilbert non séparables sera examiné au chapitre 6.

Démonstration. Soit H un espace de Hilbert séparable de dimension infinie ; on peut trouver une suite croissante $(\mathrm{E}_n)_{n\ge 0}$ de sous-espaces de dimension finie de H, telle que $\dim \mathrm{E}_n = n$ pour tout $n \ge 0$ et telle que $\bigcup_n \mathrm{E}_n$ soit dense dans H (proposition 1.7.1). On construit la suite orthonormée par récurrence de façon que pour tout $n \ge 1$, la suite $(e_1, \ldots, e_n)$ soit une base orthonormée de E_n. On commence en prenant pour e_1 un vecteur de norme un dans E_1. Supposons $e_1, \ldots, e_n$ définis, de façon que $(e_1, \ldots, e_n)$ soit une base orthonormée de E_n. Puisque $\mathrm{E}_{n+1} \ne \mathrm{E}_n$, on peut choisir un vecteur $x_{n+1} \in \mathrm{E}_{n+1}$ qui n'est pas dans E_n. Soit y la projection orthogonale de x_{n+1} sur E_n. On a $x_{n+1} \ne y$ puisque $x_{n+1} \notin \mathrm{E}_n$. Le vecteur $z = x_{n+1} - y$ est non nul et orthogonal à E_n. On prend pour e_{n+1} un multiple de norme un du vecteur z. Par construction $(e_1, \ldots, e_n, e_{n+1})$ est une suite orthonormée dans E_{n+1}, donc une base de E_{n+1} (puisque $\dim \mathrm{E}_{n+1} = n + 1$).

La suite $(e_n)_{n\ge 0}$ est totale dans H puisque l'espace vectoriel qu'elle engendre contient la réunion $\bigcup_{n\ge 0} \mathrm{E}_n$ qui est dense dans H.

//

Exemples 2.2.5.

1. La suite canonique $(\mathbf{e}_n)_{n\ge 0}$ de l'espace ℓ_2 est évidemment une base orthonormée de ℓ_2.

2. Considérons l'espace de Hilbert $\mathrm{L}_2(0, 2\pi)$ des fonctions complexes de carré sommable pour la mesure $dx/2\pi$. Pour chaque entier relatif $n \in \mathbb{Z}$, soit f_n la fonction définie par

$$f_n(s) = \mathrm{e}^{ins}$$

pour tout $s \in [0, 2\pi]$. Il est facile de vérifier que les fonctions $(f_n)_{n\in\mathbb{Z}}$ forment une suite orthonormée dans $\mathrm{L}_2(0, 2\pi)$. En revanche, il faut une petite démonstration pour voir que ce système est total (voir la section 3.6). Il s'agit donc d'une base orthonormée de $L_2(0, 2\pi)$.

2.3. Théorème de projection

Théorème 2.3.1 : théorème de projection. *Soient* H *un espace de Hilbert et* C *une partie convexe fermée non vide de* H *; pour tout* $x \in H$, *il existe un et un seul point* y_0 *de* C *en lequel la fonction* $y \to \|y - x\|$ *atteint son minimum sur* C. *On a de plus*

$$\forall y \in C, \quad \mathrm{Re}\langle x - y_0, y - y_0\rangle \le 0.$$

Démonstration. En translatant le convexe C, on peut se ramener au cas où $x = 0_H$. Notons alors

$$d = \inf\{d(y, 0_H) : y \in C\} = \inf\{\|y\| : y \in C\}$$

la distance de 0_H à C. Si y et z sont deux points de C, on a $(y + z)/2 \in C$ puisque C est convexe, donc $\|(y + z)/2\| \ge d$; de plus la relation du parallélogramme

$$\|(y + z)/2\|^2 + \|(y - z)/2\|^2 = (\|y\|^2 + \|z\|^2)/2$$

implique pour tous $y, z \in C$

$$(*) \qquad 0 \le \|(y - z)/2\|^2 \le (\|y\|^2 + \|z\|^2)/2 - d^2.$$

Pour tout entier $n \ge 1$, posons

$$C_n = \{y \in C : \|y\|^2 \le d^2 + 1/n\}.$$

L'ensemble C_n est une partie fermée non vide de H ; d'après la relation $(*)$, on a $\|(y - z)/2\|^2 \le 1/n$ pour tous $y, z \in C_n$. Le diamètre de C_n est donc inférieur ou égal à $2/\sqrt{n}$, et il tend donc vers 0. Comme l'espace H est complet, l'intersection des fermés emboîtés C_n qui est égale à $\{y \in C : \|y\| = d\}$, contient un et un seul point, qui est le point y_0 cherché.

Compte tenu de notre translation simplificatrice, la relation à démontrer ensuite devient $\mathrm{Re}(\langle -y_0, y - y_0\rangle) \le 0$ pour tout $y \in C$; pour $t \in [0, 1]$, on a $y_0 + t(y - y_0) \in C$, donc $\|y_0 + t(y - y_0)\| \ge \|y_0\|$, ce qui donne en développant le carré de la norme

$$2t\,\mathrm{Re}(\langle y_0, y - y_0\rangle) + t^2\|y - y_0\|^2 \ge 0$$

pour $0 \le t \le 1$; pour finir on divise par $t > 0$ que l'on fait ensuite tendre vers 0, et on obtient $\mathrm{Re}(\langle y_0, y - y_0\rangle) \ge 0$.

//

Un cas particulier important est celui où C est un sous-espace vectoriel fermé F de H. Dans ce cas on a $\langle x - y_0, z\rangle = 0$ pour tout vecteur $z \in F$, c'est à dire que $x - y_0 \perp F$. Dans le cas de la projection sur un sous-espace vectoriel fermé F, la projection y_0 de x sur F est entièrement caractérisée par les deux conditions suivantes

- le vecteur y_0 appartient à F ;
- le vecteur $x - y_0$ est orthogonal à F.

En effet, si ces conditions sont vérifiées et si y est un élément quelconque de F, on aura

$$(*) \qquad \|x - y\|^2 = \|(x - y_0) + (y_0 - y)\|^2 = \|x - y_0\|^2 + \|y_0 - y\|^2$$

parce que $y_0 - y \in \mathrm{F}$ est orthogonal à $x - y_0$. Cette relation montre que $\|x-y\|^2 \geq \|x-y_0\|^2$ pour tout $y \in \mathrm{F}$, c'est à dire que y_0 est bien le point de F le plus proche du point x.

On notera $\mathrm{P}_\mathrm{F}(x) = y_0$ la projection orthogonale de x sur F. La caractérisation ci-dessus montre que $\mu \mathrm{P}_\mathrm{F}(x) + \mu' \mathrm{P}_\mathrm{F}(x')$ est la projection de $\mu x + \mu' x'$, autrement dit l'application P_F est une application linéaire. L'égalité $(*)$ ci-dessus donne aussi $\|x-y\| \geq \|\mathrm{P}_\mathrm{F}(x) - y\|$ pour tout $y \in \mathrm{F}$, donc $\|x\| \geq \|\mathrm{P}_\mathrm{F}(x)\|$ en prenant $y = 0$; on a donc $\|\mathrm{P}_\mathrm{F}\| \leq 1$.

Si F est un sous-espace vectoriel fermé d'un espace hilbertien H on appelle *projecteur orthogonal* sur F l'opérateur borné $\mathrm{P}_\mathrm{F} : \mathrm{H} \to \mathrm{H}$ qui associe à tout vecteur $x \in \mathrm{H}$ sa projection sur F.

Exemple 2.3.1. Espérance conditionnelle. Si $(\Omega, \mathcal{A}, \mu)$ est un espace de probabilité et si $\mathcal{F}$ est une sous-tribu de $\mathcal{A}$, on peut considérer le sous-espace vectoriel F de L_2 formé de toutes les fonctions qui sont $\mathcal{F}$-mesurables. Le sous-espace F est fermé, et la projection orthogonale de L_2 sur F s'appelle *l'espérance conditionnelle*. Par exemple, si $\Omega = [0,1]^2$ est muni de sa tribu borélienne et de la mesure de Lebesgue, si $\mathcal{F}$ est la sous-tribu formé de tous les ensembles de la forme $\mathrm{A} \times [0,1]$, où A varie parmi les boréliens de $[0,1]$, le sous-espace F est formé des fonctions qui ne dépendent que de la première variable et la projection $\mathrm{P}_\mathrm{F} f = \mathrm{E}(f|\mathcal{F})$ d'une fonction $f \in \mathrm{L}_2$ est donnée par

$$\mathrm{E}(f|\mathcal{F})(x,y) = \int_0^1 f(x,u)\, du.$$

Corollaire 2.3.2. *Soient* H *un espace de Hilbert,* F *un sous-espace vectoriel fermé séparable de* H*, et* $(e_n)_{n \geq 0}$ *une base hilbertienne du sous-espace* F*. Pour tout vecteur* $x \in \mathrm{H}$*, la projection orthogonale de* x *sur* F *est donnée par*

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

Définition 2.3.2. On dit que des parties A et B d'un espace de Hilbert H sont *orthogonales* si tout élément de A est orthogonal à tout élément de B. Soit A une partie de H ; on appelle orthogonal de A l'ensemble $\mathrm{A}^\perp$ des éléments de H orthogonaux à A.

Il est clair que $\mathrm{A}^\perp$ est un sous-espace vectoriel fermé de H.

Proposition 2.3.3. *Soient* H *un espace de Hilbert et* F *un sous-espace vectoriel fermé de* H *; on a* $\mathrm{P}_\mathrm{F} + \mathrm{P}_{\mathrm{F}^\perp} = \mathrm{Id}_\mathrm{H}$*. Il en résulte que* $\mathrm{F} \oplus \mathrm{F}^\perp = \mathrm{H}$ *et* $\mathrm{F}^{\perp\perp} = \mathrm{F}$*.*

Démonstration. Commençons par une évidence : par définition, tout vecteur de F est orthogonal à $\mathrm{F}^\perp$, donc $\mathrm{F} \subset \mathrm{F}^{\perp\perp}$. Soit maintenant $x \in \mathrm{H}$ quelconque et écrivons $x = \mathrm{P}_\mathrm{F}(x) + (x - \mathrm{P}_\mathrm{F}(x))$; d'après les propriétés de la projection orthogonale sur le sous-espace vectoriel F, on a bien que $x - \mathrm{P}_\mathrm{F}(x) \in \mathrm{F}^\perp$, et de plus la différence $x - (x - \mathrm{P}_\mathrm{F}(x)) = \mathrm{P}_\mathrm{F}(x) \in \mathrm{F}$ est orthogonale à $\mathrm{F}^\perp$; cela montre que $x - \mathrm{P}_\mathrm{F}(x)$ est la projection orthogonale de x sur $\mathrm{F}^\perp$, c'est à dire que $\mathrm{P}_{\mathrm{F}^\perp} = \mathrm{Id}_\mathrm{H} - \mathrm{P}_\mathrm{F}$. La relation $\mathrm{Id}_\mathrm{H} = \mathrm{P}_\mathrm{F} + \mathrm{P}_{\mathrm{F}^\perp}$ implique évidemment que H est la somme de F et $\mathrm{F}^\perp$. On vérifie ensuite que la somme est directe : si $x \in \mathrm{F} \cap \mathrm{F}^\perp$ alors $\langle x, x \rangle = 0$ donc $x = 0_\mathrm{H}$.

Pour finir, si on a un vecteur $x \in \mathrm{F}^{\perp\perp}$, il est orthogonal à $\mathrm{F}^\perp$ par définition, donc 0_H est sa projection orthogonale sur $\mathrm{F}^\perp$ et la relation $\mathrm{P}_\mathrm{F}(x) = (\mathrm{Id}_\mathrm{H} - \mathrm{P}_{\mathrm{F}^\perp})(x) = x$ montre que $x \in \mathrm{F}$.

//

Corollaire 2.3.4. *Soit* H *un espace de Hilbert ;*

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

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

Démonstration. Montrons le point (i). Soit F le plus petit sous-espace vectoriel fermé de H contenant A ; on sait que tout vecteur y orthogonal à A est aussi orthogonal à l'espace vectoriel Y engendré par A (par linéarité du produit scalaire), puis à l'adhérence $\mathrm{F} = \overline{\mathrm{Y}}$ de ce sous-espace (par continuité du produit scalaire). Inversement tout vecteur orthogonal à F est évidemment orthogonal à A. On a donc $\mathrm{A}^{\perp} = \mathrm{F}^{\perp}$, donc $(\mathrm{A}^{\perp})^{\perp} = \mathrm{F}^{\perp\perp} = \mathrm{F}$.

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

//

A tout vecteur $y \in \mathrm{H}$ on a associé la forme linéaire continue ℓ_y définie par

$$(*) \qquad \forall x \in \mathrm{H}, \quad \ell_y(x) = \langle x, y \rangle$$

et on a vu que $\|\ell_y\| = \|y\|$ (proposition 2.1).

Proposition 2.3.5. *Soit* H *un espace de Hilbert ; l'application isométrique antilinéaire* $y \to \ell_y$ *de l'équation* $(*)$ *est une bijection de* H *sur le dual* H^**. En d'autres termes, pour toute forme linéaire continue* ℓ *sur* H*, il existe un vecteur* $y_\ell \in \mathrm{H}$ *unique qui représente la forme linéaire* ℓ *au sens suivant :*

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

Démonstration. Soit $\ell \in \mathrm{H}^*$; si $\ell = 0$ il suffit de (et il faut) prendre $y_\ell = 0_{\mathrm{H}}$. Si $\ell \neq 0$, notons F son noyau (fermé). Puisque $\mathrm{F} \neq \mathrm{H}$, on peut choisir un vecteur z orthogonal à F et tel que $\ell(z) = 1$. Tout vecteur $x \in \mathrm{H}$ peut s'écrire

$$x = (x - \ell(x)\, z) + \ell(x)\, z = x' + \ell(x)\, z$$

avec $x' = x - \ell(x)\, z$ qui est dans F puisque $\ell(x') = \ell(x) - \ell(x)\ell(z) = 0$. On a pour tout $x \in \mathrm{H}$, puisque $x' \perp z$

$$\langle x, z \rangle = \langle \ell(x)\, z, z \rangle = \langle z, z \rangle\, \ell(x)$$

ce qui montre que $\ell = \|z\|^{-2}\, \ell_z$. Il suffit de prendre $y_\ell = \|z\|^{-2}\, z$ pour obtenir le résultat voulu.

//

Exemple 2.3.3. Pour toute forme linéaire continue ℓ sur $\mathrm{L}_2(\Omega, \mu)$, il existe une fonction $g \in \mathrm{L}_2$ telle que

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

Une application très utile est le "petit" théorème de Radon-Nikodym. Si μ, ν sont deux mesures positives sur un espace mesurable $(\Omega, \mathcal{A})$, avec ν finie et μ σ-finie, et si $\nu(\mathrm{A}) \leq \mu(\mathrm{A})$ pour tout $\mathrm{A} \in \mathcal{A}$, il existe une fonction mesurable bornée f telle que

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

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

La démonstration fonctionne ainsi : il résulte de l'hypothèse que $\int h\, d\nu \leq \int h\, d\mu$ pour toute fonction mesurable positive h, et ceci implique que $\|g\|_{\mathrm{L}_2(\nu)} \leq \|g\|_{\mathrm{L}_2(\mu)}$ pour toute fonction $g \in \mathrm{L}_2(\mu)$, ce qui montre que $\mathrm{L}_2(\mu) \subset \mathrm{L}_2(\nu)$. Comme ν est finie, la forme linéaire $g \to \int g\, d\nu$ est définie et continue sur $\mathrm{L}_2(\nu)$, donc sur $\mathrm{L}_2(\mu)$. On peut donc la représenter par une fonction $f \in \mathrm{L}_2(\mu)$, c'est à dire que

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

pour toute fonction $g \in \mathrm{L}_2(\mu)$. En appliquant avec $g = \mathbf{1}_{\mathrm{A}}$ on obtient le résultat annoncé.

Somme hilbertienne de sous-espaces orthogonaux

Proposition 2.3.6. *Si* F_1 *et* F_2 *sont deux sous-espaces vectoriels fermés orthogonaux de l'espace de Hilbert* H*, le sous-espace* $\mathrm{F} = \mathrm{F}_1 + \mathrm{F}_2$ *est fermé et la projection orthogonale de* H *sur* F *est donnée par* $\mathrm{P}_{\mathrm{F}} = \mathrm{P}_{\mathrm{F}_1} + \mathrm{P}_{\mathrm{F}_2}$.

Supposons donnée dans un espace de Hilbert H une suite $(\mathrm{F}_n)_{n\geq 0}$ de sous-espaces vectoriels fermés, deux à deux orthogonaux. On sait que pour toute famille $(x_n)_{n\geq 0}$ de vecteurs telle que $x_n \in \mathrm{F}_n$ pour tout $n \geq 0$ et $\sum \|x_k\|^2 < +\infty$, la série $\sum x_k$ converge dans H ; le vecteur $x = \sum_{k=0}^{+\infty} x_k$, qui est limite de $y_{\mathrm{N}} = \sum_{k=0}^{\mathrm{N}} x_k$, appartient à l'espace vectoriel fermé F engendré par la famille $(\mathrm{F}_n)_{n\geq 0}$: en effet, le sous-espace F contient chaque y_{N} puisqu'il contient $\mathrm{F}_0, \ldots, \mathrm{F}_{\mathrm{N}}$ et il contient la limite x puisqu'il est fermé. Inversement

Proposition 2.3.7. *Le sous-espace vectoriel fermé* F *engendré par une famille* $(\mathrm{F}_n)_{n\geq 0}$ *de sous-espaces vectoriels fermés de* H *deux à deux orthogonaux coïncide avec*

$$\{x = \sum_{k=0}^{+\infty} x_k : \forall n \geq 0,\, x_n \in \mathrm{F}_n \quad et \quad \sum \|x_k\|^2 < +\infty\}.$$

Démonstration. Désignons par G l'ensemble ci-dessus ; on a déjà expliqué que l'espace fermé F engendré par les (F_n) doit contenir G. Inversement, soit $x \in \mathrm{F}$ et désignons par x_n, pour tout entier $n \geq 0$, la projection orthogonale de x sur F_n ; on vérifie facilement que $s_n = x_0 + \cdots + x_n$ est la projection orthogonale de x sur le sous-espace fermé $\mathrm{F}_0 + \cdots + \mathrm{F}_n$. Il en résulte que $\|s_n\| \leq \|x\|$ pour tout n, ce qui implique que $\sum \|x_k\|^2 \leq \|x\|$ et permet de définir $x' = \sum_{k=0}^{+\infty} x_k \in \mathrm{G} = \lim_n s_n$. On va montrer que $x' = x$; puisque x appartient au sous-espace fermé engendré par les F_n, on peut trouver pour tout $\varepsilon > 0$ un entier N et un vecteur $y \in \mathrm{F}_0 + \cdots + \mathrm{F}_{\mathrm{N}}$ tels que $\|x - y\| < \varepsilon$. Mais par définition de la projection orthogonale, on a $\|x - s_{\mathrm{N}}\| \leq \|x - y\|$, ce qui montre que $\lim_n s_n = x$ et $x = x'$.

//

3. Les espaces de Banach classiques

3.1. Espaces de fonctions continues ou intégrables

Soit K un espace topologique compact ; l'espace C(K) (réel ou complexe) est l'espace vectoriel des fonctions scalaires continues sur K. On sait que toute fonction réelle f continue sur K est bornée (et atteint ses bornes), ce qui permet de définir la *norme uniforme* de la fonction f en posant

$$\|f\|_\infty = \max_{t \in \mathrm{K}} |f(t)|.$$

Muni de cette norme, C(K) est un espace de Banach. Le fait qu'il soit complet est une traduction du théorème selon lequel une limite uniforme d'une suite de fonctions continues est une fonction continue.

Soit (K, d) un espace *métrique* compact. Si U est un ouvert quelconque de K, la fonction f définie par $f(s) = d(s, \mathrm{U}^c) = \inf\{d(s,t) : t \notin \mathrm{U}\}$ est une fonction réelle continue sur K qui est non nulle exactement sur U. Si (U_i) est un recouvrement ouvert fini de K, on peut donc trouver des fonctions continues ψ_i telles que ψ_i soit nulle en dehors de U_i et $\psi_i > 0$ sur U_i ; alors $\psi = \sum_j \psi_j$ est > 0 sur K (donc minorée par un $\delta > 0$). Les fonctions continues $\varphi_i = \psi_i/\psi$ réalisent une *partition de l'unité*, subordonnée au recouvrement (U_i), ce qui signifie que $\sum_j \varphi_j = 1$ sur K, $\varphi_j = 0$ hors de U_j et $0 \le \varphi_j \le 1$ pour chaque j. C'est un outil très commode pour beaucoup de questions.

Voici un exemple d'application de la notion de partition de l'unité. On dit qu'un espace topologique X est *métrisable* lorsqu'il existe une distance d sur X qui définit la topologie de X.

Théorème. *Quand* K *est un compact métrisable, l'espace de Banach* C(K) *est séparable.*

En réalité, la réciproque est vraie : si K est un espace topologique compact et si C(K) est séparable, on peut définir la topologie de K par une distance.

Démontrons le théorème. On se donne un compact métrique (K, d). Pour toute fonction continue f sur K, on introduit le *module de continuité de* f, qui est une fonction notée ω_f, définie pour tout $\delta > 0$ par

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

Dire que f est uniformément continue sur K revient à dire que $\lim_{\delta \to 0} \omega_f(\delta) = 0$. Fixons $\delta > 0$ et considérons un recouvrement fini de K par des boules ouvertes $(\mathrm{B}(s_j, \delta))_{j=1,\ldots,\mathrm{N}}$ (existence par Borel-Lebesgue). Soit $\varphi_1, \ldots, \varphi_\mathrm{N}$ une partition de l'unité associée au recouvrement de K par les ouverts $\omega_j = \mathrm{B}(s_j, \delta)$; soit F_δ le sous-espace de dimension finie de C(K) engendré par $\varphi_1, \ldots, \varphi_\mathrm{N}$.

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

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

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

Si $\varphi_j(s)(f(s)-f(s_j)) \neq 0$ pour un certain indice j, on a $\varphi_j(s) \neq 0$, donc $s \in \mathrm{B}(s_j, \delta)$, donc $|f(s)-f(s_j)| \leq \omega_f(\delta)$, donc pour tout $j = 1, \ldots, \mathrm{N}$ on a

$$\varphi_j(s)|f(s)-f(s_j)| \leq \omega_f(\delta)\,\varphi_j(s)\,;$$

en sommant en j on obtient l'inégalité $|f(s)-g(s)| \leq \omega_f(\delta)$ pour tout $s \in \mathrm{K}$, c'est à dire

$$\text{(P)} \qquad \operatorname{dist}(f, \mathrm{F}_\delta) \leq \omega_f(\delta).$$

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

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

Il en résulte que $\bigcup_n \mathrm{F}_n$ est dense dans $\mathrm{C}(\mathrm{K})$, donc $\mathrm{C}(\mathrm{K})$ est séparable d'après la proposition 1.7.1.

Beaucoup d'exemples d'espaces de Banach proviennent de la théorie de l'intégration. Un *espace mesurable* $(\Omega, \mathcal{A})$ est la donnée d'un ensemble Ω et d'une tribu $\mathcal{A}$ de parties de l'ensemble Ω. Nous supposerons donné un espace $(\Omega, \mathcal{A}, \mu)$, où $(\Omega, \mathcal{A})$ est un espace mesurable et μ une mesure positive sur $(\Omega, \mathcal{A})$. Rappelons que μ est une application de $\mathcal{A}$ dans $[0, +\infty]$ telle que $\mu(\emptyset) = 0$ et

$$\mu\Big(\bigcup_{n \geq 0} \mathrm{A}_n\Big) = \sum_{n=0}^{+\infty} \mu(\mathrm{A}_n)$$

chaque fois que les ensembles $(\mathrm{A}_n)_{n \geq 0}$ de la tribu $\mathcal{A}$ sont deux à deux disjoints (avec des conventions évidentes pour les séries dont les termes peuvent prendre la valeur $+\infty$).

Pour éviter certains désagréments nous supposerons que la mesure est *σ-finie*, ce qui veut dire qu'il existe une partition (Ω_n) de Ω en une *suite* de parties $\Omega_n \in \mathcal{A}$ telles que $\mu(\Omega_n) < +\infty$. Un exemple typique est fourni par la mesure de Lebesgue sur $\mathbb{R}^d$ (muni de la tribu borélienne), ou bien par la *mesure de comptage* μ sur $\mathbb{N}$, qui associe à tout $\mathrm{A} \subset \mathbb{N}$ le nombre $\mu(\mathrm{A})$ (fini ou $+\infty$) de ses éléments.

Pour $1 \leq p < +\infty$, l'espace $\mathrm{L}_p = \mathrm{L}_p(\Omega, \mathcal{A}, \mu)$ des (classes de) fonctions f réelles ou complexes sur Ω telles que f soit mesurable et $\int_\Omega |f|^p \, d\mu < +\infty$ est normé par

$$\|f\|_p = \Big(\int_\Omega |f(s)|^p \, d\mu(s)\Big)^{1/p}.$$

On a vu dans l'exemple 1.1.2 que la quantité ci-dessus définit une semi-norme sur $\mathcal{L}_p$, puis dans l'exemple 1.1.6 que l'on obtient une norme sur L_p en passant au quotient. Cet espace L_p est de plus complet : on peut utiliser le critère des séries normalement convergentes de la proposition 1.1.6 et quelques arguments d'intégration pour retrouver ce théorème du cours d'Intégration (appelé souvent *théorème de Fisher-Riesz*).

On a vu dans l'exemple 1.1.6 l'espace ℓ_p, qui est l'espace des suites scalaires $x = (x_n)$ telles que $\sum |x_n|^p < +\infty$. Pour unifier les arguments, on peut dire que ℓ_p est l'espace $\mathrm{L}_p(\Omega, \mu)$ pour la mesure de comptage μ sur $\Omega = \mathbb{N}$. L'espace ℓ_∞ est l'espace de Banach des suites scalaires bornées. Il existe un analogue de ℓ_∞ en théorie de l'intégration : c'est l'espace $\mathrm{L}_\infty(\Omega, \mu)$ des classes de fonctions mesurables bornées sur Ω (c'est à dire des classes qui contiennent un représentant borné). La norme $\|f\|_\infty$ est la plus petite constante M telle que l'on ait $|f(s)| \leq \mathrm{M}$ pour μ-presque tout $s \in \Omega$. L'espace L_∞ est complet pour cette norme.

3.2. Résultats de densité

On utilise très souvent le résultat suivant : les fonctions continues et à support compact sont denses dans l'espace $L_1(\mathbb{R})$. Nous allons rappeler une des voies qui conduit à ce résultat. Nous supposons que $L_1(\mathbb{R})$ a été introduit à partir de la théorie de la mesure, en commençant avec les fonctions étagées et en construisant l'intégrale de Lebesgue.

Commençons par une remarque simple. Soit $f \in L_p(\mathbb{R})$, avec $1 \le p < +\infty$; pour chaque entier $n \ge 1$, désignons par f_n la fonction $\mathbf{1}_{[-n,n]} f$, égale à f sur l'intervalle $[-n, n]$ et nulle en dehors. Il est clair que f_n tend simplement vers f sur $\mathbb{R}$, et de plus $|f(t) - f_n(t)|^p \le |f(t)|^p$ pour tout $t \in \mathbb{R}$; on a donc convergence simple vers 0 de la suite $(|f - f_n|^p)$, convergence dominée par la fonction intégrable fixe $|f|^p$. D'après le théorème de convergence dominée de Lebesgue,

$$\int_{\mathbb{R}} |f(t) - f_n(t)|^p \, dt \to 0$$

ce qui signifie que $\|f_n - f\|_p \to 0$. On obtient ainsi

Pour tout $p \in [1, +\infty[$, le sous-espace vectoriel de $L_p(\mathbb{R})$ formé des fonctions g nulles en dehors d'un compact (dépendant de g) est dense dans $L_p(\mathbb{R})$.

Soient maintenant K un espace métrique compact, $\mathcal{B}$ sa tribu borélienne et μ une mesure ≥ 0 finie sur l'espace mesurable $(K, \mathcal{B})$.

Théorème 3.2.1. *L'espace $C(K)$ est dense dans $L_p(K, \mathcal{B}, \mu)$, lorsque $1 \le p < +\infty$.*

Démonstration. On montre que les ensembles $A \in \mathcal{B}$ tels que $\mathbf{1}_A$ soit limite dans L_p d'une suite (f_n) de fonctions continues sur K telles que $0 \le f_n \le 1$, forment une tribu $\mathcal{A}$ de parties de K.

Si $A \in \mathcal{A}$ et $\mathbf{1}_A = \lim f_n$, alors $\mathbf{1}_{A^c} = 1 - \mathbf{1}_A = \lim(1 - f_n)$ montre que $A^c \in \mathcal{A}$; on a $K \in \mathcal{A}$ puisque $\mathbf{1}_K = 1$ est continue sur K. Si $A \in \mathcal{A}$ et $\mathbf{1}_A = \lim f_n$, $B \in \mathcal{A}$ et $\mathbf{1}_B = \lim g_n$, alors la suite des produits $(f_n g_n)$ tend vers $\mathbf{1}_{A \cap B} = \mathbf{1}_A \mathbf{1}_B$ dans L_p (petit exercice ; utiliser le fait que les fonctions sont toutes bornées par 1 sur K). Si (A_n) est une suite croissante d'éléments de $\mathcal{A}$ et si A désigne sa réunion, la fonction $\mathbf{1}_A$ est limite simple de la suite des $(\mathbf{1}_{A_n})$ et la convergence de $|\mathbf{1}_A - \mathbf{1}_{A_n}|^p$ vers 0 est dominée par la fonction intégrable fixe $\mathbf{1}_A$, donc $\mathbf{1}_{A_n}$ tend vers $\mathbf{1}_A$ dans L_p ; par hypothèse, chaque fonction $\mathbf{1}_{A_n}$ peut être approchée par une fonction continue f_n, de façon que $\|\mathbf{1}_{A_n} - f_n\|_p < 2^{-n}$ par exemple. On a alors $\mathbf{1}_A = \lim f_n$ dans L_p, donc $A \in \mathcal{A}$.

On vérifie ensuite que cette tribu $\mathcal{A}$ contient les ouverts de K : si U est un ouvert de K, on définit f_n en posant $f_n(t) = \min\{1, n\, d(t, U^c)\}$ pour tout $t \in K$; cette suite de fonctions continues tend simplement vers $\mathbf{1}_U$, et on montre que $\mathbf{1}_U = \lim f_n$ dans L_p par convergence dominée. Puisque $\mathcal{A}$ est une tribu contenant les ouverts de K, on a $\mathcal{B} \subset \mathcal{A}$, donc $\mathcal{A} = \mathcal{B}$.

Pour tout borélien $B \in \mathcal{B}$, on sait maintenant que $B \in \mathcal{A}$, donc il existe une suite de fonctions continues qui tend vers $\mathbf{1}_B$ en norme L_p ; par linéarité, il en résulte que toute fonction $\mathcal{B}$-étagée est limite de fonctions continues pour la norme L_p, d'où le résultat parce que les fonctions étagées sont denses dans L_p (par la construction usuelle de l'intégrale).

//

Corollaire 3.2.2. *Lorsque (K, d) est un espace métrique compact et $1 \le p < +\infty$, l'espace $L_p(K, \mathcal{B}, \mu)$ est séparable.*

Revenons maintenant à $L_p(\mathbb{R})$, $1 \le p < +\infty$. Si $f \in L_p(\mathbb{R})$, on peut déjà trouver $a > 0$, une fonction g_1 nulle en dehors de $[-a, a]$ et telle que $\|f - g_1\|_p < \varepsilon$; en appliquant le théorème qui précède au compact $[-a, a]$ et à la mesure de Lebesgue, on trouve une fonction g_2 sur $\mathbb{R}$, continue sur $[-a, a]$, nulle en dehors et telle que $\|g_1 - g_2\|_p < \varepsilon$. Si $g_2(-a) = g_2(a) = 0$, la fonction g_2 est continue sur $\mathbb{R}$ et on a atteint notre objectif. Sinon, il faut encore une petite approximation pour obtenir g_3, continue sur $\mathbb{R}$, nulle en dehors de $[-a-1, a+1]$ et telle que $\|g_3 - g_2\|_p < \varepsilon$.

Théorème 3.2.3. *L'espace vectoriel des fonctions continues et à support compact sur* $\mathbb{R}$ *est dense dans* $\mathrm{L}_p(\mathbb{R})$ *lorsque* $1 \le p < +\infty$. *Le même résultat est vrai pour* $\mathbb{R}^d$, *pour tout* $d \ge 1$.

Nous appelons *fonction en escalier* une fonction f sur $\mathbb{R}$ (ou sur un intervalle de $\mathbb{R}$) qui est combinaison linéaire de fonctions indicatrices d'intervalles. On dira que f est à support borné si elle est nulle en dehors d'un intervalle borné.

Corollaire 3.2.4. *L'espace vectoriel des fonctions en escalier à support borné sur* $\mathbb{R}$ *est dense dans* $\mathrm{L}_p(\mathbb{R})$ *lorsque* $1 \le p < +\infty$.

Démonstration. Il suffit de vérifier qu'on peut approcher, en norme L_p, toute fonction continue à support compact f sur $\mathbb{R}$ par des fonctions en escalier à support borné. C'est très facile en utilisant la continuité uniforme de f.

//

3.3. Hölder et dualité des espaces ℓ_p

Pour $p \in [1, +\infty]$, on appelle *exposant conjugué* de p le nombre $q \in [1, +\infty]$ tel que $1/p + 1/q = 1$. Cette relation est symétrique ; on dit que (p, q) est un couple *d'exposants conjugués*. On notera que si $1 < p < +\infty$, cela implique que $q(p-1) = p$ et de façon symétrique, $p(q-1) = q$; on pourra aussi noter que $(p-1)(q-1) = 1$.

Théorème 3.3.1 : inégalité de Hölder. *Soient* $p, q \in [1, +\infty]$ *tels que* $1/p + 1/q = 1$ *; si* $x = (x_n) \in \ell_p$ *et* $y = (y_n) \in \ell_q$, *alors* $(x_n y_n) \in \ell_1$ *et*

$$\Bigl|\sum_{n=0}^{+\infty} x_n y_n\Bigr| \le \|x\|_p \, \|y\|_q.$$

Si $f \in \mathrm{L}_p(\Omega, \mu)$ *et* $g \in \mathrm{L}_q(\Omega, \mu)$, *la fonction produit* fg *est intégrable et*

$$\Bigl|\int_\Omega f g \, d\mu\Bigr| \le \|f\|_p \, \|g\|_q.$$

Démonstration. On écrira la démonstration dans le cas des fonctions, où les notations sont plus agréables. Pour alléger un peu plus, on écrira simplement $\int f$ au lieu de $\int_\Omega f(s)\, d\mu(s)$ chaque fois que possible. Si $p = \infty$, alors $q = 1$; la fonction f est (presque-sûrement) bornée par $\mathrm{M} = \|f\|_\infty$ et g est intégrable ; le produit fg est mesurable et $|fg| \le \mathrm{M}\, |g|$, donc fg est intégrable et

$$\Bigl|\int fg\Bigr| \le \int |fg| \le \mathrm{M} \int |g| = \|f\|_\infty \, \|g\|_1.$$

Supposons maintenant $1 < p < +\infty$. Pour tous nombres réels $t, u \geq 0$, on a la relation

$$tu \leq \frac{1}{p}\, t^p + \frac{1}{q}\, u^q$$

(pour le voir, on pourra maximiser la fonction $t \to tu - t^p/p$). Il en résulte que pour tout $s \in \Omega$

$$|f(s)g(s)| \leq \frac{1}{p}\, |f(s)|^p + \frac{1}{q}\, |g(s)|^q,$$

ce qui montre que fg est intégrable, et que

$$\Big| \int fg \Big| \leq \frac{1}{p} \int |f|^p + \frac{1}{q} \int |g|^q.$$

L'inégalité cherchée est positivement homogène par rapport à f et à g, donc il suffit de la démontrer lorsque $\|f\|_p = \|g\|_q = 1$. Mais dans ce cas, $\int |f|^p = 1$ et $\int |g|^q = 1$, donc l'inégalité précédente donne $|\int fg| \leq 1/p + 1/q = 1$, ce qui est le résultat voulu.

//

Corollaire 3.3.2. *Soient $p, q \in [1, +\infty]$ tels que $1/p + 1/q = 1$ et $x = (x_n) \in \ell_p$; on a*

$$\|x\|_p = \sup\{\Big|\sum_{n=0}^{+\infty} x_n y_n\Big| : y = (y_n) \in \ell_q,\ \|y\|_q \leq 1\}.$$

Si $f \in \mathrm{L}_p(\Omega, \mu)$,

$$\|f\|_p = \sup\{\Big|\int_\Omega fg\, d\mu\Big| : \|g\|_q \leq 1\}.$$

Démonstration. L'inégalité de Hölder nous dit déjà que

$$\|f\|_p \geq \sup\{\Big|\int_\Omega fg\, d\mu\Big| : \|g\|_q \leq 1\},$$

le problème est de montrer l'autre direction. On va voir qu'en fait le *maximum* est atteint pour une certaine fonction $g \in \mathrm{L}_q$, $\|g\|_q \leq 1$, lorsque $1 \leq p < +\infty$. Si $f = 0$, le résultat est évident, on supposera donc $f \neq 0$, et par homogénéité on peut se ramener à $\|f\|_p = 1$. Soit $\widetilde{f}$ une “vraie” fonction mesurable de la classe f, et définissons une fonction mesurable g sur l'ensemble Ω en posant $g(s) = |\widetilde{f}(s)|^p / \widetilde{f}(s)$ sur l'ensemble mesurable $\mathrm{A} = \{s \in \Omega : \widetilde{f}(s) \neq 0\}$, et $g(s) = 0$ lorsque $s \notin \mathrm{A}$. Alors $|g(s)| = |f(s)|^{p-1}$ pour tout $s \in \mathrm{A}$; pour $p > 1$, on a $|g|^q = |f|^p$, donc $\int |g|^q = 1$, soit encore $\|g\|_q = 1$; pour $p = 1$, $g(s)$ est de module 1 quand $s \in \mathrm{A}$ donc $\|g\|_\infty = 1$. D'autre part

$$\int_\Omega fg\, d\mu = \int_\mathrm{A} |f(s)|^p\, d\mu(s) = \int_\Omega |f|^p\, d\mu = 1 = \|f\|_p.$$

//

Soit $v = (v_n) \in \ell_q$, où q est l'exposant conjugué de $p \in [1, +\infty]$. D'après ce qui précède, on peut définir une forme linéaire continue f_v sur ℓ_p en posant

$$\forall u \in \ell_p, \quad f_v(u) = \sum_{n=0}^{+\infty} u_n v_n.$$

De plus, $\|f_v\|_{\ell_p^*} = \|v\|_q$. On a ainsi défini une isométrie linéaire J_q de ℓ_q dans le dual de ℓ_p. On va maintenant voir que cette isométrie est surjective lorsque $p < +\infty$.

Notons $(\mathbf{e}_n)_{n \ge 0}$ la *suite canonique* (voir l'exemple 1.1.7). Si $x = (x_n)$ est un élément de ℓ_p, on va vérifier que la série de vecteurs $\sum x_k \mathbf{e}_k$ converge dans ℓ_p, et que sa somme $\sum_{k=0}^{+\infty}$ est le vecteur x. La somme partielle $\mathrm{U}_n = \sum_{k=0}^{n} x_k \mathbf{e}_k$ est le vecteur $\mathrm{U}_n = (x_0, x_1, \ldots, x_n, 0, \ldots)$; on voit donc que $\|x - \mathrm{U}_n\|_p^p = \sum_{k>n} |x_k|^p$, reste d'ordre n de la série numérique convergente $\sum |u_k|^p$; il en résulte que $\|x - \mathrm{U}_n\|_p$ tend vers 0 quand $n \to +\infty$, ce qui signifie précisément que $x = \sum_{k=0}^{+\infty} x_k \mathbf{e}_k$.

Soit f une forme linéaire continue sur ℓ_p, et posons $v_k = f(\mathbf{e}_k)$ pour tout $k \ge 0$; on sait que l'image linéaire continue d'une série convergente est la série convergente des images,

$$f(x) = \sum_{k=0}^{+\infty} f(x_k \mathbf{e}_k) = \sum_{k=0}^{+\infty} x_k \, v_k,$$

et $|f(x)| \le \|f\| \, \|x\|_p$. Si on choisit comme vecteur $x = x^{(n)}$ particulier celui dont les coordonnées vérifient $x_k = |v_k|^q / v_k$ si $v_k \ne 0$, $0 \le k \le n$ et $x_k = 0$ sinon, on obtiendra

$$\sum_{k=0}^{n} |v_k|^q = f(x^{(n)}) \le \|f\| \, \|x^{(n)}\|_p = \|f\| \Big(\sum_{k=0}^{n} |v_k|^q \Big)^{1/p},$$

ce qui montre que $\big(\sum_{k=0}^{n} |v_k|^q \big)^{1-1/p} \le \|f\|$ pour tout n, donc $\big(\sum_{k=0}^{+\infty} |v_k|^q \big)^{1/q} \le \|f\|$. La suite $v = (v_n)$ est donc dans ℓ_q, et $f = f_v$.

En d'autres termes,

Théorème 3.3.4. *Si $1 \le p < +\infty$ le dual de ℓ_p s'identifie à ℓ_q : l'application J_q qui associe à chaque $v \in \ell_q$ la forme linéaire $f_v \in (\ell_p)^*$ définit une bijection isométrique de ℓ_q sur le dual de ℓ_p ; de plus, J_1 définit une bijection isométrique de ℓ_1 sur le dual de c_0.*

Dans le cas des espaces L_p, l'inégalité de Hölder et son corollaire donnent aussi une isométrie j_q de L_q dans le dual de L_p ; nous allons montrer qu'elle est bijective dans certains cas. Pour cette étude, nous aurons besoin du théorème de Radon-Nikodym.

3.4. Théorème de Radon-Nikodym et dual de L_p

Une *mesure réelle* sur un espace mesurable $(\Omega, \mathcal{A})$ est une application $\mu : \mathcal{A} \to \mathbb{R}$ (pas de valeur infinie ici !) qui est σ-additive, c'est à dire que $\mu(\emptyset) = 0$ et $\mu(\bigcup_{n=0}^{+\infty} \mathrm{A}_n) = \sum_{n=0}^{+\infty} \mu(\mathrm{A}_n)$ pour toute suite (A_n) d'éléments deux à deux disjoints de $\mathcal{A}$. Une *mesure complexe* μ est une application σ-additive $\mathcal{A} \to \mathbb{C}$. Dans ce cas $\mathrm{A} \in \mathcal{A} \to \mathrm{Re}\, \mu(\mathrm{A})$ est une mesure réelle, donc une mesure complexe μ est tout simplement de la forme $\mu = \mu_1 + i\mu_2$, où μ_1 et μ_2 sont deux mesures réelles. Un résultat moins évident, le *théorème de décomposition de Hahn*, dit qu'une mesure réelle est la différence de deux mesures positives bornées.

Si μ est une mesure réelle ou complexe, elle est bornée sur $\mathcal{A}$.

Soit μ une mesure réelle ou complexe sur $(\Omega, \mathcal{A})$; on définit une nouvelle fonction $|\mu|$ sur $\mathcal{A}$ en posant pour tout $\mathrm{A} \in \mathcal{A}$

$$|\mu|(\mathrm{A}) = \sup\Bigl\{\sum |\mu(\mathrm{A}_k)| : (\mathrm{A}_k) \ \text{ disjoints et } \ \subset \mathrm{A}\Bigr\}.$$

On dit que $|\mu|$ est la *variation totale* de μ. On peut montrer (exercice) que $|\mu|$ est une mesure sur $(\Omega, \mathcal{A})$, évidemment positive, et de plus cette mesure est **finie**.

Lemme 3.4.1. *Si μ est une mesure positive sur $(\Omega, \mathcal{A})$, f une fonction réelle μ-intégrable et si $\int_{\mathrm{A}} f\, d\mu \ge 0$ pour tout $\mathrm{A} \in \mathcal{A}$, la fonction f est ≥ 0 μ-presque partout.*

Proposition 3.4.2. *Si ν est une mesure réelle ou complexe sur $(\Omega, \mathcal{A})$ et si μ est une mesure positive finie sur $(\Omega, \mathcal{A})$ telle que*

$$\forall \mathrm{A} \in \mathcal{A}, \quad |\nu(\mathrm{A})| \le \mu(\mathrm{A}),$$

il existe une fonction f mesurable bornée, réelle ou complexe, telle que

$$\forall \mathrm{A} \in \mathcal{A}, \quad \nu(\mathrm{A}) = \int_{\mathrm{A}} f\, d\mu.$$

On a $|f| \le 1$ μ-presque partout ; si ν est une mesure réelle, f est réelle μ-presque partout ; si ν est réelle positive, on a $0 \le f \le 1$ μ-presque partout.

Démonstration. Pour toute fonction $\mathcal{A}$-étagée $g = \sum_{i=1}^n c_i \mathbf{1}_{\mathrm{A}_i}$ on vérifie que l'expression $\sum_{i=1}^n c_i\, \nu(\mathrm{A}_i)$ ne dépend pas de la représentation de g (petit exercice fastidieux) ; on pose alors

$$\ell(g) = \sum_{i=1}^n c_i\, \nu(\mathrm{A}_i).$$

On vérifie que ℓ est une forme linéaire sur le sous-espace vectoriel $\mathcal{E}$ des fonctions $\mathcal{A}$-étagées, et en supposant qu'on avait utilisé une représentation de g par des ensembles (A_i) disjoints, on aura

$$|\ell(g)| \le \sum_{i=1}^n |c_i|\, |\nu(\mathrm{A}_i)| \le \sum_{i=1}^n |c_i|\, \mu(\mathrm{A}_i) = \int |g|\, d\mu \le \sqrt{\mu(\Omega)}\, \|g\|_{\mathrm{L}_2(\mu)}.$$

Cette forme linéaire continue sur le sous-espace dense $\mathcal{E} \subset \mathrm{L}_2(\mu)$ se prolonge à $\mathrm{L}_2(\mu)$, donc il existe une fonction $f_1 \in \mathrm{L}_2(\mu)$ telle que $\nu(\mathrm{A}) = \int \mathbf{1}_{\mathrm{A}}\, \overline{f}_1\, d\mu$ pour tout $\mathrm{A} \in \mathcal{A}$. Posons $f = \overline{f}_1$: on a bien la représentation annoncée, $\nu(\mathrm{A}) = \int_{\mathrm{A}} f\, d\mu$.

//

Théorème 3.4.3 : théorème de décomposition de Hahn. *Soit μ une mesure réelle sur un espace mesurable $(\Omega, \mathcal{A})$; il existe un ensemble $\mathrm{B}^+ \in \mathcal{A}$ tel qu'en posant $\mathrm{B}^- = \Omega \setminus \mathrm{B}^+$ on ait :*

pour tout $\mathrm{A} \in \mathcal{A}$, $\mu(\mathrm{A} \cap \mathrm{B}^+) \ge 0$ et $\mu(\mathrm{A} \cap \mathrm{B}^-) \le 0$.

On déduit immédiatement de l'énoncé une décomposition de μ comme différence $\mu^+ - \mu^-$ de deux mesures positives finies μ^+ et μ^-, qui sont définies par

$$\forall \mathrm{A} \in \mathcal{A}, \quad \mu^+(\mathrm{A}) = \mu(\mathrm{A} \cap \mathrm{B}^+), \quad \mu^-(\mathrm{A}) = -\mu(\mathrm{A} \cap \mathrm{B}^-).$$

Il en résulte facilement que μ^+ et μ^- peuvent être définies par des formules qui ne mentionnent pas l'ensemble B^+,

$$\forall \mathrm{A} \in \mathcal{A}, \quad \mu^+(\mathrm{A}) = \sup\{\mu(\mathrm{A}') : \mathrm{A}' \subset \mathrm{A}, \mathrm{A}' \in \mathcal{A}\}$$

et de même $\mu^-(\mathrm{A}) = \sup\{-\mu(\mathrm{A}') : \mathrm{A}' \subset \mathrm{A}, \mathrm{A}' \in \mathcal{A}\}$.

Démonstration. On applique la proposition 2 aux mesures ν et $\mu = |\nu|$; on obtient ainsi une fonction réelle f, et on pose $\mathrm{B}^+ = \{f > 0\}$.

//

Si μ, ν sont deux mesures positives sur $(\Omega, \mathcal{A})$, on dit que ν est *absolument continue* par rapport à μ, et on note $\nu << \mu$, si pour tout $\mathrm{A} \in \mathcal{A}$ la condition $\mu(\mathrm{A}) = 0$ implique $\nu(\mathrm{A}) = 0$.

Par exemple, la probabilité δ (Dirac de zéro) n'est pas absolument continue par rapport à la mesure de Lebesgue λ sur $\mathbb{R}$ puisque si $\mathrm{A} = \{0\}$, on a $\lambda(\mathrm{A}) = 0$ mais $\delta(\mathrm{A}) = 1$. Si $d\nu(x) = f(x)\,dx$ où f est Lebesgue-intégrable sur $\mathbb{R}$, la mesure finie ν sur $(\mathbb{R}, \mathcal{B})$ est absolument continue par rapport à λ. Le théorème suivant donne la réciproque.

Théorème 3.4.4 : théorème de Radon-Nikodym. *Si μ, ν sont deux mesures positives σ-finies sur $(\Omega, \mathcal{A})$, et si $\nu << \mu$, il existe une fonction mesurable $f \ge 0$ sur Ω telle que*

$$\forall \mathrm{A} \in \mathcal{A}, \quad \nu(\mathrm{A}) = \int_{\mathrm{A}} f\,d\mu.$$

On dit que f est la *densité de Radon-Nikodym* de ν par rapport à μ.

Démonstration. Il suffit de démontrer ce résultat quand μ et ν sont finies : si μ et ν sont σ-finies, on peut trouver une partition de Ω en ensembles $(\mathrm{B}_n)_{n\ge0}$ de la tribu $\mathcal{A}$ tels que $\nu(\mathrm{B}_n) < \infty$ et $\mu(\mathrm{B}_n) < \infty$. Si on pose $\mu_n = \mathbf{1}_{\mathrm{B}_n}\mu$ et $\nu_n = \mathbf{1}_{\mathrm{B}_n}\nu$, on a encore $\nu_n << \mu_n$ pour tout $n \ge 0$, et les mesures sont finies. Si on a démontré le théorème de Radon-Nikodym dans le cas fini, on sait qu'il existe une fonction mesurable $f_n \ge 0$ telle que

$$\forall \mathrm{A} \in \mathcal{A}, \quad \nu_n(\mathrm{A}) = \nu_n(\mathrm{A} \cap \mathrm{B}_n) = \int_{\mathrm{A}\cap\mathrm{B}_n} f_n\,d\mu_n = \int_{\mathrm{A}} \mathbf{1}_{\mathrm{B}_n} f_n\,d\mu$$

ce qui montre que $\mathbf{1}_{\mathrm{B}_n} f_n$ convient aussi comme densité de ν_n par rapport à μ. Maintenant, en utilisant les axiomes des mesures, Beppo-Levi et en posant $f = \sum_{n=0}^{+\infty} \mathbf{1}_{\mathrm{B}_n} f_n$, on a pour tout $\mathrm{A} \in \mathcal{A}$

$$\nu(\mathrm{A}) = \sum_{n=0}^{+\infty} \nu(\mathrm{A} \cap \mathrm{B}_n) = \sum_{n=0}^{+\infty} \int \mathbf{1}_{\mathrm{A}}\,\mathbf{1}_{\mathrm{B}_n} f_n\,d\mu = \int \mathbf{1}_{\mathrm{A}}\Big(\sum_{n=0}^{+\infty} \mathbf{1}_{\mathrm{B}_n} f_n\Big)\,d\mu = \int_{\mathrm{A}} f\,d\mu,$$

ce qui montre que f est la densité cherchée.

Montrons donc Radon-Nikodym avec l'hypothèse supplémentaire que μ et ν sont finies. La mesure positive ν est plus petite que la mesure finie $\xi = \nu + \mu$; d'après la proposition 2 appliquée à ν et ξ, il existe une fonction bornée f telle que

$$\forall \mathrm{A} \in \mathcal{A}, \quad \nu(\mathrm{A}) = \int_{\mathrm{A}} f\,d\xi = \int_{\mathrm{A}} f\,d\nu + \int_{\mathrm{A}} f\,d\mu$$

et $0 \le f \le 1$ ξ-presque partout ; quitte à modifier f sans changer les intégrales (en ν et en μ) on supposera $0 \le f \le 1$ partout. On a pour tout $\mathrm{A} \in \mathcal{A}$

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

et en passant par fonctions étagées et suites croissantes on obtient

$$\int (1-f) g\, d\nu = \int f g\, d\mu$$

pour toute fonction mesurable positive g. L'ensemble $\mathrm{B} = \{f = 1\}$ est μ-négligeable puisque

$$\int (1-f)\mathbf{1}_{\mathrm{B}}\, d\nu = 0 = \int f \mathbf{1}_{\mathrm{B}}\, d\mu = \mu(\mathrm{B})$$

donc il est aussi ν-négligeable d'après l'hypothèse d'absolue continuité. Si on prend $g_0 = 1/(1-f)$ hors de B on aura pour tout $\mathrm{A} \in \mathcal{A}$

$$\nu(\mathrm{A}) = \int_{\mathrm{A}} (1-f) g_0\, d\nu = \int_{\mathrm{A}} \frac{f}{1-f}\, d\mu,$$

ce qui donne le résultat.

//

Le dual de $\mathrm{L}_p(\Omega, \mu)$

Commençons par le cas le plus simple, celui du dual de $\mathrm{L}_1(0,1)$. On montre d'abord que toute fonction $g \in \mathrm{L}_\infty(0,1)$ permet de définir une forme linéaire continue ℓ_g sur $\mathrm{L}_1(0,1)$ en posant

$$\ell_g(f) = \int_0^1 f(t) g(t)\, dt.$$

On sait que $\|\ell_g\|_{\mathrm{L}_1^*} = \|g\|_\infty$; on rappelle que la norme de g dans $\mathrm{L}_\infty(0,1)$ est la plus petite constante M telle que l'on ait $|g| \le \mathrm{M}$ presque-partout.

Inversement, si ℓ est une forme linéaire continue sur $\mathrm{L}_1(0,1)$, on obtient par restriction à $\mathrm{L}_2(0,1) \subset \mathrm{L}_1(0,1)$ une forme linéaire continue $\widetilde{\ell}$ sur L_2, qui peut donc se représenter au moyen d'une fonction $g_1 \in \mathrm{L}_2$. On a donc en posant $g = \overline{g}_1$,

$$\forall f \in \mathrm{L}_2(0,1), \quad \widetilde{\ell}(f) = \ell(f) = \int_0^1 f(t)\overline{g_1(t)}\, dt = \int_0^1 f(t) g(t)\, dt$$

mais on sait que $|\ell(f)| \le \|\ell\|\, \|f\|_{\mathrm{L}_1}$ parce que ℓ est continue sur L_1. Cela donne pour tout $\mathrm{A} \in \mathcal{A}$, en appliquant à $f = \mathbf{1}_{\mathrm{A}}$

$$\Bigl|\int_{\mathrm{A}} g(t)\, dt\Bigr| \le \|\ell\| \int_{\mathrm{A}} dt$$

et implique que $|g| \le \|\ell\|$ presque partout d'après le lemme 1 appliqué à $\|\ell\| - g$ et $\|\ell\| + g$.

On a alors deux formes linéaires continues sur L_1, la forme ℓ et la forme ℓ_g, qui coïncident sur le sous-ensemble dense L_2 de L_1. Il en résulte que $\ell = \ell_g$.

Proposition. *L'application $g \to \ell_g$ est une isométrie linéaire surjective de* $\mathrm{L}_\infty(0,1)$ *sur le dual de* $\mathrm{L}_1(0,1)$.

Soit q le nombre tel que $1/q + 1/p = 1$; d'après l'inégalité de Hölder, on a pour toutes fonctions $f \in \mathrm{L}_p$, $g \in \mathrm{L}_q$

$$\Big| \int_\Omega f g \, d\mu \Big| \le \|f\|_p \, \|g\|_q.$$

Ceci signifie que si g est fixée dans L_q, on peut définir une forme linéaire continue ℓ_g sur L_p par la formule

$$\ell_g(f) = \int_\Omega f g \, d\mu.$$

De plus on a vu que

$$\|\ell_g\|_{\mathrm{L}_p^*} = \|g\|_q.$$

On a donc une isométrie $j_q : \mathrm{L}_q \to (\mathrm{L}_p)^*$. On a en fait le résultat suivant.

Théorème 3.4.5. *Lorsque $1 \le p < +\infty$ et que μ est σ-finie, l'application j_q est une isométrie surjective de $\mathrm{L}_q(\Omega, \mathcal{A}, \mu)$ sur le dual de $\mathrm{L}_p(\Omega, \mathcal{A}, \mu)$.*

Démonstration. On suppose que $(\Omega, \mathcal{A}, \mu)$ est un espace mesuré σ-fini et que x^* est une forme linéaire continue sur $\mathrm{L}_p = \mathrm{L}_p(\Omega, \mathcal{A}, \mu)$. Puisque μ est σ-finie, on peut trouver une suite croissante d'ensembles $\mathrm{B}_n \in \mathcal{A}$ tels que $\mu(\mathrm{B}_n) < \infty$; on va fixer l'un de ces ensembles, disons $\mathrm{C} = \mathrm{B}_{n_0}$ et on va définir une mesure réelle ou complexe γ par

$$\forall \mathrm{A} \in \mathcal{A}, \quad \gamma(\mathrm{A}) = x^*(\mathbf{1}_{\mathrm{A} \cap \mathrm{C}})$$

(petit exercice : vérifier que γ vérifie les axiomes d'une mesure complexe). On va montrer que la variation totale $|\gamma|$ est absolument continue par rapport à μ. Si les (A_k) sont des sous-ensembles de $\mathrm{B} \in \mathcal{A}$, deux à deux disjoints, on écrira

$$\sum_k |\gamma(\mathrm{A}_k)| = \sum_k u_k \gamma(\mathrm{A}_k) = x^*\big(\sum_k u_k \mathbf{1}_{\mathrm{A}_k}\big)$$

où les u_k sont des complexes de module 1 convenablement choisis ; la fonction $\sum_k u_k \mathbf{1}_{\mathrm{A}_k}$ est donc plus petite en module que $\mathbf{1}_\mathrm{B}$, et il en résulte que

$$\sum_k |\gamma(\mathrm{A}_k)| = x^*\big(\sum_k u_k \mathbf{1}_{\mathrm{A}_k}\big) \le \|x^*\| \, \|\mathbf{1}_\mathrm{B}\|_{\mathrm{L}_p(\mu)}.$$

On en déduit $|\gamma|(\mathrm{B}) \le \|x^*\| \, \mu(\mathrm{B})^{1/p}$, pour tout $\mathrm{B} \in \mathcal{A}$, ce qui implique l'absolue continuité. Il existe donc une fonction μ-intégrable $g_1 \ge 0$ telle que

$$\int f \, d|\gamma| = \int f g_1 \, d\mu$$

pour toute f mesurable bornée. De plus d'après la proposition 2 il existe une fonction g_0 de module 1 telle que $\gamma(\mathrm{A}) = \int_\mathrm{A} g_0 \, d|\gamma|$ pour tout $\mathrm{A} \in \mathcal{A}$. Finalement

$$x^*(\mathbf{1}_{\mathrm{A} \cap \mathrm{C}}) = \int_\mathrm{A} g_0 g_1 \, d\mu$$

pour tout $\mathrm{A} \in \mathcal{A}$. On passe par linéarité aux fonctions étagées positives f,

$$x^*(\mathbf{1}_\mathrm{C} \, f) = \int f \, g_0 g_1 \, d\mu$$

puis par limite croissante (qu'il faut justifier du côté de x^*) aux fonctions f mesurables ≥ 0 et bornées.

En recollant les morceaux sur les différents ensembles $\mathrm{C} = \mathrm{B}_n$ dont la réunion croissante est égale à Ω, on obtient une fonction mesurable $g \geq 0$ sur Ω telle que

$$x^*(f) = \int f\, g\, d\mu$$

pour toute fonction mesurable bornée f telle que $f = \mathbf{1}_{\mathrm{B}_n} f$ pour un certain n. On montre maintenant que $\int |g|^q < +\infty$. Posons $\mathrm{A}_n = \mathrm{B}_n \cap \{|g| \leq n\}$ et soit $f_n = \mathbf{1}_{\mathrm{A}_n} |g|^{q-1} \operatorname{sign}(g)$. On vérifie que $|f_n|^p = f_n g = \mathbf{1}_{\mathrm{A}_n} |g|^q$ est une fonction mesurable bornée, nulle en dehors de B_n, donc $f_n \in \mathrm{L}_p$ et on obtient

$$\int_{\mathrm{A}_n} |g_n|^q\, d\mu = \int f_n g\, d\mu = x^*(f_n) = |x^*(f_n)| \leq \|x^*\|\, \|f_n\|_p = \|x^*\|\, \Bigl(\int_{\mathrm{A}_n} |g_n|^q\, d\mu\Bigr)^{1/p},$$

ce qui donne $\bigl(\int_{\mathrm{A}_n} |g|^q\, d\mu\bigr)^{1/q} \leq \|x^*\|$. Il ne reste plus qu'à observer que (A_n) tend en croissant vers Ω pour obtenir que $\bigl(\int |g|^q\, d\mu\bigr)^{1/q} \leq \|x^*\|$. On sait maintenant que la forme linéaire ℓ_g définie sur L_p par $\ell_g(f) = \int fg\, d\mu$ existe, est continue, et elle coïncide avec x^* sur le sous-espace vectoriel F formé des fonctions mesurables bornées nulles en dehors d'un certain B_n ; comme F est dense dans L_p, on en déduit que $x^* = \ell_g$, et de plus $\|x^*\| = \|\ell_g\| = \|g\|_{\mathrm{L}_q}$ par Hölder.

//

En revanche, le dual de L_∞ est en général “plus grand” que L_1.

3.5. Dual de C(K)

Pour décrire le dual de C(K) il faut utiliser les mesures réelles ou complexes introduites dans la section précédente. On considérera ici le cas d'un espace **métrique** compact (K, d). Rappelons que la *tribu borélienne* $\mathcal{B}$ de K est la tribu de parties de K engendrée par les ouverts de K.

Premier exemple : mesure de Dirac. Fixons un point t_0 de K et définissons une forme linéaire continue ℓ sur C(K) (réel ou complexe) par

$$\forall f \in \mathrm{C(K)}, \quad \ell(f) = f(t_0).$$

Cette forme linéaire est *l'évaluation au point* t_0. On a $|\ell(f)| \leq \max_{t \in \mathrm{K}} |f(t)| = \|f\|_{\mathrm{C(K)}}$, donc ℓ est continue et $\|\ell\|_{\mathrm{C(K)}^*} \leq 1$. En fait $\|\ell\| = 1$ (considérer la fonction constante $f = 1$). On peut décrire cette forme linéaire ℓ comme intégrale par rapport à une mesure bien particulière, la *mesure de Dirac du point* t_0, notée δ_{t_0}. Cette mesure sur $(\mathrm{K}, \mathcal{B})$ est définie pour tout $\mathrm{A} \in \mathcal{B}$ par

$$\delta_{t_0}(\mathrm{A}) = 1 \ \text{ si } \ t_0 \in \mathrm{A}, \quad \delta_{t_0}(\mathrm{A}) = 0 \ \text{ si } \ t_0 \notin \mathrm{A}.$$

Définir la mesure δ_{t_0} consiste à placer une masse 1 au point t_0. La théorie de l'intégration par rapport à cette mesure est un peu bizarre, mais le lecteur pourra se convaincre, en passant rapidement en revue les étapes de la définition de l'intégrale par rapport à une mesure positive, que

$$\int_{\mathrm{K}} f(s)\, d\delta_{t_0}(s) = f(t_0) = \ell(f)$$

pour toute fonction continue f. On peut donc *représenter* la forme linéaire ℓ au moyen d'une mesure sur K,

$$\forall f \in \mathrm{C(K)}, \quad \ell(f) = \int f\, d\delta_{t_0}.$$

On verra un peu plus loin que c'est un résultat général.

A chaque mesure réelle ou complexe μ sur un espace $(\Omega, \mathcal{A})$ on a associé sa variation totale, ou valeur absolue $|\mu|$ qui est une mesure positive finie. On posera $\|\mu\| = |\mu|(\Omega)$. Cette expression définit une norme sur l'espace vectoriel $\mathrm{M}(\Omega, \mathcal{A})$ des mesures réelles ou complexes sur $(\Omega, \mathcal{A})$.

Soit μ une mesure réelle ou complexe sur $(\mathrm{K}, \mathcal{B})$; on a vu qu'il existe une fonction f bornée par 1 en module telle que

$$\forall \mathrm{A} \in \mathcal{B}, \quad \mu(\mathrm{A}) = \int_{\mathrm{A}} f \, d|\mu|.$$

Si g est une fonction étagée $\sum_{i=1}^n c_i \mathbf{1}_{\mathrm{A}_i}$ on pourra poser

$$\ell(g) = \int g \, d\mu = \sum_{i=1}^n c_i \mu(\mathrm{A}_i) = \int g f \, d|\mu|.$$

On a $|\int g d\mu| \le \int |gf| \, d|\mu| \le \int |g| \, d|\mu| = \|g\|_{\mathrm{L}_1}$ ce qui permet de prolonger la définition précédente précédente de $\ell(g)$ aux fonctions $g \in \mathrm{L}_1(|\mu|)$, et en particulier à $g \in \mathrm{C}(\mathrm{K})$. Dans ce cas on écrira $|\int g d\mu| \le \int |gf| \, d|\mu| \le \|g\|_\infty \int d|\mu| = \|g\|_\infty \, \|\mu\|$. On a donc associé une forme linéaire continue $\ell = \ell_\mu$ sur $\mathrm{C}(\mathrm{K})$ à la mesure réelle ou complexe μ sur $(\mathrm{K}, \mathcal{B})$. On vient de voir que la norme de dual est majorée par la norme de μ. En réalité, la norme de dual de $\mathrm{C}(\mathrm{K})$ est *égale* à la norme de mesure définie plus haut.

Les formes linéaires sur $\mathrm{C}(\mathrm{K})$ s'identifient aux mesures réelles sur $(\mathrm{K}, \mathcal{B})$ dans le cas réel, aux mesures complexes dans le cas complexe. Nous venons d'expliquer une direction. Dans l'autre direction, il faut montrer que toute forme linéaire continue sur $\mathrm{C}(\mathrm{K})$ provient d'une mesure μ sur $(\mathrm{K}, \mathcal{B})$. Ce théorème est assez long à démontrer (voir Rudin, Analyse réelle et complexe, par exemple). On indiquera seulement le premier pas, qui consiste en général à trouver d'abord la mesure positive $|\mu|$: on suppose donnée une forme linéaire continue ℓ sur $\mathrm{C}(\mathrm{K})$; on commence par définir la $|\mu|$-mesure d'un ouvert ω de K en posant

$$|\mu|(\omega) = \sup\{|\ell(f)| : |f| \le \mathbf{1}_\omega\}.$$

Enonçons le résultat.

Théorème 3.5.1. *Soit* K *un espace métrique compact ; toute forme linéaire continue* ℓ *sur* $\mathrm{C}(\mathrm{K})$ *provient d'une mesure* μ *sur* $(\mathrm{K}, \mathcal{B})$, *et de plus*

$$\|\ell\| = \|\mu\| = |\mu|(\mathrm{K}).$$

3.6. Séries de Fourier

Considérons l'espace de Hilbert $\mathrm{L}_2(0, 2\pi)$ des fonctions complexes de carré sommable pour la mesure $dx/2\pi$. Pour chaque entier relatif $n \in \mathbb{Z}$, désignons par e_n la fonction définie par

$$e_n(s) = \mathrm{e}^{ins}$$

pour tout $s \in [0, 2\pi]$. La famille $(e_n)_{n \in \mathbb{Z}}$ est une base orthonormée de $\mathrm{L}_2(0, 2\pi)$. Il est facile de vérifier que les fonctions $(e_n)_{n \in \mathbb{Z}}$ forment une suite orthonormée dans $\mathrm{L}_2(0, 2\pi)$. En revanche, il faut une petite démonstration pour voir que ce système est total. Il s'agit donc d'une base orthonormée de $L_2(0,2\pi)$. Pour toute fonction $f \in L_2(0,2\pi)$, les coefficients du développement de f dans cette base sont les coefficients de Fourier complexes

$$c_n(f) = \langle f, e_n \rangle = \int_0^{2\pi} f(s)\, \mathrm{e}^{-ins}\, \frac{ds}{2\pi}.$$

D'après Parseval, on a $\|f\|_2^2 = \int_0^{2\pi} |f(s)|^2\, ds/2\pi = \sum_{n\in\mathbb{Z}} |c_n(f)|^2$. Cette identité est la source d'une multitude d'exercices calculatoires, tels que par exemple le calcul de la somme de la série numérique $\sum_{n\geq 1} 1/n^2$.

Quand on travaille avec des fonctions réelles, on préfère parfois écrire le développement en utilisant les fonctions réelles $t \to \cos(nt)$, pour $n = 0, 1, \ldots$ et $t \to \sin(nt)$, pour $n = 1, 2, \ldots$ (pour $n = 0$, le cosinus donne la fonction constante 1). On définit classiquement les coefficients de Fourier réels de la façon suivante :

$$a_n = \frac{1}{\pi} \int_0^{2\pi} f(t) \cos(nt)\, dt; \quad b_n = \frac{1}{\pi} \int_0^{2\pi} f(t) \sin(nt)\, dt,$$

et la série de Fourier de f prend la forme

$$\frac{a_0}{2} + \sum_{k=1}^{+\infty} \big(a_k \cos(kt) + b_k \sin(kt) \big).$$

La bizarrerie du traitement de a_0 vient du fait que la fonction constante 1 n'a pas la même norme que les fonctions $t \to \cos(nt)$ pour $n \geq 1$.

On écrit souvent le développement de Fourier d'une fonction $f \in L_2$ sous la forme

$$f(s) = \sum_{n\in\mathbb{Z}} c_n(f)\, \mathrm{e}^{ins},$$

mais cette écriture est *a priori* incorrecte, car rien ne nous dit que la série numérique ci-dessus converge vers $f(s)$: ce que nous savons est que f est la somme de la série de fonctions au sens de L_2. En fait, un théorème très difficile démontré vers 1960 par le mathématicien suédois L. Carleson justifie l'écriture précédente : pour presque tout s, la série de Fourier converge au point s et sa somme est égale à $f(s)$. La convergence ponctuelle est assez facile à obtenir lorsque f est de classe C^1, et dans ce cas elle est valable pour tout s. On va obtenir un tout petit peu mieux ; rappelons qu'on dit qu'une fonction f est *lipschitzienne* s'il existe une constante M telle que $|f(t) - f(s)| \leq M\, |t - s|$ pour tous $s, t \in \mathbb{R}$.

Théorème 3.6.1. *Soit f une fonction 2π-périodique et lipschitzienne. Pour tout $s \in \mathbb{R}$, on a*

$$f(s) = \sum_{n\in\mathbb{Z}} c_n(f)\, \mathrm{e}^{ins}\,.$$

Démonstration. Posons

$$f_N(s) = S_N(f)(s) = \sum_{n=-N}^{N} c_n(f)\, \mathrm{e}^{ins} = \int_0^{2\pi} \sum_{n=-N}^{N} \mathrm{e}^{in(s-t)}\, f(t)\, \frac{dt}{2\pi}.$$

Posons
$$K_N(t) = \sum_{n=-N}^{N} e^{int} = K_N(-t) = e^{-iNt}\ \frac{1 - e^{i(2N+1)t}}{1 - e^{it}}.$$
Comme $\int_0^{2\pi} K_N(t)\,(dt/2\pi) = 1$, on aura, pour $s = 0$
$$f_n(0) - f(0) = \int_0^{2\pi} K_N(t)\,(f(t) - f(0))\,\frac{dt}{2\pi} =$$
$$= \int_0^{2\pi} (e^{-iNt} - e^{i(N+1)t})\,\frac{f(t) - f(0)}{1 - e^{it}}\,\frac{dt}{2\pi}.$$
Posons $g(t) = (f(t) - f(0))/(1 - e^{it})$. Comme f est Lipschitz et $|1 - e^{it}\,| \geq 2t/\pi$, la fonction g est bornée, donc $g \in L_2$, ce qui entraîne que les coefficients de Fourier de g tendent vers 0. Or nous avions
$$f_N(0) - f(0) = c_N(g) - c_{-N-1}(g)$$
qui tend donc vers 0 quand $N \to +\infty$. Le raisonnement est identique pour montrer que $f_N(s) \to f(s)$, pour tout $s \in \mathbb{R}$.

//

3.7. Transformation de Fourier

Pour toute fonction $f \in L_1(\mathbb{R})$ on définit la transformée de Fourier $\widehat{f}$ par
$$\forall t \in \mathbb{R}, \quad \widehat{f}(t) = \int_{\mathbb{R}} f(x)\,e^{-ixt}\ dx.$$
On définit ainsi une fonction $\widehat{f}$ sur $\mathbb{R}$; il est clair que $\widehat{f}$ est bornée ($|\widehat{f}(t)| \leq \|f\|_1$) et il est facile de voir que $\widehat{f}$ est continue sur $\mathbb{R}$ (employer le théorème de convergence dominée).

On va voir comment cette transformation se comporte quand on modifie f au moyen de certaines opérations élémentaires. Si $a \in \mathbb{R}$ et si on remplace f par la fonction translatée f_a définie par $f_a(x) = f(x - a)$ pour tout $x \in \mathbb{R}$, un changement de variable immédiat donne

$$\text{(T)} \qquad \forall t \in \mathbb{R}, \quad \widehat{f_a}(t) = e^{-iat}\,\widehat{f}(t).$$

Si on remplace f par la dilatée $f_{[\lambda]}$ définie pour $\lambda > 0$ par $f_{[\lambda]}(x) = f(\lambda x)$, on obtient par un autre changement de variable évident

$$\text{(D)} \qquad \forall t \in \mathbb{R}, \quad \widehat{f_{[\lambda]}}(t) = \frac{1}{\lambda}\,\widehat{f}\Big(\frac{t}{\lambda}\Big).$$

Ces opérations de dilatation s'appellent aussi *changement d'échelle.*

On a le résultat suivant (préliminaire à la formule d'inversion de Fourier) :

Lemme 3.7.1. *Si $a > 0$, si f est continue à support compact sur $\mathbb{R}$, et si f est linéaire sur chacun des intervalles $[(j-1)a, ja]$ pour $j \in \mathbb{Z}$, alors $\widehat{f}$ est intégrable sur $\mathbb{R}$ et on a*

$$\text{(I)} \qquad \forall x \in \mathbb{R}, \quad f(x) = \frac{1}{2\pi} \int_{\mathbb{R}} \widehat{f}(t)\, \mathrm{e}^{ixt}\, dt.$$

Démonstration. On vérifie facilement que si une fonction f vérifie la formule (I), alors les translatées de f, les dilatées de f et les combinaisons linéaires de translatées ou dilatées de f vérifient encore la formule (I). Pour démontrer le lemme, on peut d'abord se ramener à $a = 1$ par dilatation. Pour chaque $j \in \mathbb{Z}$ soit f_j la fonction continue nulle hors de $[j-1, j+1]$, égale à 1 au point j et linéaire sur les deux intervalles $[j-1, j]$ et $[j, j+1]$ (on a donc $f_j(x) = 1 - |j - x|$ sur $[j-1, j+1]$). On observe que toute fonction f continue à support compact sur $\mathbb{R}$, linéaire sur chaque intervalle $[j-1, j]$ est combinaison linéaire des fonctions f_j ; par linéarité et translation il suffit finalement de montrer le lemme pour f_0. Un calcul élémentaire (avec intégration par parties) montre que

$$\widehat{f_0}(t) = 2 \int_0^1 (1-x) \cos(xt)\, dx = 4\, \frac{\sin^2(t/2)}{t^2} \cdot$$

On voit que $\widehat{f_0}$ est intégrable sur $\mathbb{R}$, et la même chose sera vraie pour toutes les combinaisons linéaires de translatées ou dilatées, d'après les formules (T) et (D). On vérifie ensuite, par des calculs de résidus, ou bien par des calculs trigonométriques un peu pénibles, que

$$\int_{\mathbb{R}} 4\, \frac{\sin^2(t/2)}{t^2}\, \mathrm{e}^{ixt}\, dt = 2\pi f_0(x)$$

pour tout $x \in \mathbb{R}$.

//

On déduit du lemme, sous la même hypothèse sur f, en appliquant Fubini, justifié par l'intégrabilité sur $\mathbb{R}^2$ de la fonction $(x,t) \to \overline{f(x)}\, \widehat{f}(t)$:

$$\int_{\mathbb{R}} f(x)\overline{f(x)}\, dx = \frac{1}{2\pi} \int_{\mathbb{R}} \widehat{f}(t)\, \mathrm{e}^{ixt}\, \overline{f(x)}\, dx\, dt = \frac{1}{2\pi} \int_{\mathbb{R}} \widehat{f}(t)\overline{\widehat{f}(t)}\, dt$$

c'est à dire que pour une telle fonction f, on a $\|\widehat{f}\|_2 = \sqrt{2\pi}\, \|f\|_2$.

L'espace vectoriel $\mathrm{X} \subset \mathrm{L}_2(\mathbb{R})$ formé des fonctions continues à support compact et linéaires par morceaux est dense dans $\mathrm{L}_2(\mathbb{R})$, et l'application $f \to \widehat{f}$ est continue de X, muni de la norme L_2, à valeurs dans L_2. Il existe donc un prolongement linéaire continu $\mathcal{F}$ à $\mathrm{L}_2(\mathbb{R})$; en fait d'après ce qui précède $\mathrm{U} = (2\pi)^{-1/2}\mathcal{F}$ est une isométrie pour la norme L_2.

Il reste un petit point idiot à vérifier : si $f \in \mathrm{L}_1 \cap \mathrm{L}_2$, la classe de la fonction continue bornée $\widehat{f}$ coïncide bien avec la classe $\mathcal{F}(f)$ définie par prolongement à partir de X ; en effet, si $f \in \mathrm{L}_1 \cap \mathrm{L}_2$, on peut trouver une suite $(f_n) \subset \mathrm{X}$ telle que f_n tende vers f pour la norme L_1 **et** pour la norme L_2 (on pourra répéter les étapes de la section 3.2) ; alors $\widehat{f_n}$ converge uniformément vers $\widehat{f}$ et en norme L_2 vers $\mathcal{F}(f)$, d'où le résultat.

La formule inverse de Fourier du lemme précédent donne facilement que toute fonction de X est une transformée de Fourier, donc U est une isométrie à image dense, donc une isométrie surjective de $\mathrm{L}_2(\mathbb{R})$ sur lui-même. De plus, le lemme d'inversion montre que, sur le sous-espace dense $\mathcal{F}(\mathrm{X})$, l'inverse de la transformation de Fourier est donnée par $\mathcal{F}^{-1}(f)(x) = (2\pi)^{-1}\mathcal{F}(f)(-x)$. Il en résulte que cette relation est vraie pour toute $f \in \mathrm{L}_2(\mathbb{R})$: désignons par σ l'isométrie de $\mathrm{L}_2(\mathbb{R})$ définie par $(\sigma h)(x) = h(-x)$ pour toute $h \in \mathrm{L}_2$; on aura

$$\forall f \in \mathrm{L}_2(\mathbb{R}), \quad \mathcal{F}^{-1}(f) = \frac{1}{2\pi}\,\sigma \circ \mathcal{F}(f).$$

La transformée de Fourier d'une fonction f de $\mathrm{L}_2(\mathbb{R})$ ne peut pas en général s'écrire directement par la formule intégrale de Fourier (l'intégrale n'est peut-être pas absolument convergente), mais on peut toujours dire que f est la limite dans $\mathrm{L}_2(\mathbb{R})$ de la suite $f_n = \mathbf{1}_{[-n,n]} f$, donc $\mathcal{F}(f)$ est la limite dans L_2 des fonctions

$$t \to \int_{-n}^{n} f(x)\,\mathrm{e}^{-ixt}\,dx.$$

Arrivé à ce point, on peut améliorer l'énoncé de notre Fourier inverse. Cela demande un travail supplémentaire que nous ne ferons pas ici.

Proposition 3.7.2. *Si f et $\widehat{f}$ sont dans $\mathrm{L}_1(\mathbb{R})$, la fonction f "est" continue et on a*

$$\forall x \in \mathbb{R}, \quad f(x) = \frac{1}{2\pi}\int_{\mathbb{R}} \widehat{f}(t)\,\mathrm{e}^{ixt}\,dt.$$

On peut donner une démonstration directe et instructive de la formule d'inversion de Fourier, mais qui ne s'appliquera qu'aux fonctions f de classe C^2 sur $\mathbb{R}$, à support dans un intervalle compact $[-a, a]$.

Fixons $x \in \mathbb{R}$; soit N un entier quelconque tel que $\mathrm{N}\,\pi > \max(a, |x|)$; on va appliquer à la restriction de f à l'intervalle $[-\mathrm{N}\,\pi, \mathrm{N}\pi]$ la théorie des séries de Fourier, moyennant un petit changement de normalisation. Dans l'espace $\mathrm{L}_2([-\mathrm{N}\,\pi, \mathrm{N}\,\pi], \mu_{\mathrm{N}})$, où μ_{N} est la probabilité $dx/(2\pi\mathrm{N})$ sur l'intervalle $[-\mathrm{N}\,\pi, \mathrm{N}\,\pi]$, on considère la suite orthonormée $(e_n)_{n\in\mathbb{Z}}$ définie par $e_n(x) = \mathrm{e}^{inx/\mathrm{N}}$ pour tout $n \in \mathbb{Z}$ et $|x| \le \mathrm{N}\,\pi$. Les coordonnées $(c_n)_{n\in\mathbb{Z}}$ de f dans cette base sont données par

$$c_n = \int_{-\mathrm{N}\pi}^{\mathrm{N}\pi} f(x)\overline{e_n(x)}\,\frac{dx}{2\pi\mathrm{N}} = \frac{1}{2\pi\mathrm{N}}\int_{\mathbb{R}} f(x)\,\mathrm{e}^{-inx/\mathrm{N}}\,dx = \frac{1}{2\pi\mathrm{N}}\,\widehat{f}(n/\mathrm{N}).$$

D'après le théorème 6.1, on sait que

$$f(x) = \sum_{n\in\mathbb{Z}} c_n e_n(x) = \frac{1}{2\pi\mathrm{N}}\sum_{n\in\mathbb{Z}} \widehat{f}(n/\mathrm{N})\,\mathrm{e}^{inx/\mathrm{N}}.$$

Pour conclure, il nous reste à voir que

$$\lim_{\mathrm{N}} \frac{1}{2\pi\mathrm{N}}\sum_{n\in\mathbb{Z}} \widehat{f}(n/\mathrm{N})\,\mathrm{e}^{inx/\mathrm{N}} = \frac{1}{2\pi}\int_{\mathbb{R}} \widehat{f}(t)\,\mathrm{e}^{ixt}\,dt.$$

Si la somme en n s'étendait de $-\mathrm{NT}$ à NT pour un entier T fixé, cette convergence serait le résultat classique de convergence des sommes de Riemann pour une fonction continue sur un intervalle compact $[-\mathrm{T}, \mathrm{T}]$. Pour traiter le cas présent, il nous faut un petit lemme, dont la démonstration est laissée en exercice pour le lecteur.

Lemme. *Si g est continue sur $\mathbb{R}$ et si $|g(t)| \le \min(1, t^{-2})$ pour tout $t \in \mathbb{R}$, on a*

$$\int_{\mathbb{R}} g(t)\,dt = \lim_{\mathrm{N}} \frac{1}{\mathrm{N}}\sum_{n\in\mathbb{Z}} g(n/\mathrm{N}).$$

Il reste à voir que $g(t) = \widehat{f}(t)\,\mathrm{e}^{ixt}$ vérifie les hypothèses du lemme. On sait d'abord que g est continue sur $\mathbb{R}$, et bornée par $\|f\|_{\mathrm{L}_1(\mathbb{R})}$. Ensuite, deux intégrations par parties donnent $\widehat{f''}(t) = -t^2\widehat{f}(t)$, de sorte que $|g(t)| = |\widehat{f}(t)| \le t^{-2}\,\|f''\|_{\mathrm{L}_1(\mathbb{R})}$. Il en résulte que $|g(t)| \le \mathrm{C}\,\min(1, t^{-2})$ pour tout $t \in \mathbb{R}$, avec $\mathrm{C} = \max(\|f\|_{\mathrm{L}_1(\mathbb{R})}, \|f''\|_{\mathrm{L}_1(\mathbb{R})})$.

4. Les théorèmes fondamentaux

4.1. Le théorème de Baire et ses conséquences

Soit X un espace topologique ; un ouvert U de X est dense dans X si et seulement si le fermé complémentaire U^c est d'intérieur vide. Si on a un nombre fini d'ouverts denses, on vérifie facilement de proche en proche que $U_1 \cap \ldots \cap U_n$ est encore un ouvert dense. Le théorème de Baire donne un cas où cette propriété triviale d'intersection finie peut s'étendre aux *suites* d'ouverts denses.

Théorème 4.1.1 : théorème de Baire. *Soit* X *un espace métrique complet ; si* $(U_n)_{n\ge 0}$ *est une suite de parties ouvertes et* denses *dans* X, *l'intersection* $\bigcap_{n\ge 0} U_n$ *est dense dans l'espace* X.

Démonstration. Soit $(U_n)_{n\ge 0}$ une suite d'ouverts denses de X ; soit V une partie ouverte non vide de X ; on doit montrer que $\bigcap_{n\ge 0} U_n$ rencontre V. Comme U_0 est dense, U_0 rencontre V et on peut choisir un point $x_0 \in V \cap U_0$. Comme $V \cap U_0$ est ouvert, il existe un nombre $r_0 > 0$, que l'on peut choisir ≤ 1, tel que la boule ouverte $B(x_0, 2r_0)$ de centre x_0 et de rayon $2r_0$ soit contenue dans $V \cap U_0$.

Par récurrence sur $n \ge 0$ on construit une suite (x_n) d'éléments de X et une suite (r_n) de nombres réels strictement positifs tels que $r_n \le 2^{-n}$ et tels que, pour tout $n \ge 1$, la boule ouverte $B(x_n, 2r_n)$ de centre x_n et de rayon $2r_n$ soit contenue dans $U_n \cap B(x_{n-1}, r_{n-1})$: en effet, supposons x_n et r_n construits ; comme U_{n+1} est dense, il existe $x_{n+1} \in U_{n+1} \cap B(x_n, r_n)$. Comme $U_{n+1} \cap B(x_n, r_n)$ est ouvert, il existe un nombre r_{n+1} tel que $0 < r_{n+1} \le 2^{-n-1}$ et tel que la boule ouverte $B(x_{n+1}, 2r_{n+1})$ soit contenue dans $U_{n+1} \cap B(x_n, r_n)$ (on notera bien le petit jeu entre r_n et $2r_{n+1}$).

Notons maintenant B_n la boule *fermée* de centre x_n et de rayon r_n. On a

$$B_{n+1} \subset B(x_{n+1}, 2r_{n+1}) \subset B(x_n, r_n) \subset B_n.$$

Comme l'espace X est complet, que les ensembles B_n sont fermés, décroissants, non vides et que leur diamètre tend vers 0, on a $\bigcap_{n\ge 0} B_n \ne \emptyset$; or, par construction, $\bigcap_{n\ge 0} B_n \subset V \cap \bigcap_{n\ge 0} U_n$, ce qui montre que cette dernière intersection est non vide.

//

Corollaire 4.1.2. *Soient* X *un espace métrique complet* non vide *et* $(F_n)_{n\ge 0}$ *une suite de parties fermées de* X *telle que* $\bigcup_{n\ge 0} F_n = X$ *; alors l'un des fermés* F_n *a un intérieur non vide ; en réalité, on peut même dire que* $\bigcup_{n\ge 0} \mathring{F}_n$ *est dense dans* X.

Démonstration. Soit V un ouvert non vide de X ; dans l'espace métrique complet $Y = \overline{V}$, considérons les ouverts (relatifs) $U_n = \overline{V} \setminus F_n$. Puisque $\bigcup_{n\ge 0} F_n = X$, l'intersection $\bigcap_{n\ge 0} U_n$ est vide, ce qui entraîne par le théorème 1 appliqué à Y que l'un au moins des ouverts U_n n'est pas dense dans Y. Il existe donc n_0 et un ouvert non vide U de Y qui soit disjoint de U_{n_0}. Cet ouvert U doit rencontrer V ; on peut donc trouver $x \in V \cap U$ et $r > 0$ tels que $B(x, r) \subset V$ ne rencontre pas U_{n_0}, c'est à dire $B(x, r) \subset F_{n_0}$ et en particulier $x \in \mathring{F}_{n_0}$. On a ainsi montré que la réunion des intérieurs des (F_n) rencontre tout ouvert non vide V donné.

//

Exercice 4.1.1. Montrer qu'il n'existe pas de norme sur l'espace vectoriel $\mathbb{R}[X]$ des polynômes qui rende cet espace complet.

Théorème 4.1.3 : théorème des isomorphismes. *Soient* E *et* F *deux espaces de Banach ; toute application linéaire continue bijective de* E *sur* F *est un isomorphisme.*

On dit qu'une application $f : X \to Y$ est *ouverte* lorsque l'image de tout ouvert de X est ouverte dans Y. Notons que la composition de deux applications ouvertes est une application ouverte.

Théorème 4.1.5 : théorème de l'application ouverte. *Soient* E *et* F *deux espaces de Banach ; toute application linéaire, continue, surjective f de* E *sur* F *est ouverte.*

Démonstration. On considère la factorisation $f = g \circ \pi$

$$E \xrightarrow{\pi} E/\ker f \xrightarrow{g} F$$

donnée par la proposition 1.3.3 ; la première flèche π est la projection canonique de E sur le quotient par le noyau de f. Par des arguments algébriques, la deuxième flèche g est bijective, et elle est continue d'après la proposition 1.3.3. C'est donc un isomorphisme, et il en résulte que f est ouverte parce que π et g sont ouvertes.

//

Le graphe d'une application continue d'un espace topologique dans un espace topologique séparé est toujours fermé. La réciproque n'est en général pas vraie. Cependant, on a :

Théorème 4.1.6 : théorème du graphe fermé. *Soient* E *et* F *deux espaces de Banach ; toute application linéaire de* E *dans* F *dont le graphe est fermé dans* $E \times F$ *est continue.*

Démonstration. Soit f une application linéaire de E dans F dont le graphe $G \subset E \times F$ est fermé ; alors G est un espace de Banach. Tout point z du graphe G est de la forme $z = (x, f(x))$ pour un certain $x \in E$ unique ; notons $p : G \to E$ l'application définie par $p(z) = p(x, f(x)) = x \in E$. Il est clair que p est linéaire, continue et bijective (l'inverse –algébrique– étant l'application $x \to (x, f(x))$ de E dans G). D'après le théorème des isomorphismes, cet inverse $x \to (x, f(x))$ est continu de E dans G ; il en résulte que $x \to f(x)$ est continue de E dans F.

//

Nous passons maintenant à une autre conséquence du théorème de Baire, le théorème de Banach-Steinhaus ; ce théorème admet plusieurs variantes ; en voici une première, qui sort un peu de notre cadre habituel d'espaces normés.

Proposition 4.1.7. *Soit* E *un espace vectoriel muni d'une distance* d*, telle que* (E, d) *soit complet, et telle que les opérations* $(x, y) \to x + y$ *et* $(\lambda, x) \to \lambda x$ *soient continues de* $\mathrm{E} \times \mathrm{E}$ *dans* E *et* $\mathbb{K} \times \mathrm{E}$ *dans* E *respectivement ; soient d'autre part* Y *un espace normé et* A *une famille d'applications linéaires continues de* E *dans* Y*. Si pour tout* $x \in \mathrm{E}$ *la famille* $\{\mathrm{T}(x) : \mathrm{T} \in \mathrm{A}\}$ *est bornée dans* Y*, il existe un voisinage* W *de* 0_{E} *tel que*

$$\forall \mathrm{T} \in \mathrm{A}, \forall x \in \mathrm{W}, \quad \|\mathrm{T}(x)\| \leq 1.$$

Démonstration. Remarquons d'abord que pour tout $x_0 \in \mathrm{E}$ la translation $y \to x_0 + y$ est un homéomorphisme de E ; de même, pour tout $\lambda \neq 0$ l'homothétie $y \to \lambda y$ est un homéomorphisme. Il résulte du premier point que tout voisinage W de x_0 est de la forme $x_0 + \mathrm{V}$, où V est un voisinage de 0_{E}, et du second point que $\lambda \mathrm{V}$ est aussi un voisinage de 0_{E}.

Pour tout entier $n \geq 1$, posons $\mathrm{C}_n = \{x \in \mathrm{E} : \forall \mathrm{T} \in \mathrm{A},\ \|\mathrm{T}(x)\| \leq n\}$. Comme C_n est l'intersection des ensembles fermés $\mathrm{C}_{\mathrm{T},n} = \{x \in \mathrm{E} : \|\mathrm{T}(x)\| \leq n\}$ (lorsque T varie dans A), c'est un fermé de E. La réunion des C_n est égale à E : ceci n'est que la traduction de l'hypothèse $\sup_{\mathrm{T} \in \mathrm{A}} \|\mathrm{T}(x)\| < +\infty$ pour tout $x \in \mathrm{E}$.

Puisque (E, d) est métrique complet, il existe par le corollaire 2 un entier $n_0 \geq 1$ tel que C_{n_0} soit d'intérieur non vide. On peut donc trouver un point $x_0 \in \mathrm{C}_{n_0}$ et un voisinage V de 0_{E} tels que $x_0 + \mathrm{V} \subset \mathrm{C}_{n_0}$. Posons $\mathrm{M} = \sup_{\mathrm{T} \in \mathrm{A}} \|\mathrm{T}(x_0)\|_{\mathrm{Y}}$. Soient $v \in \mathrm{V}$ et $\mathrm{T} \in \mathrm{A}$ quelconques ; puisque $x_0 + v \in \mathrm{C}_{n_0}$, on a $\|\mathrm{T}(x_0) + \mathrm{T}(v)\|_{\mathrm{Y}} \leq n_0$, ce qui donne $\|\mathrm{T}(v)\|_{\mathrm{Y}} \leq n_0 + \|\mathrm{T}(x_0)\|_{\mathrm{Y}} \leq n_0 + \mathrm{M}$ par l'inégalité triangulaire. Pour terminer, on prend le voisinage $\mathrm{W} = (n_0 + \mathrm{M})^{-1} \mathrm{V}$.

//

Théorème 4.1.8 : théorème de Banach-Steinhaus. *Soient* E *un espace de Banach,* Y *un espace normé et* A *une partie de* $\mathcal{L}(\mathrm{E}, \mathrm{Y})$ *telle que* $\sup\{\|\mathrm{T}(x)\| : \mathrm{T} \in \mathrm{A}\} < +\infty$ *pour tout* $x \in \mathrm{E}$ *; alors on a aussi* $\sup\{\|\mathrm{T}\| : \mathrm{T} \in \mathrm{A}\} < +\infty$.

Démonstration. On peut appliquer la proposition précédente. Il existe un voisinage W de 0_{E} tel que $\|\mathrm{T}(x)\| \leq 1$ pour tout $x \in \mathrm{W}$ et tout $\mathrm{T} \in \mathrm{A}$. Il existe $r > 0$ tel que $\mathrm{B}(0, r) \subset \mathrm{W}$. Par homogénéité, pour tout $x \in \mathrm{B}(0, 1)$ et tout $\mathrm{T} \in \mathrm{A}$, on a $\|\mathrm{T}(x)\| \leq 1/r$; on a donc montré que pour tout $\mathrm{T} \in \mathrm{A}$, on a $\|\mathrm{T}\| \leq 1/r$.

//

Corollaire 4.1.9. *Soient* E *un espace de Banach (ou bien un espace vectoriel* (E, d) *complet comme dans la proposition 7),* Y *un espace normé et* (f_n) *une suite d'applications linéaires continues de* E *dans* Y *; on suppose que, pour tout* $x \in \mathrm{E}$*, la suite* $(f_n(x))$ *converge dans* Y *; notons* $f(x)$ *sa limite. Alors* f *est linéaire et continue.*

Démonstration. D'abord, il est évident que la limite f est linéaire. Soit $x \in \mathrm{E}$; comme la suite $(f_n(x))$ est convergente, elle est bornée ; par le théorème 8, la suite $(\|f_n\|)$ est alors bornée. Il existe alors un nombre $\mathrm{M} \geq 0$ tel que, pour tout $x \in \mathrm{E}$ et tout entier $n \geq 0$ on ait $\|f_n(x)\| \leq \mathrm{M}\, \|x\|$. Passant à la limite on trouve $\|f(x)\| \leq \mathrm{M}\, \|x\|$, pour tout $x \in \mathrm{E}$.

//

Corollaire 4.1.10. *Soient* E *un espace de Banach (ou bien un* (E, d) *complet comme dans la proposition 7),* Y *un espace normé et* (u_k) *une suite d'applications linéaires continues de* E *dans* Y *; on suppose que, pour tout* $x \in \mathrm{E}$*, la série* $\sum_k u_k(x)$ *converge dans* Y *; notons* $\mathrm{T}(x)$ *sa somme. Alors* T *est linéaire et continue.*

Commentaire. Ceux qui feront des distributions verront ressortir ce principe à propos des séries de distributions : les *distributions tempérées* sont les formes linéaires continues sur un espace de fonctions $\mathcal{S}$ qui est un (E, d) du bon type. Pour vérifier qu'une série $\sum \mathrm{T}_k$ de distributions tempérées définit une nouvelle distribution tempérée, il suffit de vérifier que la série numérique $\sum_k \mathrm{T}_k(\varphi)$ converge pour toute fonction $\varphi \in \mathcal{S}$.

4.2. Théorème de Hahn-Banach

Le premier résultat que nous allons énoncer est purement algébrique, et ne fait pas référence à une topologie sur l'espace vectoriel (réel) X. On dit que $q : \mathrm{X} \to \mathbb{R}$ est *sous-linéaire* si elle est positivement homogène et sous-additive, c'est à dire qu'elle vérifie

(i) pour tout $x \in \mathrm{X}$, on a $q(\lambda x) = \lambda q(x)$ pour tout $\lambda \geq 0$
(ii) pour tous $x, y \in \mathrm{X}$, on a $q(x + y) \leq q(x) + q(y)$.

Exemples 4.2.1. Les semi-normes sont des fonctions sous-linéaires. Une forme linéaire (réelle) est une fonction sous-linéaire. Une fonction sous-linéaire sur $\mathbb{R}$ est linéaire par morceaux : elle vaut $q(t) = at$ pour $t < 0$ et $q(t) = bt$ pour $t \geq 0$, avec $a \leq b$ (en effet, on doit avoir $-a + b = q(-1) + q(1) \geq q(0) = 0$).

Théorème 4.2.1 : théorème de prolongement de Hahn-Banach. *Soient* X *un espace vectoriel réel,* Y *un sous-espace vectoriel de* X *et* q *une fonction sous-linéaire sur* X *; pour toute forme linéaire* ℓ *sur* Y*, telle que* $\ell(y) \leq q(y)$ *pour tout* $y \in \mathrm{Y}$*, il existe une forme linéaire* m *sur* X *qui prolonge* ℓ*, c'est à dire telle que* $m(y) = \ell(y)$ *pour tout* $y \in \mathrm{Y}$ *et telle que* $m(x) \leq q(x)$ *pour tout* $x \in \mathrm{X}$.

Petite remarque évidente avant de commencer la démonstration : dans le cas $\mathrm{X} = \mathbb{R}$ et $\mathrm{Y} = \{0\}$, on a vu à quoi ressemble le graphe des fonctions sous-linéaires sur $\mathbb{R}$ et on voit bien pourquoi le résultat est vrai : on sait que $q(t) = a\,t$ pour $t < 0$, $q(t) = b\,t$ pour $t \geq 0$ et de plus $a \leq b$; il suffit de prendre n'importe quelle fonction linéaire $m(t) = c\,t$ avec $a \leq c \leq b$.

Démonstration. Le point crucial est de montrer qu'on peut prolonger à une dimension de plus : si m est linéaire, définie sur un sous-espace vectoriel Z de X, de façon que $m \leq q$ et si $x \notin \mathrm{Z}$, on peut étendre m en $\widetilde{m}$ définie sur $\mathrm{Z} + \mathbb{R}x$ en gardant $\widetilde{m} \leq q$; le reste n'est que formalité "zornique".

Lemme 4.2.2. *Soient* Z *un sous-espace vectoriel de* X *et* g *une forme linéaire définie sur* Z*, telle que* $g(z) \leq q(z)$ *pour tout* $z \in \mathrm{Z}$ *; soit* $x \in \mathrm{X}$ *tel que* $x \notin \mathrm{Z}$ *; il existe une forme linéaire* $\widetilde{g}$ *sur* $\mathrm{Z} + \mathbb{R}x$ *telle que* $\widetilde{g}$ *prolonge* g *et* $\widetilde{g} \leq q$ *sur* $\mathrm{Z} + \mathbb{R}x$.

Démonstration du lemme. Bien entendu, prolonger g à $\mathrm{Z} + \mathbb{R}x$ demande seulement de définir $\gamma = \widetilde{g}(x)$. Pour que le prolongement soit convenable, il faut (et il suffit) que $g(z)+t\,\widetilde{g}(x)=\widetilde{g}(z+tx)\le q(z+tx)$ pour tout nombre réel t et tout $z\in \mathrm{Z}$. C'est automatique si $t=0$, et nous allons découper la propriété voulue en deux, selon le signe de $t\ne 0$:

$$g(z)+\lambda\gamma\le q(z+\lambda x),\quad g(z')-\mu\gamma\le q(z'-\mu x)$$

pour tous $z,z'\in \mathrm{Z}$ et $\lambda,\mu>0$. En utilisant l'homogénéité de q (et celle de g, qui est linéaire) on peut faire entrer les facteurs positifs λ^{-1} et μ^{-1} à l'intérieur des expressions, et on obtient ainsi les conditions équivalentes

$$g(z_1)+\gamma\le q(z_1+x),\quad g(z_2)-\gamma\le q(z_2-x)$$

pour tous $z_1,z_2\in \mathrm{Z}$ (z_1 remplace $\lambda^{-1}z$ et z_2 remplace $\mu^{-1}z'$). Le nombre γ doit donc vérifier les deux inégalités

$$\sup\{g(z_2)-q(z_2-x) : z_2\in \mathrm{Z}\}=\mathrm{S}\le\gamma\le\mathrm{I}=\inf\{q(z_1+x)-g(z_1) : z_1\in \mathrm{Z}\}.$$

Notons que I n'est pas $+\infty$, parce que l'inf porte sur un ensemble non vide de valeurs finies, et de même S n'est pas $-\infty$. Pour que le choix de γ soit possible, il faut et il suffit que $\mathrm{S}\le\mathrm{I}$, ce qui garantira que I et S sont finis, et il suffira de prendre pour γ n'importe quel nombre réel compris entre le sup et l'inf (bien sûr, si $\mathrm{S}=\mathrm{I}$ on n'a pas le choix : il faut prendre pour γ la valeur commune). Il reste donc à vérifier que

$$g(z_2)-q(z_2-x)\le q(z_1+x)-g(z_1)$$

pour tous $z_1,z_2\in \mathrm{Z}$. On réécrit la propriété voulue sous la forme

$$g(z_1+z_2)=g(z_1)+g(z_2)\le q(z_1+x)+q(z_2-x)$$

et il est alors clair que cette propriété est vraie :

$$g(z_1+z_2)\le q(z_1+z_2)=q\big((z_1+x)+(z_2-x)\big)\le q(z_1+x)+q(z_2-x).$$

Le lemme est donc établi.

Le lemme de Zorn

Le lemme de Zorn est assez directement équivalent à un axiome de la théorie des ensembles, *l'axiome du choix*. Il permet de valider certains types de raisonnements où on cherche à garantir l'existence d'objets maximaux.

Soit I *un ensemble ordonné dans lequel tout sous-ensemble totalement ordonné* T *possède des majorants ; l'ensemble* I *admet alors des éléments maximaux.*

Un élément maximal $i\in \mathrm{I}$ est un élément tel que $(j\ge i)\Rightarrow j=i$ pour tout $j\in \mathrm{I}$. Le lemme de Zorn n'a d'intérêt que pour les ensembles ordonnés qui ne sont pas totalement ordonnés. On dit qu'un ensemble ordonné est *inductif* lorsqu'il vérifie l'hypothèse du lemme de Zorn.

Venons-en à l'application du lemme de Zorn pour terminer la démonstration du théorème. On désigne par I l'ensemble des couples (Z,g) où Z est un sous-espace vectoriel de X tel que $\mathrm{Y}\subset \mathrm{Z}$, et g une forme linéaire sur Z qui prolonge ℓ, et telle que $g(z)\le q(z)$ pour tout $z\in \mathrm{Z}$.

On définit l'ordre sur l'ensemble I par $(\mathrm{Z},g)\le(\mathrm{Z}',g')$ si $\mathrm{Z}\subset \mathrm{Z}'$ et si g' est un prolongement de g à Z'. Le lemme préliminaire 2 dit que si (Z,g) est un élément maximal de I, alors $\mathrm{Z}=\mathrm{X}$: sinon, si $\mathrm{Z}\ne \mathrm{X}$, on peut choisir $x\notin \mathrm{Z}$ et considérer l'extension (Z',g') à $\mathrm{Z}'=\mathrm{Z}+\mathbb{R}x$ donnée par le lemme 2, qui est un majorant strict de (Z,g). Cela signifie que l'existence d'éléments maximaux dans I implique qu'on a réussi à prolonger ℓ à l'espace X tout entier, avec une extension linéaire m qui vérifie $m\le q$ sur X.

Il reste à vérifier que l'ensemble I vérifie l'hypothèse du lemme de Zorn : si (Z_i, g_i) est une famille totalement ordonnée dans I, on verra que $Z = \bigcup_i Z_i$ est un sous-espace vectoriel et qu'il y a une façon naturelle de définir g sur Z, qui prolonge toutes les g_i. Ainsi l'ensemble $(Z_i, g_i)_i$ admet le majorant (Z, g) dans I.

Théorème 4.2.3 : théorème de séparation de Hahn-Banach. *Soient* X *un espace normé réel,* A *un convexe ouvert non vide et* B *un convexe non vide tels que* A *et* B *soient disjoints. Il existe alors une forme linéaire continue f sur* X *telle que*

$$f(a) < \inf f(B)$$

pour tout $a \in A$. Autrement dit, il existe un nombre c tel que $f(a) < c$ pour tout $a \in A$ et $c \le f(b)$ pour tout $b \in B$.

Une façon de voir le résultat est de dire que la forme linéaire f sépare l'espace X en deux demi-espaces affines $H_- = \{f < c\}$ et $H_+ = \{f \ge c\}$, dont la frontière commune est l'hyperplan affine $H = \{f = c\}$. L'énoncé nous dit que $A \subset H_-$ et $B \subset H_+$.

Rappelons que X^* désigne le dual topologique d'un espace normé X.

Exercices 4.2.2.

Si C est fermé, K compact, C et K convexes non vides et disjoints d'un espace normé X, montrer qu'il existe une forme linéaire continue $x^* \in X^*$ telle que $\sup x^*(C) < \min x^*(K)$.

Soient X un espace normé et q une fonction sous-linéaire sur X ; montrer que q est continue sur X si et seulement s'il existe une constante M telle que $|q(x)| \le M\,\|x\|$ pour tout $x \in X$.

Corollaire 4.2.4. *Si* C *est un sous-ensemble convexe fermé non vide d'un espace normé réel* X*, alors* C *est l'intersection de demi-espaces affines fermés.*

Démonstration. Soit C un convexe fermé non vide d'un espace normé réel X. On va montrer que pour tout $x \notin C$, il existe un demi-espace affine fermé D_x tel que $C \subset D_x$ et $x \notin D_x$. Il suffira ensuite d'observer que $C = \bigcap_{x \notin C} D_x$.

Pour tout $x \notin C$, on peut trouver une boule ouverte $A = B(x, r)$ disjointe de C ; d'après le théorème de séparation il existe une forme linéaire continue x^* telle que $x^*(a) < \inf x^*(C)$ pour tout $a \in A$, et en particulier $x^*(x) < \inf x^*(C)$. On voit donc que si on pose $d = \inf x^*(C)$ et

$$D_x = \{y \in X : x^*(y) \ge d\}$$

on aura $C \subset D_x$ mais $x \notin D_x$.

//

Théorème 4.2.5 : théorème de Hahn-Banach. *Soient* X *un espace normé (réel ou complexe) et* Y *un sous-espace vectoriel de* X *; pour tout $\ell \in Y^*$, il existe $m \in X^*$ dont la restriction à* Y *soit ℓ et telle que $\|m\| = \|\ell\|$.*

Démonstration. Considérons d'abord le cas réel. Ici la fonction sous-linéaire q de l'énoncé du théorème 1 sera un multiple convenable de la norme N de X. Par définition de la norme de la forme linéaire ℓ, on a $\ell \le \|\ell\|\,N = q$ sur le sous-espace vectoriel Y. On peut donc trouver un prolongement m tel que $m \le q$ sur X, ce qui donne le résultat : on a en effet $m(x) \le \|\ell\| \, \|x\|$ pour tout $x \in \mathrm{X}$, d'où aussi $|m(x)| \le \|\ell\| \, \|x\|$ en appliquant à x et $-x$; tout ceci montre que m est continue et $\|m\| \le \|\ell\|$, mais $\|\ell\| \le \|m\|$ puisque m prolonge ℓ.

Si X est un espace vectoriel complexe, on commence par le considérer comme un espace vectoriel réel, et on considère sur Y la forme linéaire réelle $\ell_1 = \mathrm{Re}\,\ell$. On trouve alors une forme linéaire réelle m_1 sur X telle que m_1 prolonge la forme linéaire réelle ℓ_1 et $\|m_1\| = \|\ell_1\|$. Par la proposition 1.6.1, on sait que m_1 est la partie réelle d'une forme linéaire complexe m sur X, et de plus $\|m\| = \|m_1\| \le \|\ell_1\| = \|\ell\|$; d'autre part m prolonge ℓ (ici Y est un sous-espace vectoriel complexe ; si $y \in \mathrm{Y}$ on a aussi $iy \in \mathrm{Y}$ ce qui permet d'écrire $m(y) = m_1(y) - i m_1(iy)$, et alors $m(y) = \ell_1(y) - i\ell_1(iy) = \ell(y)$).

//

Corollaire 4.2.6. *Soient* X *un espace normé et* Y *un sous-espace vectoriel fermé ; soient* $x \notin \mathrm{Y}$ *et* $r = \mathrm{dist}(x, \mathrm{Y}) > 0$ *; il existe une forme linéaire continue* $x^* \in \mathrm{X}^*$ *telle que :* x^* *est nulle sur* Y, $\|x^*\| = 1$ *et* $x^*(x) = r$.

Démonstration. On a $\|y - x\| \ge r$ pour tout $y \in \mathrm{Y}$, ce qui donne par homogénéité $\|y + \lambda x\| \ge |\lambda| \, r$ pour tous $\lambda \in \mathbb{K}$, $y \in \mathrm{Y}$. Définissons une forme linéaire ℓ sur $\mathrm{Y}_1 = \mathrm{Y} \oplus \mathbb{K}x$ en posant $\ell(y + \lambda x) = \lambda\, r$ pour tous $\lambda \in \mathbb{K}$, $y \in \mathrm{Y}$. L'inégalité qui précède montre que $|\ell(z)| \le \|z\|$ pour tout $z \in \mathrm{Y}_1$, donc $\|\ell\| \le 1$. En appliquant le théorème 5 à ℓ et Y_1, on trouve une forme linéaire continue x^* sur X telle que $\|x^*\| \le 1$ et $x^*(x) = \ell(x) = r$; de plus $x^*(y) = \ell(y + 0\,x) = 0$ pour tout $y \in \mathrm{Y}$. En choisissant $y \in \mathrm{Y}$ tel que $\|y - x\| < r + \varepsilon$ on aura $r = |x^*(y - x)| \le \|x^*\| \, \|y - x\| \le \|x^*\| \, (r + \varepsilon)$, ce qui montre que nécessairement $\|x^*\| = 1$.

//

Corollaire 4.2.7. *Soient* X *un espace normé et* $x \in \mathrm{X}$ *; il existe* $x^* \in \mathrm{X}^*$ *telle que* $x^*(x) = \|x\|$ *et* $\|x^*\| \le 1$.

Démonstration. Si $x = 0_{\mathrm{X}}$ on prendra tout simplement $x^* = 0$; sinon, on applique le corollaire précédent avec $\mathrm{Y} = \{0_{\mathrm{X}}\}$. Bien entendu, on a en fait $\|x^*\| = 1$ lorsque $x \ne 0_{\mathrm{X}}$, mais le corollaire tel qu'il est énoncé a l'avantage de couvrir tous les cas.

//

Remarque 4.2.3. Le dual de X/Y est identifiable isométriquement au sous-espace de X^* formé des x^* dont la restriction à Y est nulle.

Le théorème de Hahn-Banach donne des outils pour étudier la séparabilité. Il fournit en particulier le critère suivant : pour qu'un sous-ensemble $\mathrm{D} \subset \mathrm{X}$ soit total dans X, il faut et il suffit que toute forme linéaire $x^* \in \mathrm{X}^*$, nulle sur D, soit identiquement nulle.

Proposition 4.2.9. *Soit* X *un espace normé ; si le dual* X^* *est séparable, alors* X *est séparable.*

Démonstration. Soit (x_n^*) une suite dense dans X^* ; pour chaque entier $n \ge 0$, on peut trouver un vecteur $x_n \in \mathrm{X}$ tel que $\|x_n\| \le 1$ et $x_n^*(x_n) \ge \|x_n^*\|/2$. On va montrer que la suite (x_n) est totale dans X : sinon, il existerait une forme linéaire continue x^* non nulle sur X telle que $x^*(x_n) = 0$ pour tout entier $n \geq 0$; on peut supposer $\|x^*\| = 1$. D'après la densité de la suite (x_n^*), il existe un indice n_0 tel que $\|x^* - x_{n_0}^*\| < 1/4$. On aurait alors $x_{n_0}^*(x_{n_0}) = (x_{n_0}^* - x^*)(x_{n_0}) < 1/4$, mais $\|x_{n_0}^*\| \geq \|x^*\| - \|x_{n_0}^* - x^*\| \geq 3/4$, ce qui est contradictoire avec $x_{n_0}^*(x_{n_0}) \geq \|x_{n_0}^*\|/2 \geq 3/8 > 1/4$.

//

Hahn-Banach et transposition

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

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

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

$$\|{}^t\mathrm{T}\| \geq \|{}^t\mathrm{T}(y^*)\| \geq |{}^t\mathrm{T}(y^*)(x)| = y^*(\mathrm{T}(x)) = \|\mathrm{T}(x)\| > \|\mathrm{T}\| - \varepsilon.$$

//

Exemple 4.2.4. La transposée de l'injection de $\mathrm{Y} \subset \mathrm{X}$ dans X est l'application de restriction de X^* sur Y^* (surjective par Hahn-Banach). Ces deux applications sont de norme 1 lorsque $\mathrm{Y} \neq \{0\}$.

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

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

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

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

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

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

//

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

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

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

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

Exemple 4.3.1. L'injection de c_0 dans ℓ_∞ correspond à l'injection J_{c_0} de c_0 dans son bidual, modulo les identifications habituelles entre duaux et espaces de suites.

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

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

$$\|x\| = |x^*(x)| = |J_X(x)(x^*)| \leq \|x^*\| \, \|J_X(x)\| \leq \|J_X(x)\|,$$

vu que $\|x^*\| \leq 1$; donc $\|J_X(x)\| = \|x\|$.

//

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

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

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

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

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

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

//

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

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

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

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

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

En revanche, les espaces c_0, ℓ_1 et ℓ_∞ sont des espaces de Banach non réflexifs.

Espaces normés isomorphes

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

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

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

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

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

ce qu'il fallait démontrer.

//

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

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

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

ce qui montre bien que z^{**} provient du vecteur $z \in \mathrm{Z}$.

//

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

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

//

Corollaire 4.3.6. *Si* E *est un espace de Banach et si* E^* *est réflexif, alors* E *est réflexif.*

En effet E^{**} est alors réflexif et E est isomorphe à un sous-espace fermé de E^{**}.

Pourquoi s'intéresser aux espaces réflexifs ?

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

4.4. Théorème de Riesz

Lemme 4.4.1. *Soit* Z *un espace normé de dimension* n *; pour tout* $\varepsilon \in]0,1[$, *on peut trouver dans la boule unité de* Z *une famille* A *d'au moins* ε^{-n} *points dont les distances mutuelles sont* $\geq \varepsilon$ *: si* $x, y \in A$ *et* $x \neq y$, *alors* $\|x - y\| \geq \varepsilon$, *et* $\operatorname{card} A \geq \varepsilon^{-n}$.

Démonstration. Soit A une famille maximale de points de la boule unité B_Z de Z dont les distances mutuelles soient $\geq \varepsilon$; alors les boules de rayon ε centrées aux points de A recouvrent B_Z : en effet, si $x \in B_Z$ et $x \notin A$, on ne peut pas, d'après la maximalité de A, ajouter le point x à la famille A pour former une nouvelle famille A' de points à distances mutuelles $\geq \varepsilon$; cela signifie qu'il existe un point $y \in A$ tel que $d(y, x) < \varepsilon$, donc x est bien contenu dans une boule de rayon ε centrée en un point y de A. Soit V le volume de B_Z ; puisque Z est de dimension n, les boules de rayon ε ont un volume égal à $\varepsilon^n V$ (dans le cas des scalaires réels), et puisque les boules de ce rayon centrées aux points de A recouvrent B_Z, on a $(\operatorname{card} A)\, \varepsilon^n V \geq V$, d'où le résultat.

//

Théorème 4.4.2. *Si la boule unité d'un espace normé* X *est compacte, alors* X *est de dimension finie.*

Démonstration. Si la boule unité de X est compacte, on peut la recouvrir par un nombre fini N de boules B_α de rayon $< 1/4$. Si X était de dimension infinie, on pourrait choisir un sous-espace $Z \subset X$ d'une dimension finie n telle que $2^n > N$; il existerait alors dans la boule unité de Z une famille d'au moins 2^n points tels que $\|z_i - z_j\| \geq 1/2$. Mais alors chacune des boules B_α contiendrait *au plus un* des points (z_i), donc $\mathrm{N} \geq 2^n$, contradiction.

//

On dit qu'un opérateur borné T d'un espace de Banach E dans un espace de Banach F est *compact* si l'adhérence dans F de l'image de la boule unité de E est compacte dans l'espace F.

Corollaire 4.4.3. *Soit* E *un espace de Banach, réel ou complexe ; si* $\mathrm{T} \in \mathcal{L}(\mathrm{E})$ *est compact et si* $\lambda \neq 0$*, le sous-espace* $\mathrm{F}_\lambda = \ker(\mathrm{T} - \lambda\,\mathrm{Id_E}) = \{y \in \mathrm{E} : \mathrm{T}(y) = \lambda y\}$ *est de dimension finie.*

Démonstration. Désignons par K le compact de E égal à l'adhérence de $\mathrm{T}(\mathrm{B_E})$. Pour montrer que le sous-espace $\mathrm{F} = \mathrm{F}_\lambda$ est de dimension finie, il suffit de montrer que la boule unité de F est compacte, et pour cela il suffit de voir que $\mathrm{B_F} \subset |\lambda|^{-1}\mathrm{K}$. Soit $y \in \mathrm{B_F}$; on a

$$y = \frac{1}{\lambda}\,\mathrm{T}(y) \in \mathrm{T}(\lambda^{-1}\mathrm{B_E}) = |\lambda|^{-1}\mathrm{T}(\mathrm{B_E}) \subset |\lambda|^{-1}\mathrm{K}.$$

//

5. Topologies faibles

Dans ce chapitre on introduira la notion de *suite faiblement convergente*, et on verra deux résultats importants :

– si E est un espace de Banach réflexif, toute suite bornée $(x_n) \subset \mathrm{E}$ admet des sous-suites faiblement convergentes ;

– si f est une fonction convexe continue sur un espace normé X et si $(x_n) \subset \mathrm{X}$ converge faiblement vers $x \in \mathrm{X}$, alors $f(x) \le \liminf_n f(x_n)$.

Ces deux résultats permettent de minimiser certaines fonctions convexes définies sur des espaces réflexifs.

5.1. Topologies initiales

Voir la version longue du poly.

5.2. Topologie faible sur un espace normé

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

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

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

doit être un ouvert de $\mathcal{T}$. C'est aussi un ouvert de $\mathcal{O}$. Puisque toute intersection finie d'ouverts est un ouvert, la famille d'ouverts $\mathcal{T}$ doit contenir tout ensemble B de la forme

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

pour tout entier $k \ge 1$, toutes suites finies $x_1^*, \ldots, x_k^*$ dans X^* et $a_1, b_1, \ldots, a_k, b_k$ dans $\mathbb{R}$. Notons encore que tout ensemble B est dans $\mathcal{O}$.

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

La topologie $\mathcal{T}$ sur X s'appelle la *topologie faible* sur X. La topologie faible est plus faible que la topologie de la norme, puisqu'on a vérifié que tous les ouverts faibles sont des ouverts de la topologie de la norme.

Soit X un espace vectoriel normé, réel ou complexe ; pour que $\mathrm{W} \subset \mathrm{X}$ soit un voisinage du point $x_0 \in \mathrm{X}$ pour la topologie faible (on dira que W est un *voisinage faible* de x_0), il faut et il suffit qu'il existe un nombre fini de formes linéaires continues $x_1^*, \ldots, x_n^* \in \mathrm{X}^*$ et un nombre $\varepsilon > 0$ tels que

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

La topologie faible sur X est la topologie la moins fine rendant continues toutes les applications $x \in \mathrm{X} \to x^*(x)$, où x^* décrit l'ensemble de toutes les formes linéaires continues (en norme) sur X ; bien entendu la topologie de la norme rend déjà continues toutes ces applications, donc la topologie faible $\sigma(X, X^*)$ est plus faible que la topologie de la norme. D'après le théorème de Hahn-Banach, la topologie faible est séparée : si $x_1 \neq x_2$, on peut trouver $x^* \in X^*$ telle que $x^*(x_1 - x_2) \neq 0$; les deux ensembles $U_j = \{x \in X : |x^*(x - x_j)| < |x^*(x_1 - x_2)|/2\}$, $j = 1, 2$, sont deux ouverts faibles disjoints contenant x_1 et x_2.

Proposition 5.2.1. *Soient* X *et* Y *deux espaces normés et* $T \in \mathcal{L}(X, Y)$ *; alors* T *est continue de* X *muni de la topologie* $\sigma(X, X^*)$ *dans* Y *muni de la topologie* $\sigma(Y, Y^*)$.

Théorème 5.2.2. *Soient* X *un espace normé et* C *un sous-ensemble convexe de* X *; l'ensemble* C *est fermé en norme si et seulement s'il est faiblement fermé.*

Le résultat s'applique en particulier quand Y un sous-espace vectoriel de X ; alors Y est fermé pour la topologie de la norme si et seulement s'il est fermé pour $\sigma(X, X^*)$.

Démonstration. Comme la topologie de la norme est plus fine que la topologie faible, toute partie fermée pour la topologie faible est fermée pour la topologie de la norme. Démontrons l'inverse ; si X est réel, les demi-espaces affines fermés de la forme $\{x \in X : x^*(x) \geq c\}$, où $x^* \in X^*$, sont faiblement fermés puisque x^* est faiblement continue ; dans le cas complexe la forme $\mathbb{R}$-linéaire $\operatorname{Re} x^*$ est une application faiblement continue de X dans $\mathbb{R}$; il résulte de ces considérations et du corollaire 4.2.4 que tout convexe fermé est faiblement fermé.

//

Soit X un espace vectoriel normé ; on va maintenant s'intéresser à une topologie faible sur le dual X^*, la topologie $\sigma(X^*, X)$ ou *topologie $*$-faible* sur X^*. C'est la topologie la moins fine sur X^* rendant continues toutes les applications $x^* \in X^* \to x^*(x)$, où x décrit X ; bien entendu, puisque la topologie de la norme de X^* rend continues toutes ces applications, la topologie $*$-faible est plus faible que la topologie de la norme sur X^*. La topologie $*$-faible est la *topologie de la convergence simple sur* X, topologie sur l'ensemble d'applications X^*, applications de X dans $\mathbb{K}$.

Théorème 5.2.3. *Muni de la topologie* $\sigma(X^*, X)$ *la boule unité de* X^* *est compacte.*

Ce théorème est un corollaire du théorème de Tykhonov que nous admettrons.

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

Rappelons que la topologie produit sur $\prod_{i \in I} X_i$ est la topologie la moins fine qui rende continues toutes les projections $\prod_{i \in I} X_i \to X_j$, $j \in I$.

Démonstration du théorème 3. Remarquons que X^* est un sous-ensemble de l'ensemble $\mathbb{K}^X$ de toutes les applications de X dans $\mathbb{K}$, et que la topologie $\sigma(X^*, X)$ est, par définition, la topologie induite sur X^* par la topologie produit sur $\mathbb{K}^X$. Remarquons ensuite que la boule unité B de X^* est l'intersection de deux ensembles fermés dans $\mathbb{K}^X$,

$$F_1 = \{f \in \mathbb{K}^X : \forall x \in X,\ |f(x)| \leq \|x\|\},$$

$$F_2 = \{f \in \mathbb{K}^X : \forall (x, y, \lambda, \mu) \in X \times X \times \mathbb{K} \times \mathbb{K},\ f(\lambda x + \mu y) = \lambda f(x) + \mu f(y)\}.$$

Toutes ces conditions définissent des fermés, donc $B = F_1 \cap F_2$ est un fermé de $\mathbb{K}^X$. Pour $r \geq 0$, posons $D_r = \{\lambda \in \mathbb{K} : |\lambda| \leq r\}$. On a $B \subset \prod_{x\in X} D_{\|x\|}$, qui est compact par le théorème 4 ; étant fermé dans un compact, B est compact.

//

Corollaire 5.2.5. *Si* E *est réflexif, les convexes fermés bornés de* E *sont faiblement compacts.*

5.3. Suites faiblement convergentes

Il est intéressant de revoir certaines de ces propriétés de compacité "à la main", et avec des *suites*. Rappelons qu'une suite $(x_n^*) \subset X^*$ est $*$-faiblement convergente vers un vecteur x^* si $\lim_n x_n^*(x) = x^*(x)$ pour tout $x \in X$; une suite $(x_n) \subset X$ est faiblement convergente vers $x \in X$ si $x^*(x) = \lim_n x^*(x_n)$ pour tout $x^* \in X^*$.

Exemples 5.3.1.

1. Lorsque H est un espace de Hilbert, toute forme linéaire continue sur H est de la forme $x \to \langle x, y\rangle$ pour un certain vecteur $y \in H$; il en résulte qu'une suite $(x_n) \subset H$ converge faiblement vers $x \in H$ si et seulement si

$$\forall y \in H, \quad \langle x, y\rangle = \lim_n \langle x_n, y\rangle.$$

2. Soit $(e_n)_{n\geq 0}$ une suite orthonormée dans un espace de Hilbert ; alors e_n converge faiblement vers 0.

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

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

pour tout $x \in E$ (inégalité de Bessel - théorème 6.3.2) ; la suite $(\langle e_n, x\rangle)$ est de carré sommable donc tend vers 0.

3. La suite canonique $(\mathbf{e}_n)_{n\geq 0}$ tend faiblement vers 0 dans c_0, et dans l'espace ℓ_p si $1 < p < +\infty$, mais **pas** dans ℓ_1 ou ℓ_∞.

Lemme 5.3.1. *Dans un espace normé toute suite faiblement convergente est bornée. Si* X *est complet, toute suite $*$-faiblement convergente dans* X^* *est bornée.*

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

//

Proposition 5.3.2. *Si f est une fonction réelle convexe définie sur un convexe fermé* C *d'un espace normé* X*, et si $(x_n) \subset$ C converge faiblement vers x, on a $f(x) \le \liminf_n f(x_n)$.*

Démonstration. Posons $\ell = \liminf_n f(x_n)$. Si $\ell = +\infty$ l'inégalité à démontrer est évidente. Sinon, soit $m > \ell$. Considérons l'ensemble $\mathrm{D}_m = \{y \in \mathrm{C} : f(y) \le m\}$. C'est un convexe fermé, donc faiblement fermé, et puisque $\liminf_n f(x_n) < m$ il existe une sous-suite $(x_{n_k})_{k \ge 0}$ contenue dans D_m, donc sa limite faible x reste dans D_m. On a donc $f(x) \le m$ pour tout $m > \ell$, ce qui montre que $f(x) \le \liminf_n f(x_n)$.
//

Lemme 5.3.3. *Soit (x_n^*) une suite* **bornée** *dans le dual d'un espace normé* X *; pour que cette suite soit $*$-faiblement convergente vers x^*, il suffit que $x^*(d) = \lim_n x_n^*(d)$ pour tout d d'un ensemble* D *total dans* X*. Pour que cette suite soit $*$-faiblement convergente vers une limite dans* X^**, il suffit que $\lim_n x_n^*(d)$ existe pour tout d d'un ensemble* D *total dans* X*.*

Démonstration. Montrons la deuxième variante de l'énoncé. Supposons $\|x_n^*\| \le 1$ pour tout $n \ge 0$. Si la limite existe pour tout d d'un ensemble total T, elle existe aussi, par linéarité, pour tout d de l'ensemble dense $\mathrm{D} = \mathrm{Vect}(\mathrm{T})$. Montrons que $(x_n^*(x))_{n \ge 0}$ converge pour tout $x \in \mathrm{X}$. Il suffit de montrer que cette suite de scalaires est de Cauchy. Choisissons $\varepsilon > 0$ et $d \in \mathrm{D}$ tels que $\|x - d\| < \varepsilon/3$. Pour tout entier n, on a $|x_n^*(x) - x_n^*(d)| < \varepsilon/3$, et la suite $(x_n^*(d))_{n \ge 0}$ converge vers une limite ℓ ; il en résulte que pour n assez grand, on aura $|x_n^*(x) - \ell| < \varepsilon/2$, et si n, m sont assez grands, on aura $|x_n^*(x) - x_m^*(x)| < \varepsilon$. La suite $(x_n^*(x))$ est donc de Cauchy, donc convergente. Il est alors clair que la formule

$$x^*(x) = \lim_{n \to +\infty} x_n^*(x)$$

définit une forme linéaire continue telle que $\|x^*\| \le 1$, et la suite $(x_n^*)_{n \ge 0}$ converge $*$-faiblement vers x^*.
//

Exemples 5.3.2.

1. La suite canonique $(\mathbf{e}_k)_{k \ge 0}$ est totale dans $\mathrm{X} = \ell_p$, $1 \le p < \infty$. Soit $q \in]1, +\infty]$ le conjugué de p ; pour qu'une suite $(y^{(n)})$, bornée dans $\mathrm{X}^* \simeq \ell_q$ soit $*$-faiblement convergente (vers un certain $y \in \ell_q$) il suffit que $\lim_n y_k^{(n)}$ existe pour tout k : pour une suite bornée, la convergence $*$-faible est identique à la convergence pour la topologie produit de $\mathbb{K}^{\mathbb{N}}$ (convergence simple, coordonnée par coordonnée). En particulier, la suite canonique $(\mathbf{e}_k)$ tend $*$-faiblement vers 0 dans ℓ_q, considéré comme dual de ℓ_p, pour tout $q \in]1, +\infty]$; elle tend aussi $*$-faiblement vers 0 dans ℓ_1, considéré comme dual de c_0.

2. Définissons une suite (f_n) de fonctions dans $\mathrm{L}_\infty(0,1)$ par $f_n(t) = (-1)^{[nt]}$ (partie entière). Cette suite est formée de fonctions de module un, donc de norme un dans L_∞. Elle tend vers 0 dans $\sigma(\mathrm{L}_\infty, \mathrm{L}_1)$:
prenons $\mathrm{T} = \mathrm{C}([0,1])$, qui est dense dans $\mathrm{X} = \mathrm{L}_1$. On aura pour g continue

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

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

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

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

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

Démonstration. Pour exprimer la démonstration, il est utile d'introduire une petite convention de notation. Si $M = \{n_0 < \ldots < n_j < \ldots\}$ est un sous-ensemble infini de $\mathbb{N}$, convenons de noter la sous-suite (x_{n_j}) par $(x_n)_{n \in M}$. Soit donc (y_k) une suite dense dans X, et (x_n^*) une suite bornée dans X^*, telle que par exemple $\|x_n^*\| \leq 1$ pour tout entier $n \geq 0$. La suite de scalaires $(x_n^*(y_0))$ est bornée, donc elle admet une sous-suite convergente $(x_n^*(y_0))_{n \in M_0}$. La suite $(x_n^*(y_1))_{n \in M_0}$ est encore bornée, donc on peut trouver un nouvel ensemble infini $M_1 \subset M_0$ tel que la sous-suite $(x_n^*(y_1))_{n \in M_1}$ soit convergente. En continuant ainsi, on construit une suite décroissante $M_0 \supset M_1 \supset \ldots \supset M_j \supset \ldots$ telle que $(x_n^*(y_j))_{n \in M_j}$ soit convergente pour tout $j \geq 0$.

C'est ici qu'intervient le procédé de *la suite diagonale*. Construisons un ensemble infini M formé du premier élément n_0 de M_0, puis du premier élément n_1 de M_1 qui soit $> n_0$, etc... On constate que pour tout entier $k \geq 0$, la sous-suite $(x_n^*(y_k))_{n \in M}$ est convergente : en effet, l'ensemble M est contenu dans M_k à un ensemble fini près, pour tout $k \geq 0$.

//

Exemple 5.3.3. Si (μ_n) est une suite de probabilités sur le compact $[0,1]$, il existe une sous-suite (μ_{n_j}) et une probabilité μ sur $[0,1]$ telles que $\int f\, d\mu = \lim_j \int f\, d\mu_{n_j}$ pour toute fonction continue f sur $[0,1]$; on dit que la sous-suite (μ_{n_j}) converge *vaguement* vers μ. Le résultat provient du fait que l'espace des mesures sur $[0,1]$ est le dual de l'espace *séparable* $C([0,1])$.

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

Démonstration. Soit (x_n) une suite bornée dans X ; le sous-espace fermé Y engendré par la suite (x_n) est un espace réflexif séparable, donc son dual Y^* est séparable et réflexif. On peut donc appliquer le théorème précédent à $Y^{**} \simeq Y$; en considérant (x_n) comme une suite bornée dans Y^{**}, on peut trouver une sous-suite (x_{n_j}) qui soit $*$-faiblement convergente dans Y^{**} vers un élément y^{**}, c'est à dire telle que

$$\forall y^* \in Y^*, \quad y^{**}(y^*) = \lim_j J_Y(x_{n_j})(y^*) = \lim_j y^*(x_{n_j}).$$

Puisque Y est réflexif, il existe un vecteur $y \in Y$ tel que $y^{**} = J_Y(y)$, et la relation ci-dessus nous dit que $y^*(y) = y^{**}(y^*) = \lim_j y^*(x_{n_j})$ pour tout $y^* \in Y^*$, ce qui signifie que la sous-suite (x_{n_j}) converge faiblement vers y dans Y. Si x^* est une forme linéaire continue sur X, elle n'agira sur les (x_n) et sur $y \in Y$ que par sa restriction $y^* \in Y^*$ à l'espace Y, et on aura encore $x^*(y) = y^*(y) = \lim_j y^*(x_{n_j}) = \lim_j x^*(x_{n_j})$. On a donc montré que la sous-suite (x_{n_j}) converge faiblement dans X vers le vecteur y.

//

Théorème 5.3.6. *Si* C *est un convexe fermé borné non vide d'un espace réflexif et si* f *est une fonction convexe continue sur* C, *elle atteint son minimum sur* C.

Démonstration. On peut trouver une suite $(x_n) \subset \mathrm{C}$ telle que $(f(x_n))$ converge en décroissant vers $\inf f(\mathrm{C})$ (peut-être $-\infty$) ; la suite (x_n) est bornée puisque C est borné ; quitte à passer à une sous-suite on peut supposer que (x_n) converge faiblement vers $x \in \mathrm{E}$; comme C est faiblement fermé, on sait que $x \in \mathrm{C}$. Fixons momentanément un entier $m \geq 0$ et considérons l'ensemble $\mathrm{D}_m = \{y \in \mathrm{C} : f(y) \leq f(x_m)\}$. C'est un convexe fermé, donc faiblement fermé, et la suite $(x_n)_{n \geq m}$ est contenue dans D_m, donc sa limite faible x reste dans D_m. On a donc $f(x) \leq f(x_m)$ pour tout m, ce qui montre que f atteint son minimum au point x.

//

Corollaire 5.3.7. *Si* C *est un convexe fermé borné non vide d'un espace réflexif réel* E *et si* $x^* \in \mathrm{E}^*$, *la forme linéaire* x^* *atteint son maximum sur* C.

Ce résultat donne un moyen indirect de constater que certains espaces ne sont pas réflexifs : il suffit de trouver une forme linéaire continue qui n'atteint pas son maximum sur la boule unité fermée.

Exercice.

A. Montrer que la forme linéaire continue ℓ sur $\mathrm{L}_1(0,1)$, définie par

$$\forall f \in \mathrm{L}_1(0,1), \quad \ell(f) = \int_0^1 t f(t)\, dt$$

n'atteint pas son sup sur la boule unité de $\mathrm{L}_1(0,1)$.

B. Sur $\mathrm{E} = \mathrm{C}([0,1])$, on considère la forme linéaire continue

$$\ell(f) = \int_0^{1/2} f(t)\, dt - \int_{1/2}^1 f(t)\, dt.$$

Montrer que ℓ n'atteint pas son sup sur la boule unité fermée de $\mathrm{C}([0,1])$.

Corollaire 5.3.8. *Si* f *est convexe continue sur un convexe fermé non vide d'un espace réflexif* E *et si* $f(x)$ *tend vers* $+\infty$ *lorsque* $\|x\| \to +\infty$, *la fonction* f *atteint son minimum sur* E.

Démonstration. Soit x_0 un point de C, et soit $\mathrm{C}_0 = \{x \in \mathrm{C} : f(x) \leq f(x_0)\}$; l'ensemble C_0 est convexe fermé, non vide, et il est borné parce que $f(x) \to +\infty$ lorsque $\|x\| \to +\infty$. La fonction f atteint donc son minimum sur C_0, et il est facile de voir que ce minimum est aussi le minimum sur C tout entier.

//

6. Opérateurs bornés sur les espaces de Hilbert

On a déjà revu l'essentiel des propriétés des espaces de Hilbert dans le chapitre 3, en particulier l'existence de base orthonormée pour tout espace de Hilbert séparable. Dans ce chapitre, on va généraliser la notion de base hilbertienne au cas non séparable. On introduira aussi dans ce chapitre les principales classes d'opérateurs bornés entre espaces de Hilbert.

6.1. Applications linéaires continues entre espaces de Hilbert

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

Proposition 6.1.1. *Soient* E *et* F *deux espaces de Hilbert et* $\mathrm{T} \in \mathcal{L}(\mathrm{E}, \mathrm{F})$ *; il existe un unique* $\mathrm{T}^* \in \mathcal{L}(\mathrm{F}, \mathrm{E})$ *tel que pour tout* $x \in \mathrm{E}$ *et tout* $y \in \mathrm{F}$ *on ait*

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

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

Démonstration. Pour tout $y \in \mathrm{F}$, l'application $x \to \langle \mathrm{T}(x), y \rangle$ est linéaire et continue. Il existe donc un unique élément $\mathrm{T}^*(y) \in \mathrm{E}$ tel que pour tout $x \in \mathrm{E}$ on ait $\langle \mathrm{T}(x), y \rangle = \langle x, \mathrm{T}^*(y) \rangle$. On vérifie facilement que $\mathrm{T}^*(y) + \lambda \mathrm{T}^*(z)$ vérifie la propriété caractéristique de $\mathrm{T}^*(y + \lambda z)$, pour tous $y, z \in \mathrm{F}$ et $\lambda \in \mathbb{K}$, d'où l'on déduit que T^* est linéaire. On a, par définition de $\|\mathrm{T}\|$ et par la proposition 2.2.1

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

//

Définition 6.1.1. Soient E et F deux espaces de Hilbert et $\mathrm{T} \in \mathcal{L}(\mathrm{E}, \mathrm{F})$; l'unique application linéaire $\mathrm{T}^* \in \mathcal{L}(\mathrm{F}, \mathrm{E})$ tel que pour tout $x \in \mathrm{E}$ et tout $y \in \mathrm{F}$ on ait $\langle \mathrm{T}(x), y \rangle = \langle x, \mathrm{T}^*(y) \rangle$ est appelée *adjointe* de T.

Regroupons dans la proposition suivante quelques propriétés des adjoints.

Proposition 6.1.2. *Soient* E *et* F *deux espaces de Hilbert ; l'application* $\mathrm{T} \to \mathrm{T}^*$ *est antilinéaire et isométrique de* $\mathcal{L}(\mathrm{E}, \mathrm{F})$ *sur* $\mathcal{L}(\mathrm{F}, \mathrm{E})$ *; pour tout* $\mathrm{T} \in \mathcal{L}(\mathrm{E}, \mathrm{F})$ *on a* $(\mathrm{T}^*)^* = \mathrm{T}$ *et* $\|\mathrm{T}^* \circ \mathrm{T}\| = \|\mathrm{T}\|^2$*. Pour tout espace de Hilbert* H*, tout* $\mathrm{S} \in \mathcal{L}(\mathrm{E}, \mathrm{F})$ *et tout* $\mathrm{T} \in \mathcal{L}(\mathrm{F}, \mathrm{H})$ *on a* $(\mathrm{T} \circ \mathrm{S})^* = \mathrm{S}^* \circ \mathrm{T}^*$.

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

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

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

//

Exemples 6.1.2.

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

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

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

Si f est une fonction complexe, mesurable bornée sur (Ω, μ), on définit l'opérateur de multiplication M_f par $\mathrm{M}_f(g) = fg$ pour toute $g \in \mathrm{L}_2(\Omega, \mu)$. On vérifie que M_f est borné sur $\mathrm{L}_2(\Omega, \mu)$, et $\mathrm{M}_f^* = \mathrm{M}_{\overline{f}}$.

2. Shift S sur $\mathrm{H} = \ell_2(\mathbb{Z})$ ou bien $\mathrm{H} = \ell_2(\mathbb{N})$. Pour tout vecteur $x \in \mathrm{H}$ on définit un nouveau vecteur $\mathrm{S}x$ obtenu par *décalage à droite*, défini par $(\mathrm{S}x)_n = x_{n-1}$ pour tout n dans le cas $\mathbb{Z}$ et dans le cas $\mathbb{N}$, on pose $(\mathrm{S}x)_n = x_{n-1}$ pour tout $n \geq 1$, et $(\mathrm{S}x)_0 = 0$; dans le cas de $\ell_2(\mathbb{Z})$, on trouve $(\mathrm{S}^*y)_n = y_{n+1}$ pour tout $n \in \mathbb{Z}$, on vérifie que $\mathrm{S}^*\mathrm{S} = \mathrm{S}\mathrm{S}^* = \mathrm{Id}$. Dans le cas $\ell_2(\mathbb{N})$, on a aussi $(\mathrm{S}^*y)_n = y_{n+1}$, mais pour $n \geq 0$; dans ce cas S^* n'est plus l'inverse de S, mais on a encore $\mathrm{S}^*\mathrm{S} = \mathrm{Id}$.

Proposition 6.1.3. *Soient* E *et* F *deux espaces de Hilbert et soit* $\mathrm{T} \in \mathcal{L}(\mathrm{E}, \mathrm{F})$ *; alors* $\ker \mathrm{T}^* = (\mathrm{T}(\mathrm{E}))^\perp$ *et l'adhérence de* $\mathrm{T}^*(\mathrm{F})$ *est* $(\ker \mathrm{T})^\perp$.

Démonstration. Si $y \in \mathrm{F}$, on voit que $y \in \ker \mathrm{T}^*$ si et seulement si pour tout $x \in \mathrm{E}$, on a $0 = \langle \mathrm{T}^*(y), x \rangle = \langle y, \mathrm{T}(x) \rangle$; clairement, ceci équivaut à dire que $y \in (\mathrm{T}(\mathrm{E}))^\perp$, d'où la première assertion. Il en résulte (par la proposition 2.3.4) que $\overline{\mathrm{T}(\mathrm{E})} = (\ker \mathrm{T}^*)^\perp$, d'où la deuxième assertion en remplaçant T par son adjoint.

//

Définition 6.1.3. Soient E et F deux espaces de Hilbert ; un élément $\mathrm{U} \in \mathcal{L}(\mathrm{E}, \mathrm{F})$ est appelé *unitaire* si $\mathrm{U}^* \circ \mathrm{U} = \mathrm{Id}_\mathrm{E}$ et $\mathrm{U} \circ \mathrm{U}^* = \mathrm{Id}_\mathrm{F}$. Un élément $\mathrm{T} \in \mathcal{L}(\mathrm{E})$ est appelé *normal* si $\mathrm{T}^* \circ \mathrm{T} = \mathrm{T} \circ \mathrm{T}^*$, *hermitien* ou *autoadjoint* si $\mathrm{T} = \mathrm{T}^*$ et *positif* s'il est hermitien et si $\langle \mathrm{T}(x), x \rangle$ est réel ≥ 0 pour tout $x \in \mathrm{E}$.

Exemples 6.1.4.

1. Soient H un espace de Hilbert, $\mathrm{P} \in \mathcal{L}(\mathrm{H})$ un projecteur orthogonal ; notons F son image. Pour $x, x' \in \mathrm{F}$ et $y, y' \in \mathrm{F}^\perp$ on a $\langle \mathrm{P}(x+y), x'+y' \rangle = \langle x, x' \rangle = \langle x+y, \mathrm{P}(x'+y') \rangle$, donc $\mathrm{P} = \mathrm{P}^*$. De plus, $\langle \mathrm{P}(x+y), x+y \rangle = \langle x, x \rangle$ est réel ≥ 0, donc P est positif.

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

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

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

3. On dit que $\mathrm{U} \in \mathcal{L}(\mathrm{E}, \mathrm{F})$ est *isométrique* si $\|\mathrm{U}(x)\| = \|x\|$ pour tout $x \in \mathrm{E}$. On vérifie facilement que U est isométrique si et seulement si $\mathrm{U}^*\,\mathrm{U} = \mathrm{Id}_\mathrm{E}$. En effet, si $\mathrm{U}^* \circ \mathrm{U} = \mathrm{Id}_\mathrm{E}$, alors pour tout $x \in \mathrm{E}$ on a $\|\mathrm{U}(x)\|^2 = \langle \mathrm{U}(x), \mathrm{U}(x) \rangle = \langle x, \mathrm{U}^*(\mathrm{U}(x)) \rangle = \|x\|^2$. La réciproque utilise la formule de polarisation (voir la démonstration ci-dessous). Le shift de l'exemple 2 est isométrique ; dans le cas de $\ell_2(\mathbb{Z})$ il est aussi unitaire.

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

Proposition 6.1.4. *Soient* E *et* F *deux espaces de Hilbert et* $T \in \mathcal{L}(E, F)$ *; les conditions suivantes sont équivalentes :*

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

Démonstration. Si T est unitaire, comme $T \circ T^* = \mathrm{Id}_F$, l'opérateur T est surjectif, donc $(i) \Rightarrow (ii)$. Si $T^* \circ T = \mathrm{Id}_E$, alors pour tout $x \in E$ on a $\|T(x)\|^2 = \|x\|^2$, donc $(ii) \Rightarrow (iii)$. Enfin, supposons que T soit une isométrie de E sur F, c'est à dire que pour tout $x \in E$ on ait $\langle x, x\rangle = \langle T(x), T(x)\rangle$; comme

$$(x, y) \to \langle x, T^*(T(y))\rangle = \langle Tx, Ty\rangle$$

est un produit scalaire sur E, il résulte de la proposition 2.1.1 que, pour tous $x, y \in E$, on a $\langle x, T^*(T(y))\rangle = \langle x, y\rangle$, ce qui implique $T^*(T(y)) = y$, c'est à dire que $T^* \circ T = \mathrm{Id}_E$. Comme par l'hypothèse (iii) l'application T est bijective, $T^* = T^{-1}$, d'où (i).
//

Lemme 6.1.5. *Si* $T \in \mathcal{L}(H)$ *est normal, on a* $\ker T = \ker T^*$.

Démonstration. En fait on a pour tout $x \in H$

$$\|Tx\|^2 = \langle Tx, Tx\rangle = \langle T^*Tx, x\rangle = \langle TT^*x, x\rangle = \|T^*(x)\|^2.$$

//

Proposition 6.1.6. *Si* $T \in \mathcal{L}(H)$ *est normal, on a* $H = \ker(T) \oplus \overline{\mathrm{im}(T)}$*, et la somme est une somme orthogonale.*

Démonstration. On sait que $\ker(T^*)^\perp = \overline{\mathrm{im}(T^{**})} = \overline{\mathrm{im}(T)}$, donc

$$H = \ker(T^*) \oplus \ker(T^*)^\perp = \ker(T) \oplus \overline{\mathrm{im}(T)}.$$

//

6.2. Familles sommables dans un espace de Banach

Définition 6.2.1. Soient X un espace normé, I un ensemble et $(x_i)_{i\in I}$ une famille d'éléments de X ; on dit que la famille $(x_i)_{i\in I}$ est *sommable* de somme $S \in X$ et on écrit $S = \sum_{i\in I} x_i$ si, pour tout $\varepsilon > 0$, il existe une partie finie J de I telle que, pour toute partie finie K de I contenant J on ait $\left\| S - \sum_{i\in K} x_i \right\| < \varepsilon$. La somme S est unique. Il est facile de vérifier que si $(x_i)_{i\in I}$ et $(y_i)_{i\in I}$ sont deux familles sommables, la famille $(x_i + y_i)_{i\in I}$ est elle aussi sommable, avec une somme égale à la somme des deux sommes. Considérons pour commencer le cas des familles sommables de nombres réels, et d'abord de réels ≥ 0.

Proposition 6.2.1. *Une famille de nombres réels à termes positifs est sommable si et seulement si les sommes finies sont majorées ; sa somme est alors la borne supérieure de l'ensemble des sommes finies. Une famille $(x_i)_{i\in\mathrm{I}}$ à termes réels est sommable si et seulement si elle est absolument sommable, c'est à dire si la famille $(|x_i|)_{i\in\mathrm{I}}$ est sommable.*

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

Proposition 6.2.3.

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

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

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

Lemme 6.2.5. *Soit $(x_i)_{i\in\mathrm{I}}$ un système orthogonal dans un espace de Hilbert ; la famille (x_i) est sommable si et seulement si la famille $(\|x_i\|^2)$ est sommable ; dans ce cas, on a*

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

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

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

//

6.3. Bases hilbertiennes

Disons quelques mots sur les espaces de Hilbert non séparables, qui demandent une généralisation de la notion de *suite* orthonormée. Soit I un ensemble d'indices non dénombrable, et soit H l'espace des familles $x = (x_i)_{i\in\mathrm{I}}$ de scalaires telles que l'ensemble des $i \in \mathrm{I}$ tels que $x_i \neq 0$ soit un ensemble dénombrable $\mathrm{J}(x)$, et telles que $\sum_{i\in\mathrm{J}(x)} |x_i|^2 < +\infty$. Si x et y sont de telles familles, le produit x_iy_i est nul sauf pour au plus un ensemble dénombrable d'indices J, et $|x_iy_i| \le \frac{1}{2}(x_i^2 + y_i^2)$, ce qui permet de poser

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

le résultat ne dépendant pas de l'ensemble dénombrable J qui contient tous les indices i tels que $x_iy_i \neq 0$. On obtient ainsi un exemple d'espace de Hilbert non séparable, qui sera traité plus en détail à la section 6.4.

On a vu au chapitre 3 que l'espace $L_2(\Omega, \mu)$ est un espace de Hilbert pour le produit scalaire $\langle f, g\rangle = \int_\Omega f(s)\overline{g(s)}\, d\mu(s)$. Il est possible que cet espace soit non séparable, même si la mesure μ est une probabilité.

Définition 6.3.1. Soient E un espace de Hilbert et $(x_i)_{i\in \mathrm{I}}$ un système de vecteurs de E ; on dit que le système $(x_i)_{i\in \mathrm{I}}$ est *orthogonal* si les x_i sont deux à deux orthogonaux ; on dit que c'est un *système orthonormal* si de plus, pour tout $i \in \mathrm{I}$, on a $\|x_i\| = 1$; on appelle *base hilbertienne* de E un système orthonormal total dans E.

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

Théorème 6.3.1. *Tout espace de Hilbert admet une base hilbertienne.*

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

Soit B un élément maximal de U ; on veut montrer que B est total, et pour cela, on montre que $\mathrm{B}^\perp = \{0\}$; sinon, il existerait un vecteur x non nul et orthogonal à B (en particulier $x \notin \mathrm{B}$), et quitte à multiplier x par un scalaire convenable on peut supposer $\|x\| = 1$; alors $\mathrm{B} \cup \{x\} \in \mathrm{U}$, ce qui contredirait la maximalité de B. Donc $\mathrm{B}^\perp = \{0\}$, ce qui entraîne que $(\mathrm{B}^\perp)^\perp = \mathrm{H}$. Par la proposition 2.3.4, B est total. C'est donc une base hilbertienne.

//

Théorème 6.3.2 : inégalité de Bessel. *Soient* E *un espace de Hilbert et* $(e_i)_{i\in \mathrm{I}}$ *un système orthonormal dans* E *; pour tout* $x \in \mathrm{E}$ *la famille* $(|\langle x, e_i\rangle|^2)_{i\in \mathrm{I}}$ *est sommable et*

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

Démonstration. Par la proposition 2.1, il suffit de montrer que, pour toute partie finie J de I, on a $\sum_{i\in \mathrm{J}} |\langle x, e_i\rangle|^2 \le \langle x, x\rangle$. Ce résultat a été vu au lemme 2.2.5.

//

Théorème 6.3.5 : identité de Parseval. *Soient* E *un espace de Hilbert,* $(e_i)_{i\in \mathrm{I}}$ *une base hilbertienne de* E *et* $x \in \mathrm{E}$ *; la famille de nombres réels* $(|\langle x, e_i\rangle|^2)_{i\in \mathrm{I}}$ *est sommable, la famille de vecteurs* $(\langle x, e_i\rangle\, e_i)_{i\in \mathrm{I}}$ *est sommable dans* E *et*

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

Démonstration. Comme $(e_i)_{i\in \mathrm{I}}$ est un système orthonormal, il résulte du théorème 2 que la famille de réels $(|\langle x, e_i\rangle|^2)_{i\in \mathrm{I}}$ est sommable. Par le lemme 2.5, la famille

$(\langle x, e_i \rangle\, e_i)_{i \in \mathrm{I}}$ est sommable dans E et, si on note y sa somme, on a $\sum_{i \in \mathrm{I}} |\langle x, e_i \rangle|^2 = \langle y, y \rangle$. Pour tout $j \in \mathrm{I}$, appliquant la forme linéaire continue $z \to \langle z, e_j \rangle$, on trouve $\langle y, e_j \rangle = \sum_{i \in \mathrm{I}} \langle x, e_i \rangle \langle e_i, e_j \rangle = \langle x, e_j \rangle$. Donc $x - y$ est orthogonal aux e_i, donc à l'espace vectoriel engendré par les (e_i) ; comme le système (e_i) est total, $x = y$.

//

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

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

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

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

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

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

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

Proposition 6.4.1. *Muni du produit scalaire précédent, l'espace vectoriel* $\ell^2(\mathrm{I})$ *est un espace de Hilbert. La famille* $(\epsilon_i)_{i \in \mathrm{I}}$ *est une base hilbertienne de* $\ell^2(\mathrm{I})$.

Théorème 6.4.2. *Soient* H *un espace de Hilbert et* $\mathrm{B} = (e_i)_{i \in \mathrm{I}}$ *une base hilbertienne de* H *; l'application* $\mathrm{U} : x \to (\langle x, e_i \rangle)$ *est une bijection linéaire isométrique de* H *sur* $\ell^2(\mathrm{I})$.

7. Algèbres de Banach et théorie spectrale

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

7.1. Algèbres de Banach, spectre et résolvante

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

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

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

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

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

Exemples 7.1.1.

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

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

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

Définition 7.1.2. Soient A une algèbre de Banach unitaire, et $a \in \mathrm{A}$; on dit que a est *inversible dans* A s'il existe $b \in \mathrm{A}$ tel que $ab = ba = 1_\mathrm{A}$.

Exemples 7.1.3.

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

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

Lemme 7.1.1. *Soient* A *une algèbre de Banach unitaire et* $a \in \mathrm{A}$ *tel que* $\|a\| < 1$ *; alors, la série* $\sum_k a^k$ *est convergente dans* A *et sa somme est l'inverse de* $1_\mathrm{A} - a$,

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

On a de plus l'estimation

$$\bigl\|(1_\mathrm{A} - a)^{-1}\bigr\| \le \frac{1}{1 - \|a\|}.$$

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

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

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

//

Remarque 7.1.4. Le raisonnement précédent prouve ceci : si on sait simplement que la série $\sum a^k$ converge dans A, sa somme $\sum_{k=0}^{+\infty} a^k$ sera l'inverse de $1_\mathrm{A} - a$.

Proposition 7.1.2. *Soit* A *une algèbre de Banach ; l'ensemble des éléments inversibles dans* A *est un ouvert non vide* U *de* A*. L'application* $\varphi : u \to u^{-1}$ *est continue et différentiable de* U *dans* A.

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

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

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

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

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

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

//

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

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

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

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

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

Exemples 7.1.8.

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

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

Théorème 7.1.4.

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

Démonstration. Si λ n'est pas dans $\mathrm{Sp}(a)$, l'élément $a - \lambda 1_{\mathrm{A}}$ est inversible ; d'après la proposition 2, $a - \lambda' 1_{\mathrm{A}}$ sera encore inversible pour tout λ' dans un voisinage de λ, ce qui montre que le complémentaire du spectre est ouvert dans $\mathbb{C}$, donc $\mathrm{Sp}(a)$ est fermé dans $\mathbb{C}$; on a vu ci-dessus que $\mathrm{Sp}(a)$ est contenu dans le disque de rayon $\|a\|$, donc le spectre est borné.

Soit $\lambda \in \mathbb{C} \setminus \mathrm{Sp}(a)$; posons $u = a - \lambda 1_{\mathrm{A}}$; alors u est inversible, $u^{-1} = \mathrm{R}_\lambda(a)$ et on sait que pour z assez petit, $a - (\lambda + z)1_{\mathrm{A}} = u - z1_{\mathrm{A}}$ est inversible et

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

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

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

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

Il reste à montrer que $\mathrm{Sp}(a) \neq \emptyset$. Choisissons λ_0 hors du spectre ; alors $\mathrm{R}_{\lambda_0}(a)$ est non nul puisqu'inversible et on peut trouver une forme linéaire x^* continue sur

A telle que $x^*(\mathrm{R}_{\lambda_0}(a)) \neq 0$ (corollaire 4.2.7) ; l'application $g : \lambda \to x^*(\mathrm{R}_\lambda(a))$ est une fonction holomorphe scalaire définie sur $\mathbb{C} \setminus \mathrm{Sp}(a)$, telle que $g(\lambda_0) \neq 0$. Si $\mathrm{Sp}(a)$ était vide, cette fonction serait entière (holomorphe sur $\mathbb{C}$ tout entier) ; or d'après la relation (R) on voit que $g(\lambda) = x^*(\mathrm{R}_\lambda(a))$ tend vers 0 quand $\lambda \to \infty$. Par le théorème de Liouville on aurait $g(\lambda) = 0$ pour tout $\lambda \in \mathbb{C}$, ce qui n'est pas vrai, donc $\mathrm{Sp}(a) \neq \emptyset$.

//

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

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

On dit qu'une application f d'un ouvert U de $\mathbb{C}$ dans un espace de Banach complexe F, dérivable au sens complexe en tout point de l'ouvert U, est une *fonction holomorphe* de U dans F. La démonstration du théorème précédent indique que l'application $\mathrm{R}(a) : \lambda \to (a - \lambda 1_{\mathrm{A}})^{-1} = \mathrm{R}_\lambda(a)$ est holomorphe sur $\mathbb{C} \setminus \mathrm{Sp}(a)$; on peut voir que

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

(et la dérivée est donc continue).

Exemples 7.1.9.

1. Soit K un espace compact ; considérons l'espace de Banach $\mathrm{E} = C(\mathrm{K})$ des fonctions continues sur K à valeurs complexes, muni de la norme de convergence uniforme (exemples 1.1.6). Soit $f \in \mathrm{E}$; l'application $\mathrm{M}_f : g \to fg$ est linéaire de E dans E et continue puisque pour tout $g \in \mathrm{E}$, on a $\|fg\|_\infty \leq \|f\|_\infty \|g\|_\infty$. De plus la relation $\|\mathrm{M}_f(f)\|_\infty = \|f\|_\infty^2$ implique $\|\mathrm{M}_f\| \geq \|f\|_\infty$, donc $\|\mathrm{M}_f\| = \|f\|_\infty$. Soit $\lambda \in \mathbb{C}$;
– si pour tout $s \in \mathrm{K}$, on a $f(s) \neq \lambda$, alors la fonction $h : s \to (f(s) - \lambda)^{-1}$ est continue de K dans $\mathbb{C}$. On voit alors que $\mathrm{M}_f - \lambda\,\mathrm{Id}_\mathrm{E}$ est inversible et que son inverse $\mathrm{R}_\lambda(\mathrm{M}_f)$ est l'application $\mathrm{M}_h : g \to hg$;
– s'il existe $s \in \mathrm{K}$ tel que $f(s) = \lambda$, alors pour tout $g \in \mathrm{E}$, la fonction $fg - \lambda g$ s'annule au point s, donc $\mathrm{im}(\mathrm{M}_f - \lambda\,\mathrm{Id}_\mathrm{E}) \subset \{g \in \mathrm{E} : g(s) = 0\}$, qui est un sous-espace fermé de E, distinct de E. On en déduit que l'image de $\mathrm{M}_f - \lambda\,\mathrm{Id}_\mathrm{E}$ n'est pas dense, donc $\mathrm{M}_f - \lambda\,\mathrm{Id}_\mathrm{E}$ n'est pas inversible puisqu'il n'est pas surjectif. En résumé, le spectre de M_f est l'ensemble $\mathrm{Sp}(\mathrm{M}_f) = f(\mathrm{K}) = \{f(s) : s \in \mathrm{K}\}$ des valeurs de f. C'est aussi le spectre de f dans l'algèbre $C(\mathrm{K})$. On verra plus loin que ça n'est pas un hasard !

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

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

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

Avec cette notation on va chercher à exprimer la transposée de S, considérée comme endomorphisme de ℓ_q. Soit T l'opération de décalage à gauche, définie par $\mathrm{T}((y_n)_{n \geq 0}) =$

$(y_{n+1})_{n\ge 0}$. On constate sans peine que $(\mathrm{T}(y), x) = (y, \mathrm{S}(x))$ pour tous $x \in \ell_p$, $y \in \ell_q$. L'application T "est" donc la transposée de S, et de plus $\mathrm{T} - \lambda\,\mathrm{Id}_{\ell_q}$ est pour tout λ la transposée de $\mathrm{S} - \lambda\,\mathrm{Id}_{\ell_p}$. Quand un opérateur V sur ℓ_p est inversible, il est clair que sa transposée est inversible dans $\mathcal{L}(\ell_q)$, ce qui entraîne que $\mathrm{Sp}({}^t\mathrm{S}) \subset \mathrm{Sp}(\mathrm{S})$. Soit $\lambda \in \mathbb{C}$; si $|\lambda| < 1$, posons $y = (\lambda^n)_{n\ge 0}$; c'est un élément non nul de ℓ_q et ${}^t\mathrm{S}(y) = \lambda y$. Il en résulte que ${}^t\mathrm{S} - \lambda\,\mathrm{Id}_{\ell_q}$ n'est pas inversible, donc le spectre de ${}^t\mathrm{S}$ contient le disque unité ouvert, et il est contenu dans Sp(S) qui est contenu dans le disque unité fermé ; puisque le spectre est fermé,

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

7.2. Rayon spectral

Soit A une algèbre de Banach unitaire complexe ; la quantité

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

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

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

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

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

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

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

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

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

ce qui fournit les inégalités de Cauchy

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

Considérons la fonction vectorielle $f(z) = (1_A - za)^{-1}$; elle est définie pour tout complexe z tel que $1/z$ ne soit pas dans le spectre de a, ce qui est le cas lorsque $|z| < R = \rho(a)^{-1}$; la continuité de l'application $u \to u^{-1}$ sur l'ouvert des éléments inversibles (proposition 1.2) montre que $z \to f(z)$ est continue pour $|z| < R$; pour tout r tel que $0 < r < R$, la fonction $z \to \|f(z)\|$ est donc bornée par un certain $M_0(r)$ sur le cercle de rayon r (qui est compact).

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

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

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

Puisque $\|x^*\| \leq 1$, on a $|g(z)| \leq \|f(z)\|$, ce qui entraîne que $M(r, g) \leq M_0(r)$; les inégalités de Cauchy, appliquées à g, donnent $|x^*(a^n)| \leq M_0(r)/r^n$ pour tout n et toute $x^* \in A^*$ telle que $\|x^*\| = 1$; pour chaque $n \geq 1$ donné on peut choisir par Hahn-Banach (corollaire 4.2.7) une forme linéaire x^* telle que $\|x^*\| = 1$ et $x^*(a^n) = \|a^n\|$; on obtient ainsi $\|a^n\| \leq M_0(r)/r^n$ pour tout $n \geq 1$, ce qui implique $\limsup_n \|a^n\|^{1/n} \leq 1/r$, d'où $\limsup_n \|a^n\|^{1/n} \leq \rho(a)$ en faisant tendre r vers $R = 1/\rho(a)$.

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

//

Lemme 7.2.2. *Soit (u_n) une suite de nombres réels positifs ou nuls telle que, pour tous entiers $p, q \geq 1$ on ait $(u_{p+q})^{p+q} \leq u_p^p \, u_q^q$; alors la suite (u_n) converge vers $\inf_{n \geq 1} u_n$.*

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

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

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

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

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

//

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

Démonstration. Soit d'abord A un élément hermitien ; on a $\|\mathrm{A}^2\| = \|\mathrm{A}^*\mathrm{A}\| = \|\mathrm{A}\|^2$ (proposition 6.1.2) ; on en déduit par récurrence que $\|\mathrm{A}^{2^n}\| = \|\mathrm{A}\|^{2^n}$ pour tout $n \geq 0$, donc $\rho(\mathrm{A}) = \|\mathrm{A}\|$. Soit maintenant T un élément normal de $\mathcal{L}(\mathrm{H})$; par récurrence sur n, on a $(\mathrm{T}^*\mathrm{T})^n = (\mathrm{T}^*)^n\mathrm{T}^n$ donc $\|(\mathrm{T}^*\mathrm{T})^n\| = \|\mathrm{T}^n\|^2$ et $\rho(\mathrm{T}^*\mathrm{T}) = \rho(\mathrm{T})^2$. Or $\mathrm{A} = \mathrm{T}^*\mathrm{T}$ est hermitien, donc $\rho(\mathrm{T})^2 = \rho(\mathrm{T}^*\mathrm{T}) = \|\mathrm{T}^*\mathrm{T}\| = \|\mathrm{T}\|^2$.

//

Exemple 7.2.2. Posons $\mathrm{H} = \mathrm{L}_2([0,1])$; pour toute fonction $f \in \mathrm{H}$ et $s \in [0,1]$, on pose $\mathrm{V}(f)(s) = \int_0^s f(t)\,dt$. En appliquant Cauchy-Schwarz au produit $\mathbf{1}_{[0,s]}\,f$ on voit que $|\mathrm{V}(f)(s)| \leq \sqrt{s}\,\|f\|_2$, ce qui implique que

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

donc V définit une application linéaire continue notée V_2 de $\mathrm{L}_2([0,1])$ dans lui-même.

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

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

ce qui donne $\|\mathrm{V}_2^{n+1}\| \leq (n!)^{-1}$. Comme $\lim_n (n!)^{-1/n} = 0$, il s'ensuit que le rayon spectral de V_2 est nul, donc $\mathrm{Sp}(\mathrm{V}_2) = \{0\}$.

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

Définition 7.2.3. Un *homomorphisme d'algèbres de Banach unitaires* est une application linéaire continue $\varphi : \mathrm{A} \to \mathrm{B}$ entre deux algèbres de Banach unitaires A et B, telle que $\varphi(ab) = \varphi(a)\varphi(b)$ pour tous $a, b \in \mathrm{A}$ et que $\varphi(1_\mathrm{A}) = 1_\mathrm{B}$.

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

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

7.3. Décomposition du spectre d'un opérateur borné

Proposition 7.3.1. *Soient* E *et* F *deux espaces de Banach et soit* $\mathrm{T} \in \mathcal{L}(\mathrm{E},\mathrm{F})$ *; les conditions suivantes sont équivalentes :*

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

Démonstration. Si (*i*) est satisfaite, T détermine une application continue bijective T_1 de E sur l'espace de Banach im(T). Par le théorème des isomorphismes (théorème 4.1.3), T_1 est un isomorphisme : on obtient (ii) avec $c = \|\mathrm{T}_1^{-1}\|^{-1}$. Il est évident que (ii) implique (iii) ; montrons que $(iii) \Rightarrow (ii)$: si (ii) n'est pas satisfaite, il existe pour tout entier $n \ge 1$ un vecteur $y_n \in \mathrm{E}$ tel que $n^{-1}\,\|y_n\| > \|\mathrm{T}(y_n)\|$; si on pose $x_n = \|y_n\|^{-1} y_n$, on a $\|x_n\| = 1$ et $\|\mathrm{T}(x_n)\| < 1/n$, donc (iii) n'est pas satisfaite.

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

//

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

A. Il existe une constante $c > 0$ telle que $\|\mathrm{T}x\| \ge c\,\|x\|$ pour tout $x \in \mathrm{E}$.

B. On a $\overline{\mathrm{T(E)}} = \mathrm{F}$.

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

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

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

Démonstration. On a déjà vu une des directions : si T est inversible, il vérifie les deux conditions. Inversement, supposons que **A** et **B** soient vraies ; on sait alors que T(E) est fermé par la proposition 1, et dense d'après **B**, donc $\mathrm{T(E)} = \mathrm{F}$. Si $\mathrm{T}(x) = \mathrm{T}(x')$ on aura $x = x'$ puisque $0 = \|\mathrm{T}(x - x')\| \ge c\,\|x - x'\|$ d'après **A**. Cela permet de définir une application (linéaire) S de $\mathrm{F} = \mathrm{T(E)}$ sur E en posant $\mathrm{S}(y) = x \in \mathrm{E}$ si et seulement si $y \in \mathrm{F}$ et $\mathrm{T}(x) = y$. En traduisant **A**, on obtient $\|\mathrm{S}(y)\| \le c^{-1}\,\|y\|$ pour tout $y \in \mathrm{F}$, ce qui montre que S est continue. Pour finir il est clair que S est l'inverse de T.

//

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

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

Proposition 7.3.3. *Soient* E, F *deux espaces de Banach et soit* $\mathrm{T} \in \mathcal{L}(\mathrm{E}, \mathrm{F})$ *; l'opérateur transposé* ${}^t\mathrm{T} \in \mathcal{L}(\mathrm{F}^*, \mathrm{E}^*)$ *est inversible si et seulement* T *est inversible.*

Démonstration. Si T est inversible, comme $\mathrm{T}^{-1}\mathrm{T} = \mathrm{Id_E}$ et $\mathrm{T}\mathrm{T}^{-1} = \mathrm{Id_F}$, on trouve ${}^t\mathrm{T}\,{}^t(\mathrm{T}^{-1}) = \mathrm{Id}_{\mathrm{E}^*}$ et ${}^t(\mathrm{T}^{-1})\,{}^t\mathrm{T} = \mathrm{Id}_{\mathrm{F}^*}$, donc ${}^t\mathrm{T}$ est inversible et $({}^t\mathrm{T})^{-1} = {}^t(\mathrm{T}^{-1})$. Supposons inversement que T ne soit pas inversible. On sait que, ou bien T ne vérifie pas la condition **B**, ou bien il ne vérifie pas **A**.

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

//

On en déduit immédiatement :

Corollaire 7.3.4. *Soient* E *un espace de Banach complexe et* $\mathrm{T} \in \mathcal{L}(\mathrm{E})$ *; on a*

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

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

//

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

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

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

On va maintenant distinguer plusieurs sous-ensembles intéressants du spectre d'un opérateur borné, correspondant à plusieurs façons pour $\mathrm{T} - \lambda\,\mathrm{Id}_\mathrm{E}$ de ne pas être inversible. Soient E un espace de Banach complexe, $\mathrm{T} \in \mathcal{L}(\mathrm{E})$ et $\lambda \in \mathrm{Sp}(\mathrm{T})$; nous distinguerons plusieurs cas pour l'opérateur $\mathrm{T}_\lambda = \mathrm{T} - \lambda\,\mathrm{Id}_\mathrm{E}$, correspondant au valeurs vrai-faux des trois critères suivants : propriété **A**, propriété d'injectivité et propriété **B**. On remarque que **A** implique injectif, et que si λ est dans le spectre on ne peut pas avoir à la fois **A** et **B** pour T_λ. Quand $\lambda \in \mathrm{Sp}(\mathrm{T})$, il reste donc les cas suivants :

- T_λ n'est pas injectif ;
- T_λ est injectif mais **B** n'est pas vraie
- T_λ est injectif et **B** est vraie (donc **A** est fausse).

Ces trois cas correspondent aux cas suivants.

1. Le scalaire λ est une valeur propre de T ; ceci équivaut à dire que $\mathrm{T} - \lambda\,\mathrm{Id}_\mathrm{E}$ n'est pas injectif.

2. Le scalaire λ est une valeur propre du transposé ${}^t\mathrm{T}$, mais n'est pas une valeur propre de T ; autrement dit $\mathrm{T} - \lambda\,\mathrm{Id}_\mathrm{E}$ est injectif et ${}^t(\mathrm{T} - \lambda\,\mathrm{Id}_\mathrm{E})$ n'est pas injectif ; d'après le lemme 4.2.11, cela se produit si et seulement si $\mathrm{T} - \lambda\,\mathrm{Id}_\mathrm{E}$ est injectif mais n'a pas une image dense dans E.

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

Définition 7.3.1. Soient E un espace de Banach complexe et $T \in \mathcal{L}(E)$; on appelle *spectre ponctuel* de T l'ensemble $\mathrm{Sp}_p(T)$ des $\lambda \in \mathbb{C}$ tels que $T - \lambda\,\mathrm{Id}_E$ ne soit pas injectif (c'est l'ensemble des valeurs propres de T). On appelle *spectre résiduel* de T l'ensemble $\mathrm{Sp}_r(T)$ des $\lambda \in \mathbb{C}$ tels que $T - \lambda\,\mathrm{Id}_E$ soit injectif, mais son image ne soit pas dense. On appelle *spectre continu* de T l'ensemble $\mathrm{Sp}_c(T)$ des $\lambda \in \mathbb{C}$ tels que $T - \lambda\,\mathrm{Id}_E$ soit injectif, à image dense mais pas fermée.

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

Proposition 7.3.6. *Soient* E *un espace de Banach complexe et* $T \in \mathcal{L}(E)$ *; on a*

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

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

Démonstration. On a vu que λ est dans le spectre résiduel de T si et seulement si λ est une valeur propre de ${}^tT$, mais n'est pas une valeur propre de T, d'où la première assertion. Si $\lambda \in \mathrm{Sp}_c({}^tT)$, on sait que ${}^tT - \lambda\,\mathrm{Id}_{E^*}$ est injectif à image dense, donc $T - \lambda\,\mathrm{Id}_E$ est injectif à image dense par le lemme 4.2.11, et puisque $\mathrm{Sp}_c({}^tT) \subset \mathrm{Sp}({}^tT) = \mathrm{Sp}(T)$, on a $\lambda \in \mathrm{Sp}(T)$, par conséquent $\lambda \in \mathrm{Sp}_c(T)$.

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

//

Lemme 7.3.8. *Si un opérateur normal* $T \in \mathcal{L}(H)$ *est injectif, il est à image dense. Si l'opérateur normal* T *vérifie* **A**, *il est inversible.*

Démonstration. Si T est normal et injectif, on a $\overline{T(H)} = H$ d'après la proposition 6.1.6, c'est à dire que l'image est dense. Si T vérifie la propriété **A**, il est injectif, donc on a **A** et **B**, par conséquent T est inversible.

//

Proposition 7.3.9. *Le spectre résiduel d'un opérateur normal est vide.*

Démonstration. Soit $T \in \mathcal{L}(H)$ un opérateur normal ; pour tout scalaire $\lambda \in \mathbb{C}$, $T_\lambda = T - \lambda\,\mathrm{Id}_H$ est normal ; si λ est dans le spectre, ou bien T_λ n'est pas injectif et $\lambda \in \mathrm{Sp}_p(T)$, ou bien T_λ est injectif, donc à image dense et $\lambda \in \mathrm{Sp}_c(T)$.

//

Exemples 7.3.2.

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

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

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

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

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

Valeurs propres approchées

Lemme 7.3.10. *Soient* E *un espace de Banach complexe,* $\mathrm{T} \in \mathcal{L}(\mathrm{E})$ *et soit* $\lambda \in \partial\,\mathrm{Sp}(\mathrm{T})$ *(la frontière du spectre de* T*). Il existe une suite* $(x_n) \subset \mathrm{E}$ *de vecteurs de norme* 1 *telle que* $(\mathrm{T} - \lambda\,\mathrm{Id}_\mathrm{E})(x_n)$ *tende vers* 0*.*

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

0 ; alors $S - \mu_n \,\mathrm{Id_E}$ est inversible pour tout n. Il existe donc un vecteur $z_n \in E$ tel que $(S - \mu_n \,\mathrm{Id_E})(z_n) = y$. Si (z_n) était bornée, on aurait $\mu_n z_n \to 0$ et y serait limite de la suite $(S(z_n)) \subset S(E)$, ce qui est impossible puisque $S(E)$ est supposé fermé et $y \notin S(E)$. Il existe donc une sous-suite (z'_n) telle que $\|z'_n\|$ tende vers $+\infty$; en posant $x_n = \|z'_n\|^{-1} z'_n$, on voit que $S(x_n) - \mu_n x_n = \|z'_n\|^{-1} y$ tend vers 0, donc $S(x_n) \to 0$.

//

Exemple 7.3.3. Exemple de valeurs propres approchées : soit S le shift à droite sur $\ell_2(\mathbb{N})$; on sait que le spectre de S est égal au disque unité fermé, sa frontière est donc le cercle unité $\mathbb{T}$. Soit λ de module 1 un point quelconque de $\partial\,\mathrm{Sp(S)}$; on considère pour tout $n \geq 1$ le vecteur de norme 1 de ℓ_2

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

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

Remarque 7.3.4. Voici une méthode plus orthodoxe pour traiter le cas hilbertien. Soient E un espace de Hilbert, T un opérateur normal et $\lambda \in \mathrm{Sp(T)}$; alors $S = T - \lambda\,\mathrm{Id_E}$ est normal non inversible, donc S ne vérifie pas **A** (lemme 8) : *toutes les valeurs spectrales d'un opérateur normal sont valeurs propres approchées.*

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

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

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

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

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

//

8. Quelques classes d'opérateurs

8.1. Compacité dans un espace de Banach

Rappelons qu'une partie A d'un espace topologique séparé X est dite *relativement compacte* dans X si son adhérence $\overline{\mathrm{A}}$ dans X est compacte. Une partie A d'un espace métrique (X, d) est dite *précompacte* si pour tout $\varepsilon > 0$, il existe un recouvrement fini de A par des parties de diamètre $\le \varepsilon$. Cela revient à dire que pour tout $\varepsilon > 0$, on peut trouver un entier N et des points $x_1, \ldots, x_\mathrm{N} \in \mathrm{X}$ tels que A soit contenu dans la réunion des boules $\mathrm{B}(x_i, \varepsilon)$, $i = 1, \ldots, \mathrm{N}$.

Théorème 8.1.1. *Dans un espace métrique complet* (X, d), *une partie* A *est relativement compacte si et seulement si elle est précompacte. En particulier, un espace métrique est compact si et seulement s'il est précompact et complet.*

Démonstration. Si $\overline{\mathrm{A}}$ est compacte dans X, il est facile de montrer qu'il existe pour tout $\varepsilon > 0$ un ensemble fini $x_1, \ldots, x_\mathrm{N}$ de points de A tel que A soit recouvert par les boules $\mathrm{B}(x_i, \varepsilon)$, $1 \le i \le \mathrm{N}$.

Inversement supposons A précompact et soit (x_n) une suite de points de $\overline{\mathrm{A}}$. On va trouver une sous-suite (x_{n_k}) de Cauchy, donc convergente puisque X est complet. On construit à cet effet une suite décroissante (M_k) de sous-ensembles infinis de $\mathbb{N}$ tels que $d(x_m, x_n) \le 2^{-k}$ pour tous $m, n \in \mathrm{M}_k$; il suffit ensuite d'appliquer le procédé de la sous-suite diagonale pour obtenir une sous-suite de Cauchy.

Supposons donc M_k choisi ; puisque A est précompact, il existe un ensemble fini $\mathrm{B} \subset \mathrm{X}$ tel que tout point x de A vérifie $d(x, y) < 2^{-k-2}$ pour au moins un point $y \in \mathrm{B}$. Il en résulte que $\overline{\mathrm{A}}$ est contenu dans la réunion finie de boules fermées $\{x \in \mathrm{X} : d(x, y) \le 2^{-k-2}\}$, pour $y \in \mathrm{B}$; en particulier l'ensemble infini M_k est recouvert par la famille **finie** des ensembles $\mathrm{N}_y = \{m \in \mathrm{M}_k : d(x_m, y) \le 2^{-k-2}\}$, indexée par les points $y \in \mathrm{B}$; il existe donc au moins un $y_0 \in \mathrm{B}$ tel que l'ensemble $\mathrm{M}_{k+1} = \mathrm{N}_{y_0} \subset \mathrm{M}_k$ soit infini. Si $m, n \in \mathrm{M}_{k+1}$, on aura

$$d(x_m, x_n) \le d(x_m, y_0) + d(y_0, x_n) \le 2 \,.\, 2^{-k-2}.$$

//

Notons une conséquence facile : pour qu'une partie A d'un espace métrique complet X soit relativement compacte, il suffit que pour tout $\varepsilon > 0$, il existe une partie compacte K_ε de X telle que tout point de A soit à une distance $< \varepsilon$ de l'ensemble K_ε :

$$\forall x \in \mathrm{A}, \quad d(x, \mathrm{K}_\varepsilon) < \varepsilon.$$

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

a. l'ensemble A est borné ;

b. pour tout $\varepsilon > 0$, il existe un sous-espace vectoriel $\mathrm{L}_\varepsilon \subset \mathrm{E}$ *de dimension finie* tel que tout point de A soit à une distance $< \varepsilon$ de L_ε :

$$\forall x \in \mathrm{A}, \quad \mathrm{dist}(x, \mathrm{L}_\varepsilon) < \varepsilon.$$

Si l'adhérence de A est compacte il est facile de vérifier que le critère est satisfait : en effet $\overline{A}$ est borné parce que compact (la fonction continue $x \to \|x\|$ atteint son maximum sur le compact $\overline{A}$) et la deuxième condition est évidemment impliquée par la précompacité : il suffit de prendre l'espace vectoriel L_ε engendré par un ensemble fini F_ε qui approche A à moins de ε.

Dans l'autre direction, supposons les deux conditions du critère vérifiées, et montrons que A est approchable arbitrairement bien par des compacts de E ; soit M une borne pour les normes des éléments de A ; soient $\varepsilon > 0$ et L_ε un sous-espace vectoriel de dimension finie qui approche A à moins de ε. Désignons par K_ε le compact de E formé par les points de L_ε de norme $\le M + \varepsilon$. Si $x \in A$, il existe $y \in L_\varepsilon$ tel que $\|x - y\| \le \varepsilon$; puisque $\|x\| \le M$, on aura $\|y\| \le M + \varepsilon$, d'où $y \in K_\varepsilon$, et le résultat est démontré.

Proposition 8.1.2. *Si* K_1 *et* K_2 *sont compacts dans l'espace de Banach* E*, l'ensemble* $K_1 + K_2$ *est compact ; si* A_1 *et* A_2 *sont relativement compacts dans l'espace de Banach* E*, l'ensemble* $A_1 + A_2$ *est relativement compact dans* E.

Démonstration. Il est clair que $K_1 + K_2$ est borné. Si L_j, $j = 1, 2$, est un sous-espace vectoriel de dimension finie qui approche K_j à moins de $\varepsilon/2$, il est facile de vérifier que le sous-espace de dimension finie $L_1 + L_2$ approche $K_1 + K_2$ à moins de ε. De plus $K_1 + K_2$ est fermé, donc compact, comme image du compact $K_1 \times K_2$ par l'application continue $(x, y) \to x + y$. La deuxième affirmation résulte facilement de la première, car l'adhérence de la somme $A_1 + A_2$ est contenue dans $\overline{A_1} + \overline{A_2}$.

//

Théorème d'Ascoli

Un ensemble A de fonctions scalaires sur un espace topologique X est dit *équicontinu au point* $t \in X$ si pour tout $\varepsilon > 0$, il existe un voisinage V de t dans lequel **toutes** les fonctions de A sont proches à ε près de leur valeur au point t,

$$\forall f \in A,\ \forall s \in V, \quad |f(s) - f(t)| < \varepsilon.$$

Lorsque X est un espace métrique compact (K, d), on montre que si A est équicontinu en tout point t de K, alors A est *uniformément équicontinu*, c'est à dire que pour tout $\varepsilon > 0$, il existe $\delta > 0$ tel que pour toute fonction $f \in A$ et tous $s, t \in K$,

$$\big(d(s,t) < \delta\big) \Rightarrow |f(s) - f(t)| < \varepsilon.$$

Théorème 8.1.3. *Soient* (K, d) *un espace métrique compact et* A *un sous-ensemble de* C(K) *; l'ensemble* A *est relativement compact dans* C(K) *si et seulement si les deux conditions suivantes sont vérifiées :*

1. *l'ensemble* A *est borné (pour la norme de* C(K)*) ;*
2. *l'ensemble* A *est uniformément équicontinu.*

Démonstration. On va montrer la partie la plus intéressante, celle qui dit que les conditions **1 + 2** entraînent que A est relativement compact. On doit montrer pour l'ensemble A les conditions (*a*) et (*b*) du critère de compacité ; comme (*a*) est identique à **1**, il suffit de montrer que **2** entraîne (*b*). On se donne donc $\varepsilon > 0$ et on cherche un sous-espace vectoriel

L_ε de dimension finie dans $C(K)$, tel que toute fonction de A soit à distance (uniforme) $< \varepsilon$ d'un point de L_ε. A cette valeur de ε correspond par la propriété **2** une valeur de $\delta > 0$ telle que $|f(s) - f(t)| < \varepsilon$ pour toute $f \in A$ et tous $s, t \in K$ tels que $d(s,t) < \delta$. Cette propriété de A signifie que pour toute fonction $f \in A$, le module de continuité ω_f vérifie $\omega_f(\delta) \le \varepsilon$. Par compacité, on peut trouver un recouvrement ouvert fini de K par des boules $U_i = B(t_i, \delta)$, $i = 1, \ldots, N$. On a vu à la section 3.1, relation (P), que si L_ε désigne l'espace de dimension finie engendré par les fonctions $\varphi_1, \ldots, \varphi_N$ d'une partition de l'unité subordonnée au recouvrement (U_i), on a

$$d(f, L_\varepsilon) \le \omega_f(\delta)$$

pour toute $f \in C(K)$, donc $d(f, L_\varepsilon) \le \omega_f(\delta) \le \varepsilon$ pour toute fonction $f \in A$, et la condition (b) est vérifiée.

//

8.2. Applications linéaires compactes

Définition 8.2.1. Soient E et F deux espaces de Banach ; une application linéaire continue $T \in \mathcal{L}(E, F)$ est dite *compacte* si l'image $T(B_E)$ par l'application T de la boule unité fermée B_E de l'espace E est relativement compacte (en norme) dans F. On note $\mathcal{K}(E, F)$ l'ensemble des applications linéaires compactes de E dans F. On pose $\mathcal{K}(E) = \mathcal{K}(E, E)$.

Proposition 8.2.1. *Soient* E *et* F *deux espaces de Banach ; l'ensemble* $\mathcal{K}(E, F)$ *est un sous-espace vectoriel fermé de* $\mathcal{L}(E, F)$.

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

Démonstration. Il est clair que si $T \in \mathcal{K}(E, F)$ et $\lambda \in \mathbb{K}$, alors $\lambda T \in \mathcal{K}(E, F)$. Soient maintenant T_1 et T_2 deux applications linéaires compactes de E dans F, et considérons les ensembles $A_1 = T_1(B_E)$, $A_2 = T_2(B_E)$ et $A = (T_1 + T_2)(B_E)$; il est clair que A est contenu dans $A_1 + A_2$, donc il est relativement compact d'après la proposition 1.2. Ceci montre que $\mathcal{K}(E, F)$ est un sous-espace vectoriel de $\mathcal{L}(E, F)$.

Supposons que $T \in \mathcal{L}(E, F)$ soit adhérent à $\mathcal{K}(E, F)$. Pour tout $\varepsilon > 0$ donné, on peut trouver S compacte telle que $\|T - S\| < \varepsilon$; il en résulte que tout point de $T(B_E)$ est approché à ε près par un point du compact $K_\varepsilon = \overline{S(B_E)}$, donc $\overline{T(B_E)}$ est compact.

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

//

Exemples 8.2.2.

1. Il est clair que tout opérateur T de rang fini est compact : en effet, l'ensemble $T(B_E)$ est alors un ensemble borné d'un espace vectoriel de dimension finie. D'après le résultat précédent, toute limite T en norme d'opérateur d'une suite (T_n) d'opérateurs de rang fini est compacte. C'est une méthode assez efficace pour vérifier que certains

opérateurs sont compacts ; on montre par exemple que si $c_n \to 0$, l'opérateur Δ_c de ℓ_p dans ℓ_p défini par $\Delta_c((x_n)) = (c_n x_n)$ est compact :

on commence par remarquer que la norme de Δ_c dans $\mathcal{L}(\ell_p)$ est majorée par $\|c\|_\infty$ (elle est en fait égale à $\|c\|_\infty$). Ensuite, pour tout entier N on considère la suite $c^{(\mathrm{N})}$ telle que $c_n^{(\mathrm{N})} = c_n$ si $n \le \mathrm{N}$ et $c_n^{(\mathrm{N})} = 0$ sinon ; l'opérateur $\mathrm{T_N} = \Delta_{c^{(\mathrm{N})}}$ est de rang fini, et $\|\Delta_c - \mathrm{T_N}\| = \|\Delta_{c-c^{(\mathrm{N})}}\|$ est majoré par $\|c - c^{(\mathrm{N})}\|_\infty = \sup_{n>\mathrm{N}} |c_n|$ qui tend vers 0 parce que la suite (c_n) tend vers 0.

2. Pour toute fonction f intégrable sur $[0,1]$ définissons la fonction continue $\mathrm{V}(f)$ comme dans l'exemple 7.2.2,

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

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

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

Proposition 8.2.2. *Soient* E *et* F *deux espaces de Banach ; si* $\mathrm{T} \in \mathcal{L}(\mathrm{E},\mathrm{F})$ *est compacte, sa transposée* ${}^t\mathrm{T}$ *est compacte de* F^* *dans* E^*.

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

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

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

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

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

//

Proposition 8.2.3. *Soient* E *et* F *deux espaces de Banach et* $\mathrm{T} \in \mathcal{L}(\mathrm{E}, \mathrm{F})$ *; notons* $\mathrm{B_E}$ *la boule unité fermée de* E.

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

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

Démonstration. Supposons T compact, et soit K un compact de F contenant $\mathrm{T(B_E)}$; l'identité, de K muni de la topologie de la norme, dans K muni de la topologie faible est continue ; comme K est compact, c'est un homéomorphisme. Comme T est continu de $\mathrm{B_E}$ muni de la topologie faible dans K muni de la topologie faible, il en résulte que T est continu de $\mathrm{B_E}$ faible dans F muni de la norme. Si (x_n) est une suite qui converge faiblement vers 0 dans E, elle est bornée dans E (lemme 5.3.1), donc $(\mathrm{T}(x_n))$ tend vers 0 en norme par ce qui précède.

Lorsque E est réflexif, la boule $\mathrm{B_E}$ est faiblement compacte, donc son image $\mathrm{T(B_E)}$ est faiblement compacte dans F, donc faiblement fermée, donc fermée ; puisque $\mathrm{T(B_E)}$ est relativement compacte, elle est en fait compacte. Supposons encore E réflexif et que $(\mathrm{T}(x_n))$ converge vers 0 en norme dans F pour toute suite (x_n) qui tend faiblement vers 0_E ; soit (x_n) une suite dans $\mathrm{B_E}$; d'après le théorème 5.3.5, il existe une sous-suite (x_{n_k}) qui converge faiblement vers un point $x \in \mathrm{B_E}$; alors $(x_{n_k} - x)$ converge faiblement vers 0, donc $\mathrm{T}(x_{n_k}) - \mathrm{T}(x)$ converge en norme vers 0 d'après l'hypothèse ; on a ainsi montré que pour toute suite $(x_n) \subset \mathrm{B_E}$, il existe une sous-suite $(\mathrm{T}(x_{n_k}))$ qui converge en norme, donc T est compact.

//

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

Théorème 8.2.4. *Soient* E *un espace de Hilbert,* F *un espace de Banach et* $\mathrm{T} \in \mathcal{L}(\mathrm{E}, \mathrm{F})$ *; notons* $\mathrm{B_E}$ *la boule unité fermée de* E. *Les propriétés suivantes sont équivalentes :*

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

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

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

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

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

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

Démonstration. Puisque E est réflexif, on sait que (i), (ii), (iii) et (iv) sont équivalents. De plus, (v) $\Rightarrow$ (i) en général, et on sait que (iv) $\Rightarrow$ (vi) parce que les suites orthonormées tendent faiblement vers 0 (exemples 5.3.1).

Supposons que (v) ne soit pas vérifiée. Il existe alors $\varepsilon > 0$ tel que pour toute application linéaire continue de rang fini R on ait $\|\mathrm{T} - \mathrm{R}\| > \varepsilon$. Construisons alors par récurrence sur n un système orthonormal $(e_n)_{n \geq 0}$ tel que $\|\mathrm{T}(e_n)\| > \varepsilon$ pour tout

$n \geq 0$: comme $\|T\| > \varepsilon$, il existe $e_0 \in E$ tel que $\|e_0\| = 1$ et $\|T(e_0)\| > \varepsilon$; supposons e_k construit pour $k < n$ et soit P le projecteur orthogonal sur le sous-espace de E engendré par $\{e_k : k < n\}$; alors TP est de rang fini donc $\|T - TP\| > \varepsilon$; il existe donc $y_n \in E$ tel que $\|T(\mathrm{Id}_E - P)(y_n)\| > \varepsilon \|y_n\| \geq \varepsilon \|(\mathrm{Id}_E - P)(y_n)\|$; on pose alors $z_n = (\mathrm{Id}_E - P)(y_n)$, puis $e_n = \|z_n\|^{-1} z_n$. On a alors $\|T(e_n)\| = \|z_n\|^{-1} \|T(z_n)\| > \varepsilon$, donc (vi) n'est pas vérifiée. On a ainsi montré que $(vi) \Rightarrow (v)$.

//

8.3. Théorie spectrale des opérateurs compacts

Cette théorie est pour l'essentiel la création du mathématicien hongrois F. Riesz, aux alentours de 1910. Le théorème 4.4.2 (avec son corollaire) est l'un des points-clés de cette théorie.

Lemme 8.3.1. *Soit* E *un espace de Banach ; pour tout sous-espace vectoriel* L *de dimension finie de* E, *il existe un projecteur continu* P *de* E *sur* L, *c'est à dire qu'il existe un sous-espace fermé* F *tel que* $E = L \oplus F$.

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

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

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

//

Lemme 8.3.2. *Soit* $K \in \mathcal{L}(E)$ *un opérateur compact, et posons* $T = \mathrm{Id}_E - K$ *; si* F *est un sous-espace fermé de* E *tel que* T *soit injectif de* F *dans* E, *il existe une constante* $c > 0$ *telle que* $\|T(x)\| \geq c\,\|x\|$ *pour tout* $x \in F$ *; il en résulte que l'image* T(F) *est fermée.*

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

Désignons par T_1 la restriction de T à F ; on a vu dans la proposition 7.3.1 que la minoration $\|T_1(x)\| \geq c\|x\|$ (pour tout $x \in F$, et avec $c > 0$) implique que $\mathrm{im}(T_1) = T(F)$ est fermée.

//

Proposition 8.3.3. *Soit* $K \in \mathcal{L}(E)$ *un opérateur compact, et posons* $T = \mathrm{Id}_E - K$ *; le noyau de* T *est de dimension finie et l'image* T(E) *est fermée.*

On remarquera, en utilisant la formule du binôme et la propriété d'idéal de $\mathcal{K}(E)$, que $T^n = (\mathrm{Id}_E - K)^n$ est de la forme $\mathrm{Id}_E - K_n$, avec K_n compact, donc les images de T^n sont fermées pour tout $n \geq 0$ (et leurs noyaux sont de dimension finie).

Démonstration. Le noyau de T est le sous-espace propre de l'opérateur compact K pour la valeur propre 1, il est donc de dimension finie d'après le théorème de Riesz (corollaire du théorème 4.4.2). Soit F un sous-espace fermé de E tel que $\mathrm{E} = \ker(\mathrm{T}) \oplus \mathrm{F}$; alors T est injectif sur F, donc $\mathrm{T(E)} = \mathrm{T(F)}$ est fermé.

//

Lemme 8.3.4. *Si* F *et* G *sont deux sous-espaces vectoriels de* E*, avec* F *fermé et* $\mathrm{F} \subset \mathrm{G}$, $\mathrm{F} \neq \mathrm{G}$*, on peut trouver pour tout* $\varepsilon > 0$ *un vecteur* $y \in \mathrm{G}$ *tel que* $\|y\| = 1$ *et* $d(y, \mathrm{F}) > 1 - \varepsilon$*.*

Démonstration. Puisque $\mathrm{F} \neq \mathrm{G}$, on peut trouver un premier vecteur $y_0 \in \mathrm{G} \setminus \mathrm{F}$. Puisque F est fermé et $y_0 \notin \mathrm{F}$, on a $\delta = d(y_0, \mathrm{F}) > 0$. On peut trouver $x_0 \in \mathrm{F}$ tel que $\alpha = \|y_0 - x_0\| < \delta/(1-\varepsilon)$. Alors $y = \alpha^{-1}(y_0 - x_0) \in \mathrm{G}$ convient.

//

Lemme 8.3.5. *Soit* $\mathrm{K} \in \mathcal{L}(\mathrm{E})$ *un opérateur compact, et posons* $\mathrm{T} = \mathrm{Id_E} - \mathrm{K}$ *; il n'existe pas de chaîne infinie* $(\mathrm{F}_n)_{n \geq 0}$ *(resp :* $(\mathrm{F}_n)_{n \leq 0}$*) de sous-espaces vectoriels fermés de* E *telle que*

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

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

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

//

Corollaire 8.3.6. *Soit* $\mathrm{K} \in \mathcal{L}(\mathrm{E})$ *un opérateur compact, et posons* $\mathrm{T} = \mathrm{Id_E} - \mathrm{K}$ *; la suite croissante des noyaux* $(\ker(\mathrm{T}^n))_{n \geq 0}$ *est stationnaire. La suite décroissante des images* $(\mathrm{im}(\mathrm{T}^n))_{n \geq 0}$ *est stationnaire.*

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

//

Corollaire 8.3.7. *Soit* $\mathrm{K} \in \mathcal{L}(\mathrm{E})$ *un opérateur compact, et posons* $\mathrm{T} = \mathrm{Id}_{\mathrm{E}} - \mathrm{K}$ *; si* T *est surjectif, alors* $\ker(\mathrm{T}) = \{0\}$ *; si* T *est injectif, alors* $\mathrm{im}(\mathrm{T}) = \mathrm{E}$.

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

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

//

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

Théorème 8.3.8 : Alternative de Fredholm. *Soient* E *un espace de Banach et* $\mathrm{T} \in \mathcal{L}(\mathrm{E})$ *un opérateur borné de la forme* $\mathrm{T} = \mathrm{Id}_{\mathrm{E}} - \mathrm{K}$*, avec* K *compact ; l'image de* T *est fermée et de codimension finie et l'on a*

$$\operatorname{codim} \mathrm{im}(\mathrm{T}) = \dim \ker(\mathrm{T}).$$

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

Démonstration. On a vu que $\ker(\mathrm{T})$ est de dimension finie et $\mathrm{im}(\mathrm{T})$ fermée. On doit montrer de plus que $\dim \ker(\mathrm{T}) = \operatorname{codim} \mathrm{T}(\mathrm{E})$, c'est à dire que l'indice de T est nul. On va procéder par récurrence sur la dimension de $\ker(\mathrm{T})$. Si $\dim \ker(\mathrm{T}) = 0$, on sait que T est surjectif d'après le corollaire 7, donc l'indice est nul dans ce cas ; on suppose donc que n est un entier > 0 et que $\mathrm{ind}(\mathrm{T}') = 0$ pour tout opérateur $\mathrm{T}' = \mathrm{Id}_{\mathrm{E}} - \mathrm{K}'$, où K' est compact et $\dim \ker(\mathrm{T}') < n$. Soit $\mathrm{T} = \mathrm{Id}_{\mathrm{E}} - \mathrm{K}$ avec K compact et $\dim \ker(\mathrm{T}) = n > 0$; d'après le corollaire 7, on a $\mathrm{im}(\mathrm{T}) \neq \mathrm{E}$; soit donc $y_0 \notin \mathrm{im}(\mathrm{T})$; on note que $\mathbb{K} y_0 \oplus \mathrm{T}(\mathrm{E})$ est une somme directe. On va construire T' de la forme $\mathrm{Id}_{\mathrm{E}} - \mathrm{K}'$ tel que $\mathrm{ind}(\mathrm{T}') = \mathrm{ind}(\mathrm{T})$ et $\dim \ker(\mathrm{T}') < \dim \ker(\mathrm{T})$; d'après l'hypothèse de récurrence, on aura $0 = \mathrm{ind}(\mathrm{T}') = \mathrm{ind}(\mathrm{T})$, ce qui donnera le résultat.

On écrit $\mathrm{E} = \ker(\mathrm{T}) \oplus \mathrm{E}_1$ en utilisant le lemme 1 ; soit $x_1, \ldots, x_n$ une base de $\ker(\mathrm{T})$. On définit un opérateur $\mathrm{T}' \in \mathcal{L}(\mathrm{E})$ en posant pour tout $x \in \mathrm{E}$, représenté sous la forme $x = \lambda_1 x_1 + \cdots + \lambda_n x_n + y$, avec $y \in \mathrm{E}_1$

$$\mathrm{T}'(\lambda_1 x_1 + \cdots + \lambda_n x_n + y) = \lambda_1 y_0 + \mathrm{T}(y).$$

Si $\mathrm{T}'(x) = 0$, il en résulte que $\mathrm{T}(y) = 0_{\mathrm{E}}$ et $\lambda_1 y_0 = 0_{\mathrm{E}}$, donc $y \in \ker(\mathrm{T}) \cap \mathrm{E}_1$ entraîne $y = 0_{\mathrm{E}}$; d'autre part $\lambda_1 y_0 = 0_{\mathrm{E}}$ entraîne $\lambda_1 = 0$ puisque le vecteur y_0 est non nul. Il en résulte que $\ker(\mathrm{T}') = \mathrm{Vect}(x_2, \ldots, x_n)$ est de dimension $n - 1$. Par ailleurs, l'opérateur $\mathrm{R} = \mathrm{T}' - \mathrm{T}$ est de rang un : en effet $(\mathrm{T}' - \mathrm{T})(x) = \lambda_1 y_0$ pour tout x, donc l'image de R est contenue dans $\mathbb{K} y_0$; on peut écrire par conséquent $\mathrm{T}' = \mathrm{Id}_{\mathrm{E}} - \mathrm{K}'$ avec $\mathrm{K}' = \mathrm{K} - \mathrm{R}$ compact, et on a alors $\mathrm{ind}(\mathrm{T}') = 0$ d'après l'hypothèse de récurrence, ce qui montre déjà que $\operatorname{codim} \mathrm{im}(\mathrm{T}')$ est finie. Il est clair que $\mathrm{im}(\mathrm{T}') = \mathbb{K} y_0 \oplus \mathrm{T}(\mathrm{E})$ a exactement une dimension de plus que $\mathrm{T}(\mathrm{E})$, donc $\operatorname{codim} \mathrm{im}(\mathrm{T}) = \operatorname{codim} \mathrm{im}(\mathrm{T}') + 1$, et $\mathrm{ind}(\mathrm{T}) = \mathrm{ind}(\mathrm{T}') = 0$.

//

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

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

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

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

Théorème 8.3.9. *Soient* E *un espace de Banach complexe et* $\mathrm{K} \in \mathcal{K}(\mathrm{E})$ *un opérateur compact ; le spectre de* K *est fini ou formé d'une suite tendant vers* 0*. Chaque valeur* $\lambda \neq 0$ *dans* $\mathrm{Sp}(\mathrm{K})$ *est une valeur propre de* K*, de multiplicité finie.*

Démonstration. On va montrer que si $\lambda \neq 0$ est dans le spectre de K, alors λ est valeur propre de K et λ est isolé dans le spectre de K. En remplaçant K par $\lambda^{-1}\mathrm{K}$ on se ramène à traiter $\lambda = 1$. Posons $\mathrm{T} = \mathrm{Id_E} - \mathrm{K}$; si 1 n'est pas valeur propre de K, l'opérateur T est injectif, donc surjectif d'après le corollaire 7, donc $\mathrm{Id_E} - \mathrm{K}$ est inversible et 1 n'est pas dans le spectre de K. Supposons que $1 \in \mathrm{Sp}(\mathrm{K})$, donc 1 est valeur propre ; remarquons que $\mathrm{T}^n = (\mathrm{Id_E} - \mathrm{K})^n = \mathrm{Id_E} - \mathrm{K}_n$ avec K_n compact (utiliser la formule du binôme), donc on sait que $\dim\ker(\mathrm{T}^n) = \operatorname{codim}\operatorname{im}(\mathrm{T}^n)$ pour tout $n \geq 0$ (pour $n = 0$, c'est une évidence). On a vu qu'il existe un entier k tel que $\ker(\mathrm{T}^k) = \ker(\mathrm{T}^{k+1})$, et on peut prendre pour k le plus petit entier vérifiant cette propriété ; on a $k \geq 1$ puisque 1 est valeur propre de K ; alors $\ker(\mathrm{T}) \cap \operatorname{im}(\mathrm{T}^k) = \{0\}$, sinon $\ker(\mathrm{T}^k) \neq \ker(\mathrm{T}^{k+1})$; on a *a fortiori* $\ker(\mathrm{T}^k) \cap \operatorname{im}(\mathrm{T}^k) = \{0\}$, et d'après l'égalité dimension-codimension il en résulte que

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

L'espace E se trouve décomposé en deux sous-espaces fermés T-invariants. La restriction T_2 de T à $\operatorname{im}(\mathrm{T}^k)$ est injective, donc c'est un isomorphisme de $\operatorname{im}(\mathrm{T}^k)$ sur $\operatorname{im}(\mathrm{T}^k)$ d'après le théorème 8. La restriction T_1 de T à $\ker(\mathrm{T}^k)$ est un endomorphisme en dimension finie, dont la seule valeur propre est 0 ; pour tout $\lambda \neq 0$, $\mathrm{T}_1 - \lambda$ est donc bijective de $\ker(\mathrm{T}^k)$ sur $\ker(\mathrm{T}^k)$, et pour λ assez petit, $\mathrm{T}_2 - \lambda$ est encore un isomorphisme ; il en résulte que $\mathrm{T} - \lambda$ est un isomorphisme pour $\lambda \neq 0$ et assez petit, ce qui signifie que 0 est isolé dans le spectre de T, ou encore que 1 est isolé dans le spectre de K. On en déduit que pour tout $\varepsilon > 0$ il y a un nombre fini de valeurs spectrales telles que $|\lambda| \geq \varepsilon$, ce qui permet de ranger les valeurs spectrales non nulles de K dans une suite qui tend vers 0, à moins que le spectre ne soit fini. //

Théorème 8.3.10. *Pour toute application linéaire compacte normale* T *d'un espace de Hilbert complexe* H *dans lui-même, l'espace* H *est somme directe hilbertienne (orthogonale) de la famille des sous-espaces propres de* T*. Il en résulte que* H *admet une base hilbertienne formée de vecteurs propres de* T*.*

Démonstration. Commençons par une remarque : si E est un Hilbert complexe *non nul* et si S est normal compact sur E, il existe $x \neq 0$ dans E et $\mu \in \mathbb{C}$ tels que $\mathrm{S}x = \mu x$. En effet, on peut appliquer la formule du rayon spectral à l'algèbre unitaire $\mathcal{L}(\mathrm{E})$ (parce que $\mathrm{E} \neq \{0\}$) : il existe une valeur spectrale μ de S telle que $|\mu| = \rho(\mathrm{S}) = \|\mathrm{S}\|$ (proposition 7.2.3). Si $\mu = 0$, on a $\mathrm{S} = 0$ et tout vecteur $x \in \mathrm{E}$ non nul répond à la question. Si $\mu \neq 0$, on sait que μ est valeur propre d'après le théorème 9.

Soient H un espace de Hilbert complexe et $\mathrm{T} \in \mathcal{L}(\mathrm{H})$ une application linéaire compacte normale ; soit K son spectre ; c'est un ensemble fini ou dénombrable. Pour $\lambda \in \mathrm{K}$ notons $\mathrm{E}_\lambda = \ker(\mathrm{T} - \lambda\,\mathrm{Id}_\mathrm{H})$ l'espace propre de T associé. On va démontrer que les E_λ, $\lambda \in \mathrm{K}$, sont deux à deux orthogonaux, et que le sous-espace engendré par les E_λ, $\lambda \in \mathrm{K}$, est dense dans H. On pourra alors considérer la somme hilbertienne F des sous-espaces deux à deux orthogonaux (E_λ), et on aura $\mathrm{F} = \mathrm{H}$ d'après la densité de la somme des (E_λ).

On rappelle que $\ker \mathrm{S} = \ker \mathrm{S}^*$ quand S est normal ; comme $\mathrm{S} = \mathrm{T} - \lambda\,\mathrm{Id}$ est normal et $\mathrm{S}^* = \mathrm{T}^* - \overline{\lambda}\,\mathrm{Id}$, on voit que $\mathrm{E}_\lambda = \ker(\mathrm{T} - \lambda\,\mathrm{Id}) = \ker(\mathrm{T}^* - \overline{\lambda}\,\mathrm{Id})$; il en résulte que chaque E_λ est stable par T et par T^*. Si $x \in \mathrm{E}_\lambda$ et $y \in \mathrm{E}_\mu$ alors $\langle \mathrm{T}(x), y\rangle = \mu\langle x, y\rangle = \langle x, \mathrm{T}^* y\rangle = \lambda\langle x, y\rangle$ ce qui montre que $\langle x, y\rangle = 0$ si $\lambda \neq \mu$: les sous-espaces propres de T sont donc deux à deux orthogonaux.

Notons F le sous-espace fermé de H engendré par les E_λ, pour λ valeur propre de T (ces espaces sont de dimension finie si $\lambda \neq 0$; le sous-espace $\mathrm{E}_0 = \ker \mathrm{T}$ peut être réduit à $\{0\}$, ou bien de dimension finie, ou infinie). Puisque chaque E_λ est stable par T et T^*, on a $\mathrm{T}(\mathrm{F}) \subset \mathrm{F}$ et $\mathrm{T}^*(\mathrm{F}) \subset \mathrm{F}$. Il s'ensuit que $\mathrm{T}(\mathrm{F}^\perp) \subset \mathrm{F}^\perp$ et $\mathrm{T}^*(\mathrm{F}^\perp) \subset \mathrm{F}^\perp$. Notons $\mathrm{T}_1 \in \mathcal{L}(\mathrm{F}^\perp)$ la restriction de T à l'orthogonal de F. Si on avait $\mathrm{E} = \mathrm{F}^\perp \neq \{0\}$, T_1 serait un opérateur normal compact sur E, qui aurait, d'après la remarque préliminaire, au moins un vecteur propre $x \in \mathrm{F}^\perp$, $x \neq 0$ et $\mathrm{T}_1(x) = \mathrm{T}(x) = \mu\, x$ pour un certain $\mu \in \mathbb{C}$; mais alors on devrait avoir $x \in \mathrm{F}$, puisque F contient tous les vecteurs propres de T ; on a donc $x \in \mathrm{F} \cap \mathrm{F}^\perp$ ce qui implique $x = 0_\mathrm{H}$, contradiction. On a donc bien $\mathrm{F} = \mathrm{H}$. Pour obtenir une base orthonormée de H formée de vecteurs propres de T, on rassemble des bases orthonormées de chaque espace E_λ, $\lambda \neq 0$, qui sont des bases finies, et s'il y a lieu, une base orthonormée du noyau E_0.

//

Remarque 8.3.1. Il s'agit ici d'un théorème qui demande que le corps de base soit $\mathbb{C}$. Cependant, si H est un espace de Hilbert réel et $\mathrm{T} \in \mathcal{L}(\mathrm{H})$ un opérateur **hermitien** compact, il existe une base orthonormée de H formée de vecteurs propres de T.

8.4. Opérateurs de Hilbert-Schmidt

Lemme 8.4.1. *Soient* E *et* F *deux espaces de Hilbert,* B *une base hilbertienne de* E *et* B′ *une base hilbertienne de* F *; pour tout* $\mathrm{T} \in \mathcal{L}(\mathrm{E}, \mathrm{F})$ *on a :*

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

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

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

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

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

//

Pour $\mathrm{T} \in \mathcal{L}(\mathrm{E},\mathrm{F})$ on pose $\|\mathrm{T}\|_2 = \big(\sum_{b\in\mathrm{B}} \|\mathrm{T}(b)\|^2\big)^{1/2}$ où B est une base hilbertienne quelconque de E. Posons $\mathcal{L}^2(\mathrm{E},\mathrm{F}) = \{\mathrm{T} \in \mathcal{L}(\mathrm{E},\mathrm{F}) : \|\mathrm{T}\|_2 < +\infty\}$. D'après le point (vi) du théorème 2.4, tout $\mathrm{T} \in \mathcal{L}^2(\mathrm{E},\mathrm{F})$ est compact.

Définition 8.4.1. Soient E et F deux espaces de Hilbert ; un opérateur $\mathrm{T} \in \mathcal{L}^2(\mathrm{E},\mathrm{F})$ est dit de *Hilbert-Schmidt.*

Exemples 8.4.2.

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

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

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

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

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

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

On montre que $\mathrm{T_K}$ est bien défini, et agit continûment de $\mathrm{L}_2(\mathrm{Y},\nu)$ dans $\mathrm{L}_2(\mathrm{X},\mu)$:

avec Cauchy-Schwarz, on a

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

ce qui donne en réintégrant

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

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

Si $(f_n)_{n\geq 0}$ est une base hilbertienne de $\mathrm{L}_2(\mathrm{X},\mu)$ et $(g_n)_{n\geq 0}$ une base hilbertienne de $\mathrm{L}_2(\mathrm{Y},\nu)$, il en résulte que les fonctions $(s,t) \to f_m(s)g_n(t)$ (où m, n prennent toutes les valeurs entières ≥ 0) donnent une base orthonormée de l'espace $\mathrm{L}_2(\mathrm{X}\times\mathrm{Y},\mu\otimes\nu)$.

L'opérateur $\mathrm{T_K}$ est de Hilbert-Schmidt :

si on écrit

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

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

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

On a donc vérifié que la norme Hilbert-Schmidt de $\mathrm{T_K}$ est *égale* à la norme L_2 du noyau K dans l'espace $\mathrm{L}_2(\mathrm{X}\times\mathrm{Y},\mu\otimes\nu)$. En particulier, l'application $\mathrm{K} \to \mathrm{T_K}$ est injective : si l'opérateur $\mathrm{T_K}$ est l'opérateur nul, le noyau K est nul $\mu\otimes\nu$-presque partout sur $\mathrm{X}\times\mathrm{Y}$.

Exercice 8.4.3. Montrer que la composition de deux opérateurs $\mathrm{T_{K_1}}$ et $\mathrm{T_{K_2}}$ de la forme précédente est un opérateur $\mathrm{T_K}$, avec

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

On peut vérifier que l'adjoint de $\mathrm{T_K}$ est l'opérateur de noyau $\mathrm{K}^*(t,s) = \overline{\mathrm{K}(s,t)}$. Supposons que $\mathrm{X}=\mathrm{Y}$, $\mu=\nu$ et que K soit un *noyau hermitien*, c'est à dire que $\mathrm{K}(t,s) = \overline{\mathrm{K}(s,t)}$ pour tous $(s,t) \in \mathrm{X}^2$; il existe alors une base orthonormée (f_n) de $\mathrm{L}_2(\mathrm{X},\mu)$ formée de vecteurs propres de l'opérateur hermitien compact $\mathrm{T_K}$, c'est à dire telle que $\mathrm{T_K}(f_n) = \lambda_n f_n$ pour tout $n \geq 0$. Si on exprime le noyau K dans la base orthonormée de l'espace $\mathrm{L}_2(\mathrm{X}^2,\mu\otimes\mu)$ formée des fonctions $h_{m,n}(s,t) = f_m(s)\overline{f_n(t)}$, on obtient une expression $\mathrm{K}(s,t) = \sum_{m,n} c_{m,n} f_m(s)\overline{f_n(t)}$, et on voit que

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

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

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

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

Proposition 8.4.3. *Soient* E, F *et* H *des espaces de Hilbert ; pour tout* $\mathrm{S} \in \mathcal{L}(\mathrm{E},\mathrm{F})$ *et* $\mathrm{T} \in \mathcal{L}(\mathrm{F},\mathrm{H})$ *on a :*

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

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

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

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

9. Calcul fonctionnel continu

L'un des objectifs du chapitre est de construire un homomorphisme isométrique φ_{T} de $\mathrm{C}(\mathrm{Sp}(\mathrm{T}))$ dans $\mathcal{L}(\mathrm{H})$ lorsque T est un opérateur hermitien (borné) sur un espace de Hilbert complexe H.

9.1. Calcul fonctionnel polynomial

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

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

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

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

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

Proposition 9.1.1. *Soient* A *une algèbre de Banach unitaire et* $a \in \mathrm{A}$ *; il existe un unique homomorphisme d'algèbres unitaires* φ_a *de* $\mathbb{K}[\mathrm{X}]$ *dans* A *tel que* $\varphi_a(\mathrm{X}) = a$ *; cet homomorphisme est donné par* $\varphi_a(\mathrm{P}) = \mathrm{P}(a)$.

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

Lemme 9.1.2. *Si* $a_1\, a_2 = a_2\, a_1$ *est inversible dans* A, *alors* a_1 *est inversible dans* A.

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

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

//

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

Démonstration. Montrons le pour $a - \mu_1 1_{\mathrm{A}}$ par exemple ; considérons le produit $a_2 = c(a - \mu_2 1_{\mathrm{A}}) \ldots (a - \mu_k 1_{\mathrm{A}})$; alors $a_1 = a - \mu_1 1_{\mathrm{A}}$ et a_2 commutent, et a_1a_2 est inversible, donc a_1 est inversible.

//

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

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

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

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

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

Si $\lambda \in \mathrm{P}(\mathrm{Sp}(a))$, il existe $z \in \mathrm{Sp}(a)$ tel que $\mathrm{P}(z) = \lambda$; le polynôme $\mathrm{P} - \lambda$ s'annule en z, donc z est l'une des racines (μ_j), par exemple $z = \mu_1$; puisque $\mu_1 = z \in \mathrm{Sp}(a)$, l'élément $a - \mu_1 1_\mathrm{A}$ est non inversible. Alors $\mathrm{P}(a) - \lambda 1_\mathrm{A} = c \prod_{i=1}^{k} (a - \mu_i 1_\mathrm{A})$ est non inversible d'après le corollaire qui précède, et $\lambda \in \mathrm{Sp}(\mathrm{P}(a))$.

//

Exemples 9.1.1.

1. S'il existe une base orthonormée $(e_i)_{i \in \mathrm{I}}$ d'un Hilbert H réel ou complexe et un opérateur borné $\mathrm{T} \in \mathcal{L}(\mathrm{H})$ tel que $\mathrm{T}(e_i) = \lambda_i e_i$ pour tout $i \in \mathrm{I}$ (l'opérateur T est *diagonal* dans la base (e_i)), il est facile de voir que pour tout polynôme $\mathrm{P} \in \mathbb{K}[\mathrm{X}]$ l'opérateur $\mathrm{P}(\mathrm{T})$ est l'opérateur diagonal dont les coefficients diagonaux sont les $\mathrm{P}(\lambda_i)$.

2. Si on considère sur $\mathrm{H} = \mathrm{L}_2(0, 1)$ l'opérateur M_f de multiplication par $f \in \mathrm{L}_\infty(0, 1)$, on voit que $\mathrm{P}(\mathrm{M}_f)$ est l'opérateur de multiplication par la fonction $t \in [0, 1] \to \mathrm{P}(f(t))$, c'est à dire la multiplication par la fonction $\mathrm{P} \circ f$.

9.2. Calcul fonctionnel continu pour les opérateurs hermitiens

Polynômes et adjoints

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

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

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

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

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

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

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

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

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

d'après la proposition 7.2.3.

//

Remarque 9.2.1. Le résultat précédent est à peu près évident lorsque T est normal et compact. Dans ce cas, il existe une base orthonormée $(e_i)_{i \in \mathrm{I}}$ 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 $|\lambda_i| \geq \varepsilon$. Supposons I infini dénombrable pour fixer les idées. L'opérateur P(T) est l'opérateur diagonal dont les coefficients diagonaux sont les $\mathrm{P}(\lambda_i)$, la norme de P(T) est donc le sup des $|\mathrm{P}(\lambda_i)|$, qui est majoré par le sup de $|\mathrm{P}|$ 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 $\|\mathrm{P(T)}\| \geq \|\mathrm{P(T)}(e_i)\| = |\mathrm{P}(\lambda_i)| = |\mathrm{P}(\lambda)|$, ou bien $\lambda = 0$ est limite d'une suite de λ_i, ce qui conduit au même résultat puisque P définit une fonction continue sur K. On a donc bien $\|\mathrm{P(T)}\| = \|\mathrm{P}\|_{\mathrm{C(K)}}$.

Lemme 9.2.2. *Soient* K *un compact non vide et* φ *un homomorphisme isométrique de* C(K) *dans une algèbre de Banach unitaire* B *; alors* $f \in \mathrm{C(K)}$ *est inversible dans* C(K) *si et seulement si* $\varphi(f)$ *est inversible dans* B.

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

//

Corollaire 9.2.3. *Pour tout homomorphisme isométrique* φ *de* C(K) *(complexe) dans une algèbre de Banach unitaire complexe* B, *on a*

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

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

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

//

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

Théorème 9.2.4. *Soient* H *un espace de Hilbert complexe et* $\mathrm{T} \in \mathcal{L}(\mathrm{H})$ *hermitien ; posons* $\mathrm{K} = \mathrm{Sp}(\mathrm{T})$. *Il existe un et un seul homomorphisme d'algèbres de Banach unitaires complexes* $\varphi_{\mathrm{T}} : \mathrm{C}(\mathrm{K}) \rightarrow \mathcal{L}(\mathrm{H})$ *tel que* $\varphi_{\mathrm{T}}(i_{\mathrm{K}}) = \mathrm{T}$.

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

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

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

Démonstration. On a vu que $\mathrm{K} = \mathrm{Sp}(\mathrm{T})$ est contenu dans $\mathbb{R}$. Désignons par A l'ensemble des fonctions continues f sur K de la forme $f : s \rightarrow \mathrm{P}(s)$ pour un $\mathrm{P} \in \mathbb{C}[\mathrm{X}]$ (fonctions polynomiales). D'après le théorème de Weierstrass, les fonctions polynomiales à coefficients complexes sont uniformément denses dans l'espace $\mathrm{C}([-a, a])$ des fonctions complexes continues sur $[-a, a]$, pour tout $a > 0$; il en résulte que l'ensemble A est dense dans $\mathrm{C}(\mathrm{K})$, puisque $\mathrm{K} \subset [-a, a]$ lorsque par exemple $a = \|\mathrm{T}\|$.

Montrons d'abord l'unicité de φ_{T}. Si φ est un homomorphisme d'algèbres unitaires de $\mathrm{C}(\mathrm{K})$ dans $\mathcal{L}(\mathrm{H})$ tel que $\varphi(i_{\mathrm{K}}) = \mathrm{T}$, on aura nécessairement par les propriétés d'homomorphisme que l'image de la fonction $t \in \mathrm{K} \rightarrow \mathrm{P}(t)$ est égale à $\mathrm{P}(\mathrm{T})$: par définition, on a $\varphi(i_{\mathrm{K}}^0) = \varphi(1) = \mathrm{Id}_{\mathrm{H}} = \mathrm{T}^0$, et $\varphi(i_{\mathrm{K}}^k) = \mathrm{T}^k$ pour tout $k \geq 1$ (la fonction i_{K}^k est la fonction monôme $s \rightarrow s^k$) ; il en résulte puisque φ est de plus linéaire que pour toute fonction polynomiale $f : s \rightarrow \mathrm{P}(s)$, l'image $\varphi(f)$ est $\mathrm{P}(\mathrm{T}) = \varphi_{\mathrm{T}}(f)$. Par conséquent, φ est uniquement déterminé sur A. Comme un homomorphisme d'algèbres de Banach est continu par définition, et que A est dense dans $\mathrm{C}(\mathrm{K})$, il en résulte que φ, s'il existe, est uniquement défini sur $\mathrm{C}(\mathrm{K})$: si (P_n) tend uniformément vers f sur K, on aura $\varphi(f) = \lim_n \varphi(\mathrm{P}_n) = \lim_n \mathrm{P}_n(\mathrm{T})$.

Montrons maintenant l'existence, en commençant par la définition d'un homomorphisme ψ sur A : pour toute $f \in \mathrm{A}$ l'élément $\psi(f)$ peut être défini de façon unique puisque si $f = \mathrm{P}_1 = \mathrm{P}_2$ sur K,

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

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

L'ensemble A est dense dans $\mathrm{C}(\mathrm{K})$, et on a un homomorphisme isométrique ψ de A dans $\mathcal{L}(\mathrm{H})$; d'après le lemme 1.4.1, il existe un prolongement unique φ_{T} de ψ en application linéaire continue de $\mathrm{C}(\mathrm{K})$ dans $\mathcal{L}(\mathrm{H})$. Posons $f(\mathrm{T}) = \varphi_{\mathrm{T}}(f)$ pour toute $f \in \mathrm{C}(\mathrm{K})$. Pour toute suite (P_n) de polynômes qui converge uniformément sur K vers la fonction f, la suite $(\mathrm{P}_n(\mathrm{T}))$ tend en norme dans $\mathcal{L}(\mathrm{H})$ vers $f(\mathrm{T})$, puisque φ_{T} est continu. Il en résulte par continuité de la norme que

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

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

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

Si ST = TS, on en déduit que $\mathrm{SP}_n(\mathrm{T}) = \mathrm{P}_n(\mathrm{T})\mathrm{S}$ pour tout n, donc $\mathrm{S}f(\mathrm{T}) = f(\mathrm{T})\mathrm{S}$ par continuité du produit par S, à droite et à gauche. Ainsi $f(\mathrm{T})$ commute avec tout opérateur borné S qui commute avec T.

Posons $\varphi_1(f) = \overline{f}(\mathrm{T})^*$ pour toute $f \in \mathrm{C(K)}$. On vérifie que φ_1 est un homomorphisme d'algèbres de Banach unitaires de C(K) dans $\mathcal{L}(\mathrm{H})$, et $\varphi_1(i_{\mathrm{K}}) = i_{\mathrm{K}}(\mathrm{T})$ (parce que $\mathrm{K} \subset \mathbb{R}$, on a $\overline{i_{\mathrm{K}}} = i_{\mathrm{K}}$) donc $\varphi_1(i_{\mathrm{K}}) = \mathrm{T}^* = \mathrm{T}$ parce que T est hermitien. D'après l'unicité, on déduit $\varphi_1 = \varphi_{\mathrm{T}}$, ce qui signifie que $\overline{f}(\mathrm{T}) = f(\mathrm{T})^*$ pour toute $f \in \mathrm{C(K)}$. Il en résulte que $f(\mathrm{T})^* f(\mathrm{T}) = (\overline{f} f)(\mathrm{T}) = (f\overline{f})(\mathrm{T}) = f(\mathrm{T}) f(\mathrm{T})^*$ donc $f(\mathrm{T})$ est normal.

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

//

Corollaire 9.2.5. *Soient* H *un espace de Hilbert complexe,* $\mathrm{T} \in \mathcal{L}(\mathrm{H})$ *hermitien et* f *une fonction continue sur* $\mathrm{K} = \mathrm{Sp(T)}$ *; si* f *est réelle sur* K*, alors* $f(\mathrm{T})$ *est hermitien ; si* f *est réelle et* **positive** *sur* K*, alors* $f(\mathrm{T})$ *est hermitien positif. Si* $|f| = 1$ *sur* K*,* $f(\mathrm{T})$ *est unitaire.*

Démonstration. On a déjà vu le premier point : quand f est réelle, on peut écrire $f(\mathrm{T})^* = \overline{f}(\mathrm{T}) = f(\mathrm{T})$. Si de plus $f \geq 0$ sur K, on peut considérer $g(s) = \sqrt{f(s)}$ qui est une fonction réelle continue sur K. Alors $g(\mathrm{T})$ est hermitien et $f(\mathrm{T}) = (g(\mathrm{T}))^2$ est hermitien positif. Pour finir, supposons que $|f| = 1$ sur K et posons $\mathrm{U} = f(\mathrm{T})$; on a $\mathrm{U}^*\mathrm{U} = \overline{f}(\mathrm{T}) f(\mathrm{T}) = (\overline{f} f)(\mathrm{T}) = \varphi_{\mathrm{T}}(\mathbf{1}) = \mathrm{Id}_{\mathrm{H}}$, et le même calcul donne $\mathrm{UU}^* = \mathrm{Id}_{\mathrm{H}}$, donc U est unitaire.

//

Exemples 9.2.2.

1. Supposons que T soit diagonal dans une base orthonormée, avec coefficients diagonaux (λ_n) réels ; l'opérateur T est alors hermitien. 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 polynomial.

2. Supposons que T soit l'opérateur $\mathrm{M}_\varphi : \mathrm{L}_2(0,1) \to \mathrm{L}_2(0,1)$ de multiplication par une fonction φ réelle continue. On voit que pour tout polynôme P l'opérateur $\mathrm{P}(\mathrm{M}_\varphi)$ est l'opérateur de multiplication par la fonction $s \in [0,1] \to \mathrm{P}(\varphi(t))$, donc à la limite $f(\mathrm{M}_\varphi)$ est l'opérateur de multiplication par $s \to f(\varphi(s))$, c'est à dire que $f(\mathrm{M}_\varphi) = \mathrm{M}_{f \circ \varphi}$. On peut aussi raisonner en disant que $f \to \mathrm{M}_{f \circ \varphi}$ est bien l'unique homomorphisme décrit dans le théorème 4.

Le cas hermitien sur un espace réel. Complexification

Soit H un espace de Hilbert réel, dont le produit scalaire sera noté $x \,.\, y$ pour éviter les confusions avec le produit scalaire dans le complexifié ; le complexifié de H est l'espace $\mathrm{H}_{\mathbb{C}} = \mathrm{H} + i\mathrm{H}$ de tous les vecteurs $z = x + iy$ où $x, y \in \mathrm{H}$. Si $\lambda = a + ib \in \mathbb{C}$, on pose $\lambda z = (ax - by) + i(by + ax)$.

On définit le produit scalaire (complexe) sur $\mathrm{H}_{\mathbb{C}}$ en posant

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

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

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

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

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

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

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

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

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

Corollaire 9.2.6. *Soient* H *un espace de Hilbert réel et* $\mathrm{T} \in \mathcal{L}(\mathrm{H})$ *hermitien ; désignons par* K *le spectre de* T. *Il existe un et un seul homomorphisme d'algèbres de Banach unitaires réelles* $\varphi_{\mathrm{T}} : \mathrm{C}_{\mathbb{R}}(\mathrm{K}) \to \mathcal{L}(\mathrm{H})$ *tel que* $\varphi_{\mathrm{T}}(i_{\mathrm{K}}) = \mathrm{T}$.

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

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

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

9.3. Application aux hermitiens positifs. La racine carrée

Lemme 9.3.1. *Soient* H *un espace de Hilbert* **complexe** *et* $\mathrm{T} \in \mathcal{L}(\mathrm{H})$ *; si pour tout* $x \in \mathrm{H}$, *le scalaire* $\langle \mathrm{T}(x), x\rangle$ *est réel, il en résulte que* T *est hermitien.*

Démonstration. Supposons que $\langle \mathrm{T}(x), x\rangle$ soit réel pour tout $x \in \mathrm{H}$. L'application $(x, y) \to \langle \mathrm{T}(x), y\rangle$ est sesquilinéaire. Par le corollaire 2.1.2, on a $\langle \mathrm{T}(y), x\rangle = \overline{\langle \mathrm{T}(x), y\rangle}$, pour tous $x, y \in \mathrm{H}$, donc T est hermitien.

//

Théorème 9.3.2. *Soient* H *un espace de Hilbert (réel ou complexe) et* $\mathrm{T} \in \mathcal{L}(\mathrm{H})$ *; les conditions suivantes sont équivalentes :*

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

Démonstration. Supposons (ii) vérifiée ; alors l'opérateur $\mathrm{T} = \mathrm{S}^*\mathrm{S}$ est hermitien et on a $\langle \mathrm{S}^*\mathrm{S}(x), x\rangle = \langle \mathrm{S}(x), \mathrm{S}(x)\rangle \geq 0$ pour tout $x \in \mathrm{H}$, donc (ii) $\Rightarrow$ (i). L'implication (iii) $\Rightarrow$ (ii) est évidente. Supposons ensuite que T soit hermitien et que son spectre $\mathrm{K} = \mathrm{Sp(T)}$ soit contenu dans $[0, +\infty[$; notons $f \in \mathrm{C(K)}$ l'application $t \to \sqrt{t}$; par le théorème 2.4 ou le corollaire 2.6, on a $f(\mathrm{T}) = f(\mathrm{T})^*$; de plus $f^2 = i_\mathrm{K}$, donc $f(\mathrm{T})^2 = \mathrm{T}$, donc ($iv$) $\Rightarrow$ (iii). Enfin, par la proposition 7.3.11, on sait que (i) implique que $\mathrm{Sp(T)} \subset [0, +\infty[$.

//

Un élément hermitien de $\mathcal{L}(\mathrm{H})$ satisfaisant aux conditions équivalentes du théorème 2 est appelé *positif* (définition 6.1.3). On note $\mathcal{L}(\mathrm{H})_+$ l'ensemble des éléments positifs de $\mathcal{L}(\mathrm{H})$. Pour $\mathrm{T} \in \mathcal{L}(\mathrm{H})_+$ et $\alpha > 0$, on pose $\mathrm{T}^\alpha = f(\mathrm{T})$, où $f \in \mathrm{C}_\mathbb{R}(\mathrm{Sp(T)})$ est l'application $t \to t^\alpha$. Pour $\alpha, \beta > 0$ on a $\mathrm{T}^{\alpha+\beta} = \mathrm{T}^\alpha\mathrm{T}^\beta$ et, par la dernière partie du théorème 2.4 (ou du corollaire 2.6), on a $(\mathrm{T}^\alpha)^\beta = \mathrm{T}^{\alpha\beta}$.

Proposition 9.3.3. *Pour* $\mathrm{T} \in \mathcal{L}(\mathrm{H})_+$, *il existe un et seul* $\mathrm{S} \in \mathcal{L}(\mathrm{H})_+$ *tel que* $\mathrm{S}^2 = \mathrm{T}$.

Démonstration. On a déjà vu ci-dessus l'existence d'une racine hermitienne positive, passons maintenant à la démonstration de l'unicité. Soit S un opérateur hermitien positif tel que $\mathrm{S}^2 = \mathrm{T}$; considérons le spectre $\mathrm{K} = \mathrm{Sp(S)} \subset [0, +\infty[$, et considérons sur K la fonction $f : s \to s^2$, puis sur $\mathrm{L} = f(\mathrm{K}) \subset [0, +\infty[$ la fonction $g(t) = \sqrt{t}$. Du fait que $\mathrm{K} \subset [0, +\infty[$, on vérifie que $g(f(s)) = \sqrt{s^2} = s$ pour tout $s \in \mathrm{K}$, donc le résultat de composition nous donne, puisque $g \circ f = i_\mathrm{K}$

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

//

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

Décomposition polaire

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

Proposition 9.3.4. *Soient* E *et* F *deux espaces de Hilbert et* $\mathrm{T} \in \mathcal{L}(\mathrm{E}, \mathrm{F})$ *; il existe un et un seul* $u \in \mathcal{L}(\mathrm{E}, \mathrm{F})$, *nul sur* $\ker(\mathrm{T})$ *tel que* $\mathrm{T} = u\,|\mathrm{T}|$.

Démonstration. Pour $x \in \mathrm{E}$ on a

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

en particulier, $\ker(\mathrm{T}) = \ker(|\mathrm{T}|)$; par la proposition 6.1.6, l'adhérence G de l'image de $|\mathrm{T}|$ est l'orthogonal de $\ker(\mathrm{T})$, et on a $\mathrm{E} = \mathrm{G} \oplus \ker(\mathrm{T})$, somme directe orthogonale. On va expliquer la construction de u_0, restrition 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 E. Puisque E est complet, cette isométrie se prolonge en isométrie $\overline{u}_0$ de l'adhérence G, à valeurs dans E. Posons pour finir, si $x = y + z$, avec $y \in \mathrm{G}$ et $z \in \ker(\mathrm{T})$

$$u(x) = \overline{u}_0(y)$$

c'est à dire $u = \overline{u}_0 \circ \mathrm{P}_\mathrm{G}$. On vérifie que $u \circ |\mathrm{T}| = \mathrm{T}$, u nulle sur $\ker(\mathrm{T})$ et que $\|u\| \le 1$. //

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

Remarque 9.3.1. Si T est injectif, u est isométrique. Si T est injectif à image dense, u est unitaire.

En effet, G est égal à E 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.

9.4. Le cas général : opérateurs normaux

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

Dans le cas général d'un opérateur normal, on demandera explicitement que l'homomorphisme φ_T envoie la fonction $\overline{i_\mathrm{K}}$ sur T^*. Il faut généraliser nos polynômes : si la fonction i_K est envoyée sur T et la fonction $\overline{i_\mathrm{K}}$ sur T^*, alors l'image de $i_\mathrm{K}\overline{i_\mathrm{K}}$ doit être TT^* ; dans le cas hermitien ou unitaire, la fonction $i_\mathrm{K}\overline{i_\mathrm{K}}$ s'exprime à partir d'un polynôme en i_K (i_K^2 dans le cas hermitien et 1 dans le cas unitaire) ; ceci n'est plus vrai maintenant, et la fonction $i_\mathrm{K}\overline{i_\mathrm{K}}$ est une nouvelle fonction qui doit être gardée dans notre algèbre de "polynômes" ; bien sûr le problème ne s'arrête pas là, et nous devons considérer $i_\mathrm{K}^p\,\overline{i_\mathrm{K}}^q$

pour tous entiers $p, q \geq 0$. Nous allons donc considérer l'algèbre $\mathbb{C}[X, Y]$ des polynômes en deux variables, puis prendre l'ensemble des fonctions sur $K = \mathrm{Sp}(T)$ obtenues en remplaçant X par i_K et Y par $\overline{i_K}$. Notre algèbre de base A qui remplacera l'algèbre des polynômes sera l'algèbre de toutes les fonctions f sur K de la forme

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

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

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

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

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

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

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

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

Corollaire 9.4.2. *Soient* H *un espace de Hilbert complexe,* $T \in \mathcal{L}(H)$ *un opérateur normal et* f *une fonction continue sur* $\mathrm{Sp}(T)$ *; alors, si* $f(\mathrm{Sp}(T)) \subset \mathbb{R}$, $f(T)$ *est hermitien ; si de plus* $f(\mathrm{Sp}(T)) \subset \mathbb{R}_+$, $f(T)$ *est hermitien positif ; si* $f(\mathrm{Sp}(T)) \subset \mathbb{T}$, *alors* $f(T)$ *est unitaire.*

Démonstration. Si $f = \overline{f}$ (sur Sp(T)) alors $f(T) = \overline{f}(T) = f(T)^*$; si f est réelle ≥ 0 sur Sp(T), on peut introduire $g = \sqrt{f}$, et écrire $f(T)$ comme le carré d'un hermitien ; si $f(\mathrm{Sp}(T)) \subset \mathbb{T}$, alors $f\overline{f} = 1$, donc $f(T)^* f(T) = f(T) f(T)^* = (f\overline{f})(T) = \mathrm{Id}_H$, donc l'opérateur $f(T)$ est unitaire.

//

Exemples 9.4.1.

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

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

2. Pour tout $s \in \mathbb{R}$ considérons la fonction f_s définie sur $\mathbb{R}$ par $f_s(t) = e^{ist}$. Si T est hermitien, on peut considérer pour tout s l'opérateur $U_s = f_s(T) = e^{isT}$. 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 (dernière page du *poly long*, 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 $\mathrm{P_F}$ commute avec T et T^*, donc avec tout opérateur $f(\mathrm{T})$. Si S désigne la restriction de T à F, alors S est un opérateur normal sur F, et $\mathrm{Sp(S)} \subset \mathrm{Sp(T)}$; pour toute fonction continue f sur le spectre de T l'opérateur $f(\mathrm{S})$ est la restriction de $f(\mathrm{T})$ à F.

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

10. Décomposition spectrale des opérateurs normaux

Dans ce chapitre, tous les espaces considérés sont complexes.

10.1. Opérateurs unitairement équivalents

Définition 10.1.1. Soient H_1 et H_2 deux espaces de Hilbert, et $\mathrm{T}_1 \in \mathcal{L}(\mathrm{H}_1)$, $\mathrm{T}_2 \in \mathcal{L}(\mathrm{H}_2)$; on dit que T_1 et T_2 sont *unitairement équivalents* s'il existe un opérateur unitaire $\mathrm{U} : \mathrm{H}_1 \to \mathrm{H}_2$ tel que $\mathrm{T}_1 = \mathrm{U}^* \circ \mathrm{T}_2 \circ \mathrm{U}$.

Soient H_1 et H_2 deux espaces de Hilbert, et $\mathrm{U} : \mathrm{H}_1 \to \mathrm{H}_2$ un opérateur unitaire ; à tout opérateur $\mathrm{S}_2 \in \mathcal{L}(\mathrm{H}_2)$ associons l'opérateur $\mathrm{S}_1 = \mathrm{U}^*\mathrm{S}_2\mathrm{U} \in \mathcal{L}(\mathrm{H}_1)$. On vérifie que $\mathrm{S}_2 \to \mathrm{S}_1$ est un homomorphisme ψ d'algèbres de Banach unitaires de $\mathcal{L}(\mathrm{H}_2)$ dans $\mathcal{L}(\mathrm{H}_1)$, et de plus $\mathrm{Sp}(\mathrm{S}_2) = \mathrm{Sp}(\mathrm{S}_1)$ (parce que $\mathrm{S}_1 - \lambda \,\mathrm{Id}_{\mathrm{H}_1} = \mathrm{U}^*(\mathrm{S}_2 - \lambda \,\mathrm{Id}_{\mathrm{H}_2})\mathrm{U}$ pour tout $\lambda \in \mathbb{C}$). Si $\mathrm{T}_2 \in \mathcal{L}(\mathrm{H}_2)$ est normal, si on pose $\mathrm{K} = \mathrm{Sp}(\mathrm{T}_2) = \mathrm{Sp}(\mathrm{T}_1)$, et si on considère $\varphi = \psi \circ \varphi_{\mathrm{T}_2}$, on obtient un homomorphisme φ d'algèbres de Banach unitaires complexes de $\mathrm{C}(\mathrm{K})$ dans $\mathcal{L}(\mathrm{H}_1)$ tel que $\varphi(i_{\mathrm{K}}) = \mathrm{T}_1$. Il en résulte que $\varphi = \varphi_{\mathrm{T}_1}$ d'après l'unicité dans le théorème 9.4.1 ; pour toute fonction continue f sur K, on a donc la relation $f(\mathrm{T}_1) = \psi(f(\mathrm{T}_2))$, c'est à dire

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

Exemple 10.1.2. Le shift bilatéral sur $\ell_2(\mathbb{Z})$. A toute suite $x = (x_n)_{n\in\mathbb{Z}}$ dans $\ell_2(\mathbb{Z})$ on associe $\mathrm{S}(x) \in \ell_2(\mathbb{Z})$ définie par $\mathrm{S}(x)_n = x_{n-1}$ (décalage d'un cran vers la droite). On a vu que S est unitaire, que son inverse S^* est le décalage à gauche, et le spectre de S est le cercle unité entier. On va montrer un autre *modèle* pour l'opérateur S, qui rend son calcul fonctionnel facile à comprendre.

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

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

On verra que tout opérateur normal est unitairement équivalent à un modèle canonique : la multiplication par une fonction mesurable bornée sur un espace $\mathrm{L}_2(\Omega, \mu)$, pour lequel on a dit que le calcul fonctionnel était simple.

10.2. Opérateurs de multiplication et spectre

Soient $(\Omega, \mathcal{A}, \mu)$ un espace mesuré et $f \in \mathrm{L}_\infty(\Omega, \mathcal{A}, \mu)$; notons M_f l'application qui à $g \in \mathrm{L}_2$ associe la fonction fg. On voit que $\mathrm{M}_f \in \mathcal{L}(\mathrm{L}_2(\Omega, \mu))$ et $\|\mathrm{M}_f\| \le \|f\|_\infty$. Si f_1 et f_2 sont deux fonctions mesurables bornées, il est clair que $\mathrm{M}_{f_1}\,\mathrm{M}_{f_2} = \mathrm{M}_{f_1 f_2}$. Pour $g_1, g_2 \in \mathrm{L}_2(\Omega, \mu)$, on a

$$\langle f g_1, g_2 \rangle = \int_\Omega f(s) g_1(s) \overline{g_2(s)}\, d\mu(s) = \langle g_1, \overline{f} g_2 \rangle,$$

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

Proposition 10.2.1. *Le spectre de* M_f *est l'ensemble des* $\lambda \in \mathbb{C}$ *tels que, pour tout* $\varepsilon > 0$ *l'ensemble* $\{s \in \Omega : |f(s) - \lambda| < \varepsilon\}$ *ne soit pas* μ*-négligeable.*

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

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

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

//

10.3. Théorème de représentation. Décomposition spectrale

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

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

Proposition 10.3.1. *Soient* H *un espace de Hilbert,* $T \in \mathcal{L}(H)$ *un opérateur normal* monogène *et* $K = \mathrm{Sp}(T)$ *; il existe une probabilité* μ *sur* K *telle que* T *soit unitairement équivalent à l'opérateur* $M_{i_K} \in \mathcal{L}(L_2(K, \mu))$ *de multiplication par la fonction* i_K.

Démonstration. Soit x_0 un vecteur tel que $F_{x_0} = H$; on peut choisir $\|x_0\| = 1$ si on veut. Considérons la forme linéaire $\ell : f \to \langle f(T)(x_0), x_0 \rangle$ sur C(K) ; si $f \in C(K)$ est réelle positive, on sait par le calcul fonctionnel (corollaire 9.4.2) que $f(T)$ est hermitien positif, donc $\langle f(T)(x_0), x_0 \rangle \geq 0$. On voit ainsi que ℓ est une forme linéaire positive sur C(K) ; donc il existe une unique mesure positive μ sur K telle que, pour toute fonction $f \in C(K)$, on ait $\ell(f) = \int_K f(t)\, d\mu(t)$; on a $\mu(K) = \int 1\, d\mu = \langle 1(T)(x_0), x_0 \rangle = \|x_0\|^2 = 1$, donc μ est une probabilité. Si y est un vecteur de F_{x_0} de la forme $y = f(T)(x_0)$, avec $f \in C(K)$, on a

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

la relation précédente montre que y ne dépend que de la classe $\widehat{f}$ de f dans $L_2(K, \mu)$, et que l'application $u_0 : \widehat{f} \to f(T)(x_0)$, définie sur l'image Y de C(K) dans $L_2(K, \mu)$, est isométrique, de Y muni de la norme de $L_2(K, \mu)$ vers la norme de H ; par ailleurs, le sous-espace Y est dense dans $L_2(K, \mu)$ par un résultat général d'intégration. On peut donc prolonger u_0 en une isométrie u de $L_2(K, \mu)$ dans H. Pour tous $k, \ell \geq 0$, le vecteur $y = T^k (T^*)^\ell (x_0)$ est dans l'image de u, puisqu'il provient de la fonction continue $i_K^k (\overline{i_K})^\ell$; l'image de u est donc dense puisque T est monogène, et u est unitaire de $L_2(K, \mu)$ dans H. Si $y = u(f)$ avec f continue, on a $y = f(T)(x_0)$, donc

$$T(u(f)) = T f(T)(x_0) = (i_K f)(T)(x_0) = u(M_{i_K}(f)).$$

Cette relation $T \circ u = u \circ M_{i_K}$, vraie sur le sous-espace dense Y de $L_2(K, \mu)$, se prolonge à $L_2(K, \mu)$ tout entier et dit que T et la multiplication par i_K sont conjugués par l'opérateur unitaire u, donc unitairement équivalents.

//

Si on donne un opérateur normal $T \in \mathcal{L}(H)$ et un vecteur $x \in H$ non nul, on peut considérer la restriction S de T au sous-espace F_x ; puisque F_x est aussi stable par T^*, on voit facilement que $S^* \in \mathcal{L}(F_x)$ est la restriction de T^* à F_x, et que S est normal. De plus, S est évidemment monogène, en tant qu'opérateur de l'espace de Hilbert F_x. Le résultat général qui suit consiste simplement à décomposer H en somme directe orthogonale de tels morceaux, et à exprimer le recollement des morceaux.

Théorème 10.3.3. *Soient* H *un espace de Hilbert séparable et* $T \in \mathcal{L}(H)$ *un opérateur normal ; il existe un espace mesuré* (X, μ)*, une fonction* $f \in L_\infty(X, \mu)$ *et un isomorphisme unitaire* $u : L_2(X, \mu) \to H$ *tels que* $T = u\, M_f\, u^*$.

Soient H un espace de Hilbert séparable et $T \in \mathcal{L}(H)$ un opérateur normal ; écrivons $T = u\, M_f\, u^*$. L'application $g \to u\, M_{g \circ f}\, u^*$ est un morphisme continu d'algèbres de C(Sp(T)) dans $\mathcal{L}(H)$; par l'unicité dans le théorème 9.4.1, pour toute fonction continue $g \in C(\mathrm{Sp}(T))$, on a $g(T) = u\, M_{g \circ f}\, u^*$.

Théorème 10.3.4. *Soient* H *un espace de Hilbert séparable et* $T \in \mathcal{L}(H)$ *un opérateur normal ; notons* $K = \mathrm{Sp}(T)$ *et* $\mathcal{L}_\infty(K, \mathcal{B})$ *l'espace vectoriel des fonctions boréliennes bornées sur* K*. Il existe un unique homomorphisme d'algèbres unitaires* $\Phi : \mathcal{L}_\infty(K, \mathcal{B}) \to \mathcal{L}(H)$ *satisfaisant* $\Phi(i_K) = T$, $\Phi(\overline{i_K}) = T^*$*, et tel que pour toute suite bornée* (g_n) *dans*

$\mathcal{L}_\infty(\mathrm{K}, \mathcal{B})$ *convergeant simplement vers* $g \in \mathcal{L}_\infty(\mathrm{K}, \mathcal{B})$ *la suite* $\Phi(g_n)$ *converge fortement vers* $\Phi(g)$*, c'est à dire*

$$\forall x \in \mathrm{H}, \quad \|\Phi(g_n)(x) - \Phi(g)(x)\|_\mathrm{H} \to 0.$$

Si $g \in \mathrm{C}(\mathrm{K})$ *on a* $\Phi(g) = g(\mathrm{T})$.

Démonstration. Montrons l'existence ; soient (X, μ) un espace mesuré, f une fonction de $\mathrm{L}_\infty(\mathrm{X}, \mu)$ et soit $u : \mathrm{L}_2(\mathrm{X}, \mu) \to \mathrm{H}$ un isomorphisme tels que $\mathrm{T} = u\,\mathrm{M}_f\,u^*$ (théorème 3) ; notons $\Phi : \mathcal{L}_\infty(\mathrm{K}, \mathcal{B}) \to \mathcal{L}(\mathrm{H})$ l'application $g \to u\,\mathrm{M}_{g\circ f}\,u^*$; l'application Φ est clairement un homomorphisme d'algèbres unitaires satisfaisant $\Phi(i_\mathrm{K}) = \mathrm{T}$ et $\Phi(\overline{i_\mathrm{K}}) = \mathrm{T}^*$. Soit $g_n \in \mathcal{L}_\infty(\mathrm{K}, \mathcal{B})$ une suite bornée par un réel $\mathrm{M} \geq 0$ et convergeant simplement vers $g \in \mathcal{L}_\infty(\mathrm{K}, \mathcal{B})$; alors, pour tout $x \in \mathrm{H}$, la suite $(g_n \circ f)u^*(x)$ est une suite dans $\mathrm{L}_2(\mathrm{X}, \mu)$ dominée par $\mathrm{M}u^*(x)$ et convergeant partout vers $(g\circ f)u^*(x)$. Par le théorème de convergence dominée, la suite $(g_n \circ f)u^*(x)$ converge vers $(g\circ f)u^*(x)$ dans L_2, donc $\Phi(g_n)(x) = u(g_n \circ f)u^*(x)$ converge (en norme) vers $u(g\circ f)u^*(x) = \Phi(g)(x)$. Remarquons que si $g \in \mathrm{C}(\mathrm{K})$ on a $\Phi(g) = g(\mathrm{T})$, d'après la relation (E) du début de ce chapitre.

Montrons l'unicité. Soient Φ_1 et Φ_2 deux applications vérifiant les conditions ci-dessus ; posons $\mathrm{A} = \{g \in \mathcal{L}_\infty(\mathrm{K}, \mathcal{B}) : \Phi_1(g) = \Phi_2(g)\}$. On doit montrer que l'ensemble A est égal à $\mathcal{L}_\infty(\mathrm{K}, \mathcal{B})$. Par hypothèse, A est un sous-espace vectoriel de $\mathcal{L}_\infty(\mathrm{K}, \mathcal{B})$ et pour toute suite bornée $(g_n) \subset \mathrm{A}$ convergeant simplement vers $g \in \mathcal{L}_\infty(\mathrm{K}, \mathcal{B})$, on a $g \in \mathrm{A}$; de plus, par l'unicité dans le théorème 9.4.1, A contient toutes les fonctions continues sur K. Toute fonction borélienne bornée est limite uniforme de fonctions boréliennes prenant un nombre fini de valeurs ; celles-ci sont combinaisons linéaires de fonctions caractéristiques de sous-ensembles boréliens. Il suffit donc de prouver que A contient toutes les fonctions caractéristiques. Notons $\mathcal{C}$ l'ensemble des boréliens de K dont la fonction caractéristique est dans A. Comme A est un sous-anneau, $\mathcal{C}$ est stable par intersection finie ; comme A est stable par limite de suites bornées, $\mathcal{C}$ est stable par intersection dénombrable. De plus, si $g \in \mathrm{A}$, on a aussi $1 - g \in \mathrm{A}$, donc $\mathcal{C}$ est stable par passage au complémentaire. Donc $\mathcal{C}$ est une tribu. Pour savoir que $\mathcal{C}$ contient tous les boréliens, il suffit de montrer que $\mathcal{C}$ contient tous les fermés. Soit F un fermé non vide de K ; notons $h_0 : \mathrm{K} \to \mathbb{R}$ la fonction qui à $x \in \mathrm{K}$ associe sa distance à F. Posons $g = \sup(1 - h_0, 0)$ et $g_n = g^n$. Pour tout n on a $g_n \in \mathrm{C}(\mathrm{K}) \subset \mathrm{A}$; de plus la suite (g_n) est bornée et converge simplement vers la fonction caractéristique de F. On a montré que $\mathrm{K} \in \mathcal{C}$.

//

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

Exemple 10.3.5. Soit T un opérateur normal sur un espace de Hilbert complexe (resp : un opérateur hermitien sur un Hilbert réel ou complexe) ; si $f = \mathbf{1}_\mathrm{A}$ est l'indicatrice d'un borélien de $\mathbb{C}$ (resp : de $\mathbb{R}$) contenu dans le spectre de T, et si $\mathrm{P} = f(\mathrm{T})$, on aura $\mathrm{P}^* = \mathrm{P}$ parce que f est réelle, et $\mathrm{P}^2 = \mathrm{P}$ parce que $f^2 = f$. L'opérateur P est donc un projecteur orthogonal. On dit que P est un *projecteur spectral.* Puisque P est obtenu par calcul fonctionnel, il commute avec T et T^*, donc l'image $\mathrm{P}(\mathrm{H})$ est stable par T et T^*, ce qui permet de considérer la restriction S de T à $\mathrm{F} = \mathrm{P}(\mathrm{H})$, qui est un opérateur normal (resp : hermitien) sur le sous-espace F. On peut voir que si F n'est pas nul, le spectre de S est contenu dans l'adhérence de l'ensemble A.

11. Opérateurs autoadjoints non bornés

11.1. Opérateurs non bornés

Préliminaires algébriques

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

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

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

> soit $\mathrm{G} \subset \mathrm{X} \times \mathrm{Y}$ 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 $\mathrm{T}(p_1(x)) = p_2(x)$ pour tout $x \in \mathrm{G}$. 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.

Désormais on dira *opérateur* au lieu d'application linéaire partiellement définie. On appelle *noyau* de T le sous-espace $\ker(\mathrm{T}) = \{x \in \mathrm{dom}(\mathrm{T}) : \mathrm{T}(x) = 0\}$ de X et *image* de T le sous-espace $\mathrm{im}(\mathrm{T}) = \mathrm{T}(\mathrm{dom}(\mathrm{T}))$ de Y. On appelle *extension* d'un opérateur T tout opérateur S tel que $\mathrm{Gr}(\mathrm{T}) \subset \mathrm{Gr}(\mathrm{S})$. On écrit alors $\mathrm{T} \subset \mathrm{S}$.

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 R, S et T sont des opérateurs de X dans Y, on a clairement $\mathrm{R} + \mathrm{S} = \mathrm{S} + \mathrm{R}$ et $(\mathrm{R} + \mathrm{S}) + \mathrm{T} = \mathrm{R} + (\mathrm{S} + \mathrm{T})$.

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 $x \in \mathrm{X}$; si T est un opérateur 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 $(\mathrm{ST})(x) = \mathrm{S}(\mathrm{T}(x))$ pour tout $x \in \mathrm{dom}(\mathrm{ST})$. Si R est un opérateur de Z dans un quatrième espace vectoriel H, on a $(\mathrm{RS})\mathrm{T} = \mathrm{R}(\mathrm{ST})$. De plus, si T est un opérateur de X dans Y et si R et S sont des opérateurs de Y dans Z, on a $(\mathrm{R}+\mathrm{S})\mathrm{T} = \mathrm{RT} + \mathrm{ST}$.

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 $\mathrm{L}_1(\mathbb{R})$ par la formule intégrale usuelle ; on appelle aussi transformation de Fourier son extension par continuité à l'espace $\mathrm{L}_2(\mathbb{R})$). 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 $\mathrm{D}_1 \cap \mathrm{D}_2$. On ne cherche surtout pas à aller plus loin.

Exemples 11.1.1.

A. On prend $\mathrm{X} = \mathrm{Y} = \mathrm{L}_2(\mathbb{R})$, $\mathrm{dom}(\mathrm{T})$ est l'espace des fonctions C^1 à support compact et on pose $\mathrm{T}(f) = f'$ pour $f \in \mathrm{dom}(\mathrm{T})$.

B. Cet exemple se décline en trois variantes.

– B1 : on prend $\mathrm{X} = \mathrm{Y} = \mathrm{L}_2([0,1])$, $\mathrm{dom}(\mathrm{T}_1)$ est l'espace des fonctions C^1 sur $[0,1]$ et $\mathrm{T}_1(f) = f'$ pour $f \in \mathrm{dom}(\mathrm{T}_1)$.

– B2 : on prend $\mathrm{X} = \mathrm{Y} = \mathrm{L}_2([0,1])$, $\mathrm{dom}(\mathrm{T}_2)$ est l'espace des fonctions f qui sont C^1 sur $[0,1]$ et telles que $f(0) = f(1) = 0$, et $\mathrm{T}_2(f) = f'$ pour $f \in \mathrm{dom}(\mathrm{T}_2)$.

– B3 : on prend $\mathrm{X} = \mathrm{Y} = \mathrm{L}_2([0,1])$, $\mathrm{dom}(\mathrm{T}_3)$ est l'espace des fonctions f qui sont C^1 sur $[0,1]$ et telles que $f(0) = f(1)$, et $\mathrm{T}_3(f) = f'$ pour $f \in \mathrm{dom}(\mathrm{T}_3)$.

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

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

Exemple 11.1.2. Considérons $\mathrm{X} = \mathrm{Y} = \mathrm{L}_2(0,1)$; soit V l'opérateur borné de “primitive nulle en 0”, $(\mathrm{V}f)(t) = \int_0^t f(s)\,ds$, et posons $\mathrm{D} = \mathrm{im}(\mathrm{V})$. On a vu que V est injectif (exemples 7.3.2). On peut donc définir l'opérateur $\mathrm{T} = \mathrm{V}^{-1}$ de domaine D en posant pour tout $g \in \mathrm{D}$

$$(\mathrm{T}g = f) \Leftrightarrow (\mathrm{V}f = g).$$

Cet opérateur T est donc injectif lui-aussi.

La topologie revient

Soient E et F deux espaces de Banach ; un opérateur T de E dans F est dit *densément défini* si son domaine $\mathrm{dom}(\mathrm{T})$ est dense dans E.

Définition 11.1.3. Soient E et F deux espaces de Banach ; un opérateur de E dans F est dit *fermé* si son graphe est un sous-espace fermé de $E \times F$. Un opérateur de E dans F est dit *fermable* s'il admet une extension fermée.

Exemple 11.1.4. Reprenons l'opérateur borné V de $L_2([0,1])$ dans lui-même et son inverse non borné V^{-1} de l'exemple 2 ; le domaine de V^{-1} est $\mathrm{im}(V)$; on a vu que $\mathrm{im}(V)$ est dense (exemples 7.3.2), donc V^{-1} est densément défini. Puisque V est continu, son graphe est fermé, donc V^{-1} est fermé puisque son graphe s'obtient à partir de celui de V par l'homéomorphisme $(x, y) \to (y, x)$ de $L_2 \times L_2$ sur lui-même.

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 $\overline{T}$ tel que $\mathrm{Gr}(\overline{T}) = \overline{\mathrm{Gr}(T)}$. En particulier, pour que l'opérateur T soit fermable il faut et il suffit que l'on ait $\overline{\mathrm{Gr}(T)} \cap (\{0\} \times F) = \{(0,0)\}$. On en déduit immédiatement :

Proposition 11.1.1. *Soient* E *et* F *deux espaces de Banach et* T *un opérateur de* E *dans* F *; pour que l'opérateur* T *soit fermable il faut et il suffit que pour toute suite* (x_n) *de* dom(T) *qui converge vers* 0 *dans* E *et telle que* $T(x_n)$ *converge dans* F *vers un vecteur* y*, on ait* $y = 0$.

Exemple 11.1.5. Fermetures des opérateurs T, T_1, T_2 et T_3 de l'exemple 1. Commençons par l'opérateur T de l'exemple **A**, défini sur $L_2(\mathbb{R})$. On va montrer que T est fermable et isoler un candidat pour la fermeture.

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

$$f_n(t) = f_n(a) + \int_a^t f'_n(s)\, ds,$$

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

$$f(t) = f(a) + \int_a^t g(s)\, ds.$$

Il en résulte que f est continue, et qu'il existe une fonction $g \in L_2$ telle que

$$\forall t < u, \quad f(u) = f(t) + \int_t^u g(s)\, ds.$$

On introduit l'ensemble

$$G_{\mathbf{A}} = \{(f, g) \in L_2(\mathbb{R}) \times L_2(\mathbb{R}) : \forall t < u, \quad f(u) = f(t) + \int_t^u g(s)\, ds\}$$

(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 Gr(T) est contenue dans $G_{\mathbf{A}}$; pour savoir que T est fermable, il suffit de voir que $G_{\mathbf{A}}$ est un graphe : c'est clairement un espace vectoriel, et si $(0, g) \in G_{\mathbf{A}}$, on aura $\int_t^u g = 0$ pour tous $t < u$, ce qui signifie que g est orthogonale à toutes les fonctions en escalier, qui sont denses dans $L_2(\mathbb{R})$, donc $g = 0$, ce qu'il fallait démontrer.

Exercice 11.1.6. Montrer que l'adhérence du graphe de T est *égale* à $G_{\mathbf{A}}$.

On appelle $H_1(\mathbb{R})$ (espace de Sobolev) l'espace des fonctions $f \in L_2(\mathbb{R})$ telles qu'il existe $g \in L_2(\mathbb{R})$ telle que $(f, g) \in G_{\mathbf{A}}$. On dit que g est la *dérivée généralisée* de f, et on note simplement $g = f'$. La fermeture de T de l'exemple 1, **A** est donc l'opérateur $\overline{T}$ de $L_2(\mathbb{R})$ dans lui-même dont le domaine est $H_1(\mathbb{R})$ et qui est défini par $\overline{T}(f) = f'$ pour $f \in H_1(\mathbb{R})$.

On définit de même l'espace $H_1([0,1])$ des fonctions $f \in L_2([0,1])$ (en fait f sera continue) pour lesquelles existe une fonction $g \in L_2([0,1])$ telle que $f(t) = f(0) + \int_0^t g(s)\,ds$, pour tout $t \in [0,1]$. Si on se rappelle l'opérateur-exemple V de $L_2([0,1])$ dans lui-même qui associe à chaque $g \in L_2([0,1])$ sa "primitive" nulle en zéro, on voit que $H_1([0,1])$ est égal à $\mathrm{im}(V) + \mathbb{K}1$. On peut vérifier que les fermetures des variantes B1, B2, B3 sont définies sur les domaines

– 1 : $f \in H_1([0,1])$
– 2 : $f \in H_1([0,1])$ et $f(0) = f(1) = 0$
– 3 : $f \in H_1([0,1])$ et $f(0) = f(1)$.

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

Revenons sur la notion de dérivée généralisée. Avec Fubini on montre que si $f \in H^1(\mathbb{R})$ et si φ est C^1 à support compact, on a

$$\text{(D)} \qquad \int_{\mathbb{R}} f\varphi' = -\int_{\mathbb{R}} f'\varphi.$$

En modifiant très légèrement ce qui précède, on obtient la formule d'intégration par parties dans $H^1([a,b])$: si $f, g \in H^1([a,b])$, on a

$$\int_a^b f(t)g'(t)\,dt = \Big[f(t)g(t)\Big]_a^b - \int_a^b f'(t)g(t)\,dt.$$

C'est la propriété (D) précédente qui permet d'étendre la définition de H^1 au cas de plusieurs dimensions. Par exemple, on dit que $f \in H^1(\mathbb{R}^2)$ si $f \in L_2(\mathbb{R}^2)$ et s'il existe deux fonctions $g_1, g_2 \in L_2(\mathbb{R}^2)$ qui seront les *dérivées partielles faibles* de f, ce qui signifie que

$$\int_{\mathbb{R}^2} f(x)\frac{\partial\varphi}{\partial x_j}(x)\,dx = -\int_{\mathbb{R}^2} g_j(x)\varphi(x)\,dx$$

pour $j = 1, 2$ et pour toute fonction φ qui soit C^1 à support compact sur $\mathbb{R}^2$. Les fonctions de cet espace $H^1(\mathbb{R}^2)$ ne sont plus nécessairement continues, ni même bornées sur les compacts de $\mathbb{R}^2$.

Proposition 11.1.2. *L'inverse d'un opérateur injectif fermé est fermé.*

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

//

Exemple 11.1.7. L'opérateur V^{-1} de l'exemple 2 est fermé et densément défini. Les opérateurs fermés et densément définis forment la classe la plus intéressante dans cette théorie.

11.2. Spectre des opérateurs fermés

Définition 11.2.1. Soient T un opérateur d'un espace de Banach complexe E dans lui même et $\lambda \in \mathbb{C}$; on dit que λ est une *valeur régulière* de T si $T - \lambda\, \mathrm{Id}_E$ est une application linéaire bijective de dom(T) sur E et si l'application linéaire réciproque définit une application linéaire *continue* de E dans lui même. On appelle *spectre* de T le complémentaire Sp(T) dans $\mathbb{C}$ de l'ensemble des valeurs régulières de T.

Soit T un opérateur sur un espace de Banach complexe E ; désignons par Ω_T l'ensemble des $\lambda \in \mathbb{C}$ qui sont valeur régulière de T ; pour $\lambda \in \Omega_T$, on pose

$$R_\lambda(T) = (T - \lambda)^{-1} \in \mathcal{L}(E)$$

et on appelle $R_\lambda(T)$ 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 $(T - \lambda\, \mathrm{Id}_E)^{-1}$ est continu donc à graphe fermé ; on en déduit que son inverse $T - \lambda\, \mathrm{Id}_E$ 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 $\mathrm{Sp}(T) = \mathbb{C}$.

Soit T un opérateur fermé d'un espace de Banach E dans lui-même ; remarquons que pour tout $\lambda \in \mathbb{C}$, l'opérateur $T - \lambda\, \mathrm{Id}_E$ est fermé. Si $T - \lambda\, \mathrm{Id}_E$ est bijectif de dom(T) sur E, alors λ est une valeur régulière car $(T - \lambda\, \mathrm{Id}_E)^{-1}$ est fermé (proposition 1.2), donc continu par le théorème du graphe fermé (théorème 4.1.6).

Exemples 11.2.2.

1. Soit μ une mesure sur $\mathbb{C}$, positive et non nulle, donnant une mesure finie à tout compact ; on considère dans $L_2(\mu) = L_2(\mathbb{C}, \mu)$ l'application de multiplication par z, définie sur le domaine

$$D = \{f \in L_2(\mu) : \int_{\mathbb{C}} |z|^2\, |f(z)|^2\, d\mu(z) < +\infty\},$$

ce qui donne un opérateur, en général non borné, qu'on notera M, qui agit sur $f \in D$ par $(Mf)(z) = zf(z)$, et $Mf \in L_2(\mu)$. On peut décrire l'appartenance de f au domaine D en une seule formule,

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

On suppose d'abord que $\lambda \in \mathbb{C}$ est tel que pour tout $\varepsilon > 0$, on ait que $B = B(\lambda, \varepsilon)$ vérifie $\mu(B) > 0$. On peut considérer la fonction $f = \mathbf{1}_B$, qui est dans le domaine D, et qui n'est pas dans la classe nulle de $L_2(\mu)$ puisque $\mu(B) > 0$. On a $|(M - \lambda) f| = |z - \lambda|\, \mathbf{1}_B \le \varepsilon\, \mathbf{1}_B$. Ceci montre que $\|(M - \lambda) f\|_2 \le \varepsilon\, \|f\|_2$; si l'inverse $R_\lambda(M)$ de $M - \lambda\, \mathrm{Id}$ existait, il devrait vérifier $\|R_\lambda(M)\| \ge 1/\varepsilon$, pour tout $\varepsilon > 0$, ce qui est impossible. Il en résulte que $\lambda \in \mathrm{Sp}(T)$.

On suppose inversement que $\lambda \in \mathbb{C}$ est tel qu'il existe $\varepsilon_0 > 0$ tel que $\mu(B(\lambda, \varepsilon_0)) = 0$. Considérons la fonction mesurable bornée g définie sur $\mathbb{C}$ par $g(z) = (z - \lambda)^{-1}$ si $|z - \lambda| \ge$

ε_0 et $g(z) = 0$ sinon. La multiplication M_g est bornée sur $\mathrm{L}_2(\mu)$ puisque g est bornée, et on va voir que $\mathrm{M}_g = \mathrm{R}_\lambda(\mathrm{M})$. Si $f \in \mathrm{dom}(\mathrm{M})$, on voit que $\mathrm{M}_g(\mathrm{M}(f) - \lambda f) = g(z-\lambda) f$ est égale à f en dehors de B, et à 0 dans B ; mais puisque $\mu(\mathrm{B}) = 0$, on a bien $\mathrm{M}_g(\mathrm{M}(f) - \lambda f) = f$ en tant que classe. Inversement, si $h \in \mathrm{L}_2(\mu)$, on vérifie que $\mathrm{M}_g(h) \in \mathrm{dom}(\mathrm{M})$ (en effet,

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

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

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

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

2. Nous allons donc montrer maintenant que le spectre de l'opérateur $\mathrm{T} = \mathrm{V}^{-1}$ de l'exemple 1.2 est vide :

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

Le raisonnement utilisé dans l'exemple précédent montre que

Lemme 11.2.1. *Soient* T *un opérateur injectif fermé d'un espace de Banach* E *dans lui même et* λ *une valeur régulière de* T *non nulle ; alors* λ^{-1} *est une valeur régulière de* T^{-1} *et on a*

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

Proposition 11.2.2. *Le spectre d'un opérateur fermé* T *d'un espace de Banach complexe* E *dans lui même est une partie fermée de* $\mathbb{C}$*, et l'application* $\lambda \to \mathrm{R}_\lambda(\mathrm{T})$ *est continue du complémentaire du spectre dans* $\mathcal{L}(\mathrm{E})$.

> *Démonstration.* Désignons par Ω_T l'ensemble des valeurs régulières pour T, et montrons que cet ensemble est ouvert. Si Ω_T est vide, il est ouvert ; sinon, supposons que $\lambda_0 \in \Omega_\mathrm{T}$, et montrons que les valeurs voisines de λ_0 sont elles aussi régulières et $\lambda \to \mathrm{R}_\lambda(\mathrm{T})$ continue dans ce voisinage. En remplaçant T par $\mathrm{T}_0 = \mathrm{T} - \lambda_0 \, \mathrm{Id}_\mathrm{E}$ on se ramène à $\lambda_0 = 0$. On supposera donc que $\mathrm{T} = \mathrm{T} - 0 \, \mathrm{Id}_\mathrm{E}$ est une bijection de $\mathrm{dom}(\mathrm{T})$ sur E, d'inverse $\mathrm{S} = \mathrm{R}_0(\mathrm{T})$ continu ; on veut alors montrer qu'il existe $\varepsilon > 0$ tel que λ soit valeur régulière de T quand $|\lambda| < \varepsilon$.
>
> Etant donné $y \in \mathrm{E}$ quelconque, on veut résoudre en $x \in \mathrm{dom}(\mathrm{T})$, et avec solution unique, l'équation
>
> $$\mathrm{T}(x) - \lambda x = y.$$

Posons $z = \mathrm{T}(x)$, ce qui équivaut à $x = \mathrm{S}(z)$. L'équation précédente devient alors $z - \lambda \mathrm{S}(z) = y$, ou encore $(\mathrm{S} - \lambda^{-1}\,\mathrm{Id_E})(z) = -\lambda^{-1} y$. Lorsque $|\lambda| < \rho(\mathrm{S})^{-1}$, on sait que $\mathrm{S} - \lambda^{-1}\,\mathrm{Id_E}$ est inversible, donc z est uniquement défini par

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

et puisque $x = \mathrm{S}(z)$ ceci montre que $\mathrm{R}_\lambda(\mathrm{T}) = -\lambda^{-1}\mathrm{S}\,\mathrm{R}_{\lambda^{-1}}(\mathrm{S})$ existe et est borné, pour tout $\lambda \in \mathbb{C}$ tel que $|\lambda| < \rho(\mathrm{T}^{-1})^{-1}$. Si on réécrit

$$\mathrm{R}_\lambda(\mathrm{T}) = -\lambda^{-1}\mathrm{S}\,\mathrm{R}_{\lambda^{-1}}(\mathrm{S}) = -\mathrm{S}(\mathrm{Id_E} - \lambda \mathrm{S})^{-1}$$

on voit que $\lambda \to \mathrm{R}_\lambda(\mathrm{T})$ est continue au voisinage de $\lambda_0 = 0$; mais λ_0 est en fait un point quelconque de Ω_T.

//

Exemple 11.2.4. Opérateur diagonal. Pour toute suite scalaire $(\mu_n)_{n \ge 0}$, on définit un opérateur (en général non borné) sur $\ell_2(\mathbb{N})$ dont le domaine est

$$\mathrm{D} = \{x \in \ell_2 : \sum |\mu_n x_n|^2 < +\infty\}$$

et qui est défini pour $x \in \mathrm{D}$ par $(\mathrm{T}x)_n = \mu_n x_n$. Le spectre de T est l'adhérence dans $\mathbb{C}$ de l'ensemble des valeurs $(\mu_n)_{n \ge 0}$. Comme toute partie fermée non vide F de $\mathbb{C}$ admet une suite dense, on retrouve le fait que pour toute partie fermée non vide F de $\mathbb{C}$, on peut construire un opérateur T d'un espace de Hilbert H dont le spectre $\mathrm{Sp}(\mathrm{T})$ soit égal à F. L'opérateur $\mathrm{T} = \mathrm{V}^{-1}$ de l'exemple 7.2.2 fournit un cas où $\mathrm{Sp}(\mathrm{T}) = \emptyset$.

11.3. Transposés et adjoints

Soient E et F deux espaces de Banach et T un opérateur densément défini de E dans F ; on définit *le transposé* de T, qui est un opérateur de F^* dans E^*, de la façon suivante : le domaine de ${}^t\mathrm{T}$ est l'ensemble des $y^* \in \mathrm{F}^*$ telles que la forme linéaire $x \in \mathrm{dom}(\mathrm{T}) \to y^*(\mathrm{T}(x))$ soit continue (en ayant muni l'espace vectoriel $\mathrm{dom}(\mathrm{T})$ de la norme induite par celle de E). Dans le cas où $y^* \in \mathrm{dom}({}^t\mathrm{T})$, cette forme linéaire continue, définie sur le sous-espace dense $\mathrm{dom}(\mathrm{T}) \subset \mathrm{E}$, se prolonge de façon unique en une forme linéaire $x^* \in \mathrm{E}^*$ continue sur E. On pose alors ${}^t\mathrm{T}(y^*) = x^*$. On a donc

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

pour tous $x \in \mathrm{dom}(\mathrm{T})$ et $y^* \in \mathrm{dom}({}^t\mathrm{T})$.

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

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

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

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

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

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

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

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

Exemple 11.3.1. Donnons un exemple simple d'opérateur autoadjoint. On considère $\mathrm{H} = \mathrm{L}_2(\mathbb{R})$, $\mathrm{D} = \{f \in \mathrm{L}_2(\mathbb{R}) : \int_{\mathbb{R}} x^2 |f(x)|^2 \, dx < +\infty\}$ et on définit $\mathrm{M}f$ pour toute $f \in \mathrm{D}$ par

$$\forall x \in \mathbb{R}, \quad (\mathrm{M}f)(x) = x f(x).$$

On va vérifier que l'opérateur M est autoadjoint. On voit facilement que D est dense dans $\mathrm{L}_2(\mathbb{R})$ (parce que D contient toutes les fonctions de $\mathrm{L}_2(\mathbb{R})$ à support borné). Il est à peu près évident que M est symétrique,

$$\langle \mathrm{M}f, g\rangle = \int_{\mathbb{R}} \bigl(x f(x)\bigr) \overline{g(x)} \, dx = \int_{\mathbb{R}} f(x) \overline{x g(x)} \, dx = \langle f, \mathrm{M}g\rangle.$$

On en déduit $\mathrm{dom}(\mathrm{M}) \subset \mathrm{dom}(\mathrm{M}^*)$. Inversement, supposons que $g \in \mathrm{dom}(\mathrm{M}^*)$, et considérons pour tout $n \ge 0$ la fonction $f_n \in \mathrm{dom}(\mathrm{M})$ définie par $f_n(x) = x \mathbf{1}_{[-n,n]}(x) g(x)$; 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

$$\int_{-n}^{n} x^2 |g(x)|^2 \, dx \le \mathrm{C} \Bigl(\int_{-n}^{n} x^2 |g(x)|^2 \, dx\Bigr)^{1/2}$$

d'où résulte que $\int_{\mathbb{R}} x^2 |g(x)|^2 \, dx \le \mathrm{C}^2 < +\infty$, soit $g \in \mathrm{dom}(\mathrm{T})$. La vérification est finie.

Exercice 11.3.2. Adjoints des fermetures des exemples B1, B2, B3 de 1.1. Posons $\mathrm{E} = \mathrm{L}_2([0,1])$. La fermeture de l'exemple B3 est l'opérateur T_c dont le domaine est

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

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

Avant tout on vérifie que D_c est dense dans E : les fonctions affines par morceaux, nulles en 0 et 1, sont dans D_c, et elles sont denses dans $\mathrm{L}_2(0,1)$; en effet, la fonction indicatrice de tout intervalle $]a, b[$ contenu dans $[0,1]$ peut être obtenue comme limite croissante d'une suite (f_n) de fonctions affines par morceaux, nulles en 0 et 1, et telles que $0 \le f_n \le 1$ pour tout n. Ensuite, les fonctions indicatrices d'intervalles forment une partie totale de $\mathrm{L}_2(0,1)$.

On montre maintenant que S_c est symétrique, c'est à dire que $\langle f_1, \mathrm{S}_c(f_2)\rangle = \langle \mathrm{S}_c(f_1), f_2\rangle$ pour toutes $f_1, f_2 \in \mathrm{D}_c$. On a en effet en utilisant l'intégration par parties dans $\mathrm{H}^1([0,1])$

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

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

On sait donc déjà que $\mathrm{D}_c \subset \mathrm{dom}(\mathrm{S}_c^*)$, et que $\mathrm{S}_c^*(f) = \mathrm{S}_c(f) = if'$ pour $f \in \mathrm{D}_c$. Il reste à voir que $\mathrm{dom}(\mathrm{S}_c^*) \subset \mathrm{D}_c$. Dire que (g,h) est dans le graphe de S_c^* signifie que $g \in \mathrm{E}$ est dans le domaine de S_c^* et que $h = \mathrm{S}_c^*(g) \in \mathrm{E}$ vérifie

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

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

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

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

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

ce qui donne

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

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

Proposition 11.3.1. *Soient* E *et* F *deux espaces de Hilbert et* T *un opérateur densément défini de* E *dans* F *; alors* T^* *est fermé. Pour que* T *soit fermable, il faut et il suffit que* T^* *soit densément défini. Dans ce cas, on a* $\overline{\mathrm{T}} = (\mathrm{T}^*)^*$.

Démonstration. Sur l'espace $\mathrm{E} \times \mathrm{F}$ on introduit le produit scalaire

$$\langle (x,y), (x',y')\rangle = \langle x, x'\rangle + \langle y, y'\rangle$$

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

Supposons T fermable. Pour montrer que T^* est densément défini, on va montrer que $y = 0_\mathrm{F}$ est le seul vecteur de F orthogonal à $\mathrm{dom}(\mathrm{T}^*)$. Dans ce cas le couple $(y, 0_\mathrm{E})$ est orthogonal à $\mathrm{Gr}(\mathrm{T}^*)$, donc il est dans $\mathrm{U}_0^*(\mathrm{Gr}(\mathrm{T}))^{\perp\perp}$; mais puisque T est fermable, $\mathrm{Gr}(\mathrm{T}))^{\perp\perp} = \overline{\mathrm{Gr}(\mathrm{T})}$ est le graphe d'un opérateur $\overline{\mathrm{T}}$, et on obtient $(y, 0_\mathrm{E}) \in \mathrm{U}_0^*(\mathrm{Gr}(\overline{\mathrm{T}}))$, c'est à dire $(0_\mathrm{E}, y) \in \mathrm{Gr}(\overline{\mathrm{T}})$, d'où $y = \overline{\mathrm{T}}(0_\mathrm{E}) = 0_\mathrm{F}$.

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

//

Proposition 11.3.2. *Soit* T *un opérateur densément défini d'un espace de Hilbert* E *dans un espace de Hilbert* F *; alors* $\ker(\mathrm{T}^*) = \mathrm{im}(\mathrm{T})^\perp$.

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

//

Proposition 11.3.3. *Soient* E *et* F *deux espaces de Hilbert et* T *un opérateur densément défini, fermé de* E *dans* F *;*

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

Proposition 11.3.4. *Soient* E *et* F *deux espaces de Hilbert et* T *un opérateur fermé densément défini de* E *dans* F *; l'opérateur* $(\mathrm{Id}_\mathrm{E} + \mathrm{T}^*\mathrm{T})$ *est injectif, son image est égale à* E *et* $(\mathrm{Id}_\mathrm{E} + \mathrm{T}^*\mathrm{T})^{-1}$ *est un élément positif de* $\mathcal{L}(\mathrm{E})$. *L'opérateur* $\mathrm{T}^*\mathrm{T}$ *est autoadjoint et son spectre est contenu dans* $[0, +\infty[$.

11.4. Théorème de représentation. Décomposition spectrale

Exemple 11.4.1. Soit (Ω, μ) un espace mesuré et soit $\mathrm{L}_0(\Omega, \mu)$ l'ensemble des classes d'équivalence de fonctions mesurables complexes pour la relation d'égalité μ-presque partout ; si $f, g \in \mathrm{L}_0(\Omega, \mu)$, on voit facilement que la classe de $\widetilde{f}\,\widetilde{g}$ ne dépend pas des représentants $\widetilde{f}$, $\widetilde{g}$ des deux classes, ce qui permet de parler du produit ponctuel de deux classes. On peut alors définir un opérateur M_f dont le domaine est l'ensemble des $g \in \mathrm{L}_2(\Omega, \mu)$ telles que $fg \in \mathrm{L}_2(\Omega, \mu)$ et tel que $\mathrm{M}_f(g) = fg$ pour tout $g \in \mathrm{dom}(\mathrm{M}_f)$.

Pour tout $g \in \mathrm{L}_2(\Omega, \mu)$ et tout $n > 0$, la fonction $g_n = \big(n/(n + |f|)\big)\, g$ est dans le domaine de M_f ; de plus la suite (g_n) converge partout vers g et est dominée par $|g|$. Il résulte alors du théorème de convergence dominée que le domaine de M_f est dense dans L_2. Soient $\xi, \eta \in \mathrm{L}_2(\Omega, \mu)$ tels que $f\xi \in \mathrm{L}_2(\Omega, \mu)$ et $\overline{f}\eta \in \mathrm{L}_2(\Omega, \mu)$; on a

$$\langle f\xi, \eta\rangle = \int f(t)\xi(t)\overline{\eta(t)}\, d\mu(t) = \langle \xi, \overline{f}\eta\rangle.$$

On en déduit que $\mathrm{M}_{\overline{f}} \subset \mathrm{M}_f^*$. Enfin, soient $\xi, \eta \in \mathrm{L}_2(\Omega, \mu)$; posons $\xi_1 = (\xi + \overline{f}\eta)(1 + |f|^2)^{-1}$ et $\eta_1 = (\eta - f\xi)(1 + |f|^2)^{-1}$; clairement $f\xi_1 \in \mathrm{L}_2(\Omega, \mu)$, $\overline{f}\eta_1 \in \mathrm{L}_2(\Omega, \mu)$ et on voit que $(\xi, \eta) = (\xi_1, f\xi_1) + (-\overline{f}\eta_1, \eta_1)$. On en déduit que

$$\mathrm{L}_2(\Omega, \mu) \times \mathrm{L}_2(\Omega, \mu) = \mathrm{Gr}(\mathrm{M}_f) + \{(-z, y) : (y, z) \in \mathrm{Gr}(\mathrm{M}_{\overline{f}})\}.$$

Comme $\{(-z, y) : (y, z) \in \mathrm{Gr}(\mathrm{M}_{\overline{f}})\} \subset \mathrm{Gr}(\mathrm{M}_f)^\perp$, on en déduit l'égalité de ces deux sous-espaces ; il en résulte que M_f et $\mathrm{M}_{\overline{f}}$ sont fermés et adjoint l'un de l'autre.

Posons $h = (1 + |f|)^{-1}$. Remarquons que $\mathrm{M}_h \in \mathcal{L}(\mathrm{L}_2(\Omega, \mu))$ est injective, que le domaine de M_f est l'image de M_h et que $\mathrm{M}_f\mathrm{M}_h = \mathrm{M}_{fh}$. On en déduit que $\mathrm{M}_f - \lambda\, \mathrm{Id}_{\mathrm{L}_2(\Omega, \mu)}$ et $\mathrm{M}_{fh - \lambda h}$ ont même image. Le noyau de M_f est l'ensemble des fonctions $g \in \mathrm{L}_2(\Omega, \mu)$ telles que fg soit μ-négligeable. Il coïncide avec celui de M_{fh}. Par la proposition 10.2.1, le spectre de M_f est l'ensemble des $\lambda \in \mathbb{C}$ tels que, pour tout $\varepsilon > 0$ l'ensemble $\{s \in \Omega : |f(s) - \lambda| < \varepsilon\}$ ne soit pas μ-négligeable.

Remarquons que M_f est injectif si et seulement si l'ensemble $\{s \in \Omega : f(s) = 0\}$ est μ-négligeable ; dans ce cas $\mathrm{im}(\mathrm{M}_f)$ est dense et $\mathrm{M}_f^{-1} = \mathrm{M}_{f^{-1}}$.

Théorème 11.4.1. *Soient* H *un espace de Hilbert complexe et* T *un opérateur autoadjoint sur* H *; le spectre de* T *est réel :* $\mathrm{Sp(T)} \subset \mathbb{R}$.

Démonstration. Soit $\lambda \in \mathbb{C} \setminus \mathbb{R}$; notons b sa partie imaginaire, non nulle. Pour tout $x \in \mathrm{dom(T)}$, on a $\langle \mathrm{T}(x), x\rangle = \langle x, \mathrm{T}(x)\rangle$ donc $\langle \mathrm{T}(x), x\rangle \in \mathbb{R}$; la partie imaginaire de $\langle (\mathrm{T} - \lambda\,\mathrm{Id_H})(x), x\rangle$ est donc $-b\|x\|^2$. On en déduit que

$$|b|\,\|x\|^2 \le |\langle (\mathrm{T} - \lambda\,\mathrm{Id_H})(x), x\rangle| \le \|(\mathrm{T} - \lambda\,\mathrm{Id_H})(x)\|\,\|x\|,$$

donc $\|(\mathrm{T} - \lambda\,\mathrm{Id_H})(x)\| \ge |b|\,\|x\|$, pour tout $x \in \mathrm{dom(T)}$. Il en résulte que l'image $(\mathrm{T} - \lambda\,\mathrm{Id_H})(\mathrm{dom(T)})$ est fermée dans H : supposons en effet que $(x_n) \subset \mathrm{dom(T)}$ et que $(\mathrm{T}x_n - \lambda x_n)$ converge vers $y \in \mathrm{H}$; d'après ce qui précède,

$$\|x_n - x_m\| \le |b|^{-1}\|(\mathrm{T}x_n - \lambda x_n) - (\mathrm{T}x_m - \lambda x_m)\| \to 0$$

donc (x_n) est de Cauchy, donc converge vers un $x \in \mathrm{H}$; par conséquent, $\mathrm{T}x_n$ converge vers $y+\lambda x$; puisque T est fermé (parce qu'il est autoadjoint) et puisque $(x_n, \mathrm{T}(x_n)) \in \mathrm{Gr(T)}$, on en déduit que $(x, y+\lambda x) \in \mathrm{Gr(T)}$, donc $x \in \mathrm{dom(T)}$ et $y = (\mathrm{T}-\lambda\,\mathrm{Id_H})(x)$.

Pour finir, on va voir que l'image $(\mathrm{T} - \lambda\,\mathrm{Id_H})(\mathrm{dom(T)})$ est dense dans H, en vérifiant que $y = 0_\mathrm{H}$ est le seul vecteur de H orthogonal à cette image ; si y est orthogonal à l'image, on aura

$$\langle \mathrm{T}x - \lambda x, y\rangle = 0$$

pour tout $x \in \mathrm{dom(T)}$, ce qui montre que la forme linéaire $x \in \mathrm{dom(T)} \to \langle \mathrm{T}x, y\rangle$ est continue, puisqu'elle est égale à $x \in \mathrm{dom(T)} \to \lambda\langle x, y\rangle = \langle x, \overline{\lambda}y\rangle$; on en déduit que $y \in \mathrm{dom(T^*)} = \mathrm{dom(T)}$ et $\mathrm{T}(y) = \mathrm{T}^*(y) = \overline{\lambda}y$; mais $\langle \mathrm{T}(y), y\rangle = \langle y, \mathrm{T}(y)\rangle = \lambda\langle y, y\rangle$ doit être réel, ce qui n'est possible que si $y = 0_\mathrm{H}$.

Il résulte de tout ce qui précède que $\mathrm{T} - \lambda\,\mathrm{Id_H}$ est bijective de dom(T) sur H, et que l'inverse est continue, de norme $\le 1/|b|$.

//

Exercice 11.4.2. On considère l'opérateur autoadjoint S sur $\mathrm{L}_2(0,1)$ dont le domaine est $\mathrm{dom(S)} = \mathrm{H}^1(0,1) \cap \{f : f(0) = f(1)\}$, et qui est défini par $\mathrm{S}f = -if'$ pour $f \in \mathrm{dom(S)}$. Pour tout $n \in \mathbb{Z}$ on définit la fonction e_n par $e_n(t) = \mathrm{e}^{2\pi int}$; montrer que pour tout $n \in \mathbb{Z}$, la fonction e_n est dans le domaine de S et vérifie $\mathrm{S}(e_n) = 2\pi n\, e_n$. Vérifier ensuite que le spectre de S est exactement égal à $2\pi\mathbb{Z}$, en calculant la résolvante $\mathrm{R}_\lambda(\mathrm{S})$ pour tout $\lambda \notin 2\pi\mathbb{Z}$.

Produit de résolvantes

Proposition 11.4.2. *Soient* H *un espace de Hilbert complexe,* T *un opérateur fermé sur* H *et* $\lambda, \mu \notin \mathrm{Sp(T)}$ *; posons* $r_\lambda = \mathrm{R}_\lambda(\mathrm{T})$, $r_\mu = \mathrm{R}_\mu(\mathrm{T})$ *; l'image* $r_\lambda r_\mu(\mathrm{H})$ *est égale au domaine de* T^2, *et* $r_\lambda r_\mu = r_\mu r_\lambda$.

Démonstration. Rappelons que

$$\mathrm{dom(T^2)} = \{x \in \mathrm{dom(T)} : \mathrm{T}x \in \mathrm{dom(T)}\}.$$

Considérons $x = r_\lambda(r_\mu(y))$, pour un $y \in \mathrm{H}$ quelconque. Par définition, on a $r_\lambda(\mathrm{H}) = \mathrm{dom(T)}$ donc $x \in \mathrm{dom(T)}$ et $\mathrm{T}x - \lambda x = r_\mu(y) \in \mathrm{dom(T)}$; par linéarité, $\mathrm{T}x \in \mathrm{dom(T)}$ donc $x \in \mathrm{dom(T^2)}$.

Inversement supposons $x \in \mathrm{dom(T^2)}$; alors $x \in \mathrm{dom(T)}$ et $x_1 = \mathrm{T}x - \lambda x \in \mathrm{dom(T)}$, ce qui permet de calculer $y = \mathrm{T}x_1 - \mu x_1$; on aura alors $r_\mu y = x_1$, puis $r_\lambda r_\mu y = x$.

Si $x \in \mathrm{dom}(\mathrm{T}^2)$, on vérifie immédiatement en développant que

$$(\mathrm{T}-\lambda)(\mathrm{T}-\mu)x = \mathrm{T}^2(x) - \lambda\mathrm{T}(x) - \mu\mathrm{T}(x) + \lambda\mu x = (\mathrm{T}-\mu)(\mathrm{T}-\lambda)x.$$

En prenant l'inverse de cette relation sur l'image commune $\mathrm{dom}(\mathrm{T}^2) = r_\lambda r_\mu(\mathrm{H}) = r_\mu r_\lambda(\mathrm{H})$ on obtient $r_\lambda r_\mu = r_\mu r_\lambda$.

//

Proposition 11.4.3. *Soient* H *un espace de Hilbert complexe et* T *un opérateur auto-adjoint sur* H *; alors* $\mathrm{R}_i(\mathrm{T}) = (\mathrm{T} - i\,\mathrm{Id}_\mathrm{H})^{-1}$ *est un opérateur borné normal, et son adjoint est égal à* $\mathrm{R}_{-i}(\mathrm{T})$.

Démonstration. D'après le théorème 1, les opérateurs $\mathrm{R}_i(\mathrm{T})$ et $\mathrm{R}_{-i}(\mathrm{T})$ existent et sont bornés. D'après la proposition précédente il suffit de savoir que $\mathrm{R}_{-i}(\mathrm{T})$ est l'adjoint de $\mathrm{R}_i(\mathrm{T})$. Soient y, v deux vecteurs quelconques dans H et posons $x = \mathrm{R}_i(\mathrm{T})(y)$ et $u = \mathrm{R}_{-i}(\mathrm{T})(v)$. On a

$$\langle \mathrm{R}_i(\mathrm{T})(y), v\rangle = \langle x, (\mathrm{T}+i)(u)\rangle = \langle \mathrm{T}x - ix, u\rangle = \langle y, \mathrm{R}_{-i}(\mathrm{T})(v)\rangle.$$

//

Théorème 11.4.4. *Soient* H *un espace de Hilbert séparable et* T *un opérateur auto-adjoint de* H *dans* H *; il existe un espace mesuré* (Ω, μ), *une fonction* $f : \Omega \to \mathbb{R}$ *mesurable et un isomorphisme* $u : \mathrm{L}_2(\Omega, \mu) \to \mathrm{H}$ *d'espaces de Hilbert tels que* $\mathrm{T} = u\,\mathrm{M}_f\,u^*$.

On donne à l'opérateur M_f son domaine naturel (de l'exemple 1) ; la relation ci-dessus sous-entend que u et son inverse u^* échangent les domaines de T et de M_f, c'est à dire qu'on a l'égalité $u(\mathrm{dom}(\mathrm{M}_f)) = \mathrm{dom}(\mathrm{T})$ (et inversement, $u^*(\mathrm{dom}(\mathrm{T})) = \mathrm{dom}(\mathrm{M}_f)$).

Démonstration. On va se servir de l'opérateur normal $\mathrm{S} = (\mathrm{T} - i\,\mathrm{Id}_\mathrm{H})^{-1}$ et de sa représentation obtenue au chapitre 9. Il existe un espace mesuré σ-fini $(\Omega, \mathcal{A}, \mu)$, un unitaire U de $\mathrm{H}_1 = \mathrm{H}$ sur $\mathrm{H}_2 = \mathrm{L}_2(\Omega, \mu)$, une fonction $h \in \mathrm{L}_\infty(\Omega, \mu)$ tels que

$$\mathrm{S} = \mathrm{U}^*\,\mathrm{M}_h\,\mathrm{U},$$

où M_h désigne l'opérateur borné sur $\mathrm{L}_2(\Omega, \mu)$ défini par la multiplication par h,

$$\forall g \in \mathrm{L}_2(\Omega, \mu), \quad \mathrm{M}_h(g) = hg.$$

Puisque S est injectif, il en résulte que M_h est injectif aussi ; cela implique que l'ensemble $\mathrm{A} = \{h = 0\}$ est μ-négligeable (sinon on pourrait trouver, puisque μ est σ-finie, un $\mathrm{B} \subset \mathrm{A}$ tel que $0 < \mu(\mathrm{B}) < +\infty$, et alors $\mathbf{1}_\mathrm{B} \in \mathrm{L}_2(\Omega, \mu)$ vérifierait $\mathrm{M}_h(\mathbf{1}_\mathrm{B}) = 0$ et $\|\mathbf{1}_\mathrm{B}\| > 0$, ce qui n'est pas possible).

Si on comprend la traduction de S sur $\mathrm{L}_2(\Omega, \mu)$, il n'est pas bien difficile de comprendre celle de $(\mathrm{T} - i)$ puis celle de T. L'opérateur $(\mathrm{T} - i)$, qui est l'inverse de $\mathrm{S} = \mathrm{R}_i(\mathrm{T})$, se traduit sur $\mathrm{L}_2(\Omega, \mu)$ par l'inverse de la traduction de S : c'est l'opérateur M_2 de multiplication par la fonction $1/h$ (fonction qui est μ-presque partout définie). Le domaine D_2 de M_2 est l'image de M_h, l'ensemble des fonctions g de la forme $g = hk$ pour une $k \in \mathrm{L}_2$. Comme $h \neq 0$ presque partout, cela revient à dire que

$$\mathrm{D}_2 = \{g \in \mathrm{L}_2 : \int_\Omega |g/h|^2\,d\mu < +\infty\}.$$

Pour finir on définit l'opérateur non borné T_2 sur $L_2(\Omega, \mu)$ par son domaine D_2 et la formule

$$\forall g \in D_2, \quad T_2(g) = g/h + ig.$$

Autrement dit, si on pose $f = i + 1/h$ et $u = U^*$, on obtient bien la représentation voulue : l'opérateur de multiplication $M_f = T_2$ vérifie $T = u\, M_f\, u^*$.

//

Exemple 11.4.3. Soit $H = L_2(\mathbb{R})$ et définissons T sur H par $\mathrm{dom}(T) = H^1(\mathbb{R})$ et $Tf = -if'$ pour toute $f \in H^1(\mathbb{R})$. On vérifie que T est autoadjoint. Sa représentation par une multiplication est obtenue au moyen de la transformation de Fourier sur $L_2(\mathbb{R})$. On va obtenir ainsi que T est conjugué à l'opérateur M de multiplication par $t \to t$ sur $L_2(\mathbb{R})$. Si on pose pour $f \in L_1(\mathbb{R})$

$$\forall t \in \mathbb{R}, \quad (\mathcal{F}f)(t) = \int_{\mathbb{R}} e^{-ixt}\, f(x)\, dx$$

on obtient par une intégration par parties facile que pour f à support compact de classe C^1, on a $\mathcal{F}(-if')(t) = t\,\mathcal{F}(f)(t)$. Par ailleurs on sait montrer que $U = (2\pi)^{-1/2}\,\mathcal{F}$ se prolonge en opérateur unitaire de $L_2(\mathbb{R})$, et on a alors la représentation

$$T = U^* \circ M \circ U.$$

Remarque 11.4.4. Si H est un espace de Hilbert séparable et T un opérateur autoadjoint de H dans H, l'opérateur $U = (T - i\,\mathrm{Id}_H)(T + i\,\mathrm{Id}_H)^{-1}$ est un élément unitaire de $\mathcal{L}(H)$. Cette transformation, appelée *transformation de Cayley*, permet de relier les autoadjoints non bornés aux unitaires.

11.5. Le théorème de Stone

Soit H un espace de Hilbert ; on appelle *groupe à un paramètre d'unitaires* une famille $(v_t)_{t\in\mathbb{R}}$ d'éléments unitaires de $\mathcal{L}(H)$ telle que :

(*i*) pour tous $s, t \in \mathbb{R}$ on a $v_{s+t} = v_s v_t$;
(*ii*) pour tout $x \in H$ l'application $t \to v_t(x)$ est continue.

On a vu dans le chapitre sur le calcul fonctionnel continu une façon d'obtenir des exemples particuliers de tels groupes d'opérateurs. En effet, étant donné un opérateur hermitien borné A sur un espace de Hilbert *complexe* H, on considère la famille $(f_s)_{s\in\mathbb{R}}$ de fonctions continues sur $\mathrm{Sp}(T) \subset \mathbb{R}$ donnée par

$$\forall t \in \mathbb{R}, \quad f_s(t) = e^{ist}\,.$$

On voit immédiatement que $f_{s_1} f_{s_2} = f_{s_1+s_2}$ pour tous s_1, s_2 réels, $f_0 = 1$, et toutes les fonctions sont de module un sur le spectre de A. D'après le corollaire 9.2.5 et les propriétés d'homomorphisme du calcul fonctionnel, la famille

$$v_s = f_s(A) = e^{isA}$$

fournit un groupe d'unitaires, qui a en fait une propriété plus forte que la propriété (*ii*) de la définition ci-dessus : on a ici

$$\forall s, t \in \mathbb{R}, \quad \|v_s - v_t\| = \|f_s - f_t\|_{C(\mathrm{Sp}(A))} \le \|A\|\,|s - t|.$$

Le théorème qui suit permet d'aller beaucoup plus loin.

Théorème 11.5.2 : Théorème de Stone. *Soit* H *un espace de Hilbert complexe séparable ;*

(*i*) *soit* $(v_t)_{t\in\mathbb{R}}$ *un groupe à un paramètre d'opérateurs d'unitaires ; il existe un opérateur autoadjoint* T *sur* H *dont le graphe est l'ensemble des couples* $(x, y) \in \mathrm{H}\times\mathrm{H}$ *tels que la fonction* $t \to v_t(x)$ *soit dérivable en* 0*, de dérivée* iy*. Pour* $x \in \mathrm{dom(T)}$ *et* t *réel, on a* $v_t(x) \in \mathrm{dom(T)}$ *et* $\mathrm{T}(v_t(x)) = v_t(\mathrm{T}(x))$.

On dit que T *est le générateur infinitésimal de* $(v_t)_{t\in\mathbb{R}}$ *;*

(*ii*) *tout opérateur autoadjoint* T *est le générateur infinitésimal d'un unique groupe à un paramètre d'unitaires* $(v_t)_{t\in\mathbb{R}}$.

Index

Index des notations