ANALYSE FONCTIONNELLE ET

THÉORIE SPECTRALE

MT404

Année 1999-2000

première partie

0. Introduction.

L'Analyse Fonctionnelle est née au début du 20ème siècle pour fournir un cadre plus général et abstrait 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 naissance de la théorie moderne de l'intégration (Lebesgue, 1904) 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 plus clair et plus simple.

Donnons ici un seul exemple, hyper-classique, celui du *problème de Dirichlet*. Etant donné un ouvert borné Ω du plan, et étant donnée une fonction continue f sur la frontière $\partial\Omega$, on veut trouver une fonction u définie sur le compact $\overline{\Omega}$, de classe C^2 dans Ω et qui vérifie

$$\text{(D)} \qquad \begin{cases} \Delta u = 0 \quad \text{dans} \quad \Omega, \\ u = f \quad \text{sur la frontière} \quad \partial\Omega. \end{cases}$$

L'expression Δu désigne le *laplacien* de u,

$$\Delta u = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2}.$$

On savait depuis longtemps que si u_0 est solution du problème (D), la fonction u_0 minimise une certaine *fonctionnelle d'énergie* I, qui associe à chaque fonction u la quantité

$$\mathrm{I}(u) = \int_\Omega |\nabla u(x,y)|^2 \, dx\, dy,$$

le minimum étant calculé sur l'ensemble des fonctions u qui coïncident avec f au bord ; l'expression ∇u désigne le *gradient* de u, fonction vectorielle définie sur Ω dont les composantes sont les deux dérivées partielles premières de u, et $|\nabla u|^2$ est la fonction égale en chaque point $(x, y) \in \Omega$ au carré de la norme euclidienne du gradient de u au point (x, y). L'une des méthodes de l'Analyse Fonctionnelle consiste à introduire un espace de fonctions X adapté (ici : l'espace de Sobolev $\mathrm{H}^1(\Omega)$; c'est un espace de Hilbert), et de développer des théorèmes abstraits d'existence de minimum pour des classes de fonctions définies sur X (analogue, disons, au théorème qui dit que toute fonction réelle continue définie sur un compact atteint son minimum). Le problème est que l'espace des fonctions continues sur $\overline{\Omega}$ et C^2 à l'intérieur n'est justement pas adapté à cette approche ; les bons espaces ne sont pas formés de fonctions vraiment dérivables. La deuxième partie de cette approche, qui ne sera pas du tout évoquée dans ce cours, consiste à montrer que sous certaines hypothèses, la solution trouvée dans l'espace généralisé X est bien une solution au sens classique.

La première partie du poly contient les éléments de base de l'Analyse Fonctionnelle : espaces normés et espaces de Banach, applications linéaires continues, dualité, topologies faibles, espaces de Hilbert.

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.

Ce polycopié provient en très 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.

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 $x \in \mathrm{X}$. De temps en temps on notera 0_{X} le vecteur nul de l'espace vectoriel X, quand il semblera que cette notation lourde lève toute ambiguïté. Le signe $\blacksquare$ indique la fin d'une démonstration (ou une démonstration omise). La plupart des points du cours sont numérotés, par chapitre-section-type, par exemple le "théorème 2.3.1" serait le premier théorème de la section 3 du chapitre 2 ; à l'intérieur de cette section, le théorème sera appelé "théorème 1", ailleurs dans le chapitre 2 il sera désigné par "théorème 3.1" et dans un autre chapitre "théorème 2.3.1". 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.

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* ; autrement dit une semi-norme p sur X est une norme lorsque $\{x \in \mathrm{X} : p(x) = 0\} = \{0_{\mathrm{X}}\}$.

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 peut déduire une deuxième forme : on a $p(x) = p\big((x-y)+y\big) \le p(x-y) + p(y)$, donc $p(x) - p(y) \le p(x-y)$; en échangeant les rôles de x et y et en utilisant $p(x-y) = p(y-x)$ on trouve :

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 $x, y \in \mathrm{X}$.*

Il est facile de vérifier que cette deuxième forme de l'inégalité triangulaire implique la première forme (ii).

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 $[0, 1]$. Plus généralement, une combinaison convexe d'éléments d'un sous-ensemble $\mathrm{A} \subset \mathrm{X}$ est un vecteur z de la forme $z = \sum \lambda_\alpha x_\alpha$, où la somme est finie, $\lambda_\alpha \ge 0$ pour chaque α et $\sum \lambda_\alpha = 1$. L'ensemble de toutes les combinaisons convexes d'éléments de A est un ensemble convexe appelé *l'enveloppe convexe* de A, noté co(A) dans ce poly. C'est aussi le plus petit ensemble convexe contenant A.

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.

Lemme 1.1.2. *Soient* X *un espace vectoriel sur* $\mathbb{K}$ *et q une fonction* **à valeurs** ≥ 0 *et positivement homogène sur* X *; pour que q soit sous-additive sur* X, *il faut et il suffit que l'ensemble*

$$\mathrm{C}_q = \{x \in \mathrm{X} : q(x) \le 1\}$$

soit convexe.

Démonstration. Si q est positivement homogène et sous-additive, on a vu que q est convexe, ce qui implique immédiatement que C_q est convexe. Inversement, supposons que C_q soit convexe, et déduisons la sous-additivité de q ; soient x et y deux vecteurs de X, et supposons d'abord que $a = q(x) > 0$ et $b = q(y) > 0$; considérons les deux vecteurs de C_q définis par $x_1 = a^{-1}x$ et $y_1 = b^{-1}y$, puis formons la combinaison convexe

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

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

Si $a > 0$ et $b = 0$, on doit procéder un peu différemment. Prenons t tel que $0 < t < 1$, qui tendra ensuite vers 0, et posons maintenant $y_1 = (1-t)a^{-1}t^{-1}y$ (et x_1 inchangé). Par homogénéité on a $q(y_1) = 0$, donc $y_1 \in \mathrm{C}_q$, et par convexité $z = (1-t)x_1 + ty_1 \in \mathrm{C}_q$, donc $q(z) \le 1$. Mais maintenant $z = (1-t)a^{-1}(x+y)$, donc $(1-t)q(x+y) \le a = q(x)+q(y)$, et le résultat en découle lorsque $t \to 0$. Le dernier cas, quand $a = b = 0$, est laissé au lecteur.

■

Corollaire 1.1.1. *Pour que p 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.*

On appelle *espace normé* un espace vectoriel X muni d'une norme p. Si X 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. Quand il n'y aura pas d'ambiguïté sur la norme considérée, on notera ces boules simplement $\mathrm{B}(x,r)$.

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 $\mathrm{B}(x,r) \subset \mathrm{U}$. Dans un espace vectoriel normé, la *boule fermée* de centre x et de rayon $r > 0$, qui est l'ensemble $\{y \in \mathrm{X} : p(y-x) \le r\}$, est bien égale à l'adhérence de la boule ouverte $\mathrm{B}_p(x,r)$ (çà n'est pas toujours vrai dans un espace métrique général) ; on pourra donc la noter $\overline{\mathrm{B}}_p(x,r)$.

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

Cela résulte facilement 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 1.1.2. 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 1.1.2. *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 $(x, y) \to x + y$ en un point quelconque $(u, v) \in \mathrm{X}^2$. Puisque la topologie provient d'une métrique, nous utiliserons des suites convergentes. Soient donc (x_n) une suite qui converge vers u et (y_n) une suite qui converge vers v ; on aura

$$d(x_n+y_n, u+v) = \big\|(x_n+y_n)-(u+v)\big\| = \|(x_n-u)+(y_n-v)\| \le \|x_n-u\|+\|y_n-v\| \to 0.$$

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

$$d(\lambda_n x_n, \mu\, u) \le d(\lambda_n x_n, \lambda_n u) + d(\lambda_n u, \mu\, u) = |\lambda_n|\, \|x_n - u\| + |\lambda_n - \mu|\, \|u\|$$

qui tend vers 0 (noter que la suite (λ_n) est bornée puisque convergente).

∎

Définition 1.1.3. Un espace de Banach est un espace vectoriel normé, complet pour la distance associée.

Si Y est un sous-espace vectoriel fermé d'un espace de Banach X, il est lui aussi complet pour la norme induite par celle de X, donc Y est 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 $n \ge 0$ 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 donc 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.

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 cette condition, le reste de la série des normes

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

est une suite numérique qui tend vers 0, 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. Si 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 1.1.3. *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.

Exemples 1.1.1.

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

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

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.

Beaucoup d'exemples viennent de la théorie de l'intégration. Nous supposerons donné un espace $(\Omega, \mathcal{A}, \mu)$, où $\mathcal{A}$ est une tribu de parties de Ω et μ une mesure positive sur $(\Omega, \mathcal{A})$. 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 $\mu(\Omega_n) < +\infty$.

Pour $1 \leq p < +\infty$, l'espace $\mathrm{L}_p = \mathrm{L}_p(\Omega, \mathcal{A}, \mu)$ des (classes de) fonctions f complexes sur $[0,1]$ telles que f soit mesurable et $\int_\Omega |f|^p \, d\mu < +\infty$ est normé par

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

Pour vérifier qu'il s'agit d'une semi-norme, on voit d'abord que $\|\lambda f\|_p = |\lambda|\, \|f\|_p$ (facile), puis on montre que l'ensemble $\{f : \|f\|_p \le 1\}$ est convexe. Cela provient de la convexité sur $\mathbb{R}$ de la fonction $u \to |u|^p$; on a alors si f, g sont deux éléments de L_p tels que $\|f\|_p \le 1$, $\|g\|_p \le 1$ et si $0 \le t \le 1$,

$$|(1-t)f(s) + tg(s)|^p \le (1-t)|f(s)|^p + t|g(s)|^p$$

pour tout $s \in \Omega$, donc

$$\int_\Omega \bigl|(1-t)f(s)+tg(s)\bigr|^p \, d\mu(s) \le (1-t) \int_\Omega |f(s)|^p \, d\mu(s) + t \int_\Omega |g(s)|^p \, d\mu(s) \le (1-t)+t = 1.$$

Cet espace L_p est de plus complet : on peut utiliser le critère des séries normalement convergentes de la proposition 3 et quelques arguments d'intégration pour retrouver ce théorème du cours d'Intégration (appelé souvent *théorème de Fisher-Riesz*).

De façon analogue, on désigne par ℓ_p l'espace des suites scalaires $x = (x_n)$ telles que $\sum |x_n|^p < +\infty$. C'est un espace de Banach pour la norme

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

Pour unifier les arguments, on peut dire que c'est l'espace L_p pour la mesure de comptage sur $\Omega = \mathbb{N}$.

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.

Il existe aussi 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 $|f(s)| \le \mathrm{M}$ pour μ-presque tout $s \in \Omega$. L'espace L_∞ est complet pour cette norme.

L'espace de Sobolev H^1 est facile à définir dans le cas d'un intervalle borné $[a, b]$ de $\mathbb{R}$. Il s'agit de l'espace $\mathrm{H}^1[a, b]$ des fonctions f sur $[a, b]$ telles qu'il existe une fonction $g \in \mathrm{L}_2[a, b]$ pour laquelle

$$f(y) - f(x) = \int_x^y g(s)\, ds$$

pour tous $x, y \in [a, b]$. On dit que g est la *dérivée généralisée* de f, et on la note souvent simplement f'. La norme est alors

$$\|f\|_{\mathrm{H}^1} = \bigl(\|f\|_2^2 + \|f'\|_2^2\bigr)^{1/2}.$$

Le cas de $\mathbb{R}^n$, $n \ge 2$ ou d'un ouvert de $\mathbb{R}^n$ ne sera pas traité ici.

Remarque 1.1.1.

Certains espaces importants tels que l'espace $\mathrm{C}^\infty(\mathbb{R})$ des fonctions de classe C^∞ sur $\mathbb{R}$ n'admettent pas de norme satisfaisante. Dans les bons cas, on peut tout de même définir sur un espace vectoriel X une distance qui garde certaines des bonnes propriétés des distances déduites d'une norme, à savoir :

(i) : pour tous $x, y, u \in \mathrm{X}$, on a $d(x+u, y+u) = d(x, y)$ (invariance par translation) ;
(ii) : pour tout $x \in \mathrm{X}$, on a $d(\lambda x, 0) \le d(x, 0)$ si $|\lambda| \le 1$;
(iii) : pour chaque vecteur $x \in \mathrm{X}$, on a $\lim_{\lambda \to 0} d(\lambda x, 0) = 0$.

(et bien entendu on garde les propriétés usuelles des distances, en particulier l'inégalité triangulaire). Muni de la topologie associée à une telle distance, l'espace reste un espace vectoriel topologique (exercice).

C'est le cas pour l'espace C^∞. Pour tout entier $n \ge 1$, on peut considérer la semi-norme p_n qui contrôle le caractère C^n de la fonction f sur l'intervalle $[-n, n]$,

$$p_n(f) = \max\{|f^{(k)}(s)| : 0 \le k \le n,\ |s| \le n\}.$$

On pose ensuite

$$d(f, g) = \sum_{n=0}^{+\infty} \min(p_n(f-g), 2^{-n}).$$

L'espace C^∞ est alors complet pour cette distance. On dit que c'est un *espace de Fréchet*.

Un autre exemple est l'espace L_0 des (classes) de fonctions mesurables. Supposons que $\mu(\Omega) = 1$ pour simplifier. On peut définir une distance pour la convergence en mesure en posant par exemple

$$d(f, g) = \int_\Omega \min(1, |f(s) - g(s)|)\, d\mu(s).$$

L'espace L_0 est complet pour cette distance.

1.2. Applications linéaires continues

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

(i) *l'application* f *est continue sur* X *;*
(ii) *l'application* f *est continue au point* 0_X *;*
(iii) *il existe un nombre* $k \ge 0$ *tel que, pour tout* $x \in \mathrm{X}$ *on ait*

$$\|f(x)\|_\mathrm{Y} \le k\, \|x\|_\mathrm{X}.$$

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

$$d(f(x_n), f(u)) = \|f(x_n) - f(u)\|_\mathrm{Y} = \|f(x_n - u)\|_\mathrm{Y} \le k\, \|x_n - u\|_\mathrm{X} \to 0,$$

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

∎

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 $k > 0$ et $\ell \ge 0$ tels que $k\,p \le q \le \ell\,p$.

Corollaire 1.2.1. *Deux normes* p *et* q *sur un espace vectoriel* X *définissent la même topologie si et seulement si elles sont équivalentes.*

Démonstration. On applique le théorème 1, (i) $\iff$ (iii), à l'identité de X vue comme application de (X, p) dans (X, q), puis de (X, q) dans (X, p).

∎

Soient X et Y deux espaces normés et f, g deux applications linéaires continues de X dans Y ; on sait que l'application $f+g$, qui associe à tout $x \in \mathrm{X}$ l'image $f(x)+g(x) \in \mathrm{Y}$, est linéaire. Vérifions rapidement sa continuité en 0 : si (x_n) tend vers 0 dans X, alors $(f(x_n))$ et $(g(x_n))$ tendent vers 0 dans Y donc $(f+g)(x_n) = f(x_n) + g(y_n)$ 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* 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 $f : \mathrm{X} \to \mathrm{Y}$ une application linéaire continue ; d'après le théorème 1, il existe une constante k telle que $\|f(x)\|_{\mathrm{Y}} \leq k$ pour tout vecteur x de X tel que $\|x\|_{\mathrm{X}} \leq 1$. On peut donc considérer la quantité (finie)

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

qui s'appelle la *norme de l'application linéaire* f. Si x est un vecteur non nul de X, le vecteur $z = \|x\|_{\mathrm{X}}^{-1} x$ vérifie $\|z\|_{\mathrm{X}} \leq 1$, donc $\|f(z)\|_{\mathrm{Y}} \leq \|f\|$, d'où par homogénéité $\|x\|^{-1}\|f(x)\| \leq \|f\|$, ou encore

$$\|f(x)\|_{\mathrm{Y}} \leq \|f\|_{\mathcal{L}(\mathrm{X},\mathrm{Y})}\|x\|_{\mathrm{X}};$$

si x est le vecteur nul, la relation ci-dessus est encore vraie, elle est donc vraie pour tout vecteur $x \in \mathrm{X}$. Résumons ce qui vient d'être dit.

Proposition 1.2.1. *Soient* X *et* Y *deux espaces normés et* $f : \mathrm{X} \to \mathrm{Y}$ *une application linéaire continue ; on pose*

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

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

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

La constante $\|f\|_{\mathcal{L}(\mathrm{X},\mathrm{Y})}$ *est le plus petit nombre* M *tel que l'inégalité* $\|f(x)\|_{\mathrm{Y}} \leq \mathrm{M}\,\|x\|_{\mathrm{X}}$ *soit vraie pour tout* $x \in \mathrm{X}$. *L'application* $f \to \|f\|_{\mathcal{L}(\mathrm{X},\mathrm{Y})}$ *est une norme sur* $\mathcal{L}(\mathrm{X}, \mathrm{Y})$.

Démonstration. Vérifions que $\|f\|$ est une norme. Il est d'abord évident que $\|f\| = 0$ implique que $\|f(x)\| = 0$ pour tout $x \in \mathrm{X}$, c'est à dire $f(x) = 0_{\mathrm{Y}}$ pour tout $x \in \mathrm{X}$ puisque Y est normé, donc f est l'application nulle. Montrons ensuite que $f \to \|f\|$ est une semi-norme ; il est facile de vérifier que $\|\lambda f\| = |\lambda|\,\|f\|$ pour tout $\lambda \in \mathbb{K}$; ensuite, pour tout x tel que $\|x\| \leq 1$,

$$\|(f+g)(x)\| = \|f(x) + g(x)\| \leq \|f(x)\| + \|g(x)\| \leq \|f\| + \|g\|,$$

d'où l'inégalité $\|f+g\| \leq \|f\| + \|g\|$ en passant au sup sur x.

∎

La proposition suivante est facile mais importante.

Proposition 1.2.2. *Soient* X, Y *et* Z *des espaces normés,* $f : \mathrm{X} \to \mathrm{Y}$ *et* $g : \mathrm{Y} \to \mathrm{Z}$ *des applications linéaires continues ; on a*

$$\|g \circ f\| \leq \|f\|\,\|g\|.$$

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

$$\|(g \circ f)(x)\|_{\mathrm{Z}} = \|g(f(x))\|_{\mathrm{Z}} \leq \|g\|\,\|f(x)\|_{\mathrm{Y}} \leq \|g\|\,\|f\|\,\|x\|_{\mathrm{X}},$$

ce qui entraîne l'inégalité voulue.

∎

Proposition 1.2.3. *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 $\|u_k(x)\| \leq \|u_k\| \, \|x\|$, 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 $x \in \mathrm{X}$

$$\|\mathrm{U}(x)\| \leq \sum_{k=0}^{+\infty} \|u_k(x)\| \leq \Big(\sum_{k=0}^{+\infty} \|u_k\|\Big) \, \|x\|,$$

ce qui montre que U est continue. 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

$$(\mathrm{U} - \mathrm{U}_n)(x) = \sum_{k>n} u_k(x)$$

d'où il résulte comme précédemment que $\|\mathrm{U} - \mathrm{U}_n\| \leq \sum_{k>n} \|u_k\|$, 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 $f : \mathrm{X} \to \mathrm{Y}$ une application linéaire continue. Alors la série $\sum f(u_k)$ converge dans Y et

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

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.*

Démonstration. Notons $\mathrm{N} : \mathrm{X} \times \mathrm{Y} \to \mathbb{R}$ l'application $(x, y) \to \|x\|_{\mathrm{X}} + \|y\|_{\mathrm{Y}}$. Il est clair que N est une norme, et que pour toute suite (x_n, y_n) dans $\mathrm{X} \times \mathrm{Y}$, la quantité $\mathrm{N}(x_n, y_n)$ tend vers 0 si et seulement si $\|x_n\| \to 0$ **et** $\|y_n\| \to 0$, c'est à dire si et seulement si (x_n) tend vers 0 dans X et (y_n) tend vers 0 dans Y, ce qui équivaut à dire que (x_n, y_n) tend vers $(0, 0)$ pour la topologie produit.

■

Remarquons que la norme N décrite dans la démonstration ci-dessus n'est pas l'unique norme définissant la topologie produit : n'importe quelle norme équivalente convient. Plusieurs normes équivalentes sont tout aussi naturelles que N. Par exemple la norme $(x, y) \to \max(\|x\|, \|y\|)$ ou la norme $(x, y) \to (\|x\|^r + \|y\|^r)^{1/r}$, pour $r > 1$.

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 $\mathrm{R_Y}$ telle que $x\,\mathrm{R_Y}\,y \iff y - x \in \mathrm{Y}$. Le quotient X/Y est muni de l'unique structure d'espace vectoriel pour laquelle l'application quotient $\mathrm{X} \to \mathrm{X/Y}$ 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* $\pi : \mathrm{X} \to \mathrm{X/Y}$ *l'application quotient. L'application* $q : \mathrm{X/Y} \to \mathbb{R}_+$ *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 $q(x_n) = \xi$ 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 $y \in \mathrm{Y}$, donc $\xi \subset \mathrm{Y}$ ce qui implique en fait $\xi = \mathrm{Y} = 0_{\mathrm{X/Y}}$.

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 $\xi \in \mathrm{X/Y}$. Soient $\xi, \xi' \in \mathrm{X/Y}$ 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.

■

Notons que la distance $d(\xi, \eta)$ de deux classes de X/Y est simplement la distance naturelle des sous-ensembles ξ et η de X, c'est à dire l'inf de $d(x, y)$, lorsque x varie dans ξ et y dans η. Notons aussi que la projection π vérifie $\|\pi\| \le 1$; on a même en général $\|\pi\| = 1$, sauf si $\mathrm{X/Y} = \{0\}$, c'est à dire si $\mathrm{Y} = \mathrm{X}$ (on suppose toujours Y fermé). Notons encore que l'image par π de la boule unité ouverte $\mathrm{B_X}(0, 1)$ de X est exactement la boule unité ouverte du quotient X/Y (l'énoncé correspondant pour la boule unité fermée n'est pas vrai en général).

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* $h \in \mathcal{L}(\mathrm{X/Y}, \mathrm{Z})$ *telle que* $g = h \circ \pi$ *(où* $\pi : \mathrm{X} \to \mathrm{X/Y}$ *est l'application quotient) ; on a* $\|g\| = \|h\|$.

Démonstration. Soit $\xi \in \mathrm{X/Y}$; on va vérifier que tous les vecteurs x de la classe ξ ont la même image z dans Z, ce qui permettra de poser $h(\xi) = z = g(x)$: si $x, x' \in \mathrm{X}$ sont dans la classe ξ, alors $x - x' \in \mathrm{Y} \subset \ker \pi$, donc $g(x - x') = 0$, soit $g(x) = g(x')$. Donc $g(x)$ ne dépend pas du choix du représentant x de la classe ξ ; notons le $h(\xi)$. Il est clair que h est linéaire, et par construction on a $g = h \circ \pi$. Comme π est surjective, il est clair que h est unique.

Notons q la norme quotient de X/Y. Pour $\xi \in \mathrm{X/Y}$ et $x \in \pi^{-1}(\{\xi\})$ on a $\|h(\xi)\| = \|g(x)\| \le \|g\|\,\|x\|$. Cela étant vrai pour tout $x \in \pi^{-1}(\{\xi\})$, on a $\|h(\xi)\| \le \|g\|\, q(\xi)$. Donc h est continue et $\|h\| \le \|g\|$. Enfin, $\|g\| = \|h \circ \pi\| \le \|h\|\,\|\pi\| \le \|h\|$, puisque $\|\pi\| \le 1$.

■

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 1.3. Soit $\sum \xi_k$ une série normalement convergente dans le quotient. Pour tout entier $k \geq 0$ on peut trouver un représentant $u_k \in \xi_k$ tel que $\|u_k\| \leq 2\|\xi_k\|$; il en résulte que 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. Complété d'un espace normé

Commençons par un lemme de prolongement :

Lemme 1.4.1. *Soient* X, Y *deux espaces normés,* F *un espace de Banach,* $u \in \mathcal{L}(\mathrm{X}, \mathrm{Y})$ *une application linéaire isométrique d'image dense ; pour tout* $f \in \mathcal{L}(\mathrm{X}, \mathrm{F})$*, il existe une unique* $g \in \mathcal{L}(\mathrm{Y}, \mathrm{F})$ *telle que* $f = g \circ u$*. On a* $\|f\| = \|g\|$*.*

Démonstration. Soit $y \in \mathrm{Y}$; d'après la densité de $u(\mathrm{X})$, il existe une suite $(x_n) \subset \mathrm{X}$ telle que $u(x_n) \to y$; alors $(u(x_n))$ est une suite de Cauchy dans Y, et puisque u est isométrique, la suite (x_n) est de Cauchy dans X. L'image $(f(x_n))$ est alors de Cauchy dans F, donc convergente dans F puisque F est complet. On va poser pour tout $y \in \mathrm{Y}$

$$g(y) = \lim_n f(x_n) \in \mathrm{F}.$$

Il est facile de vérifier que $g(y)$ ne dépend pas de la suite (x_n) choisie, et que g est linéaire de Y dans F. On a

$$\|g(y)\| = \lim_n \|f(x_n)\| \leq \|f\| \lim_n \|x_n\| = \|f\| \lim_n \|u(x_n)\| = \|f\| \, \|y\|,$$

ce qui montre que g est continue et $\|g\| \leq \|f\|$. Si $y = u(x)$, la suite (x_n) ci-dessus converge vers x dans X, donc $g(y) = \lim_n f(x_n) = f(x)$, ce qui montre que $f = g \circ u$; il en résulte que $\|f\| \leq \|g\| \, \|u\| = \|g\|$, donc $\|f\| = \|g\|$. Si g_1 est une autre application continue telle que $f = g_1 \circ u$, on aura $g_1(y) = \lim_n g_1(u(x_n))$ par continuité de g_1, mais $g_1(u(x_n)) = f(x_n)$ par hypothèse, donc $g_1(y) = \lim_n f(x_n) = g(y)$ pour tout $y \in \mathrm{Y}$, ce qui montre l'unicité de g.

■

Corollaire 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* $f : \mathrm{X}_0 \to \mathrm{F}$ *se prolonge de façon unique en application linéaire continue* $\widetilde{f} : \mathrm{X} \to \mathrm{F}$*, et* $\|\widetilde{f}\| = \|f\|$*.*

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 $\mathrm{X}_0 = \mathrm{L}_1 \cap \mathrm{L}_2$.

Théorème 1.4.1. *Soit* X *un espace normé ;*

(i) *il existe un couple* (E, u) *où* E *est un espace de Banach et* $u : \mathrm{X} \to \mathrm{E}$ *est une application linéaire isométrique d'image dense (l'espace* E *sera le* complété *de* X*) ;*

(ii) *si* (E, u) *est comme dans* (i)*, pour tout espace de Banach* F *et toute* $f \in \mathcal{L}(\mathrm{X}, \mathrm{F})$*, il existe une unique* $g \in \mathcal{L}(\mathrm{E}, \mathrm{F})$ *telle que* $f = g \circ u$ *; on a* $\|f\| = \|g\|$ *;*

(iii) *si* (E_1, u_1) *et* (E_2, u_2) *sont deux couples comme dans* (i)*, il existe une bijection linéaire isométrique* $v : \mathrm{E}_1 \to \mathrm{E}_2$ *telle que* $u_2 = v \circ u_1$ *(autrement dit : le* complété *de* X *est* unique *à* isomorphisme près*).*

Démonstration. Désignons par $\mathrm{L} = \ell_\infty(\mathrm{X})$ l'espace des suites bornées $x = (x_n)$ de vecteurs de X, muni de la norme

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

Il est facile de vérifier que l'on a ainsi défini une norme sur $\mathrm{L} = \ell_\infty(\mathrm{X})$. Les suites de Cauchy (x_n) de vecteurs de X sont un cas particulier de suite bornée ; de plus, la somme de deux suites de Cauchy est une suite de Cauchy, et si on multiplie une suite de Cauchy par un scalaire on obtient une suite de Cauchy. Donc les suites de Cauchy forment un sous-espace vectoriel noté Y de L. On peut vérifier que Y est un sous-espace fermé de L. Posons

$$\mathrm{Z} = \{y = (y_n) \in \mathrm{Y} : \lim \|y_n\| = 0\};$$

on vérifie que Z, muni de la norme $\|x\|_\infty$, est un sous-espace vectoriel fermé de Y ; notons $\mathrm{E} = \mathrm{Y}/\mathrm{Z}$ le quotient de Y par Z et $\pi : \mathrm{Y} \to \mathrm{E}$ l'application quotient. Notons q la norme quotient sur E.

Pour $x \in \mathrm{X}$, notons $\mathrm{U}(x) \in \mathrm{Y}$ la suite constante égale à x ; posons $u = \pi \circ \mathrm{U}$. Si $y = (y_n)$ est un représentant quelconque de $u(x)$, on a par définition $y - \mathrm{U}(x) \in \mathrm{Z}$, c'est à dire $\lim_n \|y_n - x\| = 0$; il en résulte que $\|y\|_\infty = \sup_n \|y_n\| \geq \lim_n \|y_n\| = \|x\|$, donc $q(u(x)) \geq \|x\|$, mais le choix du représentant $y = \mathrm{U}(x)$ donne $\|y\|_\infty = \|x\|$, donc $q(u(x)) = \|x\|$ pour tout $x \in \mathrm{X}$. Le fait que $u(\mathrm{X})$ est dense dans E résulte immédiatement du lemme suivant :

Lemme 1.4.2. *Soit $y = (y_n)_{n\in\mathbb{N}}$ un élément de* Y *; alors*

$$\pi(y) = \lim_{n\to\infty} u(y_n).$$

Démonstration. Soit $\varepsilon > 0$; il existe un entier $n \geq 0$ tel que, pour tous $k, \ell \geq n$ on ait $\|y_k - y_\ell\| \leq \varepsilon$. Fixons un entier $k \geq n$, et considérons l'élément y' de Y défini par $y'_j = y_j$ si $j < k$ et $y'_j = y_k$ si $j \geq k$. Ce que nous avons dit se traduit par $\|y - y'\|_\infty \leq \varepsilon$, donc $q(\pi(y) - \pi(y')) \leq \varepsilon$, mais y' est un représentant de $u(y_k)$; on a donc : pour tout $k \geq n$, $q(\pi(y) - u(y_k)) \leq \varepsilon$. Le lemme en résulte.

∎

Pour finir la démonstration du point (i) du théorème 1, il nous reste à démontrer que E est un espace de Banach. Soit (ξ_n) une suite de Cauchy dans E ; comme $u(\mathrm{X})$ est dense, pour tout entier $n \geq 0$, il existe $x_n \in \mathrm{X}$ tel que $q(\xi_n - u(x_n)) \leq 2^{-n}$. La suite $(u(x_n))_{n\in\mathbb{N}}$ est alors de Cauchy et, comme u est isométrique, la suite $\eta = (x_n)_{n\in\mathbb{N}}$ est de Cauchy dans X ; par le lemme 2, la suite $(u(x_n))$ converge vers $\pi(\eta)$; comme la suite $(\xi_n - u(x_n))$ converge vers 0, la suite (ξ_n) converge vers $\pi(\eta)$, d'où (i).

Le point (ii) résulte immédiatement du lemme 1. Montrons le point (iii) : par (ii), il existe $v \in \mathcal{L}(\mathrm{E}_1, \mathrm{E}_2)$ telle que $u_2 = v \circ u_1$. Echangeant les rôles de E_1 et E_2, on en déduit qu'il existe $w \in \mathcal{L}(\mathrm{E}_2, \mathrm{E}_1)$ telle que $u_1 = w \circ u_2$. Alors $w \circ v \circ u_1 = w \circ u_2 = u_1$; il s'ensuit que $w \circ v$ est l'identité sur l'image de u_1, et comme cette image est dense (et E_1 séparé), $w \circ v$ est l'identité de E_1. De même, $v \circ w$ est l'identité de E_2. Enfin, par (ii), $\|v\| = \|u_2\| \leq 1$ et $\|w\| = \|u_1\| \leq 1$, donc pour tout $x \in \mathrm{E}_1$, $\|v(x)\| \leq \|x\| = \|w(v(x))\| \leq \|v(x)\|$, donc v est isométrique.

∎

On appelle *complété* de X un couple (E, u) où E est un espace de Banach et $u : \mathrm{X} \to \mathrm{E}$ est une application linéaire isométrique d'image dense. Le plus souvent, on ne mentionne pas l'application u : on dit alors que E est un *complété* de X (ou **le** *complété* de X).

Séparé complété

Si X est un espace muni d'une semi-norme p qui n'est pas une norme, on peut lui associer de façon naturelle un espace normé ; on voit que

$$\mathrm{Z} = \{z \in \mathrm{X} : p(z) = 0\}$$

est un sous-espace vectoriel de X. On peut donc considérer le quotient (algébrique) X/Z. Si ξ est un élément de X/Z et x un représentant quelconque de ξ, on voit que la quantité $p(x)$ ne dépend pas du représentant choisi (si x' est un autre représentant de ξ, on aura $x - x' \in \mathrm{Z}$ et $|p(x) - p(x')| \le p(x - x') = 0$). On peut donc poser $q(\xi) = p(x)$ où x est un représentant quelconque de la classe ξ. Il est facile de vérifier que q est une semi-norme sur X/Z, mais c'est en fait une *norme* : si $q(\xi) = 0$, cela signifie que $p(x) = 0$ pour tout représentant x de ξ, c'est à dire que $x \in \mathrm{Z}$, donc ξ est la classe Z qui est la classe nulle du quotient, donc $\xi = 0_{\mathrm{X/Z}}$.

Ce quotient X/Z s'appelle le *séparé* de X.

C'est exactement l'opération que l'on fait en Intégration, si on définit d'abord l'espace $\mathcal{L}_p$ des "vraies" fonctions mesurables f telles que

$$p(f) = \Bigl(\int_\Omega |f|^p \, d\mu \Bigr)^{1/p} < +\infty.$$

La fonction p est une semi-norme sur $\mathrm{X} = \mathcal{L}_p$, et l'espace Z est l'espace des *fonctions négligeables*. Le quotient X/Z est l'espace L_p des *classes* de fonctions de puissance $p$ième intégrable. On prend rapidement l'habitude de laisser tomber le mot "classe", mais il faut rester vigilant : si $f \in \mathrm{L}_p$ et si t_0 est un point **fixé** de l'espace mesuré, l'expression $f(t_0)$ n'a pas de sens !

1.5. Complexifié d'un espace normé réel

Si X est un espace vectoriel réel, on peut lui associer un espace vectoriel complexe $\mathrm{X}_\mathbb{C}$ d'une manière qui rappelle le passage de $\mathbb{R}$ à $\mathbb{C}$: l'espace $\mathrm{X}_\mathbb{C}$ sera l'espace des "vecteurs complexes" de la forme $z = x + iy$, où x et y sont deux "vecteurs réels", c'est à dire que $x, y \in \mathrm{X}$, et où il faut prendre pour l'instant iy comme une notation formelle ; on définira la somme de deux vecteurs complexes z et z' par $(x+iy)+(x'+iy') = (x+x')+i(y+y')$, et la multiplication du vecteur z par un nombre complexe $\lambda = a + ib$ sera définie par

$$\lambda z = (a + ib)(x + iy) = (ax - by) + i(bx + ay).$$

Pour rendre les choses plus "carrées", on va dire que $\mathrm{X}_\mathbb{C}$ est l'ensemble $\mathrm{X} \times \mathrm{X}$ (mais on pensera au couple $(x, y) \in \mathrm{X} \times \mathrm{X}$ comme étant le vecteur complexe $z = x+iy$). L'addition de l'espace vectoriel $\mathrm{X}_\mathbb{C}$ est l'addition naturelle $(x, y) + (x', y') = (x + x', y + y')$ (qui correspond à ce qui a été dit un peu plus haut) et la multiplication par les scalaires complexes est définie par la formule

$$(a + ib)(x, y) = (ax - by, bx + ay).$$

Résumons :

Définition 1.5.1. Soit X un espace vectoriel réel ; on appelle *complexifié* de X l'espace vectoriel complexe $\mathrm{X}_\mathbb{C}$ obtenu en munissant l'espace vectoriel réel $\mathrm{X} \times \mathrm{X}$ de la structure d'espace vectoriel complexe donnée par $(a + ib)(x, y) = (ax - by, ay + bx)$ (pour $a, b \in \mathbb{R}$

et $x, y \in \mathrm{X}$). On plonge X dans $\mathrm{X}_{\mathbb{C}}$ par l'application $u : x \to (x, 0)$, de sorte que (x, y) peut s'écrire $u(x) + iu(y)$.

Si l'espace réel X était de plus un espace normé, on a envie de définir sur $\mathrm{X}_{\mathbb{C}}$ une norme p qui soit une norme *d'espace complexe*, c'est à dire telle que $p(\lambda z) = |\lambda|\, p(z)$ si λ est un nombre complexe quelconque, mais si possible de façon que les "vecteurs réels" $z = (x, 0)$ gardent la même norme, c'est à dire que $p(x, 0) = \|x\|_{\mathrm{X}}$. On doit en particulier garantir que $p(\lambda z) = |\lambda|\, p(z)$ si λ est un nombre complexe tel que $|\lambda| = 1$. Il y a de nombreuses façons de procéder ; en voici une qui est assez simple : si $z = x + iy$, on dira que $x = \operatorname{Re} z$, et on posera $p(z) = \max\{\|\operatorname{Re}(\lambda z)\|_{\mathrm{X}} : \lambda \in \mathbb{C}, |\lambda| = 1\}$, ce qui s'écrit aussi

$$p(x + iy) = \max\{\|\cos(\theta)\, x - \sin(\theta)\, y\|_{\mathrm{X}} : \theta \in [0, 2\pi]\}.$$

Le procédé du paragraphe précédent peut s'appliquer lorsque X est déjà un espace vectoriel complexe, mais muni d'une norme p d'espace vectoriel réel : autrement dit, on a $p(\lambda x) = |\lambda|\, p(x)$ pour tout $x \in \mathrm{X}$ et $\lambda \in \mathbb{R}$, mais on n'a pas nécessairement cette égalité avec λ complexe. Si la multiplication par i, c'est à dire l'application $\varphi : x \to ix$, est continue sur (X, p), on peut remplacer p par une semi-norme q équivalente à p qui vérifie de plus $q(\lambda x) = |\lambda|\, q(x)$ pour tout $x \in \mathrm{X}$ et tout $\lambda \in \mathbb{C}$, en posant pour tout $x \in \mathrm{X}$

$$q(x) = \sup\{p(\lambda x) : \lambda \in \mathbb{C},\ |\lambda| = 1\}.$$

1.6. Le dual d'un espace normé

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 $\mathrm{X}^* = \mathcal{L}(\mathrm{X}, \mathbb{K})$. Cet espace est complet par la proposition 2.3.

Définition 1.6.1. Soient X et Y deux espaces normés et $f \in \mathcal{L}(\mathrm{X}, \mathrm{Y})$; on appelle *transposée topologique* de f (ou juste *transposée*) l'application ${}^t\!f : y^* \to y^* \circ f$ de Y^* dans X^*.

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

Proposition 1.6.1. *Soient* X, Y *et* Z *des espaces normés ;*

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

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

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

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

$$|{}^t\!f(y^*)(x)| = |y^*(f(x))| \le \|y^*\|\, \|f(x)\| \le \|y^*\|\, \|f\|\, \|x\| \le \|f\|,$$

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

∎

Soit X un espace normé complexe ; c'est, en particulier, un espace normé réel. Il y a deux notions distintes 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 $a + ib$ asssocie sa partie réelle a (pour $a, b \in \mathbb{R}$).

Proposition 1.6.2. *L'application $g \to \mathrm{Re} \circ g$ est une bijection isométrique de $X^*_{\mathbb{C}}$ sur l'espace $X^*_{\mathbb{R}}$.*

Démonstration. Soit $g \in X^*_{\mathbb{C}}$; comme $\|\mathrm{Re}\| = 1$, on a $\|\mathrm{Re} \circ g\| \le \|g\|$. 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 $|g(x)| = \lambda g(x) = g(\lambda x) = (\mathrm{Re} \circ g)(\lambda x) \le \|\mathrm{Re} \circ g\|$. Donc $\|g\| = \|\mathrm{Re} \circ g\|$.

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. Donc $g \to \mathrm{Re} \circ g$ est surjective. Comme elle est isométrique, elle est bijective. ∎

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

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

Si $f \in L_p(\Omega, \mu)$ et $g \in L_q(\Omega, \mu)$, 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 $M = \|f\|_\infty$ et g est intégrable ; le produit fg est mesurable et $|fg| \le M \, |g|$, donc fg est intégrable et

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

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

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

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

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

ce qui montre que fg est intégrable, et que

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

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 $|\int fg| \le 1/p + 1/q = 1$, ce qui est le résultat voulu.

∎

Corollaire 1.6.1. *Soient $p, q \in [1, +\infty]$ tels que $1/p + 1/q = 1$ et $x = (x_n) \in \ell_p$; on a*

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

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

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

Démonstration. L'inégalité de Hölder nous dit déjà que

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

le problème est de montrer l'autre direction. On va voir qu'en fait le *maximum* est atteint pour une certaine fonction $g \in \mathrm{L}_q$, $\|g\|_q \le 1$, lorsque $1 \le p < +\infty$. Si $f = 0$, le résultat est évident, on supposera donc $f \ne 0$, 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 $\mathrm{A} = \{s \in \Omega : \widetilde{f}(s) \ne 0\}$, et posons $g(s) = 0$ 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

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

Dans le cas $p = +\infty$, le maximum n'est pas nécessairement atteint. Si $\|f\|_\infty = 1$, l'ensemble mesurable $\mathrm{A} = \{s \in \Omega : |f(s)| > 1 - \varepsilon\}$ est de mesure > 0 pour tout $\varepsilon > 0$. On choisira $\varepsilon < 1$ et on prendra $g(s) = (\mu(\mathrm{A}))^{-1}|f(s)|/f(s)$ si $s \in \mathrm{A}$, $g(s) = 0$ sinon. Alors $\|g\|_1 = 1$ et $\int_\Omega fg\, d\mu > 1 - \varepsilon$.

∎

Corollaire 1.6.2. *Soient $p, q, r \in]0, +\infty]$ tels que $1/p + 1/q = 1/r$ et $f \in \mathrm{L}_p(\Omega, \mu)$, $g \in \mathrm{L}_q(\Omega, \mu)$; alors*

$$\Big(\int_\Omega |fg|^r\, d\mu\Big)^{1/r} \le \Big(\int_\Omega |f|^p\, d\mu\Big)^{1/p}\Big(\int_\Omega |g|^q\, d\mu\Big)^{1/q}.$$

De même, si $x \in \ell_p$ et $y \in \ell_q$,

$$\Big(\sum_{n=0}^{+\infty} |x_n y_n|^r\Big)^{1/r} \le \Big(\sum_{n=0}^{+\infty} |x_n|^p\Big)^{1/p}\Big(\sum_{n=0}^{+\infty} |y_n|^q\Big)^{1/q}.$$

Démonstration. Il suffit de poser $\mathrm{P} = p/r$, $\mathrm{Q} = q/r$. Alors $1/\mathrm{P} + 1/\mathrm{Q} = 1$ et on applique l'inégalité de Hölder précédente aux fonctions $\mathrm{F} = |f|^r$ et $\mathrm{G} = |g|^r$.

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 $p < +\infty$. Munissons ℓ_p de la norme $\|\ \|_p$ et c_0 de la norme $\|\ \|_\infty$. Notons $e_n \in \mathbb{K}^{\mathbb{N}}$ la suite dont le $n^{\text{ème}}$ terme vaut 1 et tous les autres sont nuls (pour $n \in \mathbb{N}$). On dit que $u \in \ell_p$ est à *support fini* si l'ensemble $\mathrm{A} \subset \mathbb{N}$ des indices n tels que $u_n \neq 0$ est fini ; dans ce cas on peut écrire $u = \sum_{n=0}^{+\infty} u_n e_n$, étant entendu que la somme est en fait une somme finie limitée à l'ensemble A.

Soit f une forme linéaire continue sur ℓ_p, et posons $v_k = f(e_k)$ pour tout $k \geq 0$; pour tout $u \in \ell_p$ à support fini, on a $f(u) = \sum_{k=0}^{+\infty} u_k f(e_k) = \sum_{k=0}^{+\infty} u_k v_k$. On aura si u est à support dans $\{0, \ldots, n\}$

$$\Bigl|\sum_{k=0}^{n} u_k v_k\Bigr| = |f(u)| \leq \|f\|\, \|u\|_p\,;$$

si on choisit $u_k = |v_k|^q / v_k$ si $v_k \neq 0$ et $0 \leq k \leq n$ et $u_k = 0$ sinon, on obtiendra

$$\sum_{k=0}^{n} |v_k|^q \leq \|f\| \Bigl(\sum_{k=0}^{n} |v_k|^q\Bigr)^{1/p},$$

ce qui montre que $\bigl(\sum_{k=0}^{n} |v_k|^q\bigr)^{1-1/p} \leq \|f\|$ pour tout n, donc

$$\Bigl(\sum_{k=0}^{+\infty} |v_k|^q\Bigr)^{1/q} \leq \|f\|.$$

La suite $v = (v_n)$ est donc dans ℓ_q. Pour $u \in \ell_p$ et $n \geq 0$, désignons par $w = \mathrm{P}_n(u)$ la suite telle que $w_j = u_j$ si $j \leq n$ et $w_j = 0$ si $j > n$. Si $p < +\infty$, le vecteur u est limite dans ℓ_p de la suite de ses projections $\mathrm{P}_n(u)$ qui sont à support fini, donc

$$f(u) = \lim_n f(\mathrm{P}_n(u)) = \lim_n \sum_{j=0}^{n} u_j v_j = \sum_{k=0}^{+\infty} u_k v_k = f_v(u).$$

En d'autres termes,

Théorème 1.6.2. *Si $1 \leq p < +\infty$ 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.*

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

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

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

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 $j_q : \mathrm{L}_q \to (\mathrm{L}_p)^*$.

Examinons maintenant la réciproque, dans le cas plus simple où $1 < p \le 2$ et où la mesure considérée est de masse totale 1 (mesure de probabilité), comme lorsque $\Omega = [0, 1]$ muni de la mesure de Lebesgue. On sait alors dans le cas $p \le 2$ que $\mathrm{L}_2 \subset \mathrm{L}_p$ et $\|f\|_p \le \|f\|_2$ pour toute $f \in \mathrm{L}_p$ (appliquer l'inégalité de Hölder du corollaire 2 avec $1/p = 1/2 + 1/q$, $f \in \mathrm{L}_2$ et $g = 1 \in \mathrm{L}_q$), donc une forme linéaire continue ℓ sur L_p définit par restriction une forme linéaire m continue sur L_2. On sait alors (voir aussi la proposition 4.2.1) qu'il existe une fonction $g \in \mathrm{L}_2$ telle que pour tout $f \in \mathrm{L}_2$ on ait

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

Prenons $f_n = 1_{\{0<|g|\le n\}} \, |g|^q/g$. Cette fonction est dans L_2 et

$$\ell(f_n) = m(f_n) = \int_{\{|g|\le n\}} |g|^q \, d\mu \le \|\ell\| \, \|f_n\|_p,$$

où $\|\ell\|$ est la norme de ℓ dans $(\mathrm{L}_p)^*$; comme $|f_n|^p = 1_{\{0<|g|\le n\}} \, |g|^q$, on a

$$\Bigl(\int_{\{|g|\le n\}} |g|^q \, d\mu\Bigr)^{1/q} \le \|\ell\|.$$

En faisant tendre n vers $+\infty$, on déduit que $\int |g|^q \, d\mu < +\infty$. Les deux formes linéaires ℓ et ℓ_g, continues sur L_p, coïncident par construction sur le sous-espace vectoriel L_2 de L_p, qui est un sous-espace dense dans L_p (le sous-espace L_2 contient le sous-espace des fonctions étagées, qui est déjà dense dans L_p). Il en résulte que $\ell = \ell_g$.

Le résultat reste valable lorsque la mesure est σ-finie, et aussi pour $1 \le p \le 2$:

Lorsque $1 \le p < +\infty$, *l'application* j_q *est une isométrie surjective de* L_q *sur le dual de* L_p.

En revanche, le dual de L_∞ est en général “plus grand” que L_1, c'est à dire qu'il existe des formes linéaires continues sur L_∞ qui ne peuvent pas être représentées sous la forme ℓ_f, $f \in \mathrm{L}_1$; cependant, on ne peut pas vraiment décrire explicitement une telle forme linéaire : leur existence découlera de l'axiome du choix (ou du lemme de Zorn), voir la section sur le théorème de Hahn-Banach.

Dual de C(K)

Pour décrire le dual de C(K) il faut introduire des objets qui n'ont peut-être pas été vus en Intégration, où on se limite souvent aux mesures positives. 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 ; plus précisément, il existe un ensemble $\mathrm{B} \in \mathcal{A}$ tel que

$$\forall \mathrm{A} \in \mathcal{A}, \quad \mu(\mathrm{A} \cap \mathrm{B}) \geq 0 \,; \quad \mu(\mathrm{A} \setminus \mathrm{B}) \leq 0.$$

On dit que B porte la partie positive de la mesure réelle μ, et on peut décomposer μ en posant $\mu_1(\mathrm{A}) = \mu(\mathrm{A} \cap \mathrm{B})$ et $\mu_2(\mathrm{A}) = -\mu(\mathrm{A} \setminus \mathrm{B})$ pour tout $\mathrm{A} \in \mathcal{A}$. On a alors $\mu = \mu_1 - \mu_2$ et μ_1, μ_2 sont deux mesures positives bornées. Avec cette décomposition, on définit la valeur absolue de μ, qui est la mesure positive $|\mu| = \mu_1 + \mu_2$. On peut définir $|\mu|$ directement par

$$|\mu(\mathrm{A})| = \sup\{\mu(\mathrm{A}_1) - \mu(\mathrm{A}_2)\}$$

où le sup porte sur toutes les partitions de $\mathrm{A} \in \mathcal{A}$ en deux ensembles $\mathrm{A}_1, \mathrm{A}_2 \in \mathcal{A}$. On posera $\|\mu\| = |\mu(\Omega)|$. Cette expression définit une norme sur l'espace vectoriel des mesures réelles sur $(\Omega, \mathcal{A})$.

Si K est un espace métrique compact, les boules ouvertes de K engendrent une tribu qui s'appelle la *tribu borélienne* $\mathcal{B}$ de K. Les formes linéaires sur C(K) s'identifient aux mesures réelles sur $(\mathrm{K}, \mathcal{B})$ dans le cas réel, aux mesures complexes dans le cas complexe, et la norme de dual de C(K) est la norme de mesure définie plus haut. La dualité s'exprime par l'intégration d'une fonction continue par rapport à la mesure réelle (ou complexe) μ,

$$\ell(f) = \int_{\mathrm{K}} f(s)\, d\mu(s).$$

Ce théorème est assez long à démontrer (voir Rudin par exemple). On indiquera seulement le premier pas, qui consiste en général à trouver d'abord la mesure positive $|\mu|$: on se place dans le cas réel, on suppose donnée une forme linéaire continue ℓ sur C(K) ; on commence par définir la $|\mu|$-mesure d'un ouvert ω de K en posant

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

Mentionnons un cas évident, et qui a une certaine importance théorique, le cas des *mesures de Dirac*. Fixons un point x_0 de K. On peut lui associer une forme linéaire sur C(K) qui est *l'évaluation au point* x_0, notée δ_{x_0}, qui est définie par $\delta_{x_0}(f) = f(x_0)$ pour toute fonction continue f sur K. Considérons maintenant une mesure μ sur K, qui est définie pour tout borélien $\mathrm{A} \subset \mathrm{K}$ en disant que $\mu(\mathrm{A}) = 1$ si $x_0 \in \mathrm{A}$ et $\mu(\mathrm{A}) = 0$ sinon. La mesure μ consiste à placer une masse 1 au point x_0. La théorie de l'intégration par rapport à cette mesure est un peu bizarre, mais le lecteur pourra se convaincre que

$$\int_{\mathrm{K}} f(s)\, d\mu(s) = f(x_0) = \delta_{x_0}(f)$$

pour toute fonction continue f. On note en fait $\mu = \delta_{x_0}$, et on dit que δ_{x_0} est la mesure de Dirac au point x_0.

2. Les théorèmes fondamentaux

2.1. Le théorème de Baire et ses conséquences

Commençons par rappeler le théorème de Baire.

Théorème 2.1.1 : théorème de Baire. *Soit* X *un espace métrique complet ; l'intersection d'une famille dénombrable de parties ouvertes et denses de* X *est dense dans* X.

Démonstration. Soit $(\mathrm{U}_n)_{n\in\mathbb{N}}$ une suite d'ouverts denses de X ; soit V une partie ouverte non vide de X ; on doit montrer que $\bigcap_{n\in\mathbb{N}} \mathrm{U}_n$ rencontre V. Comme U_0 est dense, U_0 rencontre V et on peut choisir un point $x_0 \in \mathrm{V} \cap \mathrm{U}_0$. Comme $\mathrm{V} \cap \mathrm{U}_0$ est ouvert, il existe un nombre $r_0 > 0$, que l'on peut choisir ≤ 1, tel que la boule ouverte $\mathrm{B}(x_0, 2r_0)$ de centre x_0 et de rayon $2r_0$ soit contenue dans $\mathrm{V} \cap \mathrm{U}_0$.

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

$$\mathrm{B}_{n+1} = \overline{\mathrm{B}(x_{n+1}, r_{n+1})} \subset \mathrm{B}(x_{n+1}, 2r_{n+1}) \subset \mathrm{B}(x_n, r_n) \subset \mathrm{B}_n.$$

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 $\bigcap_{n\in\mathbb{N}} \mathrm{B}_n \neq \emptyset$; or, par construction, $\bigcap_{n\in\mathbb{N}} \mathrm{B}_n \subset \mathrm{V} \cap \bigcap_{n\in\mathbb{N}} \mathrm{U}_n$, ce qui montre que cette dernière intersection est non vide.

■

Corollaire 2.1.1. *Soient* X *un espace métrique complet* non vide *et* $(\mathrm{F}_n)_{n\in\mathbb{N}}$ *une suite de parties fermées de* X *telle que* $\bigcup_{n\in\mathbb{N}} \mathrm{F}_n = \mathrm{X}$ *; alors l'un des fermés* F_n *a un intérieur non vide (et en réalité, on peut même dire que* $\bigcup_{n\in\mathbb{N}} \mathring{\mathrm{F}}_n$ *est dense dans* X*).*

Démonstration. Si F_n est un fermé d'intérieur vide dans X, son complémentaire U_n est un ouvert dense dans X. Mais si chaque F_n était d'intérieur vide et $\bigcup_n \mathrm{F}_n = \mathrm{X}$, on aurait une suite (U_n) d'ouverts denses telle que $\bigcap_n \mathrm{U}_n = \emptyset$, ce qui est impossible d'après le théorème 1.

Le lecteur pourra obtenir le petit raffinement entre parenthèses en améliorant l'argument ci-dessus ; une autre méthode consiste à montrer qu'une boule ouverte $\mathrm{U} = \mathrm{B}(x, r)$ d'un espace métrique complet (X, d) peut être munie d'une nouvelle métrique δ qui ne change pas la topologie, et qui fait de (U, δ) un espace métrique complet (exercice) ; pour obtenir l'énoncé raffiné, il suffit d'appliquer la première partie à l'espace métrique complet (U, δ), pour toute boule ouverte U de (X, d).

■

Théorème 2.1.2 : théorème des isomorphismes. *Soient* E *et* F *deux espaces de Banach ; toute application linéaire continue bijective de* E *sur* F *est un isomorphisme.*

Démonstration. Soit f une bijection linéaire continue de E sur F ; notons $\mathrm{B_E}$ la boule unité de E. La première étape consiste à montrer que l'adhérence de $f(\mathrm{B_E})$ contient une boule ouverte centrée en 0,

$$(*) \qquad \exists r > 0, \quad \mathrm{B_F}(0, r) \subset \overline{f(\mathrm{B_E})}.$$

On a $\bigcup_{n\geq 1} n\mathrm{B_E} = \mathrm{E}$ donc $\bigcup_{n\geq 1} nf(\mathrm{B_E}) = f(\mathrm{E}) = \mathrm{F}$, donc $\bigcup_{n\geq 1} \overline{nf(\mathrm{B_E})} = \mathrm{F}$. Par le corollaire 1, il existe un entier $n \geq 1$ tel que l'ensemble fermé $\overline{nf(\mathrm{B_E})}$ soit d'intérieur non vide. Comme la multiplication par $n \geq 1$ est un homéomorphisme, on en déduit que l'intérieur de $\overline{f(\mathrm{B_E})}$ n'est pas vide. Il reste seulement à voir que 0_F est dans cet intérieur : cela provient de la convexité et de la symétrie de l'ensemble $\mathrm{C} = \overline{f(\mathrm{B_E})}$: si y_0 est intérieur à C, il existe $r > 0$ tel que $\mathrm{B}(y_0, r) \subset \mathrm{C}$; par symétrie on a aussi $\mathrm{B}(-y_0, r) \subset \mathrm{C}$, et par convexité $\mathrm{B}(0, r) \subset \mathrm{C}$. Détaillons un peu : soit $v \in \mathrm{F}$ tel que $\|v\| < r$ et soit $\varepsilon > 0$ quelconque ; puisque $y_0 + v \in \mathrm{B}(y_0, r) \subset \overline{f(\mathrm{B_E})}$, il existe un vecteur $u_1 \in \mathrm{B_E}$ tel que $\|f(u_1) - (y_0 + v)\| < \varepsilon$, et de même il existe $u_2 \in \mathrm{B_E}$ tel que $\|f(u_2) - (y_0 - v)\| < \varepsilon$. Alors $u = \frac{1}{2}(u_1 - u_2) \in \mathrm{B_E}$ et $\|f(u) - v\| < \varepsilon$. On a bien montré que $\mathrm{B}(0, r) \subset \overline{f(\mathrm{B_E})}$.

La deuxième étape consiste à montrer que la propriété $(*)$ entraîne en fait lorsque E est complet que

$$(**) \qquad \mathrm{B_F}(0, r) \subset f(\mathrm{B_E}).$$

Le principe est intéressant et se rencontre à de multiples occasions ; c'est une variante du procédé des approximations successives. Si A_1 et A_2 sont deux sous-ensembles non vides de F, la notation $\mathrm{A}_1 + \mathrm{A}_2$ désigne l'ensemble de toutes les sommes $a_1 + a_2$, lorsque a_1 varie dans A_1 et a_2 dans A_2. On dit que $\mathrm{A}_1 + \mathrm{A}_2$ est la *somme de Minkowski* des ensembles A_1 et A_2.

Lemme 2.1.1. *On suppose que* E *est complet,* $f : \mathrm{E} \to \mathrm{F}$ *linéaire continue, et que* A *est un sous-ensemble borné de* F *tel que* $\mathrm{A} \subset f(\mathrm{B_E}) + \varepsilon\mathrm{A}$, *avec* $0 < \varepsilon < 1$. *Alors*

$$\mathrm{A} \subset \frac{1}{1-\varepsilon} f(\mathrm{B_E}).$$

Soit $a_0 \in \mathrm{A}$; d'après l'hypothèse, il existe un vecteur $u_0 \in \mathrm{B_E}$ et un vecteur $a_1 \in \mathrm{A}$ tels que $a_0 = f(u_0) + \varepsilon a_1$. En recommençant avec a_1, on trouve $u_1 \in \mathrm{B_E}$ et $a_2 \in \mathrm{A}$ tels que $a_1 = f(u_1) + \varepsilon a_2$, ce qui donne

$$a_0 = f(u_0 + \varepsilon u_1) + \varepsilon^2 a_2.$$

En continuant ainsi on construit des vecteurs $u_0, u_1, \ldots, u_n, \ldots$ dans la boule unité de E et $a_1, \ldots, a_n, \ldots$ dans A tels que pour tout entier $k \geq 0$ on ait

$$(1) \qquad a_0 = f(u_0 + \varepsilon u_1 + \cdots + \varepsilon^k u_k) + \varepsilon^{k+1} a_{k+1}.$$

La série $\sum \varepsilon^k u_k$ est normalement convergente, donc convergente dans E puisque E est complet ; sa somme $x_0 = \sum_{k=0}^{+\infty} \varepsilon^k u_k \in \mathrm{E}$ est telle que $f(x_0) = \lim_k f(u_0 + \varepsilon u_1 + \cdots + \varepsilon^k u_k)$ et d'après la relation (1), on a $f(x_0) = a_0$ puisque A est borné. Par ailleurs

$$\|x_0\| \leq \sum_{k=0}^{+\infty} \varepsilon^k \|u_k\| \leq \frac{1}{1-\varepsilon}.$$

Finissons de montrer que $(*)$ entraîne $(**)$ lorsque E est complet. Si $y \in \mathrm{F}$ est tel que $\|y\| < r$, on pourra choisir $\varepsilon > 0$ de façon que le vecteur $y_0 = r\,(1-\varepsilon)^{-1}\,y$ vérifie encore $\|y_0\| < r$; puisque $\mathrm{B_F}(0, r)$ est contenu dans l'adhérence de $f(\mathrm{B_E})$, on a $\mathrm{B_F}(0, r) \subset f(\mathrm{B_E}) + \varepsilon \mathrm{B_F}(0, r)$. D'après le lemme précédent, on a $\mathrm{B_F}(0, r) \subset (1-\varepsilon)^{-1} f(\mathrm{B_E})$. Il existe donc $x_0 \in \mathrm{E}$ tel que $\|x_0\| < (1-\varepsilon)^{-1}$ et $f(x_0) = y_0$; en revenant à y, on aura trouvé $x \in \mathrm{E}$ tel que $\|x\| < 1$ et $f(x) = y$. On a ainsi montré que $\mathrm{B_F}(0, r) \subset f(\mathrm{B_E})$.

Il est clair maintenant que l'application réciproque f^{-1} est continue : en effet, l'inclusion précédente se traduit par $f^{-1}(\mathrm{B_F}(0, r)) \subset \mathrm{B_E}$, ce qui implique $f^{-1}(\mathrm{B_F}) \subset r^{-1}\mathrm{B_E}$ par homogénéité et passage à l'adhérence ; ceci montre que $\|f^{-1}\| \le 1/r$.

■

Remarque 2.1.1. Le théorème précédent a peu à voir avec les normes. Il reste vrai si E et F sont deux espaces vectoriels munis de métriques vérifiant les propriétés (*i*), (*ii*) et (*iii*) de la remarque 1.1.1, et si E et F sont **complets** pour ces métriques. En particulier le théorème des isomorphismes, et ses conséquences immédiates que nous allons énoncer dans ce qui suit, sont valables pour les espaces de Fréchet.

On dit qu'une application $f : \mathrm{X} \to \mathrm{Y}$ est *ouverte* lorsque l'image de tout ouvert de X est ouverte dans Y ; cela revient exactement à dire que pour tout point $x \in \mathrm{X}$ et tout voisinage V de x, l'image $f(\mathrm{V})$ est un voisinage de $f(x)$ dans Y.

Pour qu'une application linéaire f d'un espace normé X dans un autre Y soit ouverte, il suffit de savoir que l'image de toute boule ouverte $\mathrm{B_X}(0, r)$ est ouverte dans Y, ce qui se ramène à savoir que l'image de la boule unité ouverte $\mathrm{B_X}(0, 1)$ de X est ouverte dans Y. C'est le cas par exemple pour la projection canonique π d'un espace normé X sur un quotient X/Y par un sous-espace fermé Y ; en effet, l'image de $\mathrm{B_X}(0, 1)$ est exactement la boule unité ouverte du quotient, donc π est ouverte. Notons aussi qu'un isomorphisme entre deux espaces de Banach est une application ouverte, et que la composition de deux applications ouvertes est une application ouverte.

Théorème 2.1.3 : 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 $f = g \circ \pi$

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

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 2.1.4 : 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* $\mathrm{E} \times \mathrm{F}$*) est continue.*

Démonstration. Soit f une application linéaire de E dans F dont le graphe $\mathrm{G} \subset \mathrm{E} \times \mathrm{F}$ 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 $x \in \mathrm{E}$ unique ; notons $p : \mathrm{G} \to \mathrm{E}$ l'application définie par $p(z) = p(x, f(x)) = x \in \mathrm{E}$. 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.

■

Théorème 2.1.5 : théorème de Banach-Steinhaus. *Soient* E *un espace de Banach,* F *un espace normé et* A *une partie de* $\mathcal{L}(\mathrm{E}, \mathrm{F})$*, telle que, pour tout* $x \in \mathrm{E}$*, le sous-ensemble* $\{\|f(x)\| : f \in \mathrm{A}\}$ *soit borné (dans* $\mathbb{R}$*). Alors* $\{\|f\| : f \in \mathrm{A}\}$ *est borné.*

Démonstration. Pour tout entier $n \geq 1$, posons $\mathrm{C}_n = \{x \in \mathrm{E} : \forall f \in \mathrm{A},\ \|f(x)\| \leq n\}$. Comme C_n est l'intersection des fermés $\mathrm{C}_{f,n} = \{x \in \mathrm{E} : \|f(x)\| \leq n\}$ (lorsque f varie dans A), c'est un fermé de E. La réunion des C_n est E ; par le corollaire 1, il existe un entier $n \geq 1$ tel que $\mathring{\mathrm{C}}_n$ ne soit pas vide. On en déduit comme avant (démonstration du théorème des isomorphismes) que $\mathring{\mathrm{C}}_1 \neq \emptyset$, puis que 0_E est dans l'intérieur de C_1.

Il existe donc $r > 0$ tel que $\mathrm{B}(0, r) \subset \mathrm{C}_1$. Par homogénéité, pour tout $x \in \mathrm{B}(0, 1)$ et tout $f \in \mathrm{A}$, on a $\|f(x)\| \leq 1/r$; on a donc montré que pour tout $f \in \mathrm{A}$, on a $\|f\| \leq 1/r$.

■

Corollaire 2.1.2. *Soient* E *un espace de Banach,* F *un espace normé et* (f_n) *une suite d'applications linéaires continues de* E *dans* F *; on suppose que, pour tout* $x \in \mathrm{E}$*, la suite* $(f_n(x))$ *converge ; notons* $f(x)$ *sa limite. Alors* f *est (linéaire et) continue.*

Démonstration. Soit $x \in \mathrm{E}$; comme la suite $(f_n(x))$ est convergente, elle est bornée ; par le théorème 5, la suite $(\|f_n\|)$ est alors bornée. Il existe alors un nombre $k \geq 0$ tel que, pour tout $x \in \mathrm{E}$ et tout entier $n \geq 0$ on ait $\|f_n(x)\| \leq k\,\|x\|$. Passant à la limite on trouve $\|f(x)\| \leq k\,\|x\|$. Par ailleurs, il est clair que f est linéaire.

■

2.2. 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 $q(x + y) \leq q(x) + q(y)$.

Les semi-normes sont des fonctions sous-linéaires. Un autre exemple important de fonction sous-linéaire est donné par la *jauge* d'un ensemble convexe C contenant 0_X comme *point interne*, ce qui signifie que pour tout $x \in \mathrm{X}$ il existe $\varepsilon > 0$ tel que le segment $[-\varepsilon\, x, \varepsilon\, x]$ soit contenu dans C. On pose alors

$$j_\mathrm{C}(x) = \inf\{\lambda : \lambda > 0 \ \text{ et } \ \lambda^{-1} x \in \mathrm{C}\}.$$

On note que $j_\mathrm{C}(x)$ est un nombre fini ≥ 0. Il est possible que $j_\mathrm{C}(x) = 0$, dans le cas où la demi-droite de direction x est contenue dans C. On a les inclusions

$$\{x \in \mathrm{X} : j_\mathrm{C}(x) < 1\} \subset \mathrm{C} \subset \{x \in \mathrm{X} : j_\mathrm{C}(x) \leq 1\}.$$

Dans le cas où X est normé et où C est un convexe ouvert de X contenant 0_X, le point 0_X est interne à C et $\mathrm{C} = \{x \in \mathrm{X} : j_\mathrm{C}(x) < 1\}$.

Théorème 2.2.1 : 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* $\ell(y) \le q(y)$ *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* $m(x) \le q(x)$ *pour tout* $x \in \mathrm{X}$.

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 Z de façon que $m \le q$ et si $x \notin \mathrm{Z}$, on peut étendre m en $\widetilde{m}$ définie sur $\mathrm{Z} + \mathbb{R}x$ en gardant $\widetilde{m} \le q$; le reste n'est que formalité “zornique”.

Lemme 2.2.1. *Soient* Z *un sous-espace vectoriel de* X *et* g *une forme linéaire définie sur* Z*, telle que* $g(z) \le q(z)$ *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* $\widetilde{g} \le q$ *sur* $\mathrm{Z} + \mathbb{R}x$.

Démonstration du lemme. Bien entendu, prolonger g à $\mathrm{Z} + \mathbb{R}x$ demande seulement de définir $\mathrm{M} = \widetilde{g}(x)$. Pour que le prolongement soit convenable, il faut 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 $t \ne 0$:

$$g(z) + \lambda \mathrm{M} \le q(z + \lambda x), \quad g(z') - \mu \mathrm{M} \le q(z' - \mu x)$$

pour tous $z, z' \in \mathrm{Z}$, $\lambda, \mu > 0$. Le nombre M doit donc vérifier les deux inégalités

$$\mathrm{M} \le \mathrm{I} = \inf\{\lambda^{-1}(q(z + \lambda x) - g(z)) : z \in \mathrm{Z}, \lambda > 0\}$$

et

$$\sup\{\mu^{-1}(g(z') - q(z' - \mu x)) : z' \in \mathrm{Z}, \mu > 0\} = \mathrm{S} \le \mathrm{M}.$$

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 M 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 M n'importe quel nombre compris entre le sup et l'inf (bien sûr, si $\mathrm{S} = \mathrm{I}$ on n'a pas le choix : il faut prendre pour M la valeur commune). Il reste donc à vérifier que

$$\mu^{-1}(g(z') - q(z' - \mu x)) \le \lambda^{-1}(q(z + \lambda x) - g(z))$$

pour tous $z, z' \in \mathrm{Z}$ et $\lambda, \mu > 0$. On réécrit la propriété voulue sous la forme

$$\frac{\lambda}{\lambda + \mu}\,(g(z') - q(z' - \mu x)) \le \frac{\mu}{\lambda + \mu}\,(q(z + \lambda x) - g(z))$$

ou encore en posant $t = \mu/(\lambda + \mu)$

$$g((1 - t)\,z' + t\,z) \le (1 - t)\,q(z' - \mu x) + t\,q(z + \lambda x).$$

Puisque $g \le q$,

$$g((1 - t)\,z' + t\,z) \le q((1 - t)z' + tz) = q((1 - t)(z' - \mu x) + t(z + \lambda x)) \le$$
$$q((1 - t)(z' - \mu x)) + q(t(z + \lambda x)) = (1 - t)\,q(z' - \mu x) + t\,q(z + \lambda x).$$

Le résultat du lemme est donc établi.

Venons-en à l'application du lemme de Zorn pour terminer la démonstration du théorème. Notons Φ l'ensemble des couples (Z, g) où Z est un sous-espace vectoriel de X contenant Y et g une forme linéaire sur Z dont la restriction à Y soit ℓ et telle que $g \le q$. Munissons Φ de la relation d'ordre $\prec$ donnée par :

$$(Z_1, g_1) \prec (Z_2, g_2) \iff Z_1 \subset Z_2 \text{ et } \forall x \in Z_1 ,\ g_2(x) = g_1(x).$$

L'ordre $\prec$ est inductif : soit $\{(Z_i, g_i) : i \in I\}$ une partie totalement ordonnée de Φ ; posons $Z = \bigcup_{i \in I} Z_i$ et notons $g : Z \to \mathbb{R}$ l'application qui à $x \in Z_i$ associe $g_i(x)$ (si I n'est pas vide ; si I est vide, on pose $(Z, g) = (Y, \ell)$). On vérifie sans peine que

– g est bien défini : si $x \in Z_i$ et $x \in Z_j$ on peut supposer, quitte à intervertir i et j que $(Z_i, g_i) \prec (Z_j, g_j)$; alors $g_i(x) = g_j(x)$.

– Z est un sous-espace vectoriel de X et g est linéaire : si $x \in Z_i$, $y \in Z_j$, alors $Z_i \cup Z_j = Z_k$ où $k \in \{i, j\}$; donc, $\forall (s,t) \in \mathbb{R}^2$, $sx + ty \in Z_k \subset Z$ et $g(sx + ty) = g_k(sx + ty) = sg(x) + tg(y)$.

– $g \le q$: pour tout $x \in Z$ il existe $i \in I$ tel que $x \in Z_i$; alors $g(x) = g_i(x) \le q(x)$.

Donc $(Z, g) \in \Phi$. Il est clair que (Z, g) majore $\{(Z_i, g_i) : i \in I\}$, et on a montré que l'ordre sur Φ est inductif, ce qui permet d'appliquer le lemme de Zorn : soit donc $(Z, g) \in \Phi$ un élément maximal ; si nous montrons que $Z = X$ le théorème sera démontré. Mais si $Z \ne X$, on peut trouver un vecteur $x \notin Z$, donc $Z \ne Z + \mathbb{R}x$, et on a vu dans l'étape préliminaire qu'on peut trouver une extension convenable $(Z + \mathbb{R}x, \widetilde{g})$, qui contredirait la maximalité de (Z, g). On a donc bien $Z = X$ et le théorème est démontré.

On peut préférer un énoncé plus "affine" pour le théorème précédent : *si f est une fonction convexe (partout finie) définie sur* X*, si* Y *est un sous-espace affine de* X *et a une fonction affine définie sur* Y *et telle que $a(y) \le f(y)$ pour tout $y \in Y$, il existe un prolongement $\widetilde{a}$ de a en fonction affine sur* X *et telle que $\widetilde{a} \le f$ sur* X. Le lecteur pourra adapter la démonstration précédente pour obtenir ce nouvel énoncé.

Théorème 2.2.2 : 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* **non nulle** *sur* X *telle que*

$$\sup f(A) \le \inf f(B).$$

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

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 $B \subset H_+$.

Démonstration. On commence par le cas où B est réduit à un seul point $b_0 \ne 0$, et où le convexe ouvert A contient 0_X. La jauge j_A est alors une fonction sous-linéaire sur X, et $j_A(b_0) \ge 1$ puisque $b_0 \notin A$. On définit ℓ sur $Y = \mathbb{R}b_0$ en posant $\ell(\lambda b_0) = \lambda$ pour tout $\lambda \in \mathbb{R}$. On vérifie que $\ell \le p = j_A$ sur Y. Si m est un prolongement de ℓ, on aura $m(a) \le j_A(a) < 1$ pour tout $a \in A$, alors que $m(b_0) = \ell(b_0) = 1$. Par ailleurs, m est une forme linéaire continue puisqu'elle est majorée par 1 sur le voisinage A de 0_X, donc minorée par -1 sur $-A$ et bornée par 1 sur le voisinage $A \cap (-A)$.

Pour démontrer le cas général, on commence par remarquer que par translation, on peut se débarrasser de la condition $0_X \in A$ dans ce qui précède. Si A est un convexe ouvert non vide disjoint du convexe non vide B, on introduit le convexe ouvert

$$C = A - B = \{a - b : a \in A, b \in B\}.$$

Puisque A et B sont disjoints, ce convexe est disjoint de $\mathrm{D} = \{0_\mathrm{X}\}$. On peut donc trouver une forme linéaire m telle que $m(c) < m(d_0)$ pour tout $c \in \mathrm{C}$, où $d_0 = 0_\mathrm{X}$ est l'unique élément de D. Cela signifie que $m(a-b) < m(d_0) = 0$ pour tous $a \in \mathrm{A}$, $b \in \mathrm{B}$, soit encore $m(a) < m(b)$ pour tous $a \in \mathrm{A}$, $b \in \mathrm{B}$.

Rappelons que $\mathbb{K}$ désigne le corps $\mathbb{R}$ ou $\mathbb{C}$, et X^* le dual topologique de X.

Théorème 2.2.3 : théorème de Hahn-Banach. *Soient* X *un espace normé (réel ou complexe) et* Y *un sous-espace vectoriel de* X *; pour tout* $\ell \in \mathrm{Y}^*$*, il existe* $m \in \mathrm{X}^*$ *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 sera un multiple convenable de la norme N de X. Par définition de la norme de la forme linéaire ℓ, on a $\ell \le \|\ell\|\,\mathrm{N} = q$ sur le sous-espace 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 1.6.2, 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 $m(y) = m_1(y) - im_1(iy)$, et alors $m(y) = \ell_1(y) - i\ell_1(iy) = \ell(y)$).

Corollaire 2.2.1. *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, considérons le sous-espace vectoriel $\mathrm{Y} = \mathbb{K}x$ de E. On définit une forme linéaire $y^* \in \mathrm{Y}^*$ telle que $y^*(x) = \|x\|$ en posant $y^*(\lambda x) = \lambda\|x\|$, pour tout $\lambda \in \mathbb{K}$. Pour tout $y = \lambda x \in \mathrm{Y}$, on a $|y^*(y)| = |y^*(\lambda x)| = |\lambda|\,\|x\| \le \|y\|$, donc $\|y^*\| \le 1$, et $0 < \|x\| = y^*(x) \le \|y^*\|\,\|x\|$, donc $\|y^*\| = 1$. Par le théorème 3, il existe $x^* \in \mathrm{E}^*$ qui prolonge y^* et tel que $\|x^*\| = \|y^*\|$. Comme x^* prolonge y^*, on a $x^*(x) = y^*(x) = \|x\|$, d'où le résultat.

Bien entendu, on a en fait $\|x^*\| = 1$ lorsque $x \ne 0$, mais le corollaire tel qu'il est énoncé a l'avantage de couvrir tous les cas.

∎

Corollaire 2.2.2. *Soient* X *et* Y *deux espaces normés ; pour tout* $f \in \mathcal{L}(\mathrm{X}, \mathrm{Y})$*, on a* $\|{}^t\!f\| = \|f\|$.

Démonstration. On sait déjà que $\|{}^t\!f\| \le \|f\|$ par la proposition 1.6.1, nous allons montrer l'égalité. Pour tout $\varepsilon > 0$, on peut trouver un vecteur $x \in \mathrm{X}$ tel que $\|x\| \le 1$ et tel que $\|f(x)\| > \|f\| - \varepsilon$, puis une forme linéaire $y^* \in \mathrm{F}^*$ telle que $\|y^*\| \le 1$ et $y^*(f(x)) = \|f(x)\|$. Alors

$$\|{}^t\!f\| \ge \|{}^t\!f(y^*)\| \ge |{}^t\!f(y^*)(x)| = y^*(f(x)) = \|f(x)\| > \|f\| - \varepsilon.$$

∎

Corollaire 2.2.3. *Soient* X *un espace normé et* Y *un sous-espace fermé ; alors il existe une partie* A *de* X^* *telle que* $\mathrm{Y} = \bigcap_{\ell \in \mathrm{A}} \ker(\ell)$.

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. Posons $\mathrm{A} = \{\ell \in \mathrm{X}^* : \mathrm{Y} \subset \ker(\ell)\}$. Pour tout $\ell \in \mathrm{A}$, on a $\mathrm{Y} \subset \ker(\ell)$; donc $\mathrm{Y} \subset \bigcap_{\ell \in \mathrm{A}} \ker(\ell)$. Inversement, soit $x \in \mathrm{X} \setminus \mathrm{Y}$; notons $f : \mathrm{X} \to \mathrm{X}/\mathrm{Y}$ l'application quotient. Alors $f(x)$ n'est pas nul ; par le corollaire 1, il existe $\ell \in (\mathrm{X}/\mathrm{Y})^*$ telle que $\ell(f(x)) \neq 0$. Alors $\ell \circ f \in \mathrm{A}$ et $x \notin \ker \ell \circ f$. On a montré que $\bigcap_{\ell \in \mathrm{A}} \ker(\ell) \subset \mathrm{Y}$.

Soit maintenant C un convexe fermé non vide d'un espace normé réel X. Pour tout $x \notin \mathrm{C}$, on peut trouver une boule ouverte $\mathrm{B} = \mathrm{B}(x, r)$ disjointe de C ; l'ensemble B est convexe et ouvert, donc d'après le théorème de séparation il existe un demi-espace affine fermé H_x qui contient C mais ne rencontre pas B, donc ne contient pas x ; l'intersection de tous les H_x, lorsque x varie dans le complémentaire de C, est égale à C.

∎

Remarque 2.2.1. Le dual de X/Y est identifiable isométriquement au sous-espace de X^* formé des x^* dont la restriction à Y est nulle : notons π la projection de X sur X/Y ; à $z^* \in (\mathrm{X}/\mathrm{Y})^*$ on associe $x^* = z^* \circ \pi \in \mathrm{X}^*$, qui est nulle sur Y. Inversement, si $x^* \in \mathrm{X}^*$ est nulle sur Y, on peut l'écrire $x^* = z^* \circ \pi$ d'après la proposition 1.3.3, et $\|z^*\| = \|x^*\|$.

Proposition 2.2.1. *Soit* X *un espace normé ;*

(i) *pour que* X *soit de dimension finie il faut et il suffit que son dual soit de dimension finie ; dans ce cas, la dimension de* X^* *est égale à celle de* X *;*

(ii) *soient* X *et* Y *des espaces normés et* $\mathrm{T} \in \mathcal{L}(\mathrm{X}, \mathrm{Y})$ *; l'application* T *est de rang fini si et seulement si sa transposée* ${}^t\mathrm{T}$ *est de rang fini ; dans ce cas, le rang de* T *coïncide avec celui de sa transposée.*

Démonstration. Si X est de dimension finie, toute forme linéaire sur X est continue, donc le dual topologique est égal au dual algébrique pour lequel on connaît l'égalité des dimensions. Si X^* est de dimension finie, X^{**} est de dimension finie par ce qui précède, donc X est de dimension finie par le corollaire 4.

Passons au point *(ii)* ; notons Y_1 l'image de T, $\mathrm{T}_1 \in \mathcal{L}(\mathrm{X}, \mathrm{Y}_1)$ l'application qui à $x \in \mathrm{X}$ associe $\mathrm{T}(x)$ considéré comme élément de Y_1 et $u \in \mathcal{L}(\mathrm{Y}_1, \mathrm{Y})$ l'application $y \to y$. On a $\mathrm{T} = u \circ \mathrm{T}_1$ donc ${}^t\mathrm{T} = {}^t\mathrm{T}_1 \circ {}^t u$; on voit donc que ${}^t\mathrm{T}$ se factorise à travers l'espace vectoriel Y_1^* ; si T est de rang fini, ceci implique $\operatorname{rang} {}^t\mathrm{T} \le \dim \mathrm{Y}_1^* = \dim \mathrm{Y}_1 = \operatorname{rang} \mathrm{T}$. Inversement soient Z un sous-espace de dimension finie n de Y_1 et $(z_1, \ldots, z_n)$ une base de Z ; on peut trouver des vecteurs $(x_1, \ldots, x_n)$ dans X tels que $z_i = \mathrm{T}(x_i)$ et des formes linéaires $(z_1^*, \ldots, z_n^*)$ sur Z telles que $z_j^*(z_i) = \delta_{i,j}$ (symbole de Kronecker ; les (z_j^*) forment la base duale de la base (z_i)), puis on peut étendre par Hahn-Banach ces formes linéaires (z_j^*) en $(y_1^*, \ldots, y_n^*)$ dans Y^*. On aura alors

$$ {}^t\mathrm{T}(y_j^*)(x_i) = y_j^*(\mathrm{T}(x_i)) = z_j^*(z_i) = \delta_{i,j}, $$

ce qui montre que ${}^t\mathrm{T}(y_1^*), \ldots, {}^t\mathrm{T}(y_n^*)$ sont linéairement indépendants dans l'image de ${}^t\mathrm{T}$, qui est donc de dimension $\ge n$. Si ${}^t\mathrm{T}$ est de rang fini m, on aura nécessairement $n \le m$, ce qui montre que la dimension de Y_1 est $\le m$: on ne peut pas trouver dans Y_1 un sous-espace Z de dimension $n > m$. On vient de voir que $\operatorname{rang} \mathrm{T} \le \operatorname{rang} {}^t\mathrm{T}$.

∎

Bidual d'un espace normé

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

$\|\mathrm{I_X}(x)\| \le \|x\|$. On dit que $\mathrm{I_X} \in \mathcal{L}(\mathrm{X}, \mathrm{X}^{**})$ est l'application canonique de X dans son bidual.

Corollaire 2.2.4. *L'application canonique* $\mathrm{I_X} : \mathrm{X} \to \mathrm{X}^{**}$ *est isométrique.*

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

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

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

■

Si X est un espace normé l'application $\mathrm{I_X}$ est injective par le corollaire 4. Remarquons que si $\mathrm{I_X}$ est bijective alors $\mathrm{I_X}$ est une isométrie de X sur un espace de Banach (par la proposition 1.2.3). Il s'ensuit que si $\mathrm{I_X}$ est bijective, alors nécessairement X est un espace de Banach. Ceci explique que nous restreindrons la définition qui suit aux espaces de Banach.

Définition 2.2.1. Un espace de Banach E est dit *réflexif* si l'application canonique $\mathrm{I_E} : \mathrm{E} \to \mathrm{E}^{**}$ est bijective.

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

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

Exemples 2.2.1. 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 fait 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, comme on va le voir. D'après les résultats généraux qui suivent, il suffit de voir que ℓ_1 n'est pas réflexif.

Désignons par c le sous-espace vectoriel de ℓ_∞ formé des suites scalaires $x = (x_n)$ *convergentes* ; sur ce sous-espace est définie la forme linéaire naturelle

$$\ell(x) = \lim_{n \to +\infty} x_n.$$

Il est clair que $|\ell(x)| \le \sup_n |x_n| = \|x\|_\infty$, donc ℓ est continue. Par le théorème de Hahn-Banach il existe un prolongement $x^* \in (\ell_\infty)^*$. On va voir que cette forme linéaire ne peut pas provenir d'un élément de ℓ_1, ce qui montrera que ℓ_1 n'est pas réflexif. Soit donc $u \in \ell_1$, qui définit une forme linéaire f_u sur ℓ_∞ par $f_u(x) = \sum_{n=0}^{+\infty} u_n x_n$; considérons un vecteur $x^{(\mathrm{N})} \in c$ dont les composantes sont $x_j^{(\mathrm{N})} = 0$ si $0 \le j \le \mathrm{N}$ et $x_j^{(\mathrm{N})} = 1$ si $j > \mathrm{N}$; alors $x^*(x^{(\mathrm{N})}) = \ell(x^{(\mathrm{N})}) = 1$ mais

$$f_u(x^{(\mathrm{N})}) = \sum_{k>\mathrm{N}} u_k$$

tend vers 0 lorsque $\mathrm{N} \to +\infty$, ce qui montre que f_u ne peut pas être égale à x^*.

Si le lecteur réfléchit un peu, il aura du mal à trouver un procédé “calculatoire” pour définir cette forme linéaire x^* qui doit affecter un résultat à *toute suite bornée* $x \in \ell_\infty$, de façon que le résultat soit linéaire et égal à la limite de la suite x quand elle est convergente. Le lemme de Zorn est intimement lié au théorème de Hahn-Banach pour un espace tel que l'espace ℓ_∞ (ou L_∞).

Proposition 2.2.2. *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 $z = x^* = x^{***} \circ \mathrm{I_X}$ 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 $z^* = \mathrm{I_X}(x)$. Alors

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

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

Proposition 2.2.3. *Si* X *est réflexif, tout sous-espace fermé* Y *de* X *est réflexif.*

Démonstration. Soit π l'application de restriction définie de X^* sur Y^* (surjective par le théorème de Hahn-Banach). Soit y^{**} une forme linéaire continue sur Y^*. Alors $x^{**} = y^{**} \circ \pi$ est une forme linéaire continue sur X^*, donc il existe $x \in \mathrm{X}$ tel que $x^{**}(x^*) = x^*(x)$ pour tout x^* in X^*. Il suffit de voir que $x \in \mathrm{Y}$ pour pouvoir conclure assez facilement ; si on avait $x \notin \mathrm{Y}$, on pourrait trouver d'après le corollaire 3 une forme linéaire $x^* \in \mathrm{X}^*$ telle que $x^*(x) = 1$ mais $x^*(y) = 0$ pour tout $y \in \mathrm{Y}$. On aurait alors $\pi(x^*) = 0$, donc $x^{**}(x^*) = y^{**}(\pi(x^*)) = 0$, ce qui contredit $x^{**}(x^*) = x^*(x) = 1$.

Corollaire 2.2.5. *Si* X^* *est réflexif, alors* X *est réflexif.*

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

2.3. Parties totales. Séparabilité

On dit qu'un espace métrique (Z, d) est *séparable* s'il existe une partie *dénombrable* $\mathrm{D} \subset \mathrm{Z}$ qui soit dense dans Z.

Exemple 2.3.1. Les espaces $\mathbb{R}$ et $\mathbb{C}$ sont séparables (par exemple, $\mathbb{Q}$ est un sous-ensemble dénombrable dense dans $\mathbb{R}$), de même que tout espace normé de dimension finie. Les espaces ℓ_p et les espaces $\mathrm{L}_p([0,1])$ sont séparables pour $1 \le p < \infty$. 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 seul sous-espace fermé de X contenant D est X. Autrement dit, D est total 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).

Le théorème de Hahn-Banach donne le critère suivant : pour que D soit total dans X, il faut et il suffit que toute forme linéaire $x^* \in X^*$, nulle sur D, soit identiquement nulle.

Proposition 2.3.1. *Pour qu'un espace normé* X *soit séparable, 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 de X, totale dans X ; soit Q une partie dénombrable de $\mathbb{K}$, dense dans $\mathbb{K}$. Pour chaque entier $n \geq 0$, considérons l'ensemble $D_n = \{\sum_{k=0}^{n} s_k x_k : (s_0, \ldots, s_n) \in Q^{n+1}\}$. Comme Q^{n+1} est dénombrable, D_n est dénombrable ; donc la réunion $D = \cup_n D_n$ est dénombrable. L'adhérence de D_n contient le sous-espace Y_n engendré par x_k, $0 \leq k \leq n$; donc l'adhérence de D contient le plus petit sous-espace contenant les x_n (qui est $Y = \cup_n Y_n$) ; donc $\overline{D} \supset \overline{Y}$; comme la suite (x_n) est totale, $\overline{D} = \overline{Y} = X$, donc D est dénombrable et dense dans X.

∎

Proposition 2.3.2. *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 $n \geq 0$, on peut trouver un vecteur $x_n \in X$ tel que $\|x_n\| \leq 1$ et $x_n^*(x_n) \geq \|x_n^*\|/2$. 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 $x_{n_0}^*(x_{n_0}) \geq \|x_{n_0}^*\|/2 \geq 3/8 > 1/4$.

2.4. Théorème de Riesz

Lemme 2.4.1. *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 $\varepsilon^n V$, 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 2.4.1. *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 $N \geq 2^n$, contradiction.

On dit qu'un opérateur 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 F.

Corollaire 2.4.1. *Soit* E *un espace de Banach, réel ou complexe ; si* $T \in \mathcal{L}(E)$ *est compact et si* $\lambda \neq 0$, *le sous-espace* $F_\lambda = \ker(T - \lambda \,\mathrm{Id}_E) = \{y \in E : T(y) = \lambda y\}$ *est de dimension finie.*

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

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

3. Topologies faibles

3.1. Topologies initiales

On est souvent amené à considérer d'autres topologies que la topologie de la norme sur un espace de Banach. Commençons par un rappel de topologie générale :

Proposition 3.1.1. *Donnons-nous deux ensembles* I *et* X, *une famille* $(\mathrm{Y}_i)_{i\in\mathrm{I}}$ *d'espaces topologiques et, pour tout* $i \in \mathrm{I}$, *une application* $f_i : \mathrm{X} \to \mathrm{Y}_i$. *Il existe une topologie* $\mathcal{O}$ *sur* X *caractérisée par la propriété suivante :*

(P) *pour qu'une application* g *d'un espace topologique* Z *dans* X *muni de la topologie* $\mathcal{O}$ *soit continue il faut et il suffit que les applications* $f_i \circ g$ *soient continues pour tout* $i \in \mathrm{I}$.

La topologie sur X possédant la propriété (P) de la proposition 1 s'appelle *topologie initiale* associée à la famille f_i. C'est la topologie sur X la moins fine rendant continues les applications $f_i : \mathrm{X} \to \mathrm{Y}_i$.

La topologie initiale n'est en général pas une topologie associée à une distance - et les suites ne jouent donc pas un rôle aussi important que dans le cas métrique. Rappelons cependant qu'une suite (x_n) de points de X tend vers $x \in \mathrm{X}$ si et seulement si : pour tout $i \in \mathrm{I}$ la suite $f_i(x_n)$ converge vers $f_i(x)$.

Si x_0 est un point de X, i_0 un élément de I et V_{i_0} un ouvert de Y_{i_0} contenant $f_{i_0}(x_0)$, alors l'ensemble

$$\mathrm{W}(i_0, \mathrm{V}_{i_0}) = f_{i_0}^{-1}(\mathrm{V}_{i_0}) = \{x \in \mathrm{X} : f_{i_0}(x) \in \mathrm{V}_{i_0}\}$$

est par définition ouvert dans la topologie initiale, et il contient x_0 par notre hypothèse, donc c'est un voisinage de x_0 ; par conséquent, étant donné un ensemble fini quelconque $\mathrm{J} = \{i_1, \ldots, i_n\}$ contenu dans I et une famille $\mathrm{V}_\mathrm{J} = (\mathrm{V}_j)_{j\in\mathrm{J}}$ où V_j est un ouvert de Y_j contenant $f_j(x_0)$ pour chaque $j \in \mathrm{J}$, l'intersection $\cap_{j\in\mathrm{J}}\mathrm{W}(j, \mathrm{V}_j)$, égale à

$$\mathrm{W}(\mathrm{J}, \mathrm{V}_\mathrm{J}) = \{x \in \mathrm{X} : \forall j \in \mathrm{J},\ f_j(x) \in \mathrm{V}_j\}$$

sera un voisinage de x_0 pour la topologie initiale. Le fait que cette topologie soit la plus faible rendant continues les (f_i) se traduit par une sorte de réciproque : pour qu'un ensemble $\mathrm{W} \subset \mathrm{X}$ soit un voisinage de x_0, il faut (et on a vu qu'il suffit) qu'il existe un ensemble fini $\mathrm{J} \subset \mathrm{I}$ et une famille V_J comme ci-dessus tels que $x_0 \in \mathrm{W}(\mathrm{J}, \mathrm{V}_\mathrm{J}) \subset \mathrm{W}$.

Proposition 3.1.2. *Donnons-nous un ensemble* I, *un espace vectoriel* X, *une famille* $(\mathrm{Y}_i)_{i\in\mathrm{I}}$ *d'espaces vectoriels topologiques (cf. définition 1.1.2) et, pour tout* $i \in \mathrm{I}$, *une application linéaire* $f_i : \mathrm{X} \to \mathrm{Y}_i$. *Muni de la topologie initiale associée aux applications* f_i, *l'espace* X *est un espace vectoriel topologique.*

Démonstration. C'est une démonstration quasi-mécanique. Notons $\varphi : \mathrm{X} \times \mathrm{X} \to \mathrm{X}$ et $\varphi_i : \mathrm{Y}_i \times \mathrm{Y}_i \to \mathrm{Y}_i$ (pour $i \in \mathrm{I}$) les applications $(x, y) \to x + y$; de même, notons $\psi : \mathbb{K} \times \mathrm{X} \to \mathrm{X}$ et $\psi_i : \mathbb{K} \times \mathrm{Y}_i \to \mathrm{Y}_i$ (pour $i \in \mathrm{I}$) les applications $(\lambda, y) \to \lambda y$. Pour tout $i \in \mathrm{I}$, l'application $f_i \times f_i : \mathrm{X} \times \mathrm{X} \to \mathrm{Y}_i \times \mathrm{Y}_i$ est continue, donc l'application $f_i \circ \varphi = \varphi_i \circ (f_i \times f_i)$ est continue. Par la définition de la topologie initiale, l'application φ est continue. De même, pour tout $i \in \mathrm{I}$, l'application $\mathrm{Id}_\mathbb{K} \times f_i : \mathbb{K} \times \mathrm{X} \to \mathbb{K} \times \mathrm{Y}_i$ est continue, donc l'application $f_i \circ \psi = \psi_i \circ (\mathrm{Id}_\mathbb{K} \times f_i)$ est continue. Par la définition de la topologie initiale, l'application ψ est continue.

■

Soient X un espace vectoriel, I un ensemble et $(p_i)_{i\in \mathrm{I}}$ une famille de semi-normes sur X ; pour $i \in \mathrm{I}$, notons X_i l'espace X muni de la topologie associée à p_i et $f_i : \mathrm{X} \to \mathrm{X}_i$ l'identité de X. On appelle *topologie associée à la famille de semi-normes* $(p_i)_{i\in \mathrm{I}}$ la topologie initiale associée aux applications f_i. Muni de cette topologie, X est d'après la proposition précédente, un espace vectoriel topologique.

Soit X un espace vectoriel ; donnons-nous un ensemble I et, pour chaque $i \in \mathrm{I}$, un espace semi-normé (Y_i, q_i) et une application linéaire $g_i : \mathrm{X} \to \mathrm{Y}_i$. On vérifie sans peine que la topologie initiale associée aux g_i coïncide avec la topologie associée à la famille de semi-normes $(q_i \circ g_i)_{i\in \mathrm{I}}$.

Exemple 3.1.1. Soient X, Y des espaces vectoriels et $\mathrm{B} : \mathrm{X} \times \mathrm{Y} \to \mathbb{K}$ une forme bilinéaire ; pour $y \in \mathrm{Y}$ notons $p_y : \mathrm{X} \to \mathbb{R}_+$ l'application $x \to |\mathrm{B}(x, y)|$. C'est une semi-norme sur X. La topologie sur X associée à la famille des semi-normes p_y s'appelle la topologie *faible* associée à B, ou à Y s'il n'y a pas d'ambiguité sur la forme B ; cette topologie se note $\sigma(\mathrm{X}, \mathrm{Y})$. Cette topologie est la topologie initiale associée aux applications $f_y : \mathrm{X} \to \mathbb{K}$, où, pour $y \in \mathrm{Y}$ on a noté f_y l'application $x \to \mathrm{B}(x, y)$. En d'autres termes, la topologie faible est la topologie la plus faible pour laquelle les applications f_y sont continues.

Une suite (x_n) dans X converge vers $x \in \mathrm{X}$ pour la topologie $\sigma(\mathrm{X}, \mathrm{Y})$, si et seulement si, pour tout $y \in \mathrm{Y}$, la suite $(\mathrm{B}(x_n, y))_{n\in\mathbb{N}}$ converge vers $\mathrm{B}(x, y)$.

On démontre que les seules formes linéaires continues sur X pour la topologie $\sigma(\mathrm{X}, \mathrm{Y})$ sont les f_y (on utilise à cet effet un lemme d'algèbre linéaire : si $\ell_1, \ldots, \ell_n$ et ℓ sont des formes linéaires sur X, et si pour tout $x \in \mathrm{X}$ les conditions $\ell_1(x) = \ell_2(x) = \cdots = \ell_n(x) = 0$ entraînent $\ell(x) = 0$, alors ℓ est combinaison linéaire de $\ell_1, \ldots, \ell_n$).

Terminons ce paragraphe par une définition.

Définition 3.1.1. Soient X et Y deux espaces normés ; chaque élément x de X définit une semi-norme N_x sur $\mathcal{L}(\mathrm{X}, \mathrm{Y})$ donnée par $\mathrm{N}_x(f) = \|f(x)\|$. On appelle *topologie forte* sur $\mathcal{L}(\mathrm{X}, \mathrm{Y})$ la topologie associée à la famille de semi-normes $(\mathrm{N}_x)_{x\in\mathrm{X}}$.

Cette topologie est la topologie initiale associée aux applications $\varphi_x : \mathcal{L}(\mathrm{X}, \mathrm{Y}) \to \mathrm{Y}$, où, pour $x \in \mathrm{X}$ on a noté φ_x l'application $f \to f(x)$. En d'autres termes, la topologie forte sur $\mathcal{L}(\mathrm{X}, \mathrm{Y})$ est la topologie la plus faible pour laquelle les applications φ_x sont continues. Elle est plus faible que la topologie de la norme sur $\mathcal{L}(\mathrm{X}, \mathrm{Y})$.

Une suite (f_n) dans $\mathcal{L}(\mathrm{X}, \mathrm{Y})$ converge vers $f \in \mathcal{L}(\mathrm{X}, \mathrm{Y})$ pour la topologie forte si et seulement si, pour tout $x \in \mathrm{X}$, la suite $(f_n(x))_{n\in\mathbb{N}}$ converge vers $f(x)$ pour la topologie de la norme sur Y.

3.2. Topologie faible sur un espace normé

Soit X un espace vectoriel normé ; considérons la forme bilinéaire $(x^*, x) \to x^*(x)$ sur $\mathrm{X}^* \times \mathrm{X}$. Cette forme bilinéaire définit des topologies faibles $\sigma(\mathrm{X}, \mathrm{X}^*)$ sur X et $\sigma(\mathrm{X}^*, \mathrm{X})$ sur X^*.

La topologie $\sigma(\mathrm{X}, \mathrm{X}^*)$ sur X s'appelle aussi la *topologie faible* sur X. C'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 $\sigma(\mathrm{X}, \mathrm{X}^*)$ est plus faible que la topologie de la norme.

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

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

C'est une simple adaptation de la description des voisinages fondamentaux données dans la section précédente. Il en résulte que si X est de dimension infinie, un voisinage faible de x_0 n'est jamais borné, et la topologie faible est strictement plus faible que la topologie de la norme : en effet, on pourra trouver dans ce cas un vecteur $y \in X$ tel que $x_1^*(y) = \cdots = x_n^*(y) = 0$, et la droite affine $x_0 + \mathbb{K}y$ sera alors contenue dans le voisinage faible W.

On peut aussi considérer la boule unité fermée B(X) de X, et la munir de la topologie induite par la topologie faible ; même dans ce cadre, la topologie de $(B(X), \sigma(X, X^*))$ est strictement plus faible en général que celle de B(X) munie de la topologie de la norme : soit $X = \ell^2$, l'espace hilbertien des suites de carré sommable ; notons (e_n) sa base hilbertienne canonique. On vérifie que la suite (e_n) tend faiblement vers 0 (voir lemme 5.5.1.b). Cependant la suite (e_n) ne peut converger en norme, vu que $\|e_n - e_m\|^2 = 2$ (pour $m \neq n$). En fait, pour tout espace normé de dimension infinie la topologie de $(B(X), \sigma(X, X^*))$ est strictement plus faible que la topologie de la norme, pour la raison évoquée plus haut : si X est de dimension infinie, un voisinage faible de 0_X dans $(B(X), \sigma(X, X^*))$ contient toujours des points de la sphère unité de X.

En revanche, la topologie induite sur la *sphère unité* S_X par la topologie faible peut coïncider avec la topologie de la norme sur S_X : c'est le cas pour $X = \ell_2$ (exercice).

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

Démonstration. Une application f d'un espace topologique X dans F munie de $\sigma(F, F^*)$ est continue si et seulement si, pour tout $y^* \in F^*$, $y^* \circ f$ est continue de X dans $\mathbb{K}$; ici, X est l'espace topologique $(E, \sigma(E, E^*))$. Or, pour tout $y^* \in F^*$, on a $y^* \circ T \in E^*$, donc $y^* \circ T$ est continue pour $\sigma(E, E^*)$ par définition de la topologie faible.

∎

Théorème 3.2.1. *Soient* X *un espace normé et* 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* $\sigma(X, X^*)$. *Plus généralement, si* C *est un sous-ensemble convexe de* X*, il est fermé en norme si et seulement s'il est faiblement fermé.*

Démonstration. Comme la topologie de la norme est plus fine que la topologie faible, toute partie fermée pour la topologie faible est fermée pour la topologie de la norme. Supposons inversement que le sous-espace vectoriel Y soit fermé pour la topologie de la norme. Alors, par le corollaire 2.2.3, il existe $A \subset X^*$ tel que $Y = \bigcap_{\ell \in A} \ker(\ell)$. Or tout $\ell \in X^*$ est continue pour $\sigma(X, X^*)$, donc $\ker(\ell)$ est fermé pour cette topologie ; donc Y est une intersection de parties fermées de X pour $\sigma(X, X^*)$.

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 $\mathrm{Re}\, x^*$ est une application faiblement continue de X dans $\mathbb{R}$; il résulte de ces considérations et de la deuxième partie du corollaire 2.2.3 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 la topologie de la norme de X^* rendant déjà continues toutes ces applications, la topologie $*$-faible est plus faible que la topologie de la norme sur X^*.

Pour que $W \subset X^*$ soit un voisinage $*$-faible du point $x_0^* \in X^*$, il faut et il suffit qu'il existe un nombre fini de vecteurs $x_1, \ldots, x_n \in X$ et un nombre $\varepsilon > 0$ tels que

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

Il en résulte que si X est de dimension infinie, un voisinage $*$-faible de x_0^* n'est jamais borné. On peut aussi considérer la boule unité fermée $B(X^*)$ de X^*, et la munir de la topologie induite par la topologie $*$-faible ; pour tout espace normé de dimension infinie la topologie de $(B(X^*), \sigma(X^*, X))$ est strictement plus faible que la topologie de la norme, toujours pour les mêmes raisons : si X est de dimension infinie, un voisinage faible de 0_{X^*} dans $(B(X^*), \sigma(X^*, X))$ contient toujours des points de la sphère unité de X^*. On verra aussi une autre raison avec le théorème qui suit : l'espace topologique $(B(X^*), \sigma(X^*, X))$ est toujours compact, alors que $B(X^*)$ n'est jamais compacte en norme si la dimension est infinie (théorème 2.4.1).

Théorème 3.2.2. *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 pour l'instant (voir dernière section).

Théorème 3.2.3 : théorème de Tykhonov. *Tout produit d'espaces compacts (muni de la topologie produit) est compact.*

(Rappelons que la topologie produit sur $\prod X_i$ est la topologie initiale associée aux projections $\prod X_i \to X_i$).

Démonstration du théorème 2. 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 $\mathbb{K}^X$,

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

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

Toutes ces conditions définissent des fermés, donc $B = F_1 \cap F_2$ est un fermé de $\mathbb{K}^X$. Pour $r \ge 0$, posons $D_r = \{\lambda \in \mathbb{K} : |\lambda| \le r\}$. On a $B \subset \prod_{x \in X} D_{\|x\|}$, qui est compact par le théorème 3 ; étant fermé dans un compact, B est compact.

∎

3.3. 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 $x^* \in X^*$.

Toute **suite** faiblement convergente dans X est bornée ; toute **suite** $*$-faiblement convergente dans X^* est bornée. Ces deux résultats viennent du corollaire 2.1.2.

Proposition 3.3.1. *Si* E *est un espace normé séparable, toute suite bornée de* E^* *admet des sous-suites* $*$*-faiblement convergentes.*

Démonstration. Pour exprimer la démonstration, il est utile d'introduire une petite convention de langage. Si $\mathrm{M} = \{n_0 < \ldots < n_j < \ldots\}$ est un sous-ensemble infini de $\mathbb{N}$, convenons de noter la sous-suite (x_{n_j}) par $(x_n)_{n\in\mathrm{M}}$. Soit donc (x_k) une suite dense dans E, et (x_n^*) une suite bornée dans E^*, telle que par exemple $\|x_n^*\| \le 1$ pour tout entier $n \ge 0$. La suite de scalaires $(x_n^*(x_0))$ est bornée, donc elle admet une sous-suite convergente $(x_n^*(x_0))_{n\in\mathrm{M}_0}$. La suite $(x_n^*(x_1))_{n\in\mathrm{M}_0}$ est encore bornée, donc on peut trouver un nouvel ensemble infini $\mathrm{M}_1 \subset \mathrm{M}_0$ tel que la sous-suite $(x_n^*(x_1))_{n\in\mathrm{M}_1}$ soit convergente. En continuant ainsi, on construit une suite décroissante $\mathrm{M}_0 \supset \mathrm{M}_1 \supset \ldots \supset \mathrm{M}_j \supset \ldots$ telle que $(x_n^*(x_j))_{n\in\mathrm{M}_j}$ soit convergente pour tout $j \ge 0$.

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

Il reste à montrer que $(x_n^*(x))_{n\in\mathrm{M}}$ converge pour tout $x \in \mathrm{E}$. Il suffit de montrer que cette suite de scalaires est de Cauchy. Choisissons $\varepsilon > 0$ et un entier k tels que $\|x - x_k\| < \varepsilon/3$. Pour tout entier n, on a $|x_n^*(x) - x_n^*(x_k)| < \varepsilon/3$, et la suite $(x_n^*(x_k))_{n\in\mathrm{M}}$ converge vers une limite ℓ ; il en résulte que pour n assez grand dans M, on aura $|x_n^*(x) - \ell| < \varepsilon/2$, et si n, m sont assez grands et dans M, on aura $|x_n^*(x) - x_m^*(x)| < \varepsilon$. La suite $(x_n^*(x))_{n\in\mathrm{M}}$ est donc de Cauchy, donc convergente. Il est alors clair que la formule

$$x^*(x) = \lim_{n\in\mathrm{M}} x_n^*(x)$$

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

Théorème 3.3.1. *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 $\mathrm{Y}^{**} \simeq \mathrm{Y}$; en considérant (x_n) comme une suite bornée dans Y^{**}, on peut trouver une sous-suite (x_{n_j}) $*$-faiblement convergente dans Y^{**} vers un élément y^{**}, c'est à dire telle que

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

Puisque Y est réflexif, il existe un vecteur $y \in \mathrm{Y}$ tel que $y^{**} = \mathrm{I}_{\mathrm{Y}}(y)$, et la relation ci-dessus nous dit que $y^*(y) = y^{**}(y^*) = \lim_j y^*(x_{n_j})$ pour tout $y^* \in \mathrm{Y}^*$, 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 $y \in \mathrm{Y}$ que par sa restriction $y^* \in \mathrm{Y}^*$ 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.

Exemple 3.3.1. 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 $\mathrm{C}([0,1])$.

Théorème 3.3.2. *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 vers $\inf f(\mathrm{C})$ (peut-être $-\infty$) ; la suite (x_n) est donc 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 3.3.1. *Si f est convexe continue sur* E *réflexif 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 E, et soit $\mathrm{C} = \{x \in \mathrm{E} : f(x) \leq f(x_0)\}$. L'ensemble C 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, et il est facile de voir que ce minimum est aussi le minimum sur E tout entier.

Remarque 3.3.1. C'est ce type de résultat qui permet de minimiser les fonctionnelles telles que celle qui a été évoquée dans l'introduction.

3.4. Appendice : démonstration du Théorème de Tykhonov

On dit qu'une famille $\mathcal{F} \subset \mathcal{P}(\mathrm{X})$ de parties d'un ensemble X possède la *propriété d'intersection finie* lorsque toute sous-famille finie $\{\mathrm{F}_1, \ldots, \mathrm{F}_n\} \subset \mathcal{F}$ a une intersection $\mathrm{F}_1 \cap \ldots \cap \mathrm{F}_n$ non vide. Lorsque X est un espace topologique compact, toute famille de **fermés** $\mathcal{F}$ possédant la propriété d'intersection finie a une intersection non vide, et cette propriété d'intersection est en fait équivalente à la définition de la compacité de X.

Soit donc $(\mathrm{X}_i)_{i \in \mathrm{I}}$ une famille d'espaces topologiques compacts, et soit $\mathcal{F}$ une famille de fermés du produit $\mathrm{X}_\mathrm{I} = \prod_{i \in \mathrm{I}} \mathrm{X}_i$ possédant la propriété d'intersection finie. Nous devons montrer que son intersection est non vide. La démonstration utilise un ultra-procédé d'extraction nécessitant le lemme de Zorn (ce qui n'est pas trop surprenant, puisque dire que X_I est non vide lorsque chaque X_i est non vide, pour tout ensemble I, est une formulation de l'axiome du choix).

Soit donc $\mathcal{U}$ une famille maximale (pour l'inclusion des familles de parties de X) contenant $\mathcal{F}$ et possédant la propriété d'intersection finie. Son existence résulte facilement du lemme de Zorn. La maximalité signifie qu'aucun sous-ensemble $\mathrm{Y} \notin \mathcal{U}$ ne peut être ajouté à la famille $\mathcal{U}$ sans perdre la propriété d'intersection finie : il existe une famille finie $\mathrm{U}_1, \ldots, \mathrm{U}_n$ d'éléments de $\mathcal{U}$ telle que $\mathrm{Y} \cap \mathrm{U}_1 \cap \ldots \cap \mathrm{U}_n = \emptyset$. Si U et U′ sont deux éléments de $\mathcal{U}$, l'ensemble $\mathrm{U} \cap \mathrm{U}'$ rencontre toutes les intersections finies $\mathrm{U}_1 \cap \ldots \cap \mathrm{U}_n$ d'éléments de $\mathcal{U}$ par la propriété d'intersection finie de $\mathcal{U}$, donc $\mathrm{U} \cap \mathrm{U}' \in \mathcal{U}$ par la maximalité : en fait, les intersections finies d'éléments de $\mathcal{U}$ sont déjà dans $\mathcal{U}$; on peut donc dire que si $\mathrm{Y} \notin \mathcal{U}$, il existe un élément $\mathrm{U} \in \mathcal{U}$ tel que $\mathrm{Y} \cap \mathrm{U} = \emptyset$.

Pour tout sous-ensemble fini $\mathrm{J} \subset \mathrm{I}$, notons $\mathrm{X}_\mathrm{J} = \prod_{j \in \mathrm{J}} \mathrm{X}_j$. Cet espace est compact. Désignons par π_J la projection naturelle de X_I sur X_J, qui consiste à “oublier” toutes les coordonnées de $x = (x_i)_{i \in \mathrm{I}}$ qui ne sont pas dans J. Il est facile de vérifier que la famille

$$\mathcal{U}_\mathrm{J} = \{\overline{\pi_\mathrm{J}(\mathrm{U})} : \mathrm{U} \in \mathcal{U}\}$$

possède la propriété d'intersection finie, donc son intersection F_J est non vide par compacité. Mais en fait, cette intersection est réduite à un seul point par maximalité ; en effet, si y_1 et y_2 étaient deux points distincts de F_J, on pourrait trouver deux voisinages V_1 et V_2 de ces deux points qui soient disjoints. Puisque y_1 est adhérent à chaque $\pi_\mathrm{J}(\mathrm{U})$, $\mathrm{U} \in \mathcal{U}$, le voisinage V_1 rencontre $\pi_\mathrm{J}(\mathrm{U})$, donc $\pi_\mathrm{J}^{-1}(\mathrm{V}_1)$ rencontre U, pour tout $\mathrm{U} \in \mathcal{U}$. La maximalité nous dit alors que $\pi_{\mathrm{J}}^{-1}(\mathrm{V}_1) \in \mathcal{U}$; mais le même raisonnement donnerait $\pi_{\mathrm{J}}^{-1}(\mathrm{V}_2) \in \mathcal{U}$, ce qui est impossible puisque ces deux ensembles sont disjoints.

On a donc $\mathrm{F}_{\mathrm{J}} = \{(x_j^{(\mathrm{J})})_{j\in \mathrm{J}}\}$. Si J′ contient J, on montre que $x_j^{(\mathrm{J}')} = x_j^{(\mathrm{J})}$ pour tout $j \in \mathrm{J}$. Il en résulte que $x_j^{(\mathrm{J})}$ ne dépend pas de l'ensemble fini J qui contient $j \in \mathrm{I}$: si J_1 et J_2 contiennent j, considérons $\mathrm{J} = \mathrm{J}_1 \cup \mathrm{J}_2$ et appliquons la remarque qui précède. On peut donc poser $x_j = x_j^{(\mathrm{J})}$, où J est n'importe quel sous-ensemble fini de I qui contient j, et ceci pour tout $j \in \mathrm{I}$.

Pour terminer, il reste à voir que le point $x = (x_i)_{i\in \mathrm{I}} \in \mathrm{X}_{\mathrm{I}}$ ainsi défini appartient à l'intersection de tous les éléments de la famille initiale $\mathcal{F}$. Sinon, il existerait un ensemble $\mathrm{F} \in \mathcal{F}$ tel que $x \notin \mathrm{F}$, donc il existerait un voisinage W élémentaire de x dans X_{I}, disjoint de F. Mais un tel voisinage élémentaire est de la forme $\pi_{\mathrm{J}}^{-1}(\mathrm{V})$, pour un certain ensemble fini $\mathrm{J} \subset \mathrm{I}$ et un ouvert V de X_{J} contenant $x_{\mathrm{J}} = \pi_{\mathrm{J}}(x) = (x_j^{(\mathrm{J})})_{j\in \mathrm{J}}$. Mais alors $\pi_{\mathrm{J}}(\mathrm{F})$ serait disjoint de V, donc on aurait $x_{\mathrm{J}} \notin \overline{\pi_{\mathrm{J}}(\mathrm{F})}$, ce qui contredit la construction initiale.

Remarque 3.4.1. On a vu que la famille maximale $\mathcal{U}$ est stable par intersection finie ; on voit aussi par maximalité que si $\mathrm{U} \in \mathcal{U}$ et $\mathrm{U} \subset \mathrm{Y}$, alors $\mathrm{Y} \in \mathcal{U}$. Il en résulte que pour tout $\mathrm{Y} \subset \mathrm{X}$, ou bien $\mathrm{Y} \in \mathcal{U}$, ou bien $\mathrm{Y} \notin \mathcal{U}$. Une telle famille $\mathcal{U}$ de parties de X est appelée *ultrafiltre* sur X.

4. Espaces hilbertiens

4.1. Produits scalaires

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

Définition 4.1.1. Soit E un espace vectoriel complexe ; on appelle *forme sesquilinéaire* sur E une application $\mathrm{B} : \mathrm{E} \times \mathrm{E} \to \mathbb{C}$ telle que, pour tout $y \in \mathrm{E}$, l'application $x \to \mathrm{B}(x, y)$ soit linéaire et telle que pour tout $x \in \mathrm{E}$, l'application $y \to \mathrm{B}(x, y)$ soit antilinéaire (de E dans $\mathbb{C}$).

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

Proposition 4.1.1 : identité de polarisation.

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

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

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

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

Corollaire 4.1.1. *Soient* E *un espace vectoriel complexe et* B *une forme sesquilinéaire sur* E *; les conditions suivantes sont équivalentes :*

(*i*) *pour tous* $x, y \in \mathrm{E}$ *on a* $\mathrm{B}(y, x) = \overline{\mathrm{B}(x, y)}$ *;*
(*ii*) *pour tout* $x \in \mathrm{E}$, *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 $x \in \mathrm{E}$, on a $\mathrm{S}(x, x) = 0$.
■

Soit E un espace vectoriel complexe ; on appelle *forme hermitienne* sur E une forme sesquilinéaire vérifiant les conditions équivalentes du corollaire 1. On peut résumer ces conditions ainsi : la forme φ sur $\mathrm{E} \times \mathrm{E}$ est hermitienne si elle vérifie les deux conditions suivantes :

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

Une forme hermitienne B sur un espace vectoriel complexe E est dite *positive* si, pour tout $x \in \mathrm{E}$, le nombre $\mathrm{B}(x,x)$ est réel ≥ 0. Rappelons qu'une forme bilinéaire symétrique B sur un espace vectoriel réel E est dite *positive* si, pour tout $x \in \mathrm{E}$, 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 4.1.2 : inégalité de Cauchy-Schwarz. *Soit* E *un espace vectoriel muni d'un produit scalaire ; pour tous* $x, y \in \mathrm{E}$ *on a*

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

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

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

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

■

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

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

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

par la proposition 2.

■

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

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

4.2. Espaces hilbertiens

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

Proposition 4.2.1. *Soit* E *un espace préhilbertien ; pour tout vecteur* $x \in \mathrm{E}$ *la forme linéaire* $f_x : y \to \langle y,x \rangle$ *est continue. L'application* $x \to f_x$ *est antilinéaire et isométrique de* E *dans* E^*.

Démonstration. Pour $y \in \mathrm{E}$ on a $|f_x(y)| \leq \|x\| \, \|y\|$ par la proposition 1.2, donc $f_x \in \mathrm{E}^*$ et $\|f_x\| \leq \|x\|$. Or $\|x\|^2 = f_x(x) \leq \|f_x\| \, \|x\|$, d'où l'on déduit que $\|x\| = \|f_x\|$. On vérifie sans peine que l'application $x \to f_x$ est antilinéaire.

■

Définition 4.2.1. Soit E un espace préhilbertien ; on dit que les éléments x et y de E sont *orthogonaux* si $\langle x, y\rangle = 0$. On dit que des parties A et B sont *orthogonales* si tout élément de A est orthogonal à tout élément de B. Soit A une partie de E ; on appelle orthogonal de A l'ensemble $A^\perp$ des éléments de E orthogonaux à A.

Il est clair que $A^\perp = \bigcap_{x\in A} \ker f_x$. Donc $A^\perp$ est un sous-espace vectoriel fermé de E.

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

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

Proposition 4.2.2. *Soit* E *un espace préhilbertien ; le complété de* E *est un espace hilbertien.*

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

$$\big|\langle x_n, y_n\rangle - \langle x_m, y_m\rangle\big| \leq \big|\langle x_n - x_m, y_n\rangle\big| + \big|\langle x_m, y_n - y_m\rangle\big| \leq$$
$$\leq \|x_n - x_m\|\,\|y_n\| + \|x_m\|\,\|y_n - y_m\|$$

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

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

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

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

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

∎

Exemples 4.2.1. L'espace $L_2(\Omega, \mu)$ est un espace hilbertien pour le produit scalaire

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

L'espace ℓ_2 est un cas particulier. L'espace de Sobolev $H^1([a, b])$ de l'exemple 1.1.1 est aussi un espace hilbertien pour le produit scalaire

$$\langle f, g\rangle = \int_a^b f(s)\overline{g(s)}\, ds + \int_a^b f'(s)\overline{g'(s)}\, ds.$$

4.3. Le théorème de projection

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

Démonstration. Notons $d = \inf\{\|x - y\| : y \in \mathrm{C}\}$ la distance de x à C. Soient $y, z \in \mathrm{C}$; posons $b = x - \frac{1}{2}(y + z)$ et $c = \frac{1}{2}(y - z)$; alors $\|b\| \geq d$ vu que $\frac{1}{2}(y + z) \in \mathrm{C}$; comme $x - y = b - c$ et $x - z = b + c$, on a d'après la relation du parallélogramme

$$\|x - y\|^2 + \|x - z\|^2 = 2(\|b\|^2 + \|c\|^2) \geq 2d^2 + \frac{1}{2}\|y - z\|^2,$$

donc $\|y - z\|^2 \leq 2(\|x - y\|^2 - d^2) + 2(\|x - z\|^2 - d^2)$. Pour tout entier $n \geq 1$, posons

$$\mathrm{C}_n = \{y \in \mathrm{C} : \|x - y\|^2 \leq d^2 + 1/n\}.$$

C'est une partie fermée non vide de H ; par ce qui précède le diamètre de C_n est inférieur ou égal à $2/\sqrt{n}$, donc tend vers 0. Comme H est complet, l'intersection des fermés emboîtés C_n qui est égale à $\{y \in \mathrm{C} : \|x - y\| = d\}$, contient un et un seul point y_0.

Soit $y \in \mathrm{C}$; pour $t \in [0, 1]$, on a $y_0 + t(y - y_0) \in \mathrm{C}$, donc $\|y_0 + t(y - y_0) - x\| \geq \|y - y_0\|$. Posons $\varphi(t) = \|y_0 + t(y - y_0) - x\|^2 = \|y_0 - x\|^2 + 2t\,\mathrm{Re}(\langle y_0 - x, y - y_0\rangle) + t^2\|y - y_0\|^2$. Comme $\varphi(0) \leq \varphi(t)$ pour $t \in [0, 1]$, la dérivée de φ en 0 est positive ou nulle, donc $\mathrm{Re}(\langle y_0 - x, y - y_0\rangle) \geq 0$.

■

Un cas particulier important est celui où C est un sous-espace vectoriel fermé E de H. Dans ce cas on a $\langle x - y_0, z\rangle = 0$ pour tout vecteur $z \in \mathrm{E}$, c'est à dire que $x - y_0 \perp \mathrm{E}$; pour le voir, choisissons un scalaire $u \in \mathbb{K}$ de module 1 tel que $\langle x - y_0, u\,z\rangle = |\langle x - y_0, z\rangle|$, puis considérons le vecteur $y = u\,z + y_0 \in \mathrm{E}$ pour lequel $y - y_0 = u\,z$; la relation $\mathrm{Re}\langle x - y_0, y - y_0\rangle \leq 0$ donne le résultat.

Dans le cas de la projection sur un sous-espace vectoriel fermé E, la projection y_0 de x sur E est entièrement caractérisée par les deux conditions suivantes

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

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

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

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

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

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

Exemple 4.3.1. 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 $A \times [0,1]$, 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 $P_F f = E(f|\mathcal{F})$ d'une fonction $f \in L_2$ est donnée par

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

Proposition 4.3.1. *Soient* H *un espace hilbertien et* E *un sous-espace vectoriel fermé de* H *; on a* $E \oplus E^\perp = H$.

Démonstration. Soit $x \in H$ et écrivons $x = P_E(x) + (x - P_E(x))$; on a bien que $P_E(x) \in E$ et $x - P_E(x) \in E^\perp$ d'après les propriétés de la projection orthogonale sur un sous-espace vectoriel, ce qui montre que H est la somme de E et $E^\perp$. On vérifie ensuite que la somme est directe : si $x \in E \cap E^\perp$ alors $\langle x, x\rangle = 0$ donc $x = 0$, ce qui termine la démonstration.

■

Corollaire 4.3.1. *Soit* H *un espace hilbertien ;*

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

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

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

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

■

Proposition 4.3.2. *Soit* H *un espace hilbertien ; l'application isométrique antilinéaire* $x \to f_x$ *de la proposition 2.1 est une bijection de* H *sur* H^*.

En d'autres termes, pour toute forme linéaire continue ℓ *sur* H, *il existe un vecteur* $y_\ell \in H$ *unique qui représente la forme linéaire* ℓ *au sens suivant :*

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

Démonstration. Soit $\ell \in H^*$; notons E son noyau. Si $\ell \neq 0$, on peut choisir un vecteur y_1 tel que $\ell(y_1) = 1$. Soit $y_0 = P_E(y_1)$, et posons $y = a(y_1 - y_0)$, avec $a = \|y_1 - y_0\|^{-2}$, de sorte que $a^{-1}\|y\|^2 = 1$; alors $y \perp E$ et $\ell(y) = a\ell(y_1) - a\ell(y_0) = a\ell(y_1) = a$ puisque $y_0 \in \ker \ell$. Soit $x \in E$ et écrivons $x = (x - a^{-1}\ell(x)y) + \ell(x)a^{-1}y$; on a $\ell(x - \ell(x)a^{-1}y) = 0$ donc $x_1 = (x - \ell(x)a^{-1}y) \in E$, ce qui implique que ce vecteur est orthogonal à y. D'autre part,

$$\langle x, y\rangle = \langle x_1, y\rangle + \ell(x)a^{-1}\langle y, y\rangle = \ell(x).$$

■

Exemple 4.3.2. Pour toute forme linéaire continue ℓ sur $L_2(\Omega, \mu)$, il existe une fonction $g \in L_2$ telle que

$$\forall f \in L_2, \quad \ell(f) = \int_\Omega f(s)\overline{g(s)}\, d\mu(s).$$

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 $\nu(A) \leq \mu(A)$ pour tout $A \in \mathcal{A}$, il existe une fonction mesurable bornée f telle que

$$\nu(A) = \int_A f(s)\, d\mu(s) = \int_\Omega 1_A(s) f(s)\, d\mu(s)$$

pour tout ensemble $A \in \mathcal{A}$, 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 $\|g\|_{L_2(\nu)} \leq \|g\|_{L_2(\mu)}$ pour toute fonction $g \in L_2(\mu)$, ce qui implique $L_2(\mu) \subset L_2(\nu)$. Comme ν est finie, la forme linéaire $g \to \int g\, d\nu$ est définie et continue sur $L_2(\nu)$, donc sur $L_2(\mu)$. On peut donc la représenter par une fonction $f \in L_2(\mu)$, c'est à dire que

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

pour toute fonction $g \in L_2(\mu)$. En appliquant avec $g = 1_A$ on obtient le résultat annoncé.

Corollaire 4.3.2. *Tout espace hilbertien est réflexif.*

Démonstration. Soient H un espace hilbertien et $\ell \in H^{**}$; l'application $x \to \overline{\ell(f_x)}$ est une forme linéaire et continue sur H. Par la proposition 2, il existe $y \in H$ tel que, pour tout $x \in H$ on ait $\ell(f_x) = \overline{f_y(x)} = \langle y, x\rangle = f_x(y)$. Il résulte alors de la proposition 2 que, pour tout $f \in H^*$, on a $\ell(f) = f(y)$, c'est à dire que ℓ est l'image de y par l'application canonique de H dans H^{**}.

∎

4.4. Adjoint d'une application linéaire continue

Proposition 4.4.1. *Soient* E *et* F *deux espaces hilbertiens et* $T \in \mathcal{L}(E, F)$ *; il existe un unique* $T^* \in \mathcal{L}(F, E)$ *tel que pour tout* $x \in E$ *et tout* $y \in F$ *on ait* $\langle T(x), y\rangle = \langle x, T^*(y)\rangle$.

Démonstration. Pour tout $y \in F$, l'application $x \to \langle T(x), y\rangle$ est linéaire et continue. Il existe donc un unique élément $T^*(y) \in E$ tel que pour tout $x \in E$ on ait $\langle T(x), y\rangle = \langle x, T^*(y)\rangle$. Pour tout $x \in E$, l'application $y \to \langle T^*(y), x\rangle = \langle y, T(x)\rangle$ est linéaire d'où l'on déduit que T^* est linéaire. Pour tout $x \in E$ et tout $y \in F$ on a

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

or, pour $y \in F$ on a (par la proposition 2.1)

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

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

∎

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

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

Définition 4.4.1. Soient E et F deux espaces hilbertiens et $T \in \mathcal{L}(E, F)$; l'unique $T^* \in \mathcal{L}(E, F)$ tel que pour tout $x \in E$ et tout $y \in F$ on ait $\langle T(x), y\rangle = \langle x, T^*(y)\rangle$ est appelé *adjoint* de T.

Regroupons dans la proposition suivante quelques propriétés des adjoints :

Proposition 4.4.2. *Soient* E *et* F *deux espaces hilbertiens ; l'application* $T \to T^*$ *est antilinéaire et isométrique de* $\mathcal{L}(E, F)$ *sur* $\mathcal{L}(F, E)$ *; pour tout* $T \in \mathcal{L}(E, F)$ *on a* $(T^*)^* = T$ *et* $\|T^* \circ T\| = \|T\|^2$. *Pour tout espace hilbertien* H, *tout* $S \in \mathcal{L}(E, F)$ *et tout* $T \in \mathcal{L}(F, H)$ *on a* $(T \circ S)^* = S^* \circ T^*$.

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

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

Les autres propriétés sont laissées en exercice.

∎

Proposition 4.4.3. *Soient* E *et* F *deux espaces hilbertiens et* $T \in \mathcal{L}(E, F)$ *; alors* $\ker T^* = (T(E))^\perp$ *et l'adhérence de* $T^*(F)$ *est* $(\ker T)^\perp$.

Démonstration. Soit $y \in F$; alors $y \in \ker T^*$ si et seulement si pour tout $x \in E$, on a $0 = \langle T^*(y), x\rangle = \langle y, T(x)\rangle$; ceci équivaut à dire que $y \in (T(E))^\perp$, d'où la première assertion. Il en résulte (par le corollaire 3.1) que $\overline{T(E)} = (\ker T^*)^\perp$, d'où la deuxième assertion en remplaçant T par son adjoint.

∎

Définition 4.4.2. Soient E et F deux espaces hilbertiens ; un élément $U \in \mathcal{L}(E, F)$ est appelé *unitaire* si $U^* \circ U = \mathrm{Id}_E$ et $U \circ U^* = \mathrm{Id}_F$. Un élément $T \in \mathcal{L}(E)$ est appelé *normal* si $T^* \circ T = T \circ T^*$, *autoadjoint* si $T = T^*$ et *positif* s'il est autoadjoint et si $\langle T(x), x\rangle$ est réel ≥ 0 pour tout $x \in E$.

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

Proposition 4.4.4. *Soient* E *et* F *des espaces hilbertiens 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) 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 $\|T(x)\|^2 = \langle T(x), T(x)\rangle = \langle x, T^*(T(x))\rangle = \|x\|^2$; 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 1.1 que, pour tout $x, y \in E$, on a $\langle x, T^*(T(y))\rangle = \langle x, y\rangle$. Donc $T^*(T(y)) - y \in E^\perp = \{0\}$, 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).

∎

4.5. Familles sommables dans un espace de Banach

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

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

La proposition suivante est laissée en exercice :

Proposition 4.5.1. *Soient* E *et* F *deux espaces normés,* $f \in \mathcal{L}(\mathrm{E},\mathrm{F})$ *et* $(x_i)_{i\in \mathrm{I}}$ *une famille sommable d'éléments de* E *; alors la famille* $(f(x_i))_{i\in \mathrm{I}}$ *est sommable dans* F *et on a* $\sum_{i\in \mathrm{I}} f(x_i) = f(\sum_{i\in \mathrm{I}} x_i)$.

■

Soient E un espace normé et $(x_i)_{i\in \mathrm{I}}$ une famille d'éléments de E ; 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 4.5.2. (i) *Toute famille sommable d'un espace normé vérifie le critère de sommabilité de Cauchy.*

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

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

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

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

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

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

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

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

■

Remarque 4.5.1. Soient E un espace normé, $(x_i)_{i\in \mathrm{I}}$ une famille d'éléments de E, $\varepsilon > 0$ et J une partie finie de I telle que, pour toute partie finie L de I disjointe de J, on ait $\left\|\sum_{i\in \mathrm{L}} x_i\right\| < \varepsilon$. En considérant des parties L à un seul élément, on voit que, pour tout $i \in \mathrm{I}\setminus \mathrm{J}$, $\|x_i\| < \varepsilon$. Donc, si $(x_i)_{i\in\mathrm{I}}$ vérifie le critère de sommabilité de Cauchy, pour tout entier $n \geq 1$, il n'y a qu'un nombre fini d'indices i tels que x_i soit de norme $\geq 1/n$; donc il y a un nombre dénombrable de x_i non nuls. En d'autres termes, on peut toujours se ramener au cas $\mathrm{I} = \mathbb{N}$.

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

La réciproque n'est cependant pas vraie : il existe des séries convergentes $\sum_{k\geq 0} u_k$ telles que la famille (u_k) ne soit pas sommable. En fait, une famille (u_k) de nombres réels est sommable si et seulement si la série $\sum u_k$ est **absolument** convergente, c'est à dire que $\sum_k |u_k| < +\infty$; on a en effet le résultat suivant :

Proposition 4.5.3. *Une famille de nombres réels à termes positifs est sommable si et seulement si les sommes finies sont majorées ; sa somme est alors le plus petit des majorants des sommes finies. Une famille à termes réels est sommable si et seulement si elle est absolument sommable.*

La démonstration est laissée en exercice.

■

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

Corollaire 4.5.1. *Soient* E *un espace de Banach et* $(x_i)_{i\in\mathrm{I}}$ *une famille d'éléments de* E *;*

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

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

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

■

Remarque 4.5.2. On peut plus généralement définir les familles sommables dans un espace vectoriel topologique séparé E : une famille $(x_i)_{i\in\mathrm{I}}$ est dite sommable de somme $\mathrm{S} \in \mathrm{E}$, si, pour tout voisinage V de S dans E, il existe une partie finie J de I telle que, pour toute partie finie K de I contenant J on ait $\sum_{i\in\mathrm{K}} x_i \in \mathrm{V}$.

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 4.5.1. *Soit $(x_i)_{i\in\mathrm{I}}$ un système orthogonal dans un espace hilbertien ; 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

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

■

4.6. Bases hilbertiennes

Définition 4.6.1. Soient E un espace préhilbertien 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.

Théorème 4.6.1. *Tout espace hilbertien admet une base hilbertienne.*

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

Soit B un élément maximal de U ; on veut montrer que B est total, et pour cela, on montre que $\mathrm{B}^\perp = \{0\}$; sinon, il existerait un vecteur x non nul et orthogonal à B (en particulier $x \notin \mathrm{B}$), et quitte à multiplier x par un scalaire convenable on peut supposer $\|x\| = 1$; alors $\mathrm{B} \cup \{x\} \in \mathrm{U}$, ce qui contredirait la maximalité de B. Donc $\mathrm{B}^\perp = \{0\}$, ce qui entraîne que $(\mathrm{B}^\perp)^\perp = \mathrm{H}$. Par le corollaire 3.1, B est total. C'est donc une base hilbertienne.

■

Remarque 4.6.1. Dans le cas séparable, on peut donner une autre démonstration de ce théorème : soit H un espace hilbertien séparable ; soit $(x_n)_{n\geq 0}$ une suite totale dans H ; notons E_n le sous-espace de H engendré par $\{x_k : 0 \leq k < n\}$, et P_n le projecteur orthogonal de H sur E_n. Pour tout $n \geq 0$, posons $y_n = x_n - \mathrm{P}_n(x_n) \in \mathrm{E}_{n+1} \cap \mathrm{E}_n^\perp$. Posons $\mathrm{D} = \{n \in \mathbb{N} : \|y_n\| \neq 0\}$ et, pour $n \in \mathrm{D}$, posons $e_n = \|y_n\|^{-1} y_n$. Il est clair que E_n est engendré par $\{y_k : 0 \leq k < n\}$, donc par $\{e_k : k \in D,\ 0 \leq k < n\}$. Le système $(e_k)_{k\in D}$ est orthonormal et total ; c'est donc une base hilbertienne. Remarquons que cette démonstration marche aussi dans le cas préhilbertien.

Théorème 4.6.2 : inégalité de Bessel. *Soient* E *un espace préhilbertien et* $(e_i)_{i\in I}$ *un système orthonormal dans* E *; pour tout* $x \in E$ *la famille* $(|\langle x, e_i\rangle|^2)_{i\in I}$ *est sommable et*

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

Démonstration. Par la proposition 5.3, il suffit de montrer que, pour toute partie finie J de I, on a $\sum_{i\in J} |\langle x, e_i\rangle|^2 \leq \langle x, x\rangle$. Soit J une partie finie de I ; pour $i \in I$ posons $y_i = \langle x, e_i\rangle e_i$; posons aussi $y = x - \sum_{i\in J} y_i$ et $z = x - y$. Pour $i \in J$, on a $\langle y_i, e_i\rangle = \langle x, e_i\rangle\langle e_i, e_i\rangle = \langle x, e_i\rangle$; pour $j \in J\backslash\{i\}$, on a $\langle y_j, e_i\rangle = \langle x, e_j\rangle\langle e_j, e_i\rangle = 0$; donc $\langle z, e_i\rangle = 0$. Donc les vecteurs y_i, $i \in J$ et z sont deux à deux orthogonaux. On a alors

$$\langle x, x\rangle = \langle z + \sum_{i\in J} y_i, z + \sum_{i\in J} y_i\rangle = \langle z, z\rangle + \sum_{i\in J}\langle y_i, y_i\rangle = \langle z, z\rangle + \sum_{i\in J} |\langle x, e_i\rangle|^2,$$

d'où le résultat.

■

Proposition 4.6.1. *Le cardinal de tout système orthonormal d'un espace préhilbertien est inférieur ou égal à celui de tout système total.*

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

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

■

Corollaire 4.6.1. *Deux bases hilbertiennes d'un espace hilbertien* E *ont le même cardinal.*

Définition 4.6.2. On appelle *dimension hilbertienne* d'un espace hilbertien E le cardinal d'une base hilbertienne quelconque de E.

Théorème 4.6.3 : identité de Parseval. *Soient* E *un espace préhilbertien,* $(e_i)_{i\in I}$ *une base hilbertienne de* E *et* $x \in E$ *; la famille de nombres réels* $(|\langle x, e_i\rangle|^2)_{i\in I}$ *est sommable, la famille de vecteurs* $(\langle x, e_i\rangle\, e_i)_{i\in I}$ *est sommable dans* E *et*

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

Démonstration. Notons H le complété de E et considérons pour simplifier que E est un sous-espace vectoriel de H, avec le produit scalaire induit par celui de H. Remarquons que

$(e_i)_{i\in\mathrm{I}}$ est une base hilbertienne de H, et il suffit de démontrer le résultat en travaillant avec l'espace H. 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 5.1, la famille $(\langle x, e_i\rangle\, e_i)_{i\in\mathrm{I}}$ est sommable dans H et, si on note y sa somme, on a $\sum_{i\in\mathrm{I}} |\langle x, e_i\rangle|^2 = \langle y, y\rangle$. Pour tout $j \in \mathrm{I}$, appliquant la proposition 5.1 à la forme linéaire $z \to \langle z, e_j\rangle$, on trouve $\langle y, e_j\rangle = \sum_{i\in\mathrm{I}}\langle x, e_i\rangle\langle e_i, e_j\rangle = \langle x, e_j\rangle$. Donc $x - y$ est orthogonal aux e_i ; comme le système e_i est total, $x = y$.

■

Exemple 4.6.1. 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 un système orthonormal 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 $\mathrm{L}_2(0, 2\pi)$. Pour toute fonction $f \in \mathrm{L}_2(0, 2\pi)$, les coefficients du développement de f dans cette base sont les coefficients de Fourier complexes

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

D'après Parseval, on a $\|f\|_2^2 = \int_0^{2\pi} |f(s)|^2\, ds/2\pi = \sum_{n\in\mathbb{Z}} |c_n(f)|^2$. On écrit souvent le développement 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 délicat 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).

Corollaire 4.6.2. *Soient* H *un espace hilbertien,* F *un sous-espace vectoriel fermé de* H*, et* $(e_i)_{i\in\mathrm{I}}$ *une base hibertienne du sous-espace* F*. Pour tout vecteur* $x \in \mathrm{H}$*, la projection orthogonale de* x *sur* F *est donnée par*

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

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

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

Remarque 4.6.2. Soient E un espace préhilbertien et $(e_i)_{i\in\mathrm{I}}$ un système orthonormal de vecteurs de E ; on a l'égalité $\|x\|^2 = \sum_{i\in\mathrm{I}} |\langle x, e_i\rangle|^2$ si et seulement si

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

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

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

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

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

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

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

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

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

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

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

Proposition 4.7.1. *Muni du produit scalaire précédent, l'espace vectoriel $\ell^2(\mathrm{I})$ est un espace hilbertien. La famille $(\epsilon_i)_{i\in\mathrm{I}}$ est une base hilbertienne de $\ell^2(\mathrm{I})$.*

Démonstration. Il est clair que $(\epsilon_i)_{i\in\mathrm{I}}$ est un système orthonormal. Pour $\xi = (x_i)_{i\in\mathrm{I}}$ et $i \in \mathrm{I}$ on a $\langle \xi, \epsilon_i\rangle = x_i$, donc $\|\xi\|_2^2 = \sum_i |\langle \xi, \epsilon_i\rangle|^2$. Par la remarque 6.2, $(\epsilon_i)_{i\in\mathrm{I}}$ est une base hilbertienne de $\ell^2(\mathrm{I})$.

Montrons enfin que $\ell^2(\mathrm{I})$ est complet. Notons $u : \ell^2(\mathrm{I}) \to \mathrm{H}$ l'application isométrique de $\ell^2(\mathrm{I})$ dans son complété. Il suffit de montrer que u est surjective. Soit $x \in \mathrm{H}$; posons $\xi = (\langle x, u(\epsilon_i)\rangle)_{i\in\mathrm{I}}$; par le théorème 6.2, on sait que $\xi \in \ell^2(\mathrm{I})$. Pour tout $i \in \mathrm{I}$, on a $\langle \xi, \epsilon_i\rangle = \langle x, u(\epsilon_i)\rangle$, donc $x - u(\xi)$ est orthogonal à $u(\epsilon_i)$. Or $(\epsilon_i)_{i\in\mathrm{I}}$ est total dans $\ell^2(\mathrm{I})$; comme l'image de u est dense dans H, $(u(\epsilon_i))_{i\in\mathrm{I}}$ est total dans H ; on en déduit que $x = u(\xi)$. Il s'ensuit que u est isométrique et bijective. Alors $\ell^2(\mathrm{I})$ est isométrique à H, donc est complet.

∎

Théorème 4.7.1. *Soient* H *un espace hilbertien et* $\mathrm{B} = (e_i)_{i\in\mathrm{I}}$ *une base hilbertienne de* H *; l'application* $\mathrm{U} : x \to (\langle x, e_i\rangle)$ *est une bijection linéaire isométrique de* H *sur* $\ell^2(\mathrm{I})$.

Démonstration. Il est clair que U est une application linéaire de H dans $\mathbb{K}^{\mathrm{I}}$. Par le théorème 6.3, U est isométrique de H dans $\ell^2(\mathrm{I})$. Soit $(\lambda_i)_{i\in\mathrm{I}} \in \ell^2(\mathrm{I})$; par le lemme 5.1, la famille $(\lambda_i e_i)_{i\in\mathrm{I}}$ est sommable dans H ; posons $x = \sum_i \lambda_i e_i$. Pour tout $j \in \mathrm{I}$, appliquant la proposition 5.1 à la forme linéaire $z \to \langle z, e_j\rangle$, on trouve $\langle x, e_j\rangle = \sum_{i\in\mathrm{I}} \lambda_i \langle e_i, e_j\rangle = \lambda_j$. Donc $\mathrm{U}(x) = (\lambda_i)_{i\in\mathrm{I}}$, donc U est surjective.

■

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

Corollaire 4.7.1. *Soient* E *et* F *des espaces hilbertiens,* $(e_i)_{i\in\mathrm{I}}$ *une base hilbertienne de* E*,* $(f_j)_{j\in\mathrm{J}}$ *une base hilbertienne de* F *et* σ *une bijection de* I *sur* J *; il existe une bijection linéaire isométrique* U *de* E *sur* F *telle que, pour tout* $i \in \mathrm{I}$*, on ait* $\mathrm{U}(e_i) = f_{\sigma(i)}$.

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

■

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

Index

Index des notations