ANALYSE FONCTIONNELLE ET

THÉORIE SPECTRALE

MT404

Année 2000-2001

première partie

Sommaire de la première partie

0. Introduction

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

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

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

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. Le *calcul fonctionnel* sera l'occasion de mettre en action nombre d'objets vus en licence.

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

Quelques notations : si X est un ensemble, on note Id_X *l'application identité* de X, c'est à dire l'application de X dans X telle que $\mathrm{Id}_\mathrm{X}(x) = x$ pour tout $x \in \mathrm{X}$. Si A est un sous-ensemble de X, on notera A^c le complémentaire de A. On notera $\mathbf{1}_\mathrm{A}$ la *fonction indicatrice* de A, qui est égale à 1 en tout point de A et à 0 en tout point de A^c. Si f est une fonction réelle définie sur l'ensemble X, il est parfois rapide et agréable d'utiliser la notation des probabilistes $\{f > t\} = \{x \in \mathrm{X} : f(x) > t\}$. De temps en temps on notera 0_E le vecteur nul d'un espace vectoriel E, quand il semblera que cette notation lourde lève toute ambiguïté. La plupart des points du cours sont numérotés, par chapitre-section-type, par exemple le "théorème 2.3.4" serait le quatrième énoncé (théorème, proposition, lemme, corollaire) de la section 3 du chapitre 2 ; à l'intérieur de cette section, le théorème sera appelé "théorème 4", ailleurs dans le chapitre 2 il sera désigné par "théorème 3.4" et dans un autre chapitre "théorème 2.3.4". Les passages écrits en petits caractères contiennent des informations qui peuvent être omises en première lecture. Le poly se termine par un index terminologique et un index des notations.

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

`http://math.univ-mlv.fr/~maurey/ths.html`

en particulier le texte de ce poly, des résumés de cours, etc... Je serais content de recevoir des suggestions constructives pour améliorer l'efficacité de cet outil, que j'utilise de façon vraiment rudimentaire.

B. Maurey

1.1. Normes, semi-normes ; espaces de Banach

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

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

(i) pour tout $x \in \mathrm{X}$ et tout $\lambda \in \mathbb{K}$, on a $p(\lambda x) = |\lambda|\, p(x)$;

(ii) pour tous $x, y \in \mathrm{X}$, on a $p(x+y) \le p(x) + p(y)$.

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

La propriété $p(x+y) \le p(x) + p(y)$ s'appelle *l'inégalité triangulaire* pour la semi-norme p. De l'inégalité triangulaire ci-dessus, on 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 $\mathrm{co}(\mathrm{A})$. C'est aussi le plus petit ensemble convexe contenant A, c'est à dire l'intersection de tous les ensembles convexes qui contiennent 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. Lorsque f est une fonction convexe définie sur un ensemble convexe $\mathrm{C} \subset \mathrm{X}$, les ensembles de la forme $\mathrm{C}_t = \{x \in \mathrm{C} : f(x) \le t\}$ sont des ensembles convexes, pour tout t réel (la réciproque n'est pas vraie).

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*

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

soit convexe.

Démonstration. Si q est positivement homogène et sous-additive, on a vu que q est une fonction convexe, ce qui implique 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 de convexité, c'est à dire que $q(z) \leq 1$. Mais on vérifie immédiatement que $z = (a+b)^{-1}(x+y)$, et l'homogénéité de q transforme alors l'inégalité $q(z) \leq 1$ en $q(x+y) \leq a+b = q(x)+q(y)$.

Si $a > 0$ et $b = q(y) = 0$, on doit procéder un peu différemment. Par homogénéité, on a $q(sy) = 0$ pour tout $s \geq 0$, ce qui indique que la demi-droite $\mathbb{R}_+ \, y$ de direction y, issue de 0_{X}, est entièrement contenue dans C_q. Comme $x_1 \in \mathrm{C}_q$, on voit par convexité que pour tout t tel que $0 < t < 1$, la demi-droite $tx_1 + \mathbb{R}_+ \, y$ est contenue dans C_q (on se fera un petit dessin). En particulier, $t(a^{-1}x + a^{-1}y)$ est dans C_q, donc $q\big(t(a^{-1}x + a^{-1}y)\big) \leq 1$, ce qui donne $t\, q(x+y) \leq a = q(x)+q(y)$, et le résultat en découle en faisant tendre t vers 1. Le dernier cas, quand $a = b = 0$, est laissé au lecteur.

//

Corollaire 1.1.3. *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* $\{x \in \mathrm{X} : p(x) \leq 1\}$ *soit convexe.*

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

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

est une semi-norme ;

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

$$|(1-t)f(s) + tg(s)|^r \leq \big((1-t)|f(s)| + t|g(s)|\big)^r \leq (1-t)|f(s)|^r + t|g(s)|^r$$

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

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

On appelle *espace normé* un espace vectoriel X muni d'une norme p. Si (X, p) est un espace normé, nous en ferons un *espace métrique* en définissant la distance d sur X par $d(x, y) = p(x - y)$, et nous munirons X de la topologie associée à cette métrique, que nous appellerons *topologie de la norme.* Soient $x \in \mathrm{X}$ et $r > 0$; on appelle *boule ouverte* de centre x et de rayon r le sous-ensemble $\mathrm{B}_p(x, r) = \{y \in \mathrm{X} : p(y - x) < r\}$ de X. 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 $\mathrm{B}_p(x, r)$ (ça 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 1.1.4. *Soit* (X, p) *un espace normé ; l'application* $p : \mathrm{X} \to \mathbb{R}_+$ *est continue pour la topologie de la norme.*

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

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

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

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

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

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

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

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

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

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

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

//

Définition 1.1.4. Un espace de Banach est un espace vectoriel normé, complet pour la distance associée à la norme.

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.

Exemples 1.1.5.

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

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

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

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

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

On a déjà vu que $f \to \|f\|_p$ est une semi-norme. Si $\|f\|_p = 0$ et si $f_1 \in \mathcal{L}_p$ est un représentant quelconque de f, on a $\int_0^1 |f_1(s)|^p\,ds = 0$. Comme la fonction $|f_1|^p$ est ≥ 0, cela entraîne que $f_1 = 0$ presque partout, donc f est la classe nulle, c'est à dire que $f = 0_{\mathrm{L}_p}$. On a ainsi montré que $f \to \|f\|_p$ est une norme sur L_p.

Cet espace L_p est de plus complet (voir le chapitre 2).

3. 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

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

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

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

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

Séries de vecteurs

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

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

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

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

Sous cette condition, le reste de la série des normes

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

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

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

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

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

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

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

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

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

//

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

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

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

Exemple 1.1.6. Pour tout $n \geq 0$, désignons par e_n le vecteur de ℓ_p, ou de c_0, dont les coordonnées sont $e_{n,i} = 0$ si $i \neq n$ et $e_{n,n} = 1$. On appellera $(e_n)_{n\geq 0}$ la *suite canonique*. Soit $x = (x_n)$ un élément de c_0. Le vecteur x est la somme (dans l'espace normé c_0) de la série $\sum_{k=0}^{+\infty} x_k e_k$ (petit exercice pour le lecteur ; le même résultat vaut pour tous les ℓ_p avec $p < +\infty$).

Remarque 1.1.7.

Certains espaces importants tels que l'espace $C^\infty(\mathbb{R})$ des fonctions de classe C^∞ sur $\mathbb{R}$, ou bien l'espace produit $\mathbb{R}^\mathbb{N}$, n'admettent pas de norme satisfaisante, c'est à dire de norme qui définisse la topologie que l'on estime être la bonne sur l'espace considéré (la topologie produit, dans le cas de $\mathbb{R}^\mathbb{N}$). 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 $x, y, v \in \mathrm{X}$, on a $d(x+v, y+v) = d(x, y)$ (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*. De même, on pourra définir sur $\mathbb{R}^\mathbb{N}$ la distance

$$d(x, y) = \sum_{n=0}^{+\infty} \min(|x_n - y_n|, 2^{-n})$$

où $x = (x_n)_{n \ge 0}$ et $y = (y_n)_{n \ge 0}$ sont deux éléments quelconques de $\mathbb{R}^\mathbb{N}$.

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* $\mathrm{M} \ge 0$ *tel que, pour tout* $x \in \mathrm{X}$ *on ait*

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

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 $\mathrm{M} = \delta^{-1}$. Enfin, supposons (iii) vérifiée ; si une suite (x_n) de X tend vers un vecteur $x \in \mathrm{X}$, on aura

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

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

//

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

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

Démonstration. On utilise l'équivalence $(i) \iff (iii)$ du théorème 1, appliquée 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 borné* une application linéaire continue entre deux espaces normés. Dans le cas où $\mathrm{Y} = \mathrm{X}$, on note simplement $\mathcal{L}(\mathrm{X})$ l'espace des endomorphismes continus de X.

Soit $f : \mathrm{X} \to \mathrm{Y}$ une application linéaire continue ; d'après le théorème 1, il existe une constante M telle que $\|f(x)\|_{\mathrm{Y}} \le \mathrm{M}$ pour tout vecteur x de X tel que $\|x\|_{\mathrm{X}} \le 1$. On peut donc considérer la quantité (finie)

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

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 $\|z\|_{\mathrm{X}} \le 1$, donc $\|f(z)\|_{\mathrm{Y}} \le \|f\|$, d'où par homogénéité $\|x\|^{-1}\|f(x)\| \le \|f\|$, ou encore

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

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

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

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

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

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 $f \to \|f\|$ 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 dans la boule unité de X.

//

Exemples 1.2.1.

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

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

Il est clair que $\|fg\|_\infty \leq \|f\|_\infty\,\|g\|_\infty$ pour toute fonction g, donc $\|\mathrm{M}_f\| \leq \|f\|_\infty$. Démontrons l'égalité $\|\mathrm{M}_f\| = \|f\|_\infty$; posons $\mathrm{M} = \|f\|_\infty$, et considérons pour $\varepsilon > 0$ l'ouvert non vide $\mathrm{U} = \{s \in [0,1] : |f(s)| > \mathrm{M} - \varepsilon\}$. Si g est une fonction continue ≥ 0 et non nulle, à support dans U (par exemple $g(s) = \mathrm{dist}(s, \mathrm{U}^c)$), elle atteint son maximum en un point $s_1 \in \mathrm{U}$. Alors

$$\|\mathrm{M}_f\|\,\|g\|_\infty \geq \|fg\|_\infty \geq |f(s_1)g(s_1)| = |f(s_1)|\,\|g\|_\infty > (\mathrm{M}-\varepsilon)\,\|g\|_\infty$$

d'où $\|\mathrm{M}_f\| \geq \|f\|_\infty$ puisque $\varepsilon > 0$ est arbitraire.

Il est en général très difficile de calculer exactement la norme d'une application linéaire continue. A titre d'anecdote, on sait depuis environ 1920 que la transformée de Fourier définit un opérateur T_p continu de $\mathrm{L}_p(\mathbb{R})$ dans $\mathrm{L}_q(\mathbb{R})$ lorsque $1 < p < 2$ et $1/p + 1/q = 1$, mais la valeur exacte de la norme n'a été établie que 50 ans plus tard !

La proposition suivante est facile mais importante.

Proposition 1.2.4. *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 1.2.5. *Soient* X *et* Y *deux espaces normés ; si* Y *est un espace de Banach, l'espace* $\mathcal{L}(\mathrm{X}, \mathrm{Y})$ *est un espace de Banach.*

Démonstration. Supposons que Y soit un espace de Banach. Soit $\sum u_k$ une série normalement convergente dans $\mathcal{L}(\mathrm{X}, \mathrm{Y})$; pour tout vecteur $x \in \mathrm{X}$, on a $\|u_k(x)\| \leq \|u_k\| \, \|x\|$, 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 $x \in \mathrm{X}$ on a $\|\mathrm{U}(x)\| \leq \sum_{k=0}^{+\infty} \|u_k(x)\| \leq \big(\sum_{k=0}^{+\infty} \|u_k\|\big) \|x\|$, ce qui montre que U est continue et

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

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

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

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

//

Image d'une série convergente. Soit $\sum u_k$ une série convergente de vecteurs dans l'espace normé X et soit $f : \mathrm{X} \to \mathrm{Y}$ une application linéaire continue. Alors la série $\sum f(u_k)$ converge dans Y et

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

Démonstration. La suite des sommes partielles $\mathrm{U}_n = \sum_{k=0}^{n} u_k$ converge dans X vers la somme U de la série, on a $f(\mathrm{U}_n) = \sum_{k=0}^{n} f(u_k)$ par linéarité de f, et $f(\mathrm{U}_n)$ tend vers l'image de U par continuité de f.

//

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 $y \in \mathrm{Y}$ puisque Y est fermé, 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* $\|h\| = \|g\|$.

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 $\xi \in$ X/Y et pour tout x tel que $\pi(x) = \xi$, on a

$$\|h(\xi)\| = \|g(x)\| \le \|g\| \, \|x\|.$$

Cela étant vrai pour tout x dans la classe ξ, 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 $\|\pi\| \le 1$, et on a bien $\|h\| = \|g\|$.

//

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

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

//

1.4. Complété d'un espace normé

Commençons par un lemme de prolongement.

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

C'est par ce procédé que l'on définit par exemple la transformée de Fourier sur $X = F = L_2(\mathbb{R})$, à partir de sa définition intégrale sur le sous-espace dense $X_0 = L_1 \cap L_2$.

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

$$A_n = \{y \in X_0 : \|y - x\| < 2^{-n}\}$$

est non vide ; si $y, y' \in A_n$, on a $\|y'-y\| \le \|y'-x\|+\|y-x\| \le 2^{-n+1}$, donc le diamètre de A_n tend vers 0 (l'ensemble A_n devrait s'appeler $A_n(x)$, mais ce serait vraiment trop lourd). Puisque f est linéaire bornée, la suite $(f(A_n))$ est une suite décroissante de sous-ensembles non vides de F, de diamètres tendant vers 0. Précisément, on déduit de ce qui précède que si $v, w \in \overline{f(A_n)}$, on a $\|v - w\| \le \|f\| \, 2^{-n+1}$. Puisque F est complet, on sait que $\bigcap_n \overline{f(A_n)}$ contient exactement un point. Appelons $g(x)$ cet unique point.

Soit $(y_k)_k \subset X_0$ une suite quelconque telle que $y_k \to x$, et soit n_0 quelconque ; on aura $y_k \in A_{n_0}$ pour tout $k \ge k_0$, donc $\|f(y_k) - g(x)\| \le 2^{-n_0+1} \, \|f\|$ pour $k \ge k_0$; ceci montre que $g(x) = \lim_k f(y_k)$ pour toute suite $(y_k) \subset X_0$ telle que $x = \lim_k y_k$. Il est facile de vérifier que $x \to g(x)$ est linéaire de X dans F, à partir de cette remarque (prendre $y_k \to x$ et $y'_k \to x'$). On a aussi

$$\|g(x)\| = \lim_k \|f(y_k)\| \le \|f\| \, \lim_k \|y_k\| = \|f\| \, \|x\|,$$

ce qui montre que g est continue et $\|g\| \leq \|f\|$. Si $x \in \mathrm{X}_0$, il est clair que $f(x)$ est l'unique point commun aux ensembles $f(\mathrm{A}_n)$, donc $g(x) = f(x)$ dans ce cas, ce qui montre que g prolonge f ; il en résulte que $\|f\| \leq \|g\|$, donc $\|f\| = \|g\|$. Si g_1 est une autre application continue qui prolonge f, on aura $g_1(x) = \lim_k g_1(y_k)$ par continuité de g_1, mais $g_1(y_k) = f(y_k)$ par hypothèse, donc $g_1(x) = \lim_k f(y_k) = g(x)$ pour tout $x \in \mathrm{X}$, ce qui montre l'unicité de g. Il nous suffit pour finir de prendre $\widetilde{f} = g$.

//

Corollaire 1.4.2. *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 toute* $f \in \mathcal{L}(\mathrm{X}, \mathrm{F})$*, il existe une unique* $g \in \mathcal{L}(\mathrm{Y}, \mathrm{F})$ *telle que* $f = g \circ u$*. On a* $\|f\| = \|g\|$*.*

Démonstration. Il suffit de considérer le sous-espace $\mathrm{Y}_0 = u(\mathrm{X})$, dense dans Y, et de prolonger à Y l'application $f \circ u^{-1}$ qui est définie de Y_0 dans F.

//

Théorème 1.4.3. *Soit* X *un espace normé ; il existe un couple* $(\widehat{\mathrm{X}}, u)$ *où* $\widehat{\mathrm{X}}$ *est un espace de Banach et* $u : \mathrm{X} \to \widehat{\mathrm{X}}$ *est une application linéaire isométrique d'image dense.*

Démonstration. La somme de deux suites de Cauchy $x = (x_n)$ de vecteurs de X 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 espace vectoriel, qui sera noté Y. De plus, les suites de Cauchy sont des suites bornées. Pour tout élément $x = (x_n)$ de Y on peut donc poser

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

On vérifie sans peine que cette quantité définit une norme sur Y. Posons

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

on vérifie que Z, muni de la norme $\|x\|_\infty$, est un sous-espace vectoriel fermé de Y ; notons $\widehat{\mathrm{X}} = \mathrm{Y}/\mathrm{Z}$ le quotient de Y par Z et $\pi : \mathrm{Y} \to \widehat{\mathrm{X}}$ l'application quotient ; munissons $\widehat{\mathrm{X}}$ de la norme quotient, notée $\|.\|$. Le point crucial est que $\widehat{\mathrm{X}}$ est complet, même si X n'était pas complet. Il est commode de démontrer d'abord les autres propriétés voulues.

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 $\|u(x)\| \geq \|x\|$, mais le choix du représentant $y = \mathrm{U}(x)$ donne $\|y\|_\infty = \|x\|$, donc $\|u(x)\| = \|x\|$ pour tout $x \in \mathrm{X}$.

Vérifions que $u(\mathrm{X})$ est dense dans $\widehat{\mathrm{X}}$. Soient $\xi \in \widehat{\mathrm{X}}$ et $\varepsilon > 0$; il existe une suite de Cauchy $y = (y_n)$ telle que $\xi = \pi(y)$. Il existe un entier $n \geq 0$ tel que, pour tout $k \geq n$ on ait $\|y_k - y_n\| \leq \varepsilon$. Considérons l'élément y' de Y défini par $y'_j = y_j$ si $j < n$ et $y'_j = y_n$ si $j \geq n$. Ce que nous avons dit se traduit par $\|y - y'\|_\infty \leq \varepsilon$, donc $\|\pi(y) - \pi(y')\| \leq \varepsilon$, mais y' est un représentant de $\mathrm{U}(y_n)$, donc $\|\pi(y) - u(y_n)\| \leq \varepsilon$, ce qui montre la densité voulue.

Pour finir la démonstration, il nous reste à démontrer que $\widehat{\mathrm{X}}$ est un espace de Banach. Soit (ξ_n) une suite de Cauchy dans $\widehat{\mathrm{X}}$; comme $u(\mathrm{X})$ est dense, pour tout entier $n \geq 0$, il existe $x_n \in \mathrm{X}$ tel que $\|\xi_n - u(x_n)\| \leq 2^{-n}$. La suite $(u(x_n))_{n \geq 0}$ est alors de Cauchy et, comme u est isométrique, la suite $x = (x_n)_{n \geq 0}$ est de Cauchy dans X ; on vérifie alors que $\pi(x)$ est la limite de la suite de Cauchy (ξ_n) : Comme ci-dessus, on voit que la suite $(u(x_n))$ converge vers $\pi(x)$; comme la suite $(\xi_n - u(x_n))$ converge vers 0, la suite (ξ_n) converge vers $\pi(x)$, d'où le résultat.

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

Séparé complété

Si X est un espace vectoriel 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 $\xi = 0_{\mathrm{X}/\mathrm{Z}}$.

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

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

La fonction p est une semi-norme sur $\mathrm{X} = \mathcal{L}_p$, et l'espace Z est l'espace vectoriel 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 $X_{\mathbb{C}}$ obtenu en munissant l'espace vectoriel réel $X \times X$ de la structure d'espace vectoriel complexe donnée par $(a+ib)(x,y) = (ax-by, ay+bx)$ (pour $a, b \in \mathbb{R}$ et $x, y \in X$). On plonge X dans $X_{\mathbb{C}}$ par l'application $u : x \to (x, 0)$, de sorte que (x, y) peut s'écrire $u(x) + iu(y)$.

Si l'espace vectoriel réel X est de plus normé, on a envie de définir sur $X_{\mathbb{C}}$ une norme p qui soit une norme *d'espace vectoriel 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 $p(x, 0) = \|x\|_X$. 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 $p(z) = \max\{\|\operatorname{Re}(\lambda z)\|_X : \lambda \in \mathbb{C}, |\lambda| = 1\}$, ce qui s'écrit aussi

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

On peut aussi se trouver dans la situation où X est un espace vectoriel complexe, mais muni d'une norme q d'espace vectoriel réel : autrement dit, on a $q(\lambda x) = |\lambda|\, q(x)$ pour tout $x \in X$ 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, q), on peut remplacer q par une norme p équivalente à q qui vérifie de plus $p(\lambda x) = |\lambda|\, p(x)$ pour tout $x \in X$ et tout $\lambda \in \mathbb{C}$, en posant pour tout $x \in X$

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

C'est exactement ce qu'on a fait au paragraphe précédent : on a commencé par définir sur $X_{\mathbb{C}}$ la norme $q(x+iy) = \max(\|x\|, \|y\|)$, mais cette norme n'est pas une norme complexe. On l'a alors remplacée par p, définie comme dans ce paragraphe.

1.6. Dual d'un espace normé, application transposée

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

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

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

On a en effet $|f_x(y)| \le \|y\|_\infty \sum_{k=0}^{+\infty} |x_k| = \|x\|_1\, \|y\|_\infty$ (en utilisant la majoration $|y_k| \le \|y\|_\infty$), pour tout $y \in c_0$, qui montre que $\|f_x\| \le \|x\|_1$. Si on choisit $y = (y_n)$ de façon que pour $0 \le n \le N$, le nombre y_n soit le complexe de module un tel que $x_n y_n = |x_n|$, et que $y_n = 0$ si $n > N$, on aura $\|y\|_\infty = 1$ et $\sum_{k=0}^{N} |x_k| = \sum_{k=0}^{N} x_k y_k = f_x(y) \le \|f_x\|$. En faisant tendre N vers l'infini on obtient $\|f_x\| = \|x\|_1$.

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

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

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

Démonstration. Soit $g \in X^*_{\mathbb{C}}$; 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 injective donc bijective.

//

Définition 1.6.2. Soient X et Y deux espaces normés et $f \in \mathcal{L}(X, Y)$; 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 1.6.2. *Soient X, Y et Z des espaces normés ;*

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

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

(iii) pour tout $f \in \mathcal{L}(X, Y)$ et tout $g \in \mathcal{L}(Y, Z)$, 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 $y^* \in Y^*$ tel que $\|y^*\| \le 1$. Pour tout vecteur $x \in X$ 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.

1.7. Parties totales. Séparabilité

Définition 1.7.1. On dit qu'un espace métrique (Z, d) est *séparable* s'il existe une partie *dénombrable* $D \subset Z$ qui soit dense dans Z.

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

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

– 3. Pour que X normé soit séparable, il faut et il suffit qu'il existe une suite croissante (F_n) de sous-espaces de dimension finie de X telle que $\bigcup_n \mathrm{F}_n$ soit dense dans X :

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

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

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

Exemples 1.7.2.

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

Pour tout $n \geq 0$, désignons par e_n le $n$ième vecteur de la *suite canonique* définie à l'exemple 1.6. Le sous-espace $\mathrm{F}_n = \mathrm{Vect}(e_0, \ldots, e_{n-1})$ est formé des vecteurs y de ℓ_p dont les coordonnées y_j, $j \geq n$ sont nulles. Il est facile de voir que tout $x \in \ell_p$ est limite d'une suite de tels vecteurs y.

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

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

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

Par exemple, la suite canonique $(e_n)_{n \geq 0}$ est totale dans ℓ_p pour tout $p < \infty$. Elle n'est pas totale dans ℓ_∞, mais elle est totale dans c_0.

Proposition 1.7.1. *Pour qu'un espace normé* X *soit séparable, il faut et il suffit qu'il admette une partie dénombrable totale.*

Démonstration. Soit X un espace normé tel qu'il existe une suite (x_n) d'éléments totale dans X ; l'espace vectoriel L engendré par la suite est dense dans X, et il est égal à la réunion croissante des sous-espaces L_n de dimension finie définis par $\mathrm{L}_n = \mathrm{Vect}(x_0, \ldots, x_{n-1})$. On sait alors que X est séparable puisque $\bigcup_{n \geq 0} \mathrm{L}_n$ est dense dans X. Dans l'autre direction c'est trivial.

//

2. Les espaces de Banach classiques

2.1. Espaces de fonctions continues ou intégrables

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

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

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

Lorsque K est *métrique* compact, avec une distance d, on dispose pour pas cher de beaucoup de fonctions réelles continues sur K, fabriquées avec les fonctions $s \to d(s,t)$, où t est un point fixé de K. Par exemple, la fonction f définie sur K par la formule $f(t) = \max(0, \varepsilon - d(t, t_0))$ est non nulle dans un petit voisinage de $t_0 \in \mathrm{K}$, et seulement dans ce petit voisinage.

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

Pour un compact quelconque K, on peut encore trouver pour tout point $t_0 \in \mathrm{K}$ une fonction réelle continue, non nulle en t_0 et nulle en dehors d'un voisinage donné de t_0, mais c'est plus difficile. En revanche étant donné un ouvert U d'un compact K non métrisable, on ne peut pas toujours trouver une fonction réelle continue f sur K telle que $\mathrm{U} = \{f \neq 0\}$. Un exemple de compact non métrisable est l'espace produit $\{-1, 1\}^{\mathbb{R}}$ (pour bien dire qu'on ne verra pas ça tous les jours : c'est un produit *non dénombrable*).

Voici un exemple d'application de la notion de partition de l'unité.

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

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

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

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

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

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

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

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

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

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

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

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

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

Beaucoup d'exemples d'espaces de Banach proviennent de la théorie de l'intégration. Un *espace mesurable* $(\Omega, \mathcal{A})$ est la donnée d'un ensemble Ω et d'une tribu $\mathcal{A}$ de parties de l'ensemble Ω. Nous supposerons donné un espace $(\Omega, \mathcal{A}, \mu)$, où $(\Omega, \mathcal{A})$ est un espace mesurable et μ une mesure positive sur $(\Omega, \mathcal{A})$. Pour éviter certains désagréments nous supposerons que la mesure est *σ-finie*, ce qui veut dire qu'il existe une partition (Ω_n) de Ω en une *suite* de parties $\Omega_n \in \mathcal{A}$ telles que $\mu(\Omega_n) < +\infty$. Un exemple typique est $\mathbb{R}^d$ muni de la mesure de Lebesgue (sur la tribu borélienne), ou bien $\mathbb{N}$ muni de la *mesure de comptage* μ qui associe à tout $\mathrm{A} \subset \mathbb{N}$ le nombre $\mu(\mathrm{A})$ (fini ou $+\infty$) de ses éléments.

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

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

On a vu dans l'exemple 1.1.2 que la quantité ci-dessus définit une semi-norme sur $\mathcal{L}_p$, puis dans l'exemple 1.1.5 que l'on obtient une norme sur L_p en passant au quotient. Cet espace L_p est de plus complet : on peut utiliser le critère des séries normalement convergentes de la proposition 1.1.6 et quelques arguments d'intégration pour retrouver ce théorème du cours d'Intégration (appelé souvent *théorème de Fisher-Riesz*).

Indiquons les grandes lignes de la démonstration. Soit $\sum u_k$ une série d'éléments de L_p telle que $\mathrm{M} = \sum_{k=0}^{+\infty} \|u_k\|_p < +\infty$. Nous devons montrer que la série converge dans L_p. Posons $v_k = |u_k|$, $g_n = (\sum_{k=0}^{n} v_k)^p$, remarquons que $\|v_k\|_p = \|u_k\|_p$ pour obtenir $\int_0^1 g_n(s)\,ds = \|\sum_{k=0}^{n} v_k\|_p^p \le \mathrm{M}^p$. La suite (g_n) est une suite croissante de fonctions mesurables ≥ 0, elle converge vers une fonction mesurable g (valeur $+\infty$ admise) dont la valeur en chaque point est $g(s) = \bigl(\sum_{k=0}^{+\infty} |u_k(s)|\bigr)^p$ (valeur $+\infty$ admise à nouveau) ; on sait que $\int_0^1 g(s)\,ds = \lim_n \int_0^1 g_n(s)\,ds \le \mathrm{M}^p$. La fonction g est donc finie presque partout, donc la série $\sum u_k(s)$ converge absolument pour presque tout s. On pose alors pour presque tout s

$$\mathrm{U}(s) = \sum_{k=0}^{+\infty} u_k(s)$$

et on remarque que $|\mathrm{U}(s) - \mathrm{U}_n(s)|^p \le g(s)$ pour presque tout s, et que $|\mathrm{U}(s) - \mathrm{U}_n(s)|^p$ tend vers 0 lorsque $n \to +\infty$ pour presque tout s. Comme g est intégrable, le théorème de convergence dominée de Lebesgue permet de conclure.

On a vu dans l'exemple 1.1.5 l'espace ℓ_p, qui est l'espace des suites scalaires $x = (x_n)$ telles que $\sum |x_n|^p < +\infty$. Pour unifier les arguments, on peut dire que ℓ_p est l'espace $\mathrm{L}_p(\Omega, \mu)$ pour la mesure de comptage μ sur $\Omega = \mathbb{N}$.

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

Théorème. *L'espace* $\mathrm{C}(\mathrm{K})$ *est dense dans* $\mathrm{L}_p(\mathrm{K}, \mu)$ *(lorsque* $p < +\infty$*).*

On montre que les ensembles $\mathrm{A} \in \mathcal{B}$ tels que $\mathbf{1}_\mathrm{A}$ soit limite dans L_p d'une suite de fonctions continues (f_n) telle que $0 \le f_n \le 1$ forment une tribu $\mathcal{A}$. On vérifie que cette tribu contient les ouverts de K, donc $\mathcal{A} = \mathcal{B}$. Il en résulte que toute fonction $\mathcal{B}$-étagée est limite de fonctions continues pour la norme L_p, d'où le résultat parce que les fonctions étagées sont denses dans L_p.

Corollaire. *Lorsque* (K, d) *est métrique et* $p < +\infty$*, l'espace* $\mathrm{L}_p(\mathrm{K}, \mu)$ *est séparable.*

Un bon nombre de questions d'intégration se traitent au moyen d'un lemme technique, le *lemme des classes monotones*. On dit que M est une *classe monotone* si M est stable par réunion dénombrable croissante et intersection dénombrable décroissante. Notons $\sigma(\mathrm{A})$ la tribu (ou *σ-algèbre*) engendrée par une famille de parties d'un ensemble Ω.

Si A *est une algèbre de parties de* Ω*, contenue dans une classe monotone* M*, alors* $\sigma(\mathrm{A}) \subset \mathrm{M}$.

On va montrer qu'étant donnée une algèbre $\mathrm{B} \subset \mathrm{M}$, il existe une algèbre B' telle que $\mathrm{B} \subset \mathrm{B}' \subset \mathrm{M}$, et qui contient toutes les réunions croissantes d'éléments de B. Posons $\mathrm{B}_0 = \mathrm{B}$, puis désignons par B_{2n+1} l'ensemble des parties de Ω qui sont réunion d'une suite croissante d'éléments de B_{2n}, et désignons par B_{2n+2} l'ensemble des parties de Ω qui sont intersection d'une suite décroissante d'éléments de B_{2n+1}. En considérant les suites constantes de parties on constate que $\mathrm{B}_n \subset \mathrm{B}_{n+1}$ pour tout $n \ge 0$. On vérifie assez facilement que $\mathrm{B}' = \bigcup_n \mathrm{B}_n$ est une algèbre, et $\mathrm{B}' \subset \mathrm{M}$. Avec le lemme de Zorn on pourra conclure ainsi : si A_∞ est une algèbre *maximale* vérifiant $\mathrm{A} \subset \mathrm{A}_\infty \subset \mathrm{M}$, on aura $(\mathrm{A}_\infty)' = \mathrm{A}_\infty$ par maximalité (en appliquant ce qui précède à $\mathrm{B} = \mathrm{A}_\infty$), ce qui signifie que A_∞ est une algèbre stable par union dénombrable, donc A_∞ est une tribu qui contient A. On aura donc $\sigma(\mathrm{A}) \subset \mathrm{A}_\infty \subset \mathrm{M}$.

L'espace ℓ_∞ est l'espace de Banach des suites scalaires bornées. Il existe un analogue de ℓ_∞ en théorie de l'intégration : c'est l'espace $\mathrm{L}_\infty(\Omega, \mu)$ des classes de fonctions mesurables bornées sur Ω (c'est à dire des classes qui contiennent un représentant borné). La norme $\|f\|_\infty$ est la plus petite constante M telle que l'on ait $|f(s)| \le \mathrm{M}$ 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

$$f(u) - f(t) = \int_t^u g(s)\, ds$$

pour tous $t, u \in [a, b]$. On dit que g est la *dérivée généralisée* de f, et on la note souvent simplement f'. La norme est alors

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

Le cas de $\mathbb{R}^n$, $n \ge 2$ ou d'un ouvert de $\mathbb{R}^n$ (qui est le cas sérieux pour les applications aux EDP) ne sera pas traité ici.

2.2. Espace de Hilbert

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

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

On démontre l'inégalité de Cauchy-Schwarz,

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

(voir chapitre 5) qui implique en particulier que $x \to \sqrt{\langle x, x\rangle}$ est une semi-norme sur l'espace vectoriel E. Certains auteurs exigent qu'un produit scalaire vérifie $\langle x, x\rangle > 0$ pour tout $x \neq 0_{\mathrm{E}}$. Dans ce cas la semi-norme $x \to \sqrt{\langle x, x\rangle}$ est une norme sur E.

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

Si H est un espace de Hilbert, on notera $\|x\| = \sqrt{\langle x, x\rangle}$ pour tout $x \in \mathrm{H}$. L'inégalité de Cauchy-Schwarz s'écrit alors $|\langle x, y\rangle| \leq \|x\| \, \|y\|$, et elle montre que pour tout y fixé, la forme linéaire $x \to \langle x, y\rangle$ est continue sur H.

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

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

L'espace ℓ_2 est un cas particulier. L'espace de Sobolev $\mathrm{H}^1([a, b])$ de la section précédente est aussi un espace hilbertien pour le nouveau produit scalaire

$$\langle f, g\rangle_{\mathrm{H}^1} = \int_a^b f(s)\overline{g(s)} \, ds + \int_a^b f'(s)\overline{g'(s)} \, ds.$$

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

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

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

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

Lemme 2.2.1. *Soient* $(u_1, \ldots, u_n)$ *des vecteurs deux à deux orthogonaux d'un espace de Hilbert* H *; on a*

$$\Big\|\sum_{k=0}^{n} u_k\Big\|^2 = \sum_{k=0}^{n} \|u_k\|^2.$$

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

Démonstration. Facile, en développant les carrés scalaires.

//

Lemme 2.2.2. *Soit* $(e_1, \ldots, e_n)$ *une suite orthonormée finie dans un espace de Hilbert* H *; posons* $\mathrm{F} = \mathrm{Vect}(e_1, \ldots, e_n)$ *; pour tout vecteur* $x \in \mathrm{H}$*, le vecteur*

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

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

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

//

Lemme 2.2.3 : inégalité de Bessel. *Soient* H *un espace de Hilbert et* $(e_n)_{n \geq 0}$ *une suite orthonormée dans* H *; pour tout* $x \in \mathrm{H}$ *la série numérique* $\sum_k |\langle x, e_k \rangle|^2$ *est convergente et*

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

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

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

d'où le résultat.

//

Lemme 2.2.4. *Soit* $(u_n)_{n\geq 0}$ *une suite orthogonale ; la série de vecteurs* $\sum_k u_k$ *converge si et seulement si* $\sum \|u_k\|^2 < +\infty$, *et*

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

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

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

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

A partir de là, il est clair que la suite (U_n) est de Cauchy dans H si et seulement si la série numérique $\sum_k \|u_k\|^2$ vérifie le critère de convergence de Cauchy. La norme de la somme de la série s'obtient en passant à la limite dans l'égalité du lemme 1.
//

Lemme 2.2.5. *Soit* $(e_n)_{n\geq 0}$ *une suite orthonormée dans* H *et soit* F *le sous-espace vectoriel fermé engendré par la suite* $(e_n)_{n\geq 0}$ *; pour tout vecteur* $y \in \mathrm{F}$, *on a*

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

Démonstration. Posons $c_j = \langle y, e_j\rangle$ pour tout $j \geq 0$, et $z = \sum_{k=0}^{+\infty} c_k\, e_k$. Cette série converge d'après le lemme 3 et le lemme précédent. Pour tout $j \geq 0$, on voit en passant à la limite par la continuité de l'application $x \to \langle x, e_j\rangle$

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

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

Théorème 2.2.6. *Pour tout espace de Hilbert séparable* H *de dimension infinie, il existe une suite orthonormée* $(e_n)_{n\geq 0}$ *dans* H *telle que*

$$\forall x \in \mathrm{H}, \quad x = \sum_{k=0}^{+\infty} \langle x, e_k\rangle\, e_k.$$

On appelle *base hilbertienne* de H une suite orthonormée qui a la propriété indiquée dans le théorème précédent. Cette propriété implique clairement que la suite (e_n) doit être totale dans H ; inversement, le lemme 5 montre qu'une suite orthonormée (e_n) est une base hilbertienne de H si et seulement si elle est totale dans H. De plus on a alors $\|x\|^2 = \sum_{k=0}^{+\infty} |\langle x, e_k\rangle|^2$ pour tout $x \in \mathrm{H}$ d'après le lemme 4.

Si H est de dimension finie, l'existence de base orthonormée (bien entendu *finie*) a été vue en DEUG. Le cas des espaces de Hilbert non séparables sera examiné au chapitre 5.

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

La suite (e_n) est totale puisque l'espace vectoriel qu'elle engendre contient la réunion $\bigcup_{n\geq 0} \mathrm{E}_n$. La conclusion provient du lemme 5.

//

Exemples 2.2.4.

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

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

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

pour tout $s \in [0, 2\pi]$. Il est facile de vérifier que les fonctions $(f_n)_{n\in\mathbb{Z}}$ forment une suite orthonormée dans $\mathrm{L}_2(0, 2\pi)$. En revanche, il faut une petite démonstration pour voir que ce système est total (voir la section 2.5). Il s'agit donc d'une base orthonormée de $\mathrm{L}_2(0, 2\pi)$.

3. Autre base orthonormée : le système de Haar.

Posons $h(t) = 1$ si $0 \leq t < 1/2$, $h(t) = -1$ si $1/2 \leq t < 1$, et $h(t) = 0$ sinon. A partir de cette fonction de base, on construit par translation et changement d'échelle une famille de fonctions, pour tous $n, j \in \mathbb{Z}$:

$$\forall t \in \mathbb{R}, \quad h_{n,j}(t) = 2^{n/2}\, h(2^n t - j).$$

On a $h_{0,0} = h$. Celles des fonctions qui sont à support dans $[0, 1]$ forment une suite orthonormée dans $\mathrm{L}_2(0, 1)$. Elles correspondent aux indices $n \geq 0$ et $0 \leq j < 2^n$. Pour des raisons de cardinalité on voit que la fonction $h_0 = 1$ et les fonctions précédentes $h_{m,j}$, pour $m = 0, \ldots, n-1$ et $j = 0, \ldots, 2^{m-1}$, qui sont au nombre de $1 + \sum_{m=0}^{n-1} 2^m = 2^n$ engendrent toutes les fonctions en escalier sur la partition de $[0, 1[$ en intervalles $[k2^{-n}, (k+1)2^{-n}[$, $k = 0, \ldots, 2^n - 1$. Il en résulte que la famille formée de la fonction 1 et des $h_{n,j}$ pour $n \geq 0$ et $0 \leq j < 2^n$ est totale dans $\mathrm{L}_2(0, 1)$, donc c'est une base orthonormée de $\mathrm{L}_2(0, 1)$, appelée *base de Haar*.

Si on travaille sur $\mathbb{R}$ on peut choisir de translater la base précédente sur chaque intervalle entier, ou bien abandonner h_0 et considérer toutes les fonctions $h_{n,j}$, pour $n, j \in \mathbb{Z}$.

Le système de Haar est un exemple de *base d'ondelettes* avant l'heure.

4. Considérons sur $\mathrm{K} = \{1, -1\}^{\mathbb{N}}$ la suite des fonctions coordonnées $(\varepsilon_j)_{j\geq 0}$, définies pour tout $j \geq 0$ par $\varepsilon_j(t) = t_j$ si $t = (t_i)_{i\geq 0} \in \mathrm{K}$. Pour tout ensemble fini $\mathrm{A} \subset \mathbb{N}$ on posera $w_{\mathrm{A}} = \prod_{j\in \mathrm{A}} \varepsilon_j$. On obtient une base orthonormée de $\mathrm{L}_2(\mathrm{K})$. C'est le *système de Walsh.*

Théorème 2.2.7 : théorème de projection. *Soient* H *un espace de Hilbert 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* $y \in \mathrm{C}$,

$$\mathrm{Re}\langle x - y_0, y - y_0\rangle \leq 0.$$

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 l'espace 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 F de H. Dans ce cas on a $\langle x - y_0, z\rangle = 0$ pour tout vecteur $z \in \mathrm{F}$, c'est à dire que $x - y_0 \perp \mathrm{F}$.

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 $y = u\,z + y_0 \in \mathrm{F}$ 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 F, la projection y_0 de x sur F est entièrement caractérisée par les deux conditions suivantes

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

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

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

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

On notera $\mathrm{P_F}(x) = y_0$ la projection orthogonale de x sur F. La caractérisation ci-dessus montre que $\mu\mathrm{P_F}(x) + \mu'\mathrm{P_F}(x')$ est la projection de $\mu x + \mu' x'$, autrement dit l'application

P_F est une application linéaire. L'égalité $(*)$ ci-dessus donne aussi $\|x-y\| \geq \|P_F(x)-y\|$ pour tout $y \in F$, donc $\|x\| \geq \|P_F(x)\|$ en prenant $y = 0$; on a donc $\|P_F\| \leq 1$.

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

Exemple 2.2.5. 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

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

Corollaire 2.2.8. *Soient* H *un espace hilbertien,* F *un sous-espace vectoriel fermé séparable de* H*, et* $(e_n)_{n\geq 0}$ *une base hibertienne du sous-espace* F*. Pour tout vecteur* $x \in H$*, la projection orthogonale de* x *sur* F *est donnée par*

$$P_F(x) = \sum_{k=0}^{+\infty} \langle x, e_k \rangle\, e_k.$$

Démonstration. Posons $y = P_F(x)$; puisque $x-y$ est orthogonal à F, on a $\langle x-y, e_j \rangle = 0$ pour tout $j \geq 0$, donc $\langle x, e_j \rangle = \langle y, e_j \rangle$. D'après le lemme 5, appliqué à F et à y, on a

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

//

Définition 2.2.6. On dit que des parties A et B d'un espace de Hilbert H sont *orthogonales* si tout élément de A est orthogonal à tout élément de B. Soit A une partie de H ; on appelle orthogonal de A l'ensemble $A^\perp$ des éléments de H orthogonaux à A.

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

Proposition 2.2.9. *Soient* H *un espace hilbertien et* F *un sous-espace vectoriel fermé de* H *; on a* $F \oplus F^\perp = H$.

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

//

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

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

Il est clair que $\|\ell_y\| \le \|y\|$ par Cauchy-Schwarz, mais en fait on voit que $\|\ell_y\| = \|y\|$ en considérant $x = y$.

Proposition 2.2.10. *Soit* H *un espace de Hilbert ; l'application isométrique antilinéaire* $y \to \ell_y$ *de l'équation* $(*)$ *est une bijection de* H *sur le dual* H^**. En d'autres termes, pour toute forme linéaire continue* ℓ *sur* H*, il existe un vecteur* $y_\ell \in \mathrm{H}$ *unique qui représente la forme linéaire* ℓ *au sens suivant :*

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

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

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

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

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

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

//

Exemple 2.2.7. Pour toute forme linéaire continue ℓ sur $\mathrm{L}_2(\Omega, \mu)$, il existe une fonction $g \in \mathrm{L}_2$ telle que

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

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

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

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

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

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

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

2.3. Dualité des espaces ℓ_p et 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 2.3.1 : 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 $f \in \mathrm{L}_p(\Omega, \mu)$ et $g \in \mathrm{L}_q(\Omega, \mu)$, la fonction produit fg est intégrable et

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

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

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

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

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

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

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

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

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

Démonstration. L'inégalité de Hölder nous dit déjà que

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

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

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

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

//

Corollaire 2.3.3. *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*

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

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

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

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 $p < +\infty$.

Notons $(e_n)_{n\geq 0}$ la *suite canonique* (voir l'exemple 1.1.6). Si $x = (x_n)$ est un élément de ℓ_p, on va vérifier que la série de vecteurs $\sum x_k e_k$ converge dans ℓ_p, et que sa somme $\sum_{k=0}^{+\infty}$ est le vecteur x. La somme partielle $\mathrm{U}_n = \sum_{k=0}^{n} x_k e_k$ est le vecteur $\mathrm{U}_n = (x_0, x_1, \ldots, x_n, 0, \ldots)$; on voit donc que $\|x - \mathrm{U}_n\|_p^p = \sum_{k>n} |x_k|^p$, reste d'ordre n de la série numérique convergente $\sum |u_k|^p$; il en résulte que $\|x - \mathrm{U}_n\|_p$ tend vers 0 quand $n \to +\infty$, ce qui signifie précisément que $x = \sum_{k=0}^{+\infty} x_k e_k$.

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

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

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

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

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, et $f = f_v$.

Le raisonnement est similaire pour c_0 ; on montre d'abord que $x = \sum_{k=0}^{+\infty} x_k e_k$ (au sens de c_0) pour tout $x \in c_0$. La suite de la démonstration est la même que pour l'espace ℓ_p.

En d'autres termes,

Théorème 2.3.4. *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 $f \in \mathrm{L}_p$, $g \in \mathrm{L}_q$

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

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 3 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 la proposition 2.10) 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.

2.4. 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

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

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. Dans une direction, il est assez facile de voir qu'une mesure μ sur $(\mathrm{K}, \mathcal{B})$ permet de définir une forme linéaire continue sur C(K). Dans le cas réel, on écrira $\mu = \mu_1 - \mu_2$, avec μ_1, μ_2 mesures positives finies données par la décomposition de Hahn, et on définira une forme linéaire ℓ sur C(K) en posant

$$\ell(f) = \int_{\mathrm{K}} f(s)\, d\mu(s) = \int_{\mathrm{K}} f(s)\, d\mu_1(s) - \int_{\mathrm{K}} f(s)\, d\mu_2(s)$$

(dans le cas complexe, on fait intervenir quatre mesures positives). On aura $|\ell(f)| \le \|f\|_\infty\, (\mu_1(\mathrm{K}) + \mu_2(\mathrm{K})) = \|f\|_\infty\, \|\mu\|$ pour toute $f \in \mathrm{C}(\mathrm{K})$, ce qui montre que $\|\ell\| \le \|\mu\|$.

> Pour montrer l'inégalité dans l'autre direction (et dans le cas réel) il faut savoir que pour toute mesure positive ν sur $(\mathrm{K}, \mathcal{B})$ et pour tout $\mathrm{A} \in \mathcal{B}$, il existe un fermé F tel que $\mathrm{F} \subset \mathrm{A}$ et $\nu(\mathrm{A} \setminus \mathrm{F}) < \varepsilon$. Si $\mathrm{B} \in \mathcal{B}$ porte la partie positive μ_1 de μ, on trouvera $\mathrm{F}_1 \subset \mathrm{B}$ et $\mathrm{F}_2 \subset \mathrm{K} \setminus \mathrm{B}$ tels que $\mu_1(\mathrm{F}_1) \ge \mu_1(\mathrm{B}) - \varepsilon = \mu_1(\mathrm{K}) - \varepsilon$ et de même $\mu_2(\mathrm{F}_2) \ge \mu_2(\mathrm{K}) - \varepsilon$. On peut trouver une fonction continue f telle que $\|f\|_\infty \le 1$, $f = 1$ sur F_1 et $f = -1$ sur F_2. On aura alors $\int f\, d\mu \ge \|\mu\| - 4\,\varepsilon$.

Dans l'autre direction, il faut montrer que toute forme linéaire continue sur C(K) provient d'une mesure μ sur $(\mathrm{K}, \mathcal{B})$. 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

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

Enonçons le résultat.

Théorème 2.4.1. *Soit* K *un espace métrique compact. Toute forme linéaire continue* ℓ *sur* C(K) *provient d'une mesure* μ *sur* $(\mathrm{K}, \mathcal{B})$*, et de plus*

$$\|\ell\| = \|\mu\|$$

où $\|\mu\|$ *est défini (dans le cas réel) à partir de la décomposition de Hahn de la mesure* μ*.*

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.

Définir 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.

2.5. Séries de Fourier

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

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

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

Pour voir que la suite exponentielle est totale, on peut employer le marteau-pilon Stone-Weierstrass : l'espace vectoriel complexe L engendré se trouve être une algèbre, invariante par conjugaison complexe et qui sépare les points du compact $\mathbb{R}/2\pi\mathbb{Z}$. Il en résulte que L est dense, pour la norme uniforme, dans l'espace des fonctions continues 2π-périodiques, ce qui implique la densité dans $\mathrm{L}_2(0, 2\pi)$.

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

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

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

Il résulte aussi de cette identité que pour toute $f \in \mathrm{L}_2$, les coefficients de Fourier $c_n(f)$ tendent vers 0 quand $n \to +\infty$. Ce résultat peut s'étendre à L_1 ; pour calculer les coefficients de Fourier, il suffit que f soit intégrable (mais nous ne savons pas que $f \in \mathrm{L}_1$ est la somme (au sens de L_1) de la série de Fourier) ; on peut montrer que

Pour toute $f \in \mathrm{L}_1$, les coefficients de Fourier $c_n(f)$ tendent vers 0 *quand* $|n| \to +\infty$

et ce résultat porte le nom de *lemme de Riemann-Lebesgue*.

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

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

et

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

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

On écrit souvent le développement de Fourier d'une fonction $f \in \mathrm{L}_2$ sous la forme

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

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

Théorème 2.5.1. *Soit f une fonction 2π-périodique et lipschitzienne. Pour tout $s \in \mathbb{R}$, on a*

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

Démonstration. Posons

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

Posons

$$\mathrm{K}_{\mathrm{N}}(t) = \sum_{n=-\mathrm{N}}^{\mathrm{N}} \mathrm{e}^{int} = \mathrm{K}_{\mathrm{N}}(-t) = \mathrm{e}^{-i\mathrm{N}t}\ \frac{1 - \mathrm{e}^{i(2\mathrm{N}+1)t}}{1 - \mathrm{e}^{it}}.$$

Comme $\int_0^{2\pi} \mathrm{K}_{\mathrm{N}}(t)\,(dt/2\pi) = 1$, on aura, pour $s = 0$

$$f_n(0) - f(0) = \int_0^{2\pi} \mathrm{K}_{\mathrm{N}}(t)\,(f(t) - f(0))\,\frac{dt}{2\pi} =$$

$$= \int_0^{2\pi} (\mathrm{e}^{-i\mathrm{N}t} - \mathrm{e}^{i(\mathrm{N}+1)t})\,\frac{f(t) - f(0)}{1 - \mathrm{e}^{it}}\,\frac{dt}{2\pi}.$$

Posons $g(t) = (f(t) - f(0))/(1 - \mathrm{e}^{it})$. Comme f est Lipschitz et $|1 - \mathrm{e}^{it}\,| \geq 2t/\pi$, la fonction g est bornée, donc $g \in \mathrm{L}_2$, ce qui entraîne que les coefficients de Fourier de g tendent vers 0. Or nous avions

$$f_{\mathrm{N}}(0) - f(0) = c_{\mathrm{N}}(g) - c_{-\mathrm{N}-1}(g)$$

qui tend donc vers 0 quand $\mathrm{N} \to +\infty$. Le raisonnement est identique pour montrer que $f_{\mathrm{N}}(s) \to f(s)$, pour tout s.

//

Si on utilise le lemme de Riemann-Lebesgue pour les fonctions de L_1, on pourra étendre la démonstration précédente aux fonctions f telles que $|f(t) - f(s)| \leq M\,|t-s|^\alpha$ pour un $\alpha > 0$. En effet, la fonction g de la démonstration précédente sera encore intégrable dans ce cas.

Le résultat précédent redémontre le fait que la suite exponentielle est totale dans L_2. Oublions le pour un instant. Le théorème précédent s'applique à toute fonction continue f périodique linéaire par morceaux : les sommes de Fourier $S_N(f) = f_N$ tendent simplement vers la fonction f ; mais par ailleurs, elles tendent pour la norme L_2 vers la somme F de la série $\sum_{n\in\mathbb{Z}} c_n(f)\, e_n$, donc d'après un théorème d'intégration une certaine sous-suite tend presque partout vers F. Puisque $S_N(f)$ tend simplement vers f, on a nécessairement $F = f$ presque partout, ce qui implique que f est la limite en norme L_2 des sommes de Fourier $S_N(f)$. Ces sommes appartiennent au sous-espace vectoriel L engendré par les exponentielles, et le raisonnement précédent montre que l'adhérence de L dans L_2 contient toutes les fonctions linéaires par morceaux, qui sont denses dans $C(\mathbb{R}/2\pi\mathbb{Z})$, donc aussi denses dans L_2. Finalement L est dense dans L_2 et on a montré la totalité de la suite $(e_n)_{n\in\mathbb{Z}}$ directement, sans utiliser Stone-Weierstrass.

3. Les théorèmes fondamentaux

3.1. Le théorème de Baire et ses conséquences

Commençons par énoncer le théorème de Baire.

Théorème 3.1.1 : théorème de Baire. *Soit* X *un espace métrique complet ; si* $(\mathrm{U}_n)_{n\ge 0}$ *est une suite de parties ouvertes et* denses *dans* X, *l'intersection* $\bigcap_{n\ge 0} \mathrm{U}_n$ *est dense dans l'espace* X.

Démonstration. Soit $(\mathrm{U}_n)_{n\ge 0}$ une suite d'ouverts denses de X ; soit V une partie ouverte non vide de X ; on doit montrer que $\bigcap_{n\ge 0} \mathrm{U}_n$ 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 $n \ge 0$ on construit une suite (x_n) d'éléments de X et une suite (r_n) de nombres réels strictement positifs tels que $r_n \le 2^{-n}$ et tels que, pour tout $n \ge 1$, la boule ouverte $\mathrm{B}(x_n, 2r_n)$ 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 $0 < r_{n+1} \le 2^{-n-1}$ 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 $\bigcap_{n\ge 0} \mathrm{B}_n \ne \emptyset$; or, par construction, $\bigcap_{n\ge 0} \mathrm{B}_n \subset \mathrm{V} \cap \bigcap_{n\ge 0} \mathrm{U}_n$, ce qui montre que cette dernière intersection est non vide.

//

Corollaire 3.1.2. *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 3.1.3 : théorème des isomorphismes. *Soient* E *et* F *deux espaces de Banach ; toute application linéaire continue bijective de* E *sur* F *est un isomorphisme.*

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_F,

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

On a $\bigcup_{n\ge 1} n\mathrm{B_E} = \mathrm{E}$ donc $\bigcup_{n\ge 1} nf(\mathrm{B_E}) = f(\mathrm{E}) = \mathrm{F}$, donc $\bigcup_{n\ge 1} \overline{nf(\mathrm{B_E})} = \mathrm{F}$. Par le corollaire 2, il existe un entier $n \ge 1$ tel que l'ensemble fermé $\overline{nf(\mathrm{B_E})}$ soit d'intérieur non vide. Comme la multiplication par $n \ge 1$ 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 qui suit 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 3.1.4. *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 $k \ge 0$ 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

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

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 $\mathrm{B_F}(0,r) \subset f(\mathrm{B_E}) + \varepsilon \mathrm{B_F}(0,r)$ pour tout $\varepsilon > 0$: posons $\mathrm{A} = \mathrm{B}(0_\mathrm{F}, r)$; si $x \in \mathrm{A}$ on peut trouver $y \in f(\mathrm{B_E})$ tel que $\|x - y\| < \varepsilon\, r$ puisque x est adhérent à $f(\mathrm{B})$, donc $x - y \in \varepsilon \mathrm{A}$ et $x = y + (x - y) \in f(\mathrm{B_E}) + \varepsilon \mathrm{A}$. D'après le lemme précédent, on a $\mathrm{B_F}(0,r) = \mathrm{A} \subset (1-\varepsilon)^{-1} f(\mathrm{B_E})$. 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 3.1.1. Le théorème précédent a peu de chose 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 1.1.7, 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 3.1.5 : théorème de l'application ouverte. *Soient* E *et* F *deux espaces de Banach ; toute application linéaire, continue, surjective* f *de* E *sur* F *est ouverte.*

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

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

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 3.1.6 : théorème du graphe fermé. *Soient* E *et* F *deux espaces de Banach ; toute application linéaire de* E *dans* F *dont le graphe est fermé dans* $\mathrm{E}\times\mathrm{F}$ *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.

//

Remarquons qu'inversement, le théorème du graphe fermé implique immédiatement le théorème des isomorphismes : si f est bijective, le graphe de f^{-1} est tout simplement le symétrique dans $\mathrm{F}\times\mathrm{E}$ du graphe de f. Si f est continue, le graphe de f est fermé dans $\mathrm{E}\times\mathrm{F}$, donc celui de f^{-1} est fermé aussi, donc f^{-1} est continue d'après le théorème du graphe fermé.

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

Proposition 3.1.7. *Soit* E *un espace vectoriel muni d'une distance* d*, telle que* (E,d) *soit complet, et telle que les opérations* $(x,y)\to x+y$ *et* $(\lambda,x)\to\lambda x$ *soient continues de* $\mathrm{E}\times\mathrm{E}$ *dans* E *et* $\mathbb{K}\times\mathrm{E}$ *dans* E *respectivement ; soient d'autre part* F *un espace normé et* A *une famille d'applications linéaires continues de* E *dans* F*. Si pour tout* $x\in\mathrm{E}$ *la famille* $\{\mathrm{T}(x):\mathrm{T}\in\mathrm{A}\}$ *est bornée dans* F*, il existe un voisinage* W *de* 0_E *tel que*

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

Démonstration. Remarquons d'abord que pour tout $x_0\in\mathrm{E}$ la translation $y\to x_0+y$ est un homéomorphisme de E ; de même, pour tout $\lambda\neq 0$ l'homothétie $y\to\lambda y$ est un homéomorphisme. Il résulte du premier point que tout voisinage W de x_0 est de la forme $x_0+\mathrm{U}$, où U est un voisinage de 0_E, et du second point que $-\mathrm{U}$ est aussi un voisinage de 0_E (prendre $\lambda=-1$), donc $\mathrm{V}=\mathrm{U}\cap(-\mathrm{U})$ est un voisinage symétrique du point 0_E.

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

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

//

Théorème 3.1.8 : 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. On peut appliquer la proposition précédente. Il existe un voisinage V de 0_{E} tel que $\|f(x)\| \le 1$ pour tout $x \in \mathrm{V}$ et tout $f \in \mathrm{A}$. Il existe $r > 0$ tel que $\mathrm{B}(0,r) \subset \mathrm{V}$. Par homogénéité, pour tout $x \in \mathrm{B}(0,1)$ et tout $f \in \mathrm{A}$, on a $\|f(x)\| \le 1/r$; on a donc montré que pour tout $f \in \mathrm{A}$, on a $\|f\| \le 1/r$.
//

Corollaire 3.1.9. *Soient* E *un espace de Banach (ou bien un* (E,d) *complet comme dans la proposition 7),* 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. D'abord, il est évident que la limite f est linéaire. Soit $x \in \mathrm{E}$; comme la suite $(f_n(x))$ est convergente, elle est bornée ; par le théorème 8, la suite $(\|f_n\|)$ est alors bornée. Il existe alors un nombre $\mathrm{M} \ge 0$ tel que, pour tout $x \in \mathrm{E}$ et tout entier $n \ge 0$ on ait $\|f_n(x)\| \le \mathrm{M}\,\|x\|$. Passant à la limite on trouve $\|f(x)\| \le \mathrm{M}\,\|x\|$.
//

Corollaire 3.1.10. *Soient* E *un espace de Banach (ou bien un* (E,d) *complet comme dans la proposition 7),* F *un espace normé et* (u_k) *une suite d'applications linéaires continues de* E *dans* F *; on suppose que, pour tout* $x \in \mathrm{E}$*, la série* $\sum_k u_k(x)$ *converge dans* F *; notons* $f(x)$ *sa somme. Alors* f *est linéaire et continue.*

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

Attention ! Dans le corollaire 9, où on prouve la continuité de la limite simple f d'une suite (f_n), on n'a en général **pas** que $\|f_n - f\| \to 0$.

3.2. 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 $\lambda \ge 0$
- (ii) pour tous $x, y \in \mathrm{X}$, on a $q(x+y) \le q(x) + q(y)$.

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

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

Lemme 3.2.2. *Soient* Z *un sous-espace vectoriel de* X *et g une forme linéaire définie sur* Z*, telle que $g(z) \leq q(z)$ pour tout $z \in \mathrm{Z}$; soit $x \in \mathrm{X}$ tel que $x \notin \mathrm{Z}$; il existe une forme linéaire $\widetilde{g}$ sur $\mathrm{Z} + \mathbb{R}x$ telle que $\widetilde{g}$ prolonge g et $\widetilde{g} \leq q$ sur $\mathrm{Z} + \mathbb{R}x$.*

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

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

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

$$g(z_1) + \mathrm{M} \leq q(z_1 + x), \quad g(z_2) - \mathrm{M} \leq q(z_2 - x)$$

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

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

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

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

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

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

et il est alors clair que cette propriété est vraie :

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

Le lemme est donc établi.

Venons-en à l'application du lemme de Zorn pour terminer la démonstration du théorème. Notons Φ l'ensemble des sous-espaces vectoriels $\Gamma \subset \mathrm{X} \times \mathbb{R}$ qui sont le graphe d'une forme linéaire g définie sur un sous-espace vectoriel Z (variable) de X contenant Y, où la restriction de g à Y est ℓ, et où de plus $g \le q$. On a donc

$$\Gamma = \{(z,t) : z \in \mathrm{Z}, t = g(z)\},$$

Γ contient $\Gamma_0 = \{(y, \ell(y)) : y \in \mathrm{Y}\}$ (ce qui signifie que g prolonge ℓ), et $(z,t) \in \Gamma \Rightarrow t \le q(z)$. Munissons Φ de la relation d'ordre $\prec$ d'inclusion des sous-espaces de $\mathrm{X} \times \mathbb{R}$. L'ordre $\prec$ est inductif : si $\{\Gamma_i : i \in \mathrm{I}\}$ une famille totalement ordonnée de Φ, il est clair que $\Gamma = \bigcup_i \Gamma_i$ est un sous-espace vectoriel (le truc, c'est que si on a $i, j \in \mathrm{I}$, on a ou bien $\Gamma_i \subset \Gamma_j$, ou bien $\Gamma_j \subset \Gamma_i$, donc $\Gamma_i \cup \Gamma_j$ est un sous-espace vectoriel !) On vérifie sans peine que Γ est le graphe d'une forme linéaire g définie sur le sous-espace vectoriel

$$\mathrm{Z}_\Gamma = \{z \in \mathrm{X} : \exists t \in \mathbb{R},\ (z,t) \in \Gamma\},$$

que g prolonge ℓ et $g \le q$. On a donc $\Gamma \in \Phi$. Il est clair que Γ majore tous les Γ_i, $i \in \mathrm{I}$, et on a montré que l'ordre sur Φ est inductif, ce qui permet d'appliquer le lemme de Zorn : soit donc $\Gamma \in \Phi$ un élément maximal, et $\mathrm{Z} = \mathrm{Z}_\Gamma$ le sous-espace de X associé ; si nous montrons que $\mathrm{Z} = \mathrm{X}$ le théorème sera démontré. Mais si $\mathrm{Z} \neq \mathrm{X}$, on peut trouver un vecteur $x \notin \mathrm{Z}$, donc $\mathrm{Z} \neq \mathrm{Z} + \mathbb{R}x$, et on a vu dans l'étape préliminaire qu'on peut trouver une extension convenable $(\mathrm{Z} + \mathbb{R}x, \widetilde{g})$, qui contredirait la maximalité de Γ. On a donc bien $\mathrm{Z} = \mathrm{X}$ 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 $y \in \mathrm{Y}$, 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 3.2.3 : théorème de séparation de Hahn-Banach. *Soient* X *un espace normé réel,* A *un convexe ouvert non vide et* B *un convexe non vide tels que* A *et* B *soient disjoints. Il existe alors une forme linéaire continue f sur* X *telle que*

$$f(a) < \inf f(\mathrm{B})$$

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

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

Démonstration. On commence par le cas où B est réduit à un seul point $b_0 \neq 0$. Par translation on peut se ramener au cas où le convexe ouvert A contient 0_X. La jauge j_A est alors une fonction sous-linéaire sur X, et $j_\mathrm{A}(b_0) \ge 1$ puisque $b_0 \notin \mathrm{A}$. On définit ℓ sur la droite $\mathrm{Y} = \mathbb{R}b_0$ en posant $\ell(\lambda b_0) = \lambda$ pour tout $\lambda \in \mathbb{R}$. On vérifie que $\ell \le p = j_\mathrm{A}$ sur Y. Si m est un prolongement de ℓ majoré par p, on aura $m(a) \le j_\mathrm{A}(a) < 1$ pour tout $a \in \mathrm{A}$, 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 $-\mathrm{A}$ et bornée par 1 sur le voisinage $\mathrm{A} \cap (-\mathrm{A})$.

Passons au cas général. Si A est un convexe ouvert non vide disjoint du convexe non vide B, on introduit le convexe ouvert

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

Puisque A et B sont disjoints, ce convexe est disjoint de $\mathrm{D} = \{0_{\mathrm{X}}\}$. D'après l'étape préliminaire, on peut trouver une forme linéaire continue m sur X 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 $a \in \mathrm{A}$, $b \in \mathrm{B}$, donc $m(a) \le \inf m(\mathrm{B})$. Puisque m n'est pas nulle, on peut choisir $u \in \mathrm{X}$ tel que $m(u) > 0$, puis $\varepsilon > 0$ tel que $a + \varepsilon u \in \mathrm{A}$ puisque A est ouvert ; on termine en écrivant $m(a) < m(a + \varepsilon u) \le \inf m(\mathrm{B})$, et en prenant $f = m$.

//

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

Théorème 3.2.4 : 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 de l'énoncé du théorème 1 sera un multiple convenable de la norme N de X. Par définition de la norme de la forme linéaire ℓ, on a $\ell \le \|\ell\|\,\mathrm{N} = q$ 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 1.6.1, on sait que m_1 est la partie réelle d'une forme linéaire complexe m sur X, et de plus $\|m\| = \|m_1\| \le \|\ell_1\| = \|\ell\|$; d'autre part m prolonge ℓ (ici Y est un sous-espace vectoriel complexe ; si $y \in \mathrm{Y}$ on a aussi $iy \in \mathrm{Y}$ ce qui permet d'écrire $m(y) = m_1(y) - im_1(iy)$, et alors $m(y) = \ell_1(y) - i\ell_1(iy) = \ell(y)$).

//

Corollaire 3.2.5. *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 4, 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 3.2.6. *Soient* X *et* Y *deux espaces normés ; pour tout* $f \in \mathcal{L}(X, Y)$*, on a* $\|{}^t f\| = \|f\|$.

Démonstration. On sait déjà que $\|{}^t f\| \le \|f\|$ par la proposition 1.6.2, nous allons montrer l'égalité. Pour tout $\varepsilon > 0$, on peut trouver un vecteur $x \in X$ tel que $\|x\| \le 1$ et tel que $\|f(x)\| > \|f\| - \varepsilon$, puis une forme linéaire $y^* \in F^*$ 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 3.2.7. *Si* C *est un sous-ensemble convexe fermé non vide d'un espace normé réel* X*, alors* C *est l'intersection de demi-espaces affines fermés.*

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

Pour tout $x \notin C$, on peut trouver une boule ouverte $A = B(x, r)$ disjointe de C ; d'après le théorème de séparation il existe une forme linéaire continue x^* telle que

$$\forall a \in A, \quad x^*(a) < \inf x^*(C)$$

et en particulier $x^*(x) < \inf x^*(C)$. On voit donc que si on pose $d = \inf x^*(C)$ et

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

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

//

Corollaire 3.2.8. *Soient* X *un espace normé et* Y *un sous-espace fermé ; soient* $x \notin Y$ *et* $r = \operatorname{dist}(x, Y) > 0$ *; il existe une forme linéaire continue* $x^* \in X^*$ *telle que :* x^* *est nulle sur* Y*,* $\|x^*\| = 1$ *et* $x^*(x) = r$*.*

Démonstration. Considérons le quotient X/Y et la projection π de X sur X/Y ; par définition de la norme du quotient, on a $\|\pi(x)\| = r$. En appliquant le théorème de Hahn-Banach à X/Y, on trouve une forme linéaire continue z^* sur X/Y telle que $\|z^*\| \le 1$ et $z^*(\pi(x)) = \|\pi(x)\| = r$; alors $x^* = z^* \circ \pi$ donne la solution.

//

Il en résulte que si X est un espace normé et Y un sous-espace fermé, il existe une partie A de X^* telle que $Y = \bigcap_{x^* \in A} \ker(x^*)$.

Remarque 3.2.1. 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 $z^* \in (X/Y)^*$ on associe $x^* = z^* \circ \pi \in X^*$, qui est nulle sur Y. Inversement, si $x^* \in X^*$ est nulle sur Y, on peut l'écrire $x^* = z^* \circ \pi$ d'après la proposition 1.3.3, et $\|z^*\| = \|x^*\|$.

Proposition 3.2.9. *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 3.1.

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 $\operatorname{rang} \mathrm{T} \le \operatorname{rang} {}^t\mathrm{T}$.

//

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

Proposition 3.2.10. *Soit* X *un espace normé ; si le dual* X^* *est séparable, alors* X *est séparable.*

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

//

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

Soit X un espace normé ; le dual du dual X^* de X s'appelle le *bidual* de X et se note X^{**}. Pour $x \in X$ notons $I_X(x) : X^* \to \mathbb{K}$ la forme linéaire sur X^* qui à $x^* \in X^*$ associe $x^*(x)$. Pour tout $x^* \in X^*$, on a $|I_X(x)(x^*)| = |x^*(x)| \le \|x^*\| \, \|x\|$, donc $I_X(x) \in X^{**}$ et $\|I_X(x)\| \le \|x\|$. On dit que $I_X \in \mathcal{L}(X, X^{**})$ est l'application canonique de X dans son bidual.

Corollaire 3.3.1. *L'application canonique* $I_X : X \to X^{**}$ *est isométrique.*

Démonstration. Soit $x \in X$; par le corollaire 2.5, il existe $x^* \in X^*$ tel que $\|x^*\| \le 1$ et $x^*(x) = \|x\|$. Alors
$$\|x\| = |x^*(x)| = |I_X(x)(x^*)| \le \|x^*\| \, \|I_X(x)\| \le \|I_X(x)\|,$$
vu que $\|x^*\| \le 1$; donc $\|I_X(x)\| = \|x\|$.

//

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

Si X est un espace normé l'application I_X est injective par le corollaire 1. Remarquons que si I_X est bijective alors I_X est une isométrie de X sur un espace de Banach (par la proposition 1.2.5). Il s'ensuit que si 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 3.3.1. Un espace de Banach E est dit *réflexif* si l'application canonique $I_E : E \to E^{**}$ est bijective.

Autrement dit, un espace de Banach E est réflexif lorsque toute forme linéaire x^{**} continue sur le dual E^* provient d'un vecteur x de E de la façon expliquée précédemment,
$$\forall x^* \in E^*, \quad x^{**}(x^*) = x^*(x).$$

Proposition 3.3.2. *Tout espace de Hilbert est réflexif.*

Démonstration. Soient H un espace de Hilbert et $x^{**} \in H^{**}$; pour tout vecteur $y \in H$ soit $\ell_y \in H^*$ la forme linéaire sur H définie par $\ell_y(x) = \langle x, y \rangle$; l'application $y \to \overline{x^{**}(\ell_y)}$ est une forme linéaire et continue sur H. Par la proposition 2.2.10, il existe $x \in H$ tel que, pour tout $y \in H$ on ait $\overline{x^{**}(\ell_y)} = \langle y, x \rangle$. D'après la proposition 2.2.10, toute $f \in H^*$ est de la forme $f = \ell_y$ pour un certain $y \in H$, donc on a $x^{**}(f) = x^{**}(\ell_y) = \langle x, y \rangle = f(x)$, c'est à dire que x^{**} est l'image de x par l'application canonique de H dans H^{**}, qui est donc surjective.

//

Exemple 3.3.2. 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_∞).

Espaces normés isomorphes

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

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

Lemme 3.3.3. *Si* X *est réflexif et si* Y *est isomorphe à* X, *alors* Y *est réflexif.*

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

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

ce qu'il fallait démontrer.

//

Proposition 3.3.4. *Si* X *est réflexif, alors* X^* *est réflexif.*

Démonstration. Posons $\mathrm{Z} = \mathrm{X}^*$. Si $z^{**} = x^{***}$ est une forme linéaire sur le dual $\mathrm{Z}^* = \mathrm{X}^{**}$ de $\mathrm{Z} = \mathrm{X}^*$, elle définit une forme linéaire continue $z = x^* = x^{***} \circ \mathrm{I_X}$ 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 3.3.5. *Si* X *est réflexif, tout sous-espace fermé* Y *de* X *est réflexif.*

Démonstration. Soit π l'application de restriction définie de X^* sur Y^* (surjective par le théorème de Hahn-Banach). Soit y^{**} une forme linéaire continue sur Y^*. Alors $x^{**} = y^{**} \circ \pi$ est une forme linéaire continue sur X^*, donc il existe $x \in \mathrm{X}$ tel que $x^{**}(x^*) = x^*(x)$ pour tout x^* in X^*. Il suffit de voir que $x \in \mathrm{Y}$ pour pouvoir conclure assez facilement ; si on avait $x \notin \mathrm{Y}$, on pourrait trouver d'après le corollaire 2.8 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 3.3.6. *Si* X^* *est réflexif, alors* X *est réflexif.*

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

Pourquoi s'intéresser aux espaces réflexifs ?

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

3.4. Théorème de Riesz

Lemme 3.4.1. *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* $x, y \in \mathrm{A}$ *et* $x \neq y$, *alors* $\|x - y\| \geq \varepsilon$, *et* $\operatorname{card} \mathrm{A} \geq \varepsilon^{-n}$.

Démonstration. Soit A une famille maximale de points de la boule unité $\mathrm{B_Z}$ de Z dont les distances mutuelles soient $\geq \varepsilon$. Alors les boules de rayon ε centrées aux points de A recouvrent $\mathrm{B_Z}$: en effet, si $x \in \mathrm{B_Z}$ et $x \notin \mathrm{A}$, 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 $y \in \mathrm{A}$ 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 $\mathrm{B_Z}$; puisque Z est de dimension n, les boules de rayon ε ont un volume égal à $\varepsilon^n \mathrm{V}$ (dans le cas des scalaires réels), et puisque les boules de ce rayon centrées aux points de A recouvrent $\mathrm{B_Z}$, on a $(\operatorname{card} \mathrm{A})\, \varepsilon^n \mathrm{V} \geq \mathrm{V}$, d'où le résultat.

//

Théorème 3.4.2. *Si la boule unité d'un espace normé* X *est compacte, alors* X *est de dimension finie.*

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

//

On dit qu'un opérateur 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 3.4.3. *Soit* E *un espace de Banach, réel ou complexe ; si* $\mathrm{T} \in \mathcal{L}(\mathrm{E})$ *est compact et si* $\lambda \neq 0$, *le sous-espace* $\mathrm{F}_\lambda = \ker(\mathrm{T} - \lambda \,\mathrm{Id_E}) = \{y \in \mathrm{E} : \mathrm{T}(y) = \lambda y\}$ *est de dimension finie.*

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

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

//

4. Topologies faibles

4.1. 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 4.1.1. *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 $\mathrm{J} = \{i_1, \dots, i_n\}$ 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 4.1.2. *Donnons-nous un ensemble* I*, un espace vectoriel* X*, une famille* $(\mathrm{Y}_i)_{i\in\mathrm{I}}$ *d'espaces vectoriels topologiques (cf. définition 1.1.3) 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 4.1.1. 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 4.1.2. 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.

4.2. 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. D'après le théorème de Hahn-Banach, la topologie faible est séparée.

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

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

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 $y \in \mathrm{X}$ 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 $(\mathrm{B}(\mathrm{X}), \sigma(\mathrm{X}, \mathrm{X}^*))$ est strictement plus faible en général que celle de B(X) munie de la topologie de la norme : soit $\mathrm{X} = \ell^2$, 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 8.1.3.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 $(\mathrm{B}(\mathrm{X}), \sigma(\mathrm{X}, \mathrm{X}^*))$ 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 $(\mathrm{B}(\mathrm{X}), \sigma(\mathrm{X}, \mathrm{X}^*))$ 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 $\mathrm{X} = \ell_2$ (exercice).

Proposition 4.2.1. *Soient* X *et* Y *deux espaces normés et* $\mathrm{T} \in \mathcal{L}(\mathrm{X}, \mathrm{Y})$ *; alors* T *est continue de* X *muni de la topologie* $\sigma(\mathrm{X}, \mathrm{X}^*)$ *dans* Y *muni de la topologie* $\sigma(\mathrm{Y}, \mathrm{Y}^*)$.

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

//

Théorème 4.2.2. *Soient* X *un espace normé et* C *un sous-ensemble convexe de* X *; l'ensemble* C *est fermé en norme si et seulement s'il est faiblement fermé.*

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

Démonstration. Comme la topologie de la norme est plus fine que la topologie faible, toute partie fermée pour la topologie faible est fermée pour la topologie de la norme. Démontrons l'inverse. Si X est réel, les demi-espaces affines fermés de la forme $\{x \in \mathrm{X} : x^*(x) \geq c\}$, où $x^* \in \mathrm{X}^*$ 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 3.2.7 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

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

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 3.4.2).

Théorème 4.2.3. *Muni de la topologie $\sigma(X^*, X)$ la boule unité de X^* est compacte.*

Ce théorème est un corollaire du théorème de Tykhonov que nous admettrons pour l'instant (voir dernière section).

Théorème 4.2.4 : 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 3. Remarquons que X^* est un sous-ensemble de l'ensemble $\mathbb{K}^X$ de toutes les applications de X dans $\mathbb{K}$, et que la topologie $\sigma(X^*, X)$ est, par définition, la topologie induite sur X^* par la topologie produit sur $\mathbb{K}^X$. Remarquons ensuite que la boule unité B de X^* est l'intersection de deux ensembles fermés dans $\mathbb{K}^X$,

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

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

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 4 ; étant fermé dans un compact, B est compact.

//

Corollaire 4.2.5. *Si E est réflexif, les convexes fermés bornés de E sont faiblement compacts.*

4.3. 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 $(x_n^*) \subset \mathrm{X}^*$ est $*$-faiblement convergente vers un vecteur x^* si $\lim_n x_n^*(x) = x^*(x)$ pour tout $x \in \mathrm{X}$; une suite $(x_n) \subset \mathrm{X}$ est faiblement convergente vers $x \in \mathrm{X}$ si $x^*(x) = \lim_n x^*(x_n)$ pour tout $x^* \in \mathrm{X}^*$.

Toute **suite** faiblement convergente dans X est bornée ; si X est complet, toute **suite** $*$-faiblement convergente dans X^* est bornée. Ces deux résultats viennent du corollaire 3.1.9.

Lemme 4.3.1. *Soit (x_n^*) une suite* **bornée** *dans le dual d'un espace normé* X *; pour que cette suite soit $*$-faiblement convergente vers x^*, il suffit que $x^*(d) = \lim_n x_n^*(d)$ pour tout d d'un ensemble* D *total dans* X. *Pour que cette suite soit $*$-faiblement convergente vers une limite dans X^*, il suffit que $\lim_n x_n^*(d)$ existe pour tout d d'un ensemble* D *total dans* X.

Démonstration. Montrons la deuxième variante de l'énoncé. Supposons $\|x_n^*\| \leq 1$ pour tout $n \geq 0$. Si la limite existe pour tout d d'un ensemble total T, elle existe aussi pour tout d de l'ensemble dense $\mathrm{D} = \mathrm{Vect}(\mathrm{T})$. Montrons que $(x_n^*(x))_{n \geq 0}$ converge pour tout $x \in \mathrm{X}$. Il suffit de montrer que cette suite de scalaires est de Cauchy. Choisissons $\varepsilon > 0$ et un entier k tels que $\|x - d\| < \varepsilon/3$. Pour tout entier n, on a $|x_n^*(x) - x_n^*(d)| < \varepsilon/3$, et la suite $(x_n^*(d))_{n \geq 0}$ 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

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

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

//

Proposition 4.3.2. *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 notation. Si $\mathrm{M} = \{n_0 < \ldots < n_j < \ldots\}$ est un sous-ensemble infini de $\mathbb{N}$, convenons de noter la sous-suite (x_{n_j}) par $(x_n)_{n \in \mathrm{M}}$. Soit donc (y_k) une suite dense dans E, et (x_n^*) une suite bornée dans E^*, telle que par exemple $\|x_n^*\| \leq 1$ pour tout entier $n \geq 0$. La suite de scalaires $(x_n^*(y_0))$ est bornée, donc elle admet une sous-suite convergente $(x_n^*(y_0))_{n \in \mathrm{M}_0}$. La suite $(x_n^*(y_1))_{n \in \mathrm{M}_0}$ est encore bornée, donc on peut trouver un nouvel ensemble infini $\mathrm{M}_1 \subset \mathrm{M}_0$ tel que la sous-suite $(x_n^*(y_1))_{n \in \mathrm{M}_1}$ 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 $(x_n^*(y_j))_{n \in \mathrm{M}_j}$ soit convergente pour tout $j \geq 0$.

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

//

Exemple 4.3.1. 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 $\mathrm{C}([0,1])$.

Attention ! Une suite (x_n^*) dans un dual n'a pas toujours de *sous-suite* $*$-faiblement convergente. Ainsi la suite des fonctions $t \to \mathrm{e}^{int}$, considérée dans le dual de l'espace $\mathrm{X} = \mathrm{M}(0, 2\pi)$ des mesures bornées sur $[0, 2\pi]$, n'a pas de sous-suite $*$-faiblement convergente. En effet, une telle sous-suite serait en particulier simplement convergente vers une limite (tester la convergence-$*$ sur les mesures de Dirac) ; cela n'est pas possible (utiliser le théorème de convergence dominée, et le fait que la suite est orthonormée).

Théorème 4.3.3. *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

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

Puisque Y est réflexif, il existe un vecteur $y \in \mathrm{Y}$ tel que $y^{**} = \mathrm{I_Y}(y)$, 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.

//

Théorème 4.3.4. *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.

//

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

Corollaire 4.3.5. *Si f est convexe continue sur un convexe fermé non vide d'un espace réflexif* E *et si $f(x)$ tend vers $+\infty$ lorsque $\|x\| \to +\infty$, la fonction f atteint son minimum sur* E.

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

//

Remarque 4.3.2. C'est ce type de résultat qui permet de minimiser les fonctionnelles telles que celle qui a été évoquée dans l'introduction.

4.4. 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 finie 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 $\pi_\mathrm{J}^{-1}(\mathrm{V}_1)$ et $\pi_\mathrm{J}^{-1}(\mathrm{V}_2)$ 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

$F \in \mathcal{F}$ tel que $x \notin F$, donc il existerait un voisinage W élémentaire de x dans X_I, disjoint de F. Mais un tel voisinage élémentaire est de la forme $\pi_J^{-1}(V)$, pour un certain ensemble fini $J \subset I$ et un ouvert V de X_J contenant $x_J = \pi_J(x) = (x_j^{(J)})_{j \in J}$. Mais alors $\pi_J(F)$ serait disjoint de V, donc on aurait $x_J \notin \overline{\pi_J(F)}$, ce qui contredit la construction initiale.

Remarque 4.4.1. On a vu que la famille maximale $\mathcal{U}$ est stable par intersection finie ; on voit aussi par maximalité que si $U \in \mathcal{U}$ et $U \subset Y$, alors $Y \in \mathcal{U}$. Il en résulte que pour tout $Y \subset X$, ou bien $Y \in \mathcal{U}$, ou bien $Y \notin \mathcal{U}$. Une telle famille $\mathcal{U}$ de parties de X est appelée *ultrafiltre* sur X.

Index

Index des notations