MT404, Cours n$^{\text{o}}$ 1, Lundi 24 Septembre 2001.

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

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,

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

L'une des méthodes de l'Analyse Fonctionnelle consiste à introduire un espace de fonctions X sympathique, par exemple un espace de Hilbert, qui permette d'obtenir assez facilement des théorèmes abstraits d'existence. Le problème est que l'espace des fonctions continues sur $\overline{\Omega}$ et C^2 à l'intérieur n'est justement pas un espace bien adapté à des “méthodes douces” d'Analyse Fonctionnelle ; les bons espaces pour ce point de vue ne sont pas formés de fonctions vraiment dérivables : ce sont les *espaces de Sobolev* qui seront un peu évoqués plus tard dans le cours.

Chapitre 1. Espaces normés et applications linéaires continues

1.1. Normes, semi-normes ; espaces de Banach

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

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

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

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

La propriété $p(x+y) \le p(x) + p(y)$ s'appelle *l'inégalité triangulaire* pour la semi-norme p. De l'inégalité triangulaire ci-dessus, on 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* $|p(x) - p(y)| \leq p(x - y)$ *pour tous vecteurs* $x, y \in \mathrm{X}$.

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

Une fonction réelle f définie sur un sous-ensemble convexe C de X est dite *fonction convexe sur* C si

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

pour tous $x, y \in \mathrm{C}$ et tout $t \in [0, 1]$.

Dans le cas d'une fonction f définie sur un intervalle ouvert I de $\mathbb{R}$, on sait que si $f'' \geq 0$ en tout point de I, la fonction f est convexe sur I. Par exemple, on montre ainsi que $t \to t^r$ est convexe sur $\mathrm{I} = [0, +\infty[$ lorsque $1 < r < +\infty$.

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

Démonstration. Montrons que la condition est suffisante : supposons que $\mathrm{C}_p = \{x \in \mathrm{X} : p(x) \leq 1\}$ soit convexe, et déduisons la sous-additivité de p ; soient x et y deux vecteurs de X, et introduisons $a > p(x)$ et $b > p(y)$; considérons les deux vecteurs de C_p 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_p d'après l'hypothèse de convexité, c'est à dire que $p(z) \leq 1$. Mais on vérifie immédiatement que $z = (a + b)^{-1}(x + y)$, et l'homogénéité de p transforme alors l'inégalité $p(z) \leq 1$ en $p(x + y) \leq a + b$. En faisant tendre a vers $p(x)$ et b vers $p(y)$ on obtient $p(x + y) \leq p(x) + p(y)$.

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, c'est à dire une famille de parties possédant les propriétés suivantes : $\emptyset \in \mathcal{A}$, $\mathrm{A} \in \mathcal{A}$ implique que $\mathrm{A}^c \in \mathcal{A}$ et pour toute suite (A_n) d'éléments de $\mathcal{A}$ la réunion $\bigcup_n \mathrm{A}_n$ est un élément de $\mathcal{A}$.

Nous supposerons donné un espace $(\Omega, \mathcal{A}, \mu)$, où $(\Omega, \mathcal{A})$ est un espace mesurable et μ une mesure positive sur $(\Omega, \mathcal{A})$. Rappelons que μ est une application de $\mathcal{A}$ dans $[0, +\infty]$ telle que $\mu(\emptyset) = 0$ et

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

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

Un exemple typique est fourni par la mesure de Lebesgue sur $\mathbb{R}^d$ (muni de la tribu borélienne), ou bien par la *mesure de comptage* μ sur $\mathbb{N}$, qui associe à tout $\mathrm{A} \subset \mathbb{N}$ le nombre $\mu(\mathrm{A})$ (fini ou $+\infty$) de ses éléments.

Une fonction f de $(\Omega, \mathcal{A})$ dans $\mathbb{R}$ est dite mesurable si $\{f > c\} \in \mathcal{A}$ pour tout c réel. Il en résulte que $f^{-1}(\mathrm{B}) \in \mathcal{A}$ pour tout borélien B de $\mathbb{R}$.

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

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

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

pour le vérifier, on voit d'abord que $p(\lambda f) = |\lambda| \, p(f)$ (facile), puis on montre que l'ensemble $\{f \in \mathcal{L}_r : p(f) \leq 1\}$ 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 $0 \leq t \leq 1$,

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

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

$$\int_0^1 \bigl| (1-t)f(s) + tg(s) \bigr|^r \, d\mu(s) \leq (1-t) \int_0^1 |f(s)|^r \, d\mu(s) + t \int_0^1 |g(s)|^r \; d\mu(s) \leq$$

$$\leq (1-t) + t = 1.$$

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.*

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

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

En général, nous noterons $\|x\|$ la norme d'un vecteur x d'un espace normé X.

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

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

Exemples 1.1.6.

– 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.

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

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

On peut considérer que ℓ_p est l'espace $\mathcal{L}_p(\mathbb{N}, \mathcal{P}(\mathbb{N}), \mu_{\mathrm{comptage}})$, ce qui permet d'unifier les arguments.

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

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

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

Séries de vecteurs

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

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

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

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

Il faut bien comprendre que la notion de somme de la série n'a aucun sens si on ne mentionne pas la topologie qui a été utilisée pour définir la notion de limite.
Notons que lorsque la série $\sum u_k$ converge dans X, on a l'inégalité

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

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

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

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

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

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

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

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

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

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

Démonstration. Voir poly.

Exemple. Si H est un espace de Hilbert, une série de vecteurs 2 à 2 orthogonaux $\sum u_k$ converge dans H si et seulement si $\sum \|u_k\|^2 < +\infty$.

Exemple 1.1.7. Pour tout $n \geq 0$, désignons par $\mathbf{e}_n$ le vecteur de ℓ_p, ou de c_0, dont les coordonnées sont $\mathbf{e}_{n,i} = 0$ si $i \neq n$ et $\mathbf{e}_{n,n} = 1$. On appellera $(\mathbf{e}_n)_{n\geq 0}$ la *suite canonique*. Soit $x = (x_n)$ un élément de c_0 ; le vecteur x est la somme (dans l'espace normé c_0) de la série $\sum_{k=0}^{+\infty} x_k \mathbf{e}_k$; en effet, si $u_k = x_k\, \mathbf{e}_k$, on voit que

$$\mathrm{U}_n = (x_0, x_1, \ldots, x_n, 0, 0, \ldots)$$

et il en résulte que $\|\mathrm{U}_n - x\| = \sup_{k>n} |x_k| = \varepsilon_n \to 0$.

1.2. Applications linéaires continues

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

- (i) *l'application* T *est continue sur* X *;*
- (ii) *l'application* T *est continue au point* 0_X *;*
- (iii) *il existe un nombre* $\mathrm{M} \geq 0$ *tel que, pour tout* $x \in \mathrm{X}$ *on ait*

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

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

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

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

Soient X et Y deux espaces normés et S, T deux applications linéaires continues de X dans Y ; on sait que l'application S + T, qui associe à tout $x \in \mathrm{X}$ l'image $\mathrm{S}(x) + \mathrm{T}(x) \in \mathrm{Y}$, est linéaire. 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'espace des applications linéaires de X dans Y. Soit $\mathrm{T} : \mathrm{X} \to \mathrm{Y}$ une application linéaire continue ; d'après le théorème 1, il existe une constante M telle que $\|\mathrm{T}(x)\|_\mathrm{Y} \le \mathrm{M}$ pour tout vecteur x de X tel que $\|x\|_\mathrm{X} \le 1$. On peut donc considérer la quantité (finie)

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

qui s'appelle la *norme de l'application linéaire* T ; on a noté B_X la *boule unité fermée* de l'espace normé X, $\mathrm{B}_\mathrm{X} = \{x \in \mathrm{X} : \|x\| \le 1\}$.

Si x est un vecteur non nul de X, le vecteur $z = \|x\|_\mathrm{X}^{-1} x$ vérifie $\|z\|_\mathrm{X} \le 1$, donc $\|\mathrm{T}(z)\|_\mathrm{Y} \le \|\mathrm{T}\|$, d'où par homogénéité $\|x\|^{-1}\|\mathrm{T}(x)\| \le \|\mathrm{T}\|$, ou encore

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

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* $\mathrm{T} : \mathrm{X} \to \mathrm{Y}$ *une application linéaire continue ; on pose*

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

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

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

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

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

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

d'où l'inégalité $\|\mathrm{S}+\mathrm{T}\| \le \|\mathrm{S}\| + \|\mathrm{T}\|$.

Cours n° 2, Mercredi 26 Septembre 2001.

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 $\|fg\|_\infty \le \|f\|_\infty \, \|g\|_\infty$ pour toute fonction g, donc $\|\mathrm{M}_f\| \le \|f\|_\infty$. Inversement considérons la fonction constante $g_0 = 1$. Il est clair que $\|g_0\| = 1$, et $\mathrm{M}_f(g_0) = f$. On a donc $\|f\| = \|\mathrm{M}_f(g_0)\| \le \|\mathrm{M}_f\| \, \|g_0\| = \|\mathrm{M}_f\|$, d'où $\|\mathrm{M}_f\| = \|f\|_\infty$.

– On considère l'espace $X = L_1([0,1])$ (mesure de Lebesgue) ; pour toute fonction $f \in X$ on peut définir une nouvelle fonction $V(f)$ par la formule

$$\forall x \in [0,1], \quad (Vf)(x) = \int_0^x f(t)\,dt.$$

On sait par le cours d'intégration que Vf est une fonction continue sur $[0,1]$. On a

$$\forall x \in [0,1], \quad \big|(Vf)(x)\big| \le \int_0^x |f(t)|\,dt \le \int_0^1 |f(t)|\,dt = \|f\|_{L_1}$$

ce qui montre que $\|Vf\|_\infty \le \|f\|_{L_1}$ pour toute f ; on en déduit que V définit une application (clairement linéaire) continue de $L_1([0,1])$ dans $C([0,1])$, et $\|V\|_{\mathcal{L}(L_1,C)} \le 1$. Si on prend à nouveau $f_0 = 1$, on voit que $(Vf_0)(x) = x$ pour tout x de $[0,1]$, donc $\|Vf_0\|_\infty = 1$ et $1 = \|Vf_0\|_\infty \le \|V\|\,\|f_0\|_{L_1} = \|V\|$ montre finalement que $\|V\| = 1$.

La proposition suivante est facile mais importante.

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

$$\|T \circ S\| \le \|S\|\,\|T\|.$$

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

$$\|(T \circ S)(x)\|_Z = \|T(S(x))\|_Z \le \|T\|\,\|S(x)\|_Y \le \|T\|\,\|S\|\,\|x\|_X,$$

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* $\mathcal{L}(X,Y)$ *est un espace de Banach.*

Démonstration. Voir poly.

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

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

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

1.4. Principe de prolongement

Montrons un lemme de prolongement, facile mais très important.

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

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

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

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

est non vide ; il est clair que $\mathrm{A}_{n+1}(x) \subset \mathrm{A}_n(x)$. Dans l'espace complet F, on va considérer la suite décroissante des fermés non vides $\mathrm{F}_n = \overline{\mathrm{T}(\mathrm{A}_n(x))}$; si on montre en plus que leurs diamètres tendent vers 0, on saura que leur intersection est non vide et réduite à un point unique. Si $y, y' \in \mathrm{A}_n(x)$, on a $\|y'-y\| \leq \|y'-x\|+\|y-x\| \leq 2^{-n+1}$. Puisque T est linéaire bornée, on déduit de ce qui précède que si $v, w \in \mathrm{T}(\mathrm{A}_n(x))$, on a $\|v - w\| \leq \|\mathrm{T}\|\, 2^{-n+1}$, c'est à dire que le diamètre de F_n tend vers 0. Puisque F est complet, on sait que $\bigcap_n \overline{\mathrm{T}(\mathrm{A}_n(x))}$ contient exactement un point. Cet unique point est le point $\widetilde{\mathrm{T}}(x)$ cherché.

Pour les propriétés supplémentaires de $\widetilde{\mathrm{T}}$, voir le poly.

1.6. Dual d'un espace normé

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

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

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

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

On a en effet $|f_x(y)| \leq \|y\|_\infty \sum_{k=0}^{+\infty} |x_k| = \|x\|_1\, \|y\|_\infty$ (en utilisant pour tout k la majoration $|y_k| \leq \|y\|_\infty$), et ceci pour tout $y \in c_0$, ce qui montre que $\|f_x\| \leq \|x\|_1$.

Inversement, on va montrer que l'application $\mathrm{S} : x \to f_x$ est surjective de ℓ_1 sur c_0^*, et qu'elle est isométrique. On a dit que si $y = (y_n) \in c_0$, on a $y = \sum_{k=0}^{+\infty} y_k e_k$, série convergente dans c_0. Si f est linéaire et continue sur c_0, l'image par f de la somme y de la série est la somme de la série des images,

$$(*) \qquad f(y) = \sum_{k=0}^{+\infty} f(y_k e_k) = \sum_{k=0}^{+\infty} f(e_k)\, y_k.$$

Posons $x_n = f(e_n) = r_n\, \mathrm{e}^{i\theta_n}$ pour tout $n \geq 0$, avec $r_n = |x_n|$ réel ≥ 0 et θ_n réel ; fixons N et définissons $y = (y_n) \in c_0$ en posant $y_n = \mathrm{e}^{-i\theta_n}$ si $0 \leq n \leq \mathrm{N}$ et $y_n = 0$ si $n > \mathrm{N}$. On aura $\|y\|_\infty = 1$ et

$$\sum_{k=0}^{\mathrm{N}} |x_k| = \sum_{k=0}^{\infty} x_k y_k = f(y) \leq \|f\|.$$

En faisant tendre N vers l'infini on obtient $\sum_{k=0}^{+\infty} |x_k| \le \|f\|$, donc la suite $x = (x_n)$ est dans ℓ_1 et d'après la formule $(*)$ ci-dessus, on a que $f = f_x$. On a donc vérifié que $\mathrm{S} : x \to f_x$ est surjective ; on avait déjà $\|f_x\| \le \|x\|_1$, et on vient de trouver $\|x\|_1 \le \|f\| = \|f_x\|$, ce qui montre l'égalité des normes.

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

2. Espaces de Hilbert

2.1. Produits scalaires

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

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

Certains auteurs exigent qu'un produit scalaire vérifie $\langle x, x\rangle > 0$ pour tout $x \ne 0_{\mathrm{X}}$. Nous ne le ferons pas.

Exemple. Si $x = (x_n)$ et $y = (y_n)$ sont deux éléments de $\ell_2(\mathbb{N})$, on pose

$$\langle x, y\rangle = \sum_{k=0}^{+\infty} x_k \overline{y_k}\,;$$

on définit ainsi un produit scalaire sur ℓ_2. On remarque que $\langle x, x\rangle = \|x\|^2$ pour tout $x \in \ell_2$.

Si x, y sont deux vecteurs d'un espace vectoriel X muni d'un produit scalaire, on a

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

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

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

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

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

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

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

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

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

par la proposition 3.

2.2. Espaces de Hilbert, orthogonalité, bases

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

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

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

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

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

MT404, Cours n^o 3, Lundi 1er octobre 2001.

Exercice traité. Si Y est un sous-espace vectoriel d'un espace normé X, son adhérence $\overline{\mathrm{Y}}$ dans X est un sous-espace vectoriel.

Mise au point. Soit $(\Omega, \mathcal{A})$ un espace mesurable ; on désigne par $\mathcal{L}_0(\Omega, \mathcal{A})$ l'ensemble des applications $\mathcal{A}$-mesurables de Ω dans $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$, c'est à dire les fonctions f telles que $f^{-1}(\mathrm{U}) \in \mathcal{A}$ pour tout ouvert U de $\mathbb{K}$. C'est un espace vectoriel de fonctions. Soit de plus μ une mesure positive sur $(\Omega, \mathcal{A})$; l'ensemble $\mathcal{N}$ des fonctions μ-négligeables est le sous-ensemble de $\mathcal{L}_0$ formé des fonctions f telles que

$$\mu\{\omega \in \Omega : f(\omega) \neq 0\} = 0.$$

C'est un sous-espace vectoriel de $\mathcal{L}_0$. Ce sous-espace est aussi égal à

$$\{f \in \mathcal{L}_0 : \int_\Omega |f|\, d\mu = 0\}$$

ou à

$$\{f \in \mathcal{L}_0 : \int_\Omega |f|^2\, d\mu = 0\}.$$

Sous la dernière forme on voit que $\mathcal{N}$ est un sous-espace vectoriel de $\mathcal{L}_2(\Omega, \mathcal{A}, \mu)$, ce qui permet de considérer le quotient $\mathrm{L}_2(\Omega, \mathcal{A}, \mu) = \mathcal{L}_2/\mathcal{N}$.

Si $f, g \in \mathrm{L}_2$, si $\widetilde{f}, \widetilde{g}$ sont des représentants des classes f et g respectivement, on voit que la quantité

$$\int_\Omega \widetilde{f}(\omega)\overline{\widetilde{g}(\omega)}\, d\mu(\omega)$$

ne dépend pas des représentants choisis (si on change de représentants, la nouvelle fonction sous l'intégrale sera μ-presque partout égale à la précédente, donc l'intégrale sera la même). Cette quantité est le produit scalaire sur L_2, et on écrit simplement

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

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

Démonstration. Pour tout $x \in \mathrm{H}$ on a $|\ell_y(x)| \le \|x\|\,\|y\|$ par la proposition 1.3, donc l'application linéaire ℓ_y est continue et $\|\ell_y\| \le \|y\|$. Or $\|y\|^2 = \ell_y(y) \le \|\ell_y\|\,\|y\|$, d'où l'on déduit que $\|\ell_y\| = \|y\|$.

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

Lemme. *Si x est orthogonal à une partie* A *de* H*, alors x est orthogonal à* $\overline{\mathrm{Vect}(\mathrm{A})}$ *(le sous-espace vectoriel fermé engendré par* A*).*

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

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

Le même raisonnement s'applique à l'espace vectoriel Vect(A) engendré par une partie $\mathrm{A} \subset \mathrm{H}$ quelconque, dont tous les éléments sont orthogonaux à x. De plus, si x est orthogonal à tous les vecteurs d'un ensemble $\mathrm{B} \subset \mathrm{H}$ (par exemple $\mathrm{B} = \mathrm{Vect(A)}$), alors x est aussi orthogonal à l'adhérence de B (parce que l'application $b \to \langle b, x\rangle$ est continue).

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

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

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

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

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

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

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

Lemme 2.2.5 : inégalité de Bessel. *Soient* H *un espace de Hilbert et* $(e_n)_{n\geq 0}$ *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*

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

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

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

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. Le premier point a été vu en exercice la semaine précédente. La norme de la somme de la série s'obtient en passant à la limite dans l'égalité du lemme 3.

Lemme 2.2.7. *Soit* $(e_n)_{n\geq 0}$ *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* $y \in$ 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 5 et le lemme précédent, et $z \in$ F. Pour tout $j \geq 0$, on voit en passant à la limite grâce à la continuité de l'application $x \to \langle x, e_j\rangle$

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

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

Définition 1.7.1. On dit qu'un espace topologique Z 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 $\mathbb{R}$). Plus généralement, tout fermé de $\mathbb{C}$ est séparable (exercice).

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

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).

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

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

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

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

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

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

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

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

Exemples 2.2.5.

1. La suite canonique $(\mathbf{e}_n)_{n\geq 0}$ de l'espace ℓ_2 est 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 ci-dessous).

3.6. Séries de Fourier

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

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

pour tout $s \in [0, 2\pi]$. On a déjà vu que les fonctions $(e_n)_{n\in\mathbb{Z}}$ forment une suite orthonormée dans $\mathrm{L}_2(0, 2\pi)$. 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 $L_2(0, 2\pi)$. On va procéder d'une autre façon.

Pour toute fonction $f \in L_2(0, 2\pi)$, on introduit les coefficients de Fourier complexes

$$\forall n \in \mathbb{Z}, \quad c_n(f) = \langle f, e_n \rangle = \int_0^{2\pi} f(t)\, e^{-int}\, \frac{dt}{2\pi}.$$

D'après Bessel, on a $\|f\|_2^2 = \int_0^{2\pi} |f(s)|^2\, ds/2\pi \ge \sum_{n\in\mathbb{Z}} |c_n(f)|^2$. Il résulte de cette inégalité que pour toute $f \in L_2$, les coefficients de Fourier $c_n(f)$ tendent vers 0 quand $|n| \to +\infty$ (en fait, *a posteriori*, cette inégalité de Bessel est une égalité, puisqu'on va montrer que la famille $(e_n)_{n\in\mathbb{Z}}$ est une base hilbertienne).

La convergence ponctuelle des séries de Fourier 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 $|f(t) - f(s)| \le M\, |t - s|$ pour tous $s, t \in \mathbb{R}$.

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

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

Démonstration. Posons

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

Posons

$$K_N(t) = \sum_{n=-N}^{N} e^{int} = K_N(-t) = e^{-iNt}\, \frac{1 - e^{i(2N+1)t}}{1 - e^{it}}.$$

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

$$f_N(0) - f(0) = \int_0^{2\pi} K_N(t)\, (f(t) - f(0))\, \frac{dt}{2\pi} =$$

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

Posons $g(t) = (f(t) - f(0))/(1 - e^{it})$. Comme f est Lipschitz et $|1 - e^{it}\,| \ge 2|t|/\pi$ lorsque $|t| \le \pi$, la fonction g est bornée, donc $g \in L_2$, ce qui entraîne que les coefficients de Fourier de g tendent vers 0. Or nous avions

$$f_N(0) - f(0) = c_N(g) - c_{-N-1}(g)$$

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

Le résultat précédent montre que la suite exponentielle est totale dans L_2 : il 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.

Cours n° 4, Mercredi 3 octobre 2001.

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

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

Démonstration. Soit H un espace de Hilbert séparable de dimension infinie ; il existe une suite $(d_n)_{n\ge 0}$ qui est dense dans H ; on désigne par n_0 le plus petit entier n tel que $d_n \ne 0$ (il en existe puisque $H \ne \{0\}$) ; on pose $x_0 = d_{n_0}$, $y_0 = 0$ et et $e_0 = \|x_0 - y_0\|^{-1}(x_0 - y_0)$; on pose $F_0 = \mathrm{Vect}(e_0)$ et on remarque que $d_0, \ldots, d_{n_0}$ sont tous dans F_0 (les premiers, s'ils existent, sont nuls, et $d_{n_0} = x_0 = \|x_0 - y_0\|\, e_0 \in F_0$.

On construit la suite orthonormée par récurrence. Supposons $e_0, \ldots, e_{k-1}$ définis, deux à deux orthogonaux et de norme un, ainsi que $n_0 < n_1 < \ldots < n_{k-1}$, de façon que $d_0, \ldots, d_{n_{k-1}}$ soient tous dans $F_{k-1} = \mathrm{Vect}(e_0, \ldots, e_{k-1})$. Désignons par n_k le plus petit entier n tel que $d_n \notin F_{k-1}$; il en existe parce que F_{k-1} est fermé et différent de H, et $n_k > n_{k-1}$ d'après ce qui précède. Posons $x_k = d_{n_k}$, y_k sa projection orthogonale sur F_{k-1} ; alors $x_k - y_k \ne 0$ et on pose $e_k = \|x_k - y_k\|^{-1}(x_k - y_k)$; ce vecteur est de norme un, et orthogonal à F_{k-1}, donc à $e_0, \ldots, e_{k-1}$; on pose $F_k = \mathrm{Vect}(e_0, \ldots, e_k)$; enfin, tous les vecteurs d_n, $n < n_k$ sont dans $F_{k-1} \subset F_k$ et $d_{n_k} = \|x_k - y_k\|\, e_k + y_k \in F_k$, donc $d_0, \ldots, d_{n_k} \in F_k$.

La suite $(e_n)_{n\ge 0}$ est totale dans H puisque l'espace vectoriel qu'elle engendre contient tous les vecteurs (d_n) de la suite dense initiale.

2.3. Théorème de projection

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

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

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

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

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

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

implique pour tous $y, z \in \mathrm{C}$

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

Pour tout entier $n \geq 1$, posons

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

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

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

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

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

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

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

Si F est un sous-espace vectoriel de H il est clair que $\mathrm{F}^\perp$ est un sous-espace vectoriel fermé de H.

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

Une application très utile est le "petit" théorème de Radon-Nikodym. Si μ, ν sont deux mesures positives finies sur un espace mesurable $(\Omega, \mathcal{A})$, 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

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

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

La démonstration fonctionne ainsi : il résulte de l'hypothèse que $\int h\, d\nu \le \int h\, d\mu$ 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 $g = \mathbf{1}_{\mathrm{A}}$ on obtient le résultat annoncé.

MT404, Cours n$^{\text{o}}$ 5, Lundi 8 octobre 2001.

Exercice proposé. Si F est un sous-espace vectoriel de dimension finie d'un espace normé, alors F est fermé ; indication : utiliser une base et le théorème d'équivalence des normes en dimension finie.

Reprenons le “petit” théorème de Radon-Nikodym calmement. On rappelle que μ et ν sont deux mesures positives finies sur $(\Omega, \mathcal{A})$ et que $\nu(\mathrm{A}) \le \mu(\mathrm{A})$ pour tout $\mathrm{A} \in \mathcal{A}$. On veut trouver une fonction $\mathcal{A}$-mesurable réelle f telle que $0 \le f \le 1$ et

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

On va travailler avec les espaces de fonctions réelles ($\mathbb{K} = \mathbb{R}$). Une fonction $\mathcal{A}$-étagée $g \ge 0$ est une fonction g de la forme

$$g = \sum_{i=1}^{n} a_i \mathbf{1}_{\mathrm{A}_i}$$

où $\mathrm{A}_1, \ldots, \mathrm{A}_n$ forment une partition de Ω en ensembles de $\mathcal{A}$, et $a_i \ge 0$ pour $i = 1, \ldots, n$. Dans ce cas

$$\int g \, d\nu = \sum_{i=1}^{n} a_i \nu(\mathrm{A}_i) \le \sum_{i=1}^{n} a_i \mu(\mathrm{A}_i) = \int g \, d\mu.$$

On sait que toute fonction $\mathcal{A}$-mesurable positive g est limite croissante d'une suite (g_n) de fonctions $\mathcal{A}$-étagées positives. On obtient avec le lemme des suites croissantes

$$\int g \, d\nu = \lim_n \int g_n \, d\nu \le \lim_n \int g_n \, d\mu = \int g \, d\mu.$$

On en déduit que toute fonction $g \in \mathrm{L}_2(\mu)$ est intégrable pour ν,

$$\int |g| \, d\nu \le \int |g| \, d\mu \le \sqrt{\mu(\Omega)} \, \|g\|_{\mathrm{L}_2(\mu)}$$

par Cauchy-Schwartz appliqué au produit de 1 et de $|g|$. On peut donc poser

$$\forall g \in \mathrm{L}_2(\mu), \quad \ell(g) = \int g \, d\nu \,;$$

on définit ainsi une forme linéaire, et $|\ell(g)| \le \ell(|g|) \le \sqrt{\mu(\Omega)} \, \|g\|_{\mathrm{L}_2(\mu)}$, donc cette forme linéaire est continue. On peut par conséquent la représenter par produit scalaire avec une fonction $f \in \mathrm{L}_2(\mu)$, sans barre de conjugaison puisqu'on travaille sur les réels,

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

On applique à $g = \mathbf{1}_{\mathrm{A}}$, qui est bien élément de $\mathrm{L}_2(\mu)$ puisque μ est finie, pour obtenir

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

comme promis. On montre pour finir que $0 \le f \le 1$ en utilisant le lemme suivant (d'abord pour f, puis pour $1 - f$) :

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

Démonstration. Soient $c > 0$ et $\mathrm{A}_c = \{f \leq -c\}$, qui est un ensemble de $\mathcal{A}$ puisque f est $\mathcal{A}$-mesurable ; on aura

$$0 \leq \int_{\mathrm{A}_c} f\,d\mu \leq -c\,\mu(\mathrm{A}_c) \leq 0$$

ce qui implique $\mu(\mathrm{A}_c) = 0$. En prenant la réunion des ensembles μ-négligeables A_c sur les valeurs $c = 2^{-k}$, k entier ≥ 0 on en déduit que $\mu\bigl(\{f < 0\}\bigr) = 0$.

3. Les espaces de Banach classiques

3.1. Espaces de fonctions continues ou intégrables

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

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

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

Démontrons le théorème pour $\mathrm{K} = [0, 1]$, voir le poly pour le cas général. On considère un entier $\mathrm{N} > 1$ et $\mathrm{N}+1$ fonctions $g_0, \ldots, g_\mathrm{N}$, linéaires sur chaque intervalle $[i/\mathrm{N}, (i+1)/\mathrm{N}]$, $i = 0, \ldots, \mathrm{N}-1$, telles que $g(i/\mathrm{N}) = 1$ et que g soit nulle en dehors de $[(i-1)/\mathrm{N}, (i+1)/\mathrm{N}]$. La formule qui définit ces fonctions est

$$\forall t \in [0, 1], \quad g_i(t) = \max\{1 - \mathrm{N}\,|t - i/\mathrm{N}|, 0\}.$$

On vérifie que $\sum_{i=0}^{\mathrm{N}} g_i(t) = 1$ pour tout t, et $g_i \geq 0$. On dit que ce système est une *partition de l'unité*. Pour toute fonction continue f sur $[0, 1]$, on introduit le *module de continuité de* f, qui est une fonction notée ω_f, définie pour tout $\delta > 0$ par

$$\omega_f(\delta) = \sup\{|f(s) - f(t)| : s, t \in [0, 1],\ |s - t| \leq \delta\}.$$

Dire que f est uniformément continue sur $[0, 1]$ revient à dire que $\lim_{\delta \to 0} \omega_f(\delta) = 0$. Soit F_N le sous-espace de dimension finie de $\mathrm{C}([0, 1])$ engendré par $g_0, \ldots, g_\mathrm{N}$; on obtient alors

pour toute fonction continue f sur $[0, 1]$, *on a* $\mathrm{dist}(f, \mathrm{F}_\mathrm{N}) \leq \omega_f(1/\mathrm{N})$.

On pose $g = \sum_{j=0}^{\mathrm{N}} f(j/\mathrm{N})\, g_j \in \mathrm{F}_\mathrm{N}$; on voit que pour tout $s \in [0, 1]$,

$$f(s) - g(s) = \sum_{j=0}^{\mathrm{N}} g_j(s)(f(s) - f(j/\mathrm{N})).$$

Si $g_j(s)(f(s) - f(j/\mathrm{N})) \neq 0$ pour un certain indice j, on a $g_j(s) \neq 0$, donc s est dans l'intervalle $[(j-1)/\mathrm{N}, (j+1)/\mathrm{N}]$, donc $|f(s) - f(j/\mathrm{N})| \leq \omega_f(1/\mathrm{N})$, donc pour tout $j = 0, \ldots, \mathrm{N}$ on a

$$g_j(s)|f(s) - f(j/\mathrm{N})| \leq \omega_f(1/\mathrm{N})\, g_j(s)\,;$$

en sommant en j on obtient l'inégalité $|f(s) - g(s)| \leq \omega_f(1/\mathrm{N})$ pour tout $s \in [0,1]$, c'est à dire

$$\mathrm{dist}(f, \mathrm{F_N}) \leq \omega_f(1/\mathrm{N}).$$

Si on fait tendre N vers l'infini, on aura pour toute fonction continue f

$$\mathrm{dist}(f, \mathrm{F_N}) \leq \omega_f(1/\mathrm{N}) \to 0.$$

Il en résulte que $\bigcup_{\mathrm{N}} \mathrm{F_N}$ est dense dans $\mathrm{C}([0,1])$, donc $\mathrm{C}([0,1])$ est séparable d'après la proposition 1.7.1.

Théorème. *L'espace* $\mathrm{L}_p(\Omega, \mathcal{A}, \mu)$ *est 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 $v_k = |u_k|$, $g_n = \bigl(\sum_{k=0}^{n} v_k\bigr)^p$, remarquons que $\|v_k\|_p = \|u_k\|_p$ pour obtenir $\int g_n(s)\, d\mu(s) = \|\sum_{k=0}^{n} v_k\|_p^p \leq \bigl(\sum_{k=0}^{n} \|v_k\|_p\bigr)^p \leq \mathrm{M}^p$. 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 $\int g(s)\, d\mu(s) = \lim_n \int g_n(s)\, d\mu(s) \leq \mathrm{M}^p$. 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 $|\mathrm{U}(s) - \mathrm{U}_n(s)|^p \leq g(s)$ 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 que

$$\|\mathrm{U} - \mathrm{U}_n\|_p^p = \int |\mathrm{U} - \mathrm{U}_n|^p\, d\mu \to 0.$$

Il existe un analogue de ℓ_∞ en théorie de l'intégration : c'est l'espace $\mathrm{L}_\infty(\Omega, \mathcal{A}, \mu)$ des classes de fonctions mesurables bornées sur Ω (c'est à dire des classes qui contiennent un représentant borné). La norme $\|f\|_\infty$ est la plus petite constante M telle que l'on ait $|f(s)| \leq \mathrm{M}$ pour μ-presque tout $s \in \Omega$. L'espace L_∞ est complet pour cette norme.

3.2. Résultats de densité

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

Soient K un espace métrique compact, $\mathcal{B}$ sa tribu borélienne (c'est à dire la plus petite tribu de parties de K contenant les ouverts de K) et μ une mesure ≥ 0 finie sur l'espace mesurable $(K, \mathcal{B})$.

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

Démonstration. Prenons $p = 1$ pour simplifier. Désignons par F l'adhérence dans L_1 du sous-espace vectoriel formé par les (classes de) fonctions continues sur K. On va montrer que les ensembles $A \in \mathcal{B}$ tels que $\mathbf{1}_A \in F$, forment une tribu $\mathcal{A}$ de parties de K.

Tout d'abord, $\mathbf{1}_\emptyset = 0 \in F$ puisque F est un espace vectoriel. Remarquons que la fonction constante 1 est dans $C(K)$, donc dans F. Si $A \in \mathcal{A}$, alors $\mathbf{1}_{A^c} = 1 - \mathbf{1}_A \in F$ montre que $A^c \in \mathcal{A}$. Si $A \in \mathcal{A}$ et $\mathbf{1}_A = \lim f_n$, $B \in \mathcal{A}$ et $\mathbf{1}_B = \lim g_n$, alors la suite des fonctions $\inf(f_n, g_n)$ tend vers $\mathbf{1}_{A \cap B} = \inf(\mathbf{1}_A, \mathbf{1}_B)$ dans L_1 (petit exercice ; utiliser le fait que $|\inf(s,t) - \inf(s', t')| \leq |s - s'| + |t - t'|$, pour tous réels s, s', t, t'), donc $A \cap B \in \mathcal{A}$. Si (A_n) est une suite croissante d'éléments de $\mathcal{A}$ et si A désigne sa réunion, la fonction $\mathbf{1}_A$ est limite simple de la suite des $(\mathbf{1}_{A_n})$ et la convergence de $|\mathbf{1}_A - \mathbf{1}_{A_n}|$ vers 0 est dominée par la fonction intégrable fixe $\mathbf{1}_A$, donc $\mathbf{1}_{A_n}$ tend vers $\mathbf{1}_A$ dans L_1 ; puisque F est fermé dans L_1 et $\mathbf{1}_{A_n} \in F$ pour tout n, il en résulte que $\mathbf{1}_A \in F$. Toutes ces propriétés impliquent que $\mathcal{A}$ est une tribu.

On vérifie ensuite que cette tribu $\mathcal{A}$ contient les ouverts de K : si U est un ouvert de K, on définit f_n en posant $f_n(t) = \min\{1, n\, d(t, U^c)\}$ pour tout $t \in K$; cette suite de fonctions continues tend simplement vers $\mathbf{1}_U$, et on montre que $\mathbf{1}_U = \lim f_n$ dans L_1 par convergence dominée. Puisque $\mathcal{A}$ est une tribu contenant les ouverts de K, on a $\mathcal{B} \subset \mathcal{A}$, donc $\mathcal{A} = \mathcal{B}$.

Pour tout borélien $B \in \mathcal{B}$, on sait maintenant que $B \in \mathcal{A}$, donc $\mathbf{1}_B \in F$; par linéarité, il en résulte que F contient toutes les fonctions $\mathcal{B}$-étagées, d'où $F = L_1$ parce que les fonctions étagées sont denses dans L_1 (par la construction usuelle de l'intégrale), ce qui est le résultat de densité annoncé.

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

3.3. Hölder et dualité des espaces ℓ_p

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

Théorème 3.3.1 : inégalité de Hölder. *Soient* $p, q \in [1, +\infty]$ *tels que* $1/p + 1/q = 1$ *; si* $f \in \mathrm{L}_p(\Omega, \mu)$ *et* $g \in \mathrm{L}_q(\Omega, \mu)$*, la fonction produit* fg *est intégrable et*

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

Démonstration. Pour alléger un peu, 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

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

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 sur $[0, +\infty[$ la fonction $t \to tu - t^p/p$). Il en résulte que pour tout $s \in \Omega$

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

ce qui montre que fg est intégrable, et que

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

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

Cours n° 6, Mercredi 10 octobre 2001.

Mise au point. Soient K un espace métrique compact, $\mathcal{B}$ sa tribu borélienne et μ une mesure positive finie sur $(\mathrm{K}, \mathcal{B})$; on a une application naturelle de $\mathrm{C}(\mathrm{K})$ dans $\mathrm{L}_p(\mathrm{K}, \mathcal{B}, \mu)$ qui associe à chaque fonction continue f sur K sa μ-classe dans L_p ; cette application n'est pas injective en général : s'il existe un ouvert non vide U de K tel que $\mu(\mathrm{U}) = 0$, toute fonction f nulle en dehors de U sera dans la classe nulle pour μ.

Il aurait donc mieux valu dire quand on a parlé de la densité de $\mathrm{C}(\mathrm{K})$:

L'image de $\mathrm{C}(\mathrm{K})$ *est dense dans* $\mathrm{L}_p(\mathrm{K}, \mathcal{B}, \mu)$*, lorsque* $1 \le p < +\infty$.

Autre écriture de Hölder. Si $0 \le \alpha \le 1$ et si u, v sont deux fonctions μ-intégrables ≥ 0, alors $u^{1-\alpha} v^\alpha$ est intégrable et

$$\int u^{1-\alpha} v^\alpha\, d\mu \le \Big(\int u\, d\mu\Big)^{1-\alpha} \Big(\int v\, d\mu\Big)^\alpha.$$

Il suffit de poser $1/p = 1 - \alpha$, $1/q = \alpha$, $f = u^{1-\alpha}$ et $g = v^\alpha$ pour se ramener à la version précédente.

Corollaire 3.3.2. *Soient* $p, q \in [1, +\infty[$ *tels que* $1/p + 1/q = 1$ *; si* $f \in \mathrm{L}_p(\Omega, \mathcal{A}, \mu)$,

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

Le résultat est vrai pour $p = +\infty$ *si* μ *vérifie une petite condition supplémentaire : si* $\mathrm{B} \in \mathcal{A}$ *et* $\mu(\mathrm{B}) = +\infty$, *on peut trouver* $\mathrm{A} \subset \mathrm{B}$ *tel que* $\mathrm{A} \in \mathcal{A}$ *et* $0 < \mu(\mathrm{A}) < +\infty$.

Démonstration. L'inégalité de Hölder nous dit déjà que

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

le problème est de montrer l'autre direction. On va voir qu'en fait le *maximum* est atteint pour une certaine fonction $g \in \mathrm{L}_q$, $\|g\|_q \le 1$, lorsque $1 \le p < +\infty$. Si $f = 0$, le résultat est évident, on supposera donc $f \ne 0$, et par homogénéité on peut se ramener à $\|f\|_p = 1$. Soit $\widetilde{f}$ une "vraie" fonction mesurable de la classe f, et définissons une fonction mesurable g sur l'ensemble Ω en posant $g(s) = |\widetilde{f}(s)|^p / \widetilde{f}(s)$ sur l'ensemble mesurable $\mathrm{A} = \{s \in \Omega : \widetilde{f}(s) \ne 0\}$, et $g(s) = 0$ 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 $\varepsilon > 0$. **Il faut savoir** qu'on peut trouver $\mathrm{A} \subset \mathrm{B}$ tel que $\mu(\mathrm{A}) > 0$ et $\mu(\mathrm{A})$ **finie** (ce point n'est pas correct dans le poly). On choisira $\varepsilon < 1$ et on prendra $g(s) = (\mu(\mathrm{A}))^{-1}|f(s)|/f(s)$ si $s \in \mathrm{A}$, $g(s) = 0$ sinon. Alors $\|g\|_1 = 1$ et $\int_\Omega fg\,d\mu > 1 - \varepsilon$.

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 $f \in \mathrm{L}_p$, $g \in \mathrm{L}_q$

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

Ceci signifie que si g est fixée dans L_q, on peut définir une forme linéaire continue ℓ_g sur $\mathrm{L}_p(\Omega, \mathcal{A}, \mu)$ par la formule

$$\forall f \in \mathrm{L}_p, \quad \ell_g(f) = \int_\Omega fg\,d\mu.$$

De plus on a vu ci-dessus que

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

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

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

Remarquons que nous connaissons déjà le résultat quand $p = 2$, et il est valable pour toute mesure μ.

Démonstration. On va seulement traiter le cas μ finie et $1 \le p \le 2$. Voir le poly pour le cas général. On supposera même $\mu(\Omega) = 1$ pour simplifier. Dans ce cas on a avec Hölder (deuxième forme)

$$\int |f|^p \, d\mu = \int |f^2|^{p/2} \, 1^{1-p/2} \, d\mu \le \Bigl(\int |f|^2 \, d\mu\Bigr)^{p/2} \Bigl(\int 1 \, d\mu\Bigr)^{1-p/2} = \|f\|^p_{\mathrm{L}_2}$$

ce qui montre qu'on a une injection i de norme 1 de L_2 dans L_p. Si ℓ est une forme linéaire continue sur L_p, on obtient par restriction à $\mathrm{L}_2 \subset \mathrm{L}_p$ une forme linéaire continue $\widetilde{\ell} = \ell \circ i$ sur L_2, qui peut donc se représenter au moyen d'une fonction $g_1 \in \mathrm{L}_2$. On a donc en posant $g = \overline{g}_1$,

$$\forall f \in \mathrm{L}_2, \quad \widetilde{\ell}(f) = \ell(f) = \int f(t)\overline{g_1(t)} \, d\mu = \int f(t) g(t) \, d\mu.$$

On montre maintenant que $\int |g|^q \, d\mu < +\infty$. Posons $\mathrm{A}_n = \{|g| \le n\}$ et soit

$$f_n = \mathbf{1}_{\mathrm{A}_n} |g|^{q-1} \operatorname{sign}(g).$$

On vérifie que f_n est une fonction mesurable bornée, donc $f_n \in \mathrm{L}_2$ et on obtient

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

ce qui donne $\bigl(\int_{\mathrm{A}_n} |g|^q \, d\mu\bigr)^{1/q} \le \|\ell\|$. Il ne reste plus qu'à observer que (A_n) tend en croissant vers Ω pour obtenir que $\bigl(\int |g|^q \, d\mu\bigr)^{1/q} \le \|\ell\|$. On sait maintenant que $g \in \mathrm{L}_q$, donc la forme linéaire ℓ_g définie sur L_p par $\ell_g(f) = \int fg \, d\mu$ existe, est continue, et elle coïncide avec ℓ sur L_2 ; comme L_2 est dense dans L_p, on en déduit que $\ell = \ell_g$, et de plus $\|\ell\| = \|\ell_g\| = \|g\|_{\mathrm{L}_q}$ par Hölder.

4. Les théorèmes fondamentaux

4.1. Le théorème de Baire et ses conséquences

Soit X un espace topologique ; un ouvert U de X est dense dans X si et seulement si le fermé complémentaire U^c est d'intérieur vide. Si on a un nombre fini d'ouverts denses, on vérifie facilement de proche en proche que $\mathrm{U}_1 \cap \ldots \cap \mathrm{U}_n$ est encore un ouvert dense. Le théorème de Baire donne un cas où cette propriété triviale d'intersection finie peut s'étendre aux *suites* d'ouverts denses.

Théorème 4.1.1 : théorème de Baire. *Soit* X *un espace métrique complet ; si* $(\mathrm{U}_n)_{n\ge0}$ *est une suite de parties ouvertes et* denses *dans* X*, l'intersection* $\bigcap_{n\ge0}\mathrm{U}_n$ *est dense dans l'espace* X*.*

Démonstration. Soit $(\mathrm{U}_n)_{n\ge0}$ une suite d'ouverts denses de X ; soit V une partie ouverte non vide de X ; on doit montrer que $\bigcap_{n\ge0}\mathrm{U}_n$ 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 $n\ge0$ 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 $n\ge1$, 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

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

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 $\bigcap_{n\ge0}\mathrm{B}_n\ne\emptyset$; or, par construction, $\bigcap_{n\ge0}\mathrm{B}_n\subset\mathrm{V}\cap\bigcap_{n\ge0}\mathrm{U}_n$, ce qui montre que cette dernière intersection est non vide.

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

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

$$\forall\mathrm{T}\in\mathrm{A},\forall x\in\mathrm{W},\quad\|\mathrm{T}(x)\|\le1.$$

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 $\lambda\ne0$ l'homothétie $y\to\lambda y$ est un homéomorphisme. Il résulte du premier point que tout voisinage W de x_0 est de la forme $x_0+\mathrm{V}$, où V est un voisinage de 0_E, et du second point que $\lambda\mathrm{V}$ est aussi un voisinage de 0_E.

Pour tout entier $n\ge1$, posons $\mathrm{C}_n=\{x\in\mathrm{E}:\forall\mathrm{T}\in\mathrm{A},\ \|\mathrm{T}(x)\|\le n\}$. Comme C_n est l'intersection des ensembles fermés $\mathrm{C}_{\mathrm{T},n}=\{x\in\mathrm{E}:\|\mathrm{T}(x)\|\le n\}$ (lorsque T varie dans l'ensemble 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 $n_0\ge1$ tel que C_{n_0} soit d'intérieur non vide. On peut donc trouver un point $x_0\in\mathrm{C}_{n_0}$ et un voisinage V de 0_E tels que $x_0+\mathrm{V}\subset\mathrm{C}_{n_0}$. Soient $v\in\mathrm{V}$ et $\mathrm{T}\in\mathrm{A}$ quelconques ; puisque

$x_0+v \in \mathrm{C}_{n_0}$, on a $\|\mathrm{T}(x_0)+\mathrm{T}(v)\|_\mathrm{Y} \le n_0$, ce qui donne $\|\mathrm{T}(v)\|_\mathrm{Y} \le n_0+\|\mathrm{T}(x_0)\|_\mathrm{Y} \le 2\,n_0$ par l'inégalité triangulaire. Pour terminer, on prend le voisinage $\mathrm{W} = (2\,n_0)^{-1}\mathrm{V}$.

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

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

Exemple traité : si (x_n) est une suite numérique telle que la série $\sum x_k y_k$ converge pour tout $y = (y_k)$ dans $\ell_2 = \ell_2(\mathbb{N})$, alors $\sum |x_k|^2 < +\infty$.

Pour tout entier $\mathrm{N} \ge 0$, désignons par $\xi^{(\mathrm{N})}$ le vecteur

$$\xi^{(\mathrm{N})} = (\overline{x_0}, \ldots, \overline{x_\mathrm{N}}, 0, 0, \ldots) \in \ell_2$$

et considérons l'application linéaire continue T_N (forme linéaire) de ℓ_2 dans $\mathbb{K}$ définie par

$$\forall y = (y_k) \in \ell_2, \quad \mathrm{T}_\mathrm{N}(y) = \langle y, \xi^{(\mathrm{N})}\rangle = \sum_{k=0}^{\mathrm{N}} x_k y_k.$$

On sait que $\|\mathrm{T}_\mathrm{N}\| = \|\xi^{(\mathrm{N})}\|_2$.

D'après l'hypothèse, la suite scalaire $(\mathrm{T}_\mathrm{N}(y))_\mathrm{N}$ converge pour tout $y \in \ell_2$. Considérons donc l'espace complet $\mathrm{E} = \ell_2$ et la famille $\mathrm{A} \subset \mathcal{L}(\mathrm{E},\mathbb{K})$ donnée par

$$\mathrm{A} = \{\mathrm{T}_\mathrm{N} : \mathrm{N} \in \mathbb{N}\}.$$

Puisque la suite $(\mathrm{T}_\mathrm{N}(y))_\mathrm{N}$ converge, elle est bornée, donc

$$\sup_{\mathrm{T}\in\mathrm{A}} \|\mathrm{T}(y)\| = \sup_{\mathrm{N}\in\mathbb{N}} |\mathrm{T}_\mathrm{N}(y)| < +\infty$$

pour tout $y \in \mathrm{E} = \ell_2$. D'après le théorème précédent, on a $\sup_\mathrm{N} \|\mathrm{T}_\mathrm{N}\| < +\infty$, c'est à dire

$$\sum_{k=0}^{+\infty} |x_k|^2 = \sup_\mathrm{N} \sum_{k=0}^{\mathrm{N}} |x_k|^2 = \sup_\mathrm{N} \|\xi^{(\mathrm{N})}\|_2^2 = \sup_\mathrm{N} \|\mathrm{T}_\mathrm{N}\|^2 < +\infty.$$

MT404, Cours n^o 7, Lundi 15 octobre 2001.

Précisions sur le dual de $\mathrm{L}_p(\Omega, \mathcal{A}, \mu)$

On peut montrer que si $1 < p < +\infty$, le dual de L_p "est" L_q pour toute mesure (c'est bien ce qu'on a vu quand $p = 2$ par la théorie des espaces de Hilbert). En revanche, le dual de $\mathrm{L}_1(\Omega, \mathcal{A}, \mu)$ n'est pas toujours $\mathrm{L}_\infty(\Omega, \mathcal{A}, \mu)$, mais c'est vrai si la mesure μ est σ-finie, ce qui veut dire que Ω est réunion d'une suite de parties mesurables de mesure finie : il existe $(\mathrm{A}_n) \subset \mathcal{A}$ telle que $\Omega = \bigcup_n \mathrm{A}_n$ et $\mu(\mathrm{A}_n) < +\infty$ pour tout n.

Si μ est la mesure qui vaut toujours $+\infty$ sauf pour $\mathrm{A} = \emptyset$, on a $\mathrm{L}_1 = (0)$, et $\mathrm{L}_\infty \neq (0)$ (il contient la fonction constante $\mathbf{1}$) ne peut pas être son dual.

Dual de ℓ_p

Il est intéressant de voir la démonstration directe de la dualité dans le cas des espaces de suites. C'est la même chose que ce qui a été fait pour le dual de c_0 dans la première semaine de cours.

Supposons $1 < p < +\infty$ et $1/q = 1 - 1/p$. On montre d'abord que tout vecteur $x = (x_n)$ de ℓ_p, $1 \le p < +\infty$, est somme de la série $\sum_{k=0}^{+\infty} x_k \mathbf{e}_k$, série convergente dans ℓ_p. Si f est une forme linéaire continue sur ℓ_p, on en déduit que

$$\forall x \in \ell_p, \quad f(x) = \sum_{k=0}^{+\infty} x_k f(\mathbf{e}_k).$$

Il reste seulement à voir que la suite $(f(\mathbf{e}_k))$ est dans ℓ_q ; on écrit $f(\mathbf{e}_k) = r_k \,\mathrm{e}^{i\theta_k}$ avec $r_k \ge 0$ et θ_k réel pour tout $k \ge 0$; ensuite, pour n fixé on choisit $x \in \ell_p$ de la forme

$$x = (r_0^{q-1}\,\mathrm{e}^{-i\theta_0}, \ldots, r_n^{q-1}\,\mathrm{e}^{-i\theta_n}, 0, 0, \ldots).$$

En appliquant f à x on trouve (en utilisant la relation $p(q-1) = q$)

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

d'où $(\sum_{k=0}^{n} r_k^q)^{1-1/p} \le \|f\|$, puis quand n tend vers l'infini

$$\sum_{k=0}^{+\infty} |f(\mathbf{e}_k)|^q \le \|f\|^q.$$

La forme linéaire f coïncide donc avec la forme linéaire f_y provenant d'un certain élément y de ℓ_q, l'élément $y = (y_n)$ donné par $y_n = f(\mathbf{e}_n)$ pour tout $n \ge 0$.

On montrerait de même que le dual de ℓ_1 "est" ℓ_∞. Mais le dual de ℓ_∞ n'est pas ℓ_1, il est "plus grand" ; on le déduira de Hahn-Banach.

Banach-Steinhaus : suite et fin

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

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

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

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

Démonstration. Soit T 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 $\mathrm{T(B_E)}$ contient une boule ouverte centrée en 0_F,

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

On a $\bigcup_{n\geq1} n\mathrm{B_E} = \mathrm{E}$ donc $\bigcup_{n\geq1} n\mathrm{T(B_E)} = \mathrm{T(E)} = \mathrm{F}$, donc $\bigcup_{n\geq1} \overline{n\mathrm{T(B_E)}} = \mathrm{F}$. Par le théorème de Baire, il existe un entier $n \geq 1$ tel que l'ensemble fermé $\overline{n\mathrm{T(B_E)}}$ soit d'intérieur non vide. Comme la multiplication par $n \geq 1$ est un homéomorphisme, on en déduit que l'intérieur de $\overline{\mathrm{T(B_E)}}$ n'est pas vide. Il reste à voir que 0_F est dans cet intérieur : cela provient de la convexité et de la symétrie de l'ensemble $\mathrm{C} = \overline{\mathrm{T(B_E)}}$: 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 $\mathrm{B}(0, r) \subset \mathrm{C}$.

Si $\mathrm{T(B_E)}$ était fermé, la démonstration serait (presque) terminée et beaucoup plus simple. Dans le cas général, il faut faire une deuxième étape qui permet de se débarrasser du signe adhérence dans la formule $(*)$. La deuxième étape consiste en effet à montrer que la propriété $(*)$ entraîne en fait, parce que E est complet, que l'on a

$$(**) \qquad \mathrm{B_F}(0, r) \subset 2\,\mathrm{T(B_E)}.$$

Soit $a_0 \in \mathrm{A} = \overline{\mathrm{T(B_E)}}$; il existe un vecteur $u_0 \in \mathrm{B_E}$ tel que $\|a_0 - \mathrm{T}(u_0)\| < r/2$; on peut écrire $a_0 - \mathrm{T}(u_0) = \frac{1}{2}\,a_1$ avec $a_1 \in \mathrm{B_F}(0, r) \subset \mathrm{A}$ (d'après $(*)$), donc on peut écrire $a_0 = \mathrm{T}(u_0) + \frac{1}{2}\,a_1$ avec $a_1 \in \mathrm{A}$. En recommençant avec a_1, on trouve $u_1 \in \mathrm{B_E}$ et $a_2 \in \mathrm{A}$ tels que $a_1 = \mathrm{T}(u_1) + \frac{1}{2}\,a_2$, ce qui donne

$$a_0 = \mathrm{T}(u_0 + \frac{1}{2}\,u_1) + \frac{1}{4}\,a_2.$$

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

$$(1) \qquad a_0 = \mathrm{T}(u_0 + \frac{1}{2}\,u_1 + \cdots + \frac{1}{2^k}\,u_k) + \frac{1}{2^{k+1}}\,a_{k+1}.$$

La série $\sum 2^{-k}\,u_k$ est normalement convergente, donc convergente dans E puisque E est complet ; sa somme $x_0 = \sum_{k=0}^{+\infty} 2^{-k}\,u_k \in \mathrm{E}$ est telle que

$$\mathrm{T}(x_0) = \lim_k \mathrm{T}(u_0 + 2^{-1}u_1 + \cdots + 2^{-k}u_k)$$

et d'après la relation (1), on a $\mathrm{T}(x_0) = a_0$ puisque A est borné. Par ailleurs

$$\|x_0\| \leq \sum_{k=0}^{+\infty} 2^{-k} \|u_k\| \leq 2$$

donc $(**)$ est démontrée.

Il est clair maintenant que l'application réciproque T^{-1} est continue : en effet, l'inclusion précédente se traduit par $\mathrm{T}^{-1}(\mathrm{B_F}(0, r)) \subset 2\,\mathrm{B_E}$, ce qui implique $\mathrm{T}^{-1}(\mathrm{B_F}) \subset 2\,r^{-1}\mathrm{B_E}$ par homogénéité et passage à l'adhérence ; ceci montre que $\|\mathrm{T}^{-1}\| \leq 2/r$.

Exemple. Supposons que E soit un espace de Banach, et que l'on ait défini une nouvelle norme N sur E telle que $\mathrm{N}(x) \leq \|x\|$ pour tout $x \in \mathrm{E}$. Si le nouvel espace normé (E, N) reste complet, la norme N est équivalente à la norme initiale (appliquer le théorème précédent avec $\mathrm{T} = \mathrm{Id}$ et $\mathrm{F} = (\mathrm{E}, \mathrm{N})$).

Cela signifie que la norme d'un espace de Banach n'est pas n'importe quoi : on ne peut pas la remplacer par une strictement plus faible (c'est à dire plus petite et non équivalente) sans perdre le caractère complet.

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

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

4.2. Théorème de Hahn-Banach

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

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

Exemples 4.2.1. Les semi-normes sont des fonctions sous-linéaires.

Une fonction sous-linéaire sur $\mathbb{R}$ est linéaire par morceaux : elle vaut $q(t) = at$ pour $t < 0$ et $q(t) = bt$ pour $t \geq 0$, avec $a \leq b$ (en effet, on doit avoir $-a + b = q(-1) + q(1) \geq q(0) = 0$).

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

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

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

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

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

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

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

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

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

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

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

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

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

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

et il est alors clair que cette propriété est vraie :

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

Le lemme est donc établi.

Cours n° 8, Mercredi 17 octobre 2001.

Le lemme de Zorn

Le lemme de Zorn est assez directement équivalent à un axiome de la théorie des ensembles, *l'axiome du choix*. Il permet de valider certains types de raisonnements où on cherche à garantir l'existence d'objets maximaux. L'axiome du choix dit que le produit $\prod_{i \in \mathrm{I}} \mathrm{X}_i$ est non vide lorsque tous les ensembles X_i sont non vides ; cela revient à dire qu'il existe une famille $(x_i)_{i \in \mathrm{I}} \in \prod_{i \in \mathrm{I}} \mathrm{X}_i$, ou encore que l'on a pu *choisir* un élément x_i dans chaque ensemble non vide X_i. Voici maintenant le lemme de Zorn :

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

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

Il faut une démonstration, pas complètement évidente, pour passer de l'axiome du choix au lemme de Zorn. Indiquons seulement le point de contact entre les deux : soit I ordonné inductif, désignons par $\mathcal{I}$ l'ensemble des sous-ensembles $\mathrm{T} \subset \mathrm{I}$ totalement ordonnés et tels que T admette un majorant $j \notin \mathrm{T}$, ce qui équivaut à dire que T ne contient pas d'élément maximal. A chaque $\mathrm{T} \in \mathcal{I}$, on associe l'ensemble non vide X_T des majorants $j \notin \mathrm{T}$ de T dans I ; l'axiome du choix donne une fonction de choix φ qui sélectionne un majorant $\varphi(\mathrm{T}) \notin \mathrm{T}$ pour tout sous-ensemble totalement ordonné T qui ne contient pas d'élément maximal.

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

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

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

Dans le cas normé séparable, on peut donner une démonstration de la version suivante du théorème de Hahn-Banach qui n'utilise pas l'axiome du choix, mais seulement l'axiome ACD, *axiome du choix dépendant*, sans lequel il est à peu près impossible de faire des mathématiques classiques.

Théorème 4.2.5 : théorème de Hahn-Banach. *Soient* X *un espace normé (réel ou complexe) et* Y *un sous-espace vectoriel de* X *; pour tout* $\ell \in \mathrm{Y}^*$, *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 vectoriel Y. On peut donc trouver un prolongement m tel que $m \le q$ sur X, ce qui donne le résultat : on a en effet $m(x) \le \|\ell\|\,\|x\|$ pour tout $x \in \mathrm{X}$, d'où aussi $|m(x)| \le \|\ell\|\,\|x\|$ en appliquant à x et $-x$; tout ceci montre que m est continue et $\|m\| \le \|\ell\|$, mais on a aussi $\|\ell\| \le \|m\|$ 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 (par la première partie, cas réel) une forme linéaire réelle m_1 sur X telle que m_1 prolonge la forme linéaire réelle ℓ_1 et $\|m_1\| = \|\ell_1\|$. On pose ensuite

$$\forall x \in \mathrm{X}, \quad m(x) = m_1(x) - i m_1(ix)$$

et on vérifie que m est une $\mathbb{C}$-forme linéaire telle que $m_1 = \mathrm{Re}\,m$. De plus si on écrit $m(x) = r\,\mathrm{e}^{i\theta}$ avec $r \ge 0$, on voit que

$$|m(x)| = r = m(\mathrm{e}^{-i\theta}\,x) = m_1(\mathrm{e}^{-i\theta}\,x) \le \|m_1\|\,\|\,\mathrm{e}^{-i\theta}\,x\| \le \|\ell\|\,\|x\|\,;$$

on a donc $\|m\| \le \|\ell\|$; d'autre part m prolonge ℓ (ici Y est un sous-espace vectoriel complexe ; si $y \in \mathrm{Y}$ on a aussi $iy \in \mathrm{Y}$ ce qui permet d'écrire $m(y) = m_1(y) - i m_1(iy)$, et alors $m(y) = \ell_1(y) - i\ell_1(iy) = \ell(y)$).

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

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

Exemple. Une forme linéaire continue sur ℓ_∞ qui ne provient pas d'un $y \in \ell_1$.

Considérons $\mathrm{X} = \ell_\infty$ et son sous-espace vectoriel fermé $\mathrm{Y} = c_0$; le vecteur $x = (1, 1, \ldots, 1, \ldots)$ n'est pas dans c_0, et on voit que $\mathrm{dist}(x, c_0) = 1$. D'après le corollaire, il existe une forme linéaire continue ℓ sur ℓ_∞ telle que $\ell(x) \ne 0$ et ℓ nulle sur c_0. En particulier, cette forme linéaire non nulle est nulle sur tous les vecteurs $\mathbf{e}_n$ de la suite canonique, puisque $\mathbf{e}_n \in c_0$. Or, si on avait un vecteur $y \in \ell_1$ tel que $\ell = \ell_y$, c'est à dire

$$\forall z = (z_n) \in \ell_\infty, \quad \ell_y(z) = \sum_{k=0}^{+\infty} z_k y_k$$

on en déduirait $y_k = \ell(\mathbf{e}_k) = 0$ pour tout k, donc $y = 0$ et $\ell_y = 0$, ce qui n'est pas le cas.

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

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

Le théorème de Hahn-Banach donne des outils pour étudier la *totalité* d'un ensemble dans un espace vectoriel normé X :

pour qu'un sous-ensemble $\mathrm{T} \subset \mathrm{X}$ *soit total dans* X, *il faut et il suffit que toute forme linéaire* $x^* \in \mathrm{X}^*$, *nulle sur* T, *soit identiquement nulle.*

Si T est total, toute forme linéaire continue x^* nulle sur T est nulle sur Vect(T) par linéarité, puis sur $\mathrm{X} = \overline{\mathrm{Vect}(\mathrm{T})}$ par continuité, donc $x^* = 0$.

Inversement, si T n'est pas total, on considère le sous-espace vectoriel fermé $\mathrm{Y} = \overline{\mathrm{Vect}(\mathrm{T})} \neq \mathrm{X}$; on peut trouver un vecteur $x \notin \mathrm{Y}$, puis une forme linéaire x^* telle que $x^*(x) \neq 0$ et x^* nulle sur Y, donc en particulier nulle sur $\mathrm{T} \subset \mathrm{Y}$.

Exemple traité. On suppose donnée une suite $(a_n) \subset]0, +\infty[$ qui converge vers $a > 0$, avec $a_n \neq a$ pour tout n. On définit la fonction $f_n(x) = \mathrm{e}^{-a_n x}$ pour $x \ge 0$; cette fonction f_n est dans $\mathrm{L}_1([0, +\infty[)$ pour tout $n \ge 0$. On va voir que

l'ensemble $(f_n)_{n \ge 0}$ *est total dans* $\mathrm{L}_1([0, +\infty[)$.

Pour la vérification, on va utiliser le critère précédent : toute forme linéaire continue sur L_1 qui est nulle sur tous les f_n, est la forme nulle. On utilisera aussi un résultat sur la transformée de Laplace, qu'on justifiera la prochaine fois.

Si g est une fonction mesurable bornée sur $[0, +\infty[$, on définit une fonction $\mathcal{L}g$ (la *transformée de Laplace* de g) par

$$\forall t > 0, \quad (\mathcal{L}g)(t) = \int_0^{+\infty} \mathrm{e}^{-tx}\, g(x)\, dx.$$

On a le résultat d'injectivité suivant : si $\mathcal{L}g$ est nulle pour tout $t > t_0$, alors $g = 0$.

Soit donc ℓ une forme linéaire continue sur L_1, qui annule toutes les fonctions f_n ; on sait que ℓ est donnée par une certaine fonction g de L_∞, par la formule

$$\forall f \in \mathrm{L}_1, \quad \ell(f) = \int_0^{+\infty} f(x) g(x)\, dx.$$

La nullité de ℓ sur les f_n donne

$$(\mathcal{L}g)(a_n) = \int_0^{+\infty} \mathrm{e}^{-a_n x}\, g(x)\, dx = 0$$

pour tout $n \ge 0$. Mais la fonction $\mathcal{L}g$ se prolonge en fonction holomorphe à l'ouvert U de $\mathbb{C}$ formé des z tels que $\mathrm{Re}\, z > 0$,

$$(\mathcal{L}g)(z) = \int_0^{+\infty} \mathrm{e}^{-zx}\, g(x)\, dx.$$

Cette fonction holomorphe $\mathcal{L}$ admet donc une suite de zéros (a_n) qui s'accumule en un point $a \in \mathrm{U}$; d'après le principe des zéros isolés, on déduit que $\mathcal{L}g$ est nulle sur U ; d'après l'injectivité de la transformation de Laplace, on déduit que $g = 0$, c'est à dire $\ell = 0$. On a bien montré que 0 est la seule forme linéaire qui soit nulle sur l'ensemble (f_n). Cette suite est donc totale dans $\mathrm{L}_1([0, +\infty[)$.