Chapitre 1

Compléments de topologie

1.1 Espaces topologiques

En licence comme à l'agrégation, les concepts topologiques considérés sont toujours associés à une distance d : une suite (x_n) converge vers x si et seulement si $\lim d(x_n, x) = 0$.

Mais il faut parfois sortir de ce cadre. Voici un exemple concret. Soit H un espace de Hilbert, (x_n) une suite d'éléments de H et $x \in H$. On dit que (x_n) converge faiblement vers x si et seulement si, pour tout y dans H, on a $\lim\langle y, x_n\rangle = \langle y, x\rangle$. On peut montrer qu'il n'existe pas de distance sur H associée à cette notion de convergence. Le bon cadre est alors celui des espaces topologiques.

Définition. Soit X un ensemble. On appelle *topologie* sur X un ensemble $\mathcal{T} \subset \mathcal{P}(X)$ tel que

1. on a $\emptyset \in \mathcal{T}$ et $X \in \mathcal{T}$,
2. l'ensemble $\mathcal{T}$ est stable par union quelconque,
3. l'ensemble $\mathcal{T}$ est stable par intersection finie.

On dit que $(X, \mathcal{T})$ est un *espace topologique*. Les éléments de $\mathcal{T}$ s'appellent les *ouverts*.

Exemple. Soit (X, d) un espace métrique. On note $B(x, \varepsilon)$ (resp. $B_f(x, \varepsilon)$) la boule ouverte (resp. fermée) de centre $x \in X$ et de rayon $\varepsilon > 0$. Rappelons qu'une partie A de X est un ouvert si $\forall x \in A, \exists \varepsilon > 0 \;:\; B(x, \varepsilon) \subset A$. L'ensemble des ouverts ainsi définis forme une topologie sur X, dite topologie *associée à d*.

Une topologie sur X telle qu'il existe une distance à laquelle elle est associée est dite *métrisable*.

Exemple. Pour tout ensemble X, l'ensemble $\mathcal{P}(X)$ est une topologie. Est-elle métrisable ? Oui, par la distance discrète $d(x, y) = \mathbf{1}_{x=y}$. On dit que $\mathcal{P}(X)$ est la *topologie discrète*.

Exemple. Pour tout ensemble X, l'ensemble $\{\emptyset, X\}$ est une topologie dite *topologie grossière*. Est-elle métrisable ? Non !

Une topologie $\mathcal{T}$ est dite *séparée* si pour tous $x \neq y$ dans X, on peut trouver deux ouverts disjoints O_x et O_y tels que $x \in O_x$, $y \in O_y$. La topologie grossière sur un ensemble de cardinal $\geqslant 2$ n'est pas séparée, alors que toute topologie métrisable est séparée (choisir $O_x = B(x, r)$ et $O_y = B(y, r)$ pour $r = d(x, y)/2$).

Remarque. Si une topologie est métrisable, il existe plusieurs distances différentes à qui elle est associée. Par exemple, une topologie associée à la distance d est aussi associée à la distance $2d$, mais aussi à la distance d' définie par

$$d'(x, y) = \min(d(x, y), 1).$$

Si $(X, \mathcal{T})$ est un espace topologique et si $Y \subset X$, l'ensemble $\{O \cap Y \ : \ O \in \mathcal{T}\}$ est une topologie sur Y dite *topologie induite* par $\mathcal{T}$.

On étend au cadre des espaces topologiques le vocabulaire usuel du cadre métrique. Soit $(X, \mathcal{T})$ un espace topologique.

- On dit que $A \subset X$ est *fermé* (pour $\mathcal{T}$) si $X \setminus A \in \mathcal{T}$.
- On dit que que $A \subset X$ est un *voisinage* de $x \in X$ si il existe un ouvert V tel que $x \in V \subset A$.
- L'*intérieur* de $A \subset X$, noté $\mathring{A}$, est l'ensemble des points dont A est un voisinage ; c'est aussi la réunion des ouverts inlcus dans A.
- L'*adhérence* de $A \subset X$, notée $\bar{A}$, est $\{x \in X \ : \ \text{tout voisinage de } x \text{ intersecte } A\}$; c'est aussi l'intersection des fermés contenant A.
- Une partie $A \subset X$ est *dense* si $\bar{A} = X$. Une partie est dense si et seulement si elle intersecte tout ouvert non vide.
- Une suite (x_n) d'éléments de X converge vers $x \in X$ si, pour tout voisinage V de x, il existe un entier N tel que $n \geqslant N \Longrightarrow x_n \in V$.

Soit $(X, \mathcal{T}_X)$ et $(Y, \mathcal{T}_Y)$ deux espaces topologiques et $f : X \to Y$ une fonction

- On dit que f est *continue* si, pour tout $O \in \mathcal{T}_Y$ on a $f^{-1}(O) \in \mathcal{T}_X$.
- On dit que f est *continue en* $x \in X$ si, pour tout voisinage V de $f(x)$ pour $\mathcal{T}_Y$, l'image réciproque $f^{-1}(V)$ est un voisinage de x pour $\mathcal{T}_X$. On vérifie que f est continue si et seulement si elle est continue en tout point.
- On dit que f est un *homéomorphisme* si (1) f est bijective et (2) f et f^{-1} sont continues.

Remarque. Les concepts de suites de Cauchy et de complétude dépendent de la distance sous-jacente et ne peuvent pas être définis uniquement connaissant la topologie. Par exemple, sur $\mathbf{R}$, la topologie usuelle est associée à la distance usuelle mais aussi à la distance d définie par

$$d(x, y) = |\arctan x - \arctan y|.$$

Si $\mathbf{R}$ est complet pour la distance usuelle, il ne l'est pas pour la distance d puisque la suite (x_n) définie par $x_n = n$ est une suite de Cauchy (pour d) non convergente.

1.2 Compacité

Définition. Un espace topologique $(X, \mathcal{T})$ est dit *compact* s'il est séparé et si tout recouvrement ouvert de X admet un sous-recouvrement fini, c'est-à-dire que si une famille $(O_\alpha)_{\alpha \in A}$ d'éléments de $\mathcal{T}$ vérifie $X = \bigcup_{\alpha \in A} O_\alpha$, alors il existe un sous-ensemble fini F de A tel que $X = \bigcup_{\alpha \in F} O_\alpha$.

Voici une conséquence de la compacité.

Proposition. *Soit $(X, \mathcal{T})$ un espace topologique compact et soit (F_n) une suite décroissante de fermés non vides. Alors $\bigcap_n F_n$ est non vide.*

En effet, si $\bigcap F_n$ était vide, la suite croissante d'ouverts (O_n) définis par $O_n = X \setminus F_n$ serait un recouvrement de X. Par compacité, il existerait alors $F \subset \mathbf{N}$ fini tel que

$$X = \bigcup_{n \in F} O_n = O_{n_0}$$

pour $n_0 = \max F$, d'où $F_{n_0} = \emptyset$ et une contradiction.

Un espace topologique est dit *localement compact* s'il est séparé et si tout point admet un voisinage compact.

Un espace vectoriel normé est localement compact si (et seulement si) il est de dimension finie.

Dans un espace métrisable, la notion de compacité peut s'exprimer de façon très différente. En effet, si $(X, \mathcal{T})$ est un espace topologique métrisable, alors X est compact si et seulement si toute suite d'éléments de X admet une sous-suite convergente.

1.3 Produits d'espaces topologiques

1.3.1 Produit fini d'espaces topologiques

Soient $(X_1, \mathcal{T}_1), \ldots, (X_p, \mathcal{T}_p)$ des espaces topologiques. On note X le produit cartésien $X_1 \times \cdots \times X_p$. On appelle *pavé ouvert* un ensemble de la forme $O_1 \times \cdots \times O_p$, où O_i est un élément de $\mathcal{T}_i$. On note $\mathcal{T}$ l'ensemble des réunions quelconques de pavés ouverts. C'est une topologie sur X, que l'on appelle topologie produit.

Si $A \subset X$, on a $A \in \mathcal{T}$ si et seulement si, pour tout $x = (x_1, \ldots, x_p) \in A$, il existe $O_1 \in \mathcal{T}_1, \ldots, O_p \in \mathcal{T}_p$ tels que $x_i \in O_i$ et $O_1 \times \cdots \times O_p \subset A$.

De plus, pour tout i, l'application de $i$ème projection $\mathrm{pr}_i : X \to X_i$ définie par

$$\mathrm{pr}_i(x_1, \ldots, x_n) = x_i$$

est continue de $(X, \mathcal{T})$ vers $(X, \mathcal{T}_i)$. La topologie produit est la topologie la moins fine sur X qui rend continues les applications pr_i : si une topologie sur X a cette propriété, elle contient les pavés ouverts puisque

$$O_1 \times \cdots \times O_p = \bigcap_{i=1}^{p} \mathrm{pr}_i^{-1}(O_i)$$

et contient donc $\mathcal{T}$.

Si on suppose maintenant que pour tout i, la topologie $\mathcal{T}_i$ est métrisable et associée à une distance d_i sur X_i, alors la topologie produit est associée à la distance d sur X définie par

$$d((x_1, \ldots, x_p), (y_1, \ldots, y_p)) = \max_{1 \leqslant i \leqslant p} d(x_i, y_i)$$

1.3.2 Produit quelconque

Supposons maintenant donnée une famille quelconque d'espaces topologiques $(X_i, \mathcal{T}_i)_{i \in I}$ et considérons le produit cartésien

$$X = \prod_{i \in I} X_i = \{(x_i)_{i \in I} \ : \ \forall i \in I, x_i \in X_i\}.$$

On définit la topologie produit $\mathcal{T}$ sur X comme étant la topologie la moins fine qui rend continues les applications $\mathrm{pr}_i : X \to X_i$. De manière plus explicite, $\mathcal{T}$ est l'ensemble des réunions quelconques de *cylindres ouverts*, c'est-à-dire d'ensemble de la forme

$$\prod_{i \in I} Y_i$$

où les $Y_i \in \mathcal{T}_i$ sont tels que l'ensemble $\{i \in I \ : \ Y_i \neq X_i\}$ est fini.

Dans la suite on s'intéressera surtout au cas d'un produit dénombrable d'espaces métriques. Soit $(X_i, d_i)_{i \in \mathbf{N}}$ une suite d'espaces métriques et $X = \prod X_i$. On définit une distance sur X par la formule

$$d((x_i), (y_i)) = \sum_{i \in \mathbf{N}} \min(2^{-i}, d_i(x_i, y_i))$$

Alors d est une distance sur X. Nous allons maintenant vérifier que la topologie produit sur X (notée $\mathcal{T}$) coïncide avec la topologie associée à d (notée $\mathcal{T}_d$). Pour cela, il suffit de vérifier que tout $A \in \mathcal{T}$ est un voisinage au sens de $\mathcal{T}_d$ de chacun de ses éléments, et réciproquement.

Fin cours # 1 du 15 janvier

— Soit $A \in \mathcal{T}$ et $x = (x_i) \in A$. Il existe un cylindre ouvert C tel que $x \in C \subset A$. On peut supposer que C est de la forme

$$C = B(x_1, \varepsilon_1) \times \cdots \times B(x_p, \varepsilon_p) \times \prod_{i>p} X_i.$$

Si on pose $\varepsilon = \min(\varepsilon_1, \ldots, \varepsilon_p, 2^{-p}) > 0$, on a $B(x, \varepsilon) \subset C$. (En effet, si un point $y = (y_i)$ de X vérifie $d(x, y) < \varepsilon$, alors pour tout i on a $\min(2^{-i}, d_i(x_i, y_i)) < \varepsilon$. Puisque $\varepsilon \leqslant 2^{-p}$, si $i < p$ cela implique $d_i(x_i, y_i) < \varepsilon_i$ et donc $y_i \in B(x_i, \varepsilon_i)$.) Ainsi A est un voisinage au sens de $\mathcal{T}_d$ de chacun de ses points.

— Soit $A \in \mathcal{T}_d$ et $x = (x_i) \in A$. Il existe $\varepsilon > 0$ tel que $B(x, \varepsilon) \subset A$. Pour p assez grand, on a $\sum_{i>p} 2^{-i} < \varepsilon$ et on peut trouver $\varepsilon_1, \ldots, \varepsilon_p > 0$ tels que $\varepsilon_1 + \cdots + \varepsilon_p + \sum_{i>p} 2^{-i} < \varepsilon$, ce qui fait que le cylindre ouvert

$$C = B(x_1, \varepsilon_1) \times \cdots \times B(x_p, \varepsilon_p) \times \prod_{i>p} X_i$$

est inclus dans $B(x, \varepsilon)$, donc dans A. Ainsi A est un voisinage au sens de $\mathcal{T}$ de chacun de ses points.

Remarque. La topologie produit est la topologie de la convergence simple. En effet, soit $(X_i, \mathcal{T}_i)$ une famille d'espaces topologiques et soit $\mathcal{T}$ la topologie produit sur $X = \prod X_i$. Soit $(x^n)_{n \in \mathbf{N}}$ une suite d'éléments de X ; pour tout n, l'élément x^n de X s'écrit $x^n = (x_i^n)_{i \in I}$. La suite (x^n) converge vers $x = (x_i) \in X$ si et seulement si, pour tout $i \in I$, la suite (x_i^n) converge vers x_i.

Théorème (Théorème de Tychonoff)**.** *Soit $(X_i, \mathcal{T}_i)_{i \in I}$ une famille d'espaces topologiques compacts. Alors l'espace $\prod X_i$ muni de la topologie produit est un espace topologique compact.*

La preuve du théorème de Tychonoff tel qu'énoncé ci-dessus nécessite le recours à l'axiome du choix. Pour simplifier, on ne donnera la preuve que dans le cas d'un produit dénombrable d'espaces métrisables.

Démonstration. Supposons que $I = \mathbf{N}$ et que pour tout i, la topologie $\mathcal{T}_i$ est associée à une distance d_i sur X_i. Soit $(x^n)_n$ une suite d'éléments de X ; pour chaque n on écrit $x^n = (x_i^n)_{i \in \mathbf{N}}$. Il suffit de montrer que $(x^n)_n$ admet une sous-suite $x_n^{\tau(n)}$ convergente pour la topologie produit, c'est-à-dire telle que, pour tout i, la suite $(x_i^{\tau(n)})_n$ converge. On utilise pour cela un argument classique d'extraction diagonale.

0. Puisque X_0 est compact, il existe une sous-suite $(x^{\sigma_0(n)})_n$ de $(x^n)_n$ telle que la suite $(x_0^{\sigma_0(n)})_n$ converge vers $x_0 \in X_0$.
1. Puisque X_1 est compact, il existe une sous-suite $(x^{\sigma_1(n)})_n$ de $(x^{\sigma_0(n)})_n$ telle que la suite $(x_1^{\sigma_1(n)})_n$ converge vers $x_1 \in X_1$. Remarquons que par le point précédent, on a également la convergence de $(x_0^{\sigma_1(n)})_n$ vers x_0.

$\vdots$

i. Pour tout entier i, on construit par récurrence une sous-suite $(x^{\sigma_i(n)})_n$ de $(x^{\sigma_{i-1}(n)})_n$ telle que la suite $(x_i^{\sigma_i(n)})_n$ converge vers $x_i \in X_i$. Remarquons que par les points précédents, on a également la convergence de $(x_j^{\sigma_i(n)})_n$ vers x_j pour tout $j < i$.

$\vdots$

L'idée finale est définir une extraction $\tau : \mathbf{N} \to \mathbf{N}$ par la formule $\tau(n) = \sigma_n(n)$. Pour tout $i \in \mathbf{N}$, la suite $(x_i^{\tau(n)})_n$ converge vers x_i puisque, à un nombre fini de termes près, c'est une sous-suite de $(x_i^{\sigma_i(n)})_n$.

Nous avons bien démontré que la suite $(x^{\tau-n})_n$ converge vers $x = (x_i)$ au sens de la topologie produit. $\square$

Chapitre 2

Le théorème de Baire et ses conséquences

2.1 Le théorème de Baire

On considère un espace métrique complet (X, d).

Théorème (Théorème de Baire)**.** *Soit* $(F_n)_n$ *une suite de fermés d'intérieur vides de* (X, d)*. Alors* $\bigcup F_n$ *est d'intérieur vide.*

Notons $\operatorname{diam}(A) = \sup_{x,y \in A} d(x, y)$ le *diamètre* d'une partie $A \subset X$. Pour démontrer le théorème de Baire, on va utiliser la propriété dite des fermés emboîtés : si B_n est une suite décroissante de fermés non vides tels que $\lim \operatorname{diam}(B_n) = 0$, alors $\bigcap B_n$ est non vide. (Preuve : prendre $x_n \in B_n$ est observer que la suite (x_n) est de Cauchy donc converge vers $x \in \bigcap B_n$).

Le théorème de Baire s'énonce souvent sous la forme complémentaire.

Théorème (Théorème de Baire)**.** *Soit* $(O_n)_n$ *une suite d'ouverts dense de* (X, d)*. Alors* $\bigcap O_n$ *est dense.*

Montrons-le sous cette forme, équivalente à la précédente en considérant $F_n = X \setminus O_n$. Il est élémentaire, dans tout espace topologique, que l'intersection de deux ouverts denses est un ouvert dense ; par récurrence, cela est vrai également pour l'intersection d'un nombre fini d'ouverts denses.

Démonstration. Il suffit de montrer que tout ouvert non vide Ω de X intersecte $\bigcap O_n$. Nous allons construire une suite décroissante de boules fermées $B_n = B_f(x_n, r_n)$ vérifiant $0 < r_n \leqslant 2^{-n}$ et $B_n \subset O_n \cap \Omega$. On note également B'_n la boule ouverte $B(x_n, r_n)$.

0. Puisque O_0 est dense, l'ouvert $O_0 \cap \Omega$ est non vide et contient donc une boule $B_0 = B_f(x_0, r_0)$ avec $0 < r_0 \leqslant 1$.
1. Puisque O_1 est dense, l'ouvert $B'_0 \cap O_1 \cap \Omega$ est non vide et contient donc une boule $B_1 = B_f(x_1, r_1)$ avec $0 < r_1 \leqslant 1/2$.

$\vdots$

n. Puisque O_n est dense, l'ouvert $B'_{n-1} \cap O_n \cap \Omega$ est non vide et contient donc une boule $B_n = B_f(x_n, r_n)$ avec $0 < r_n \leqslant 1/2^n$.

$\vdots$

Puisque $\operatorname{diam}(B_n) \leqslant 2r_n$, la suite (B_n) est une suite décroissante de fermés dont le diamètre tend vers 0 ; son intersection contient donc un (unique) point x. Pour tout n, on a $x \in B_n \subset O_n \cap \Omega$, et donc Ω intersecte $\bigcap O_n$. □

Fixons quelques points de terminologie

— On dit qu'une partie $A \subset X$ est *maigre* si elle est incluse dans une réunion dénombrable de fermés d'intérieur vide.

— Le théorème de Baire dit qu'une partie maigre est d'intérieur vide ; il implique aussi qu'une réunion dénombrable de parties maigres est encore une partie maigre.

— On dit qu'une partie $A \subset X$ est *comaigre* si $X \setminus A$ est maigre. Le théorème de Baire dit qu'une partie comaigre est dense.

— On dit qu'une partie $A \subset X$ est un F_σ si on peut l'écrire comme réunion dénombrable de fermés.

— On dit qu'une partie $A \subset X$ est un G_δ si on peut l'écrire comme intersection dénombrable d'ouverts.

Exemple. Dans $\mathbf{R}$, toute partie dénombrable est maigre.

Exercice. Trouver une partition $\mathbf{R} = A \cup B$ où A est maigre et B est de mesure nulle.

2.2 Fonctions de première classe

On dit qu'une fonction $f : \mathbf{R} \to \mathbf{R}$ est de première classe si est elle limite simple d'une suite de fonctions continues.

Une fonction continue est de première classe. La fonction indicatrice d'un intervalle de $\mathbf{R}$ est de première classe.

Théorème. *Soit f une fonction de première classe. Alors l'ensemble des points de continuité de f est comaigre (et donc dense).*

Montrons d'abord le lemme suivant.

Lemme. *Soit f une fonction de première classe. Alors, pour tout fermé F de $\mathbf{R}$, l'ensemble $f^{-1}(F)$ est un G_δ.*

De manière équivalente, le lemme dit que pour tout ouvert O de $\mathbf{R}$, l'ensemble $f^{-1}(O)$ est un F_σ.

Démonstration. Écrivons $f = \lim f_n$ avec (f_n) continues. On écrit F comme un G_δ en écrivant $F = \bigcap_{k \in \mathbf{N}} O_k$ où $O_k = \{x \in \mathbf{R} \ : \ d(x,F) < 2^{-k}\}$. La conclusion du lemme découle de la formule

$$f^{-1}(F) = \bigcap_{k \in \mathbf{N}} \bigcup_{n \geqslant k} f_n^{-1}(O_k).$$

Vérifions cette formule en utilisant la remarque que $d(\cdot, F)$ s'annule exactement sur F, puisque c'est un fermé

$\boxed{\subset}$ Si $x \in f^{-1}(F)$, alors pour tout k, $f(x) \in O_k$ et donc $f_n(x) \in O_k$ pour n assez grand.

$\boxed{\supset}$ Si x appartient au membre de droite, alors pour tout k il existe n_k tel que $f_{n_k}(x) \in O_k$, donc $d(f_{n_k}(x), F) < 2^{-k}$. On prenant la limite $k \to \infty$ on obtient $d(f(x), F) = 0$ et donc $f(x) \in F$.

□

Preuve du théorème. Soit (V_n) une base dénombrable d'ouverts de $\mathbf{R}$, c'est-à-dire que tout ouvert de $\mathbf{R}$ s'écrit comme réunion d'un sous-ensemble de la famille $(V_n)_{n\in\mathbf{N}}$. (Par exemple, considérer l'ensemble des intervalles à extrémités rationnelles). On remarque que f est continue en x si et seulement si, pour tout V_n contenant $f(x)$, l'ensemble $f^{-1}(V_n)$ est un voisinage de x.

Soit C l'ensemble des points de continuité de f. Alors

$$x \in C \iff \forall n, \quad f(x) \in V_n \Longrightarrow f^{-1}(V_n) \text{ est un voisinage de } x$$

et donc

$$\mathbf{R} \setminus C = \bigcup_{n\in\mathbf{N}} \underbrace{f^{-1}(V_n)}_{F_\sigma} \setminus \text{int}[f^{-1}(V_n)]$$

et on a bien écrit $\mathbf{R} \setminus C$ comme réunion dénombrables de fermés d'intérieurs vide (si X est un F_σ, alors $X \setminus \text{int}(X)$ est un F_σ d'intérieur vide). $\square$

Fin cours # 2 du 16 janvier

Chapitre 3

Les grands théorèmes sur les espaces de BANACH

Il s'agit de plusieurs théorèmes qui sont des conséquences de la complétude et dont la preuve utilise le théorème de BAIRE.

3.1 Le théorème de BANACH–STEINHAUS

Dans le cours tous les espaces de BANACH seront sur $\mathbf{R}$ ou $\mathbf{C}$. Si X et Y sont des espaces de BANACH, on note $\mathcal{L}(X,Y)$ l'espace vectoriel des applications linéaires continues de X dans Y. Rappelons qu'une application linéaire $T : X \to Y$ est continue si et seulement si elle est bornée sur la boule-unité de X. On définit sa norme d'opérateur par la formule

$$\begin{aligned}\|T\|_{op} &= \sup\{\|Tx\|_Y \;:\; \|x\|_X \leqslant 1\}.\\ &= \inf\{C > 0 \;:\; \forall x \in X,\; \|Tx\|_Y \leqslant C\|x\|_X\}.\end{aligned}$$

L'espace $\mathcal{L}(X,Y)$ muni de la norme d'opérateur, est alors un espace de BANACH.

Dans le cas particulier ou $Y = \mathbf{R}$ (ou $\mathbf{C}$ dans le cas complexe), on trouve l'espace dual de X, noté X^*. La norme sur $X^* = \mathcal{L}(X,\mathbf{R})$ est appelée norme duale (c'est un cas particulier de la norme d'opérateur) ; si $f \in X^*$ alors

$$\|f\|_{X^*} = \sup_{x\in X, \|x\|_X \leqslant 1} |f(x)|.$$

Théorème (Théorème de BANACH–STEINHAUS)**.** *Soient X et Y des espaces de BANACH et $\Phi \subset \mathcal{L}(X,Y)$. Les deux propriété suivantes sont équivalentes*

1. *Φ est borné, c'est-à-dire*
$$\sup_{T\in\Phi} \|T\|_{op} < +\infty$$
2. *Φ est ponctuellement borné, c'est-à-dire*
$$\forall x \in X,\; \sup_{T\in\Phi} \|Tx\|_Y < +\infty$$

De plus, si Φ n'est pas borné, l'ensemble $\{x \in X \;:\; \sup_{T\in\Phi} \|Tx\|_Y = +\infty\}$ est comaigre dans X.

En anglais, ce théorème s'appelle «uniform boundedness principle».

Démonstration. L'implication 1 $\Longrightarrow$ 2 est évidente.

Supposons Φ non borné et considérons pour un entier n

$$F_n = \{x \in X \ : \ \forall T \in \Phi, \|Tx\|_Y \leqslant n\}.$$

Alors F_n est un fermé de X (intersection de fermés...) qui est d'intérieur vide [en effet, si F_n contient une boule $B(x,\varepsilon)$, étant symétrique et convexe il contiendrait $B(0,\varepsilon)$, d'où pour tout $T \in \Phi$ l'implication $\|y\|_X \leqslant \varepsilon \Longrightarrow \|Tx\|_Y \leqslant n$ et donc $\|T\|_{op} \leqslant n/\varepsilon$, contredisant le caractère borné].

Par le théorème de Baire, $\bigcup F_n$ est maigre dans X, donc non égal à X, ce qui montre que Φ n'est pas ponctuellement borné puisque

$$\bigcup_{n\in\mathbf{N}} F_n = \{x \in X \ : \ \sup_{T\in\Phi} \|Tx\|_Y = +\infty\}.$$

On a également la dernière partie du théorème. $\square$

On va donner une application du théorème de Banach–Steinhaus aux séries de Fourier. On note X l'espace des fonctions continues et 2π-périodiques, muni de $\|\cdot\|_\infty$. C'est un espace de Banach. Pour $f \in X$ et $k \in \mathbf{Z}$, on considère les coefficients de Fourier

$$\hat{f}(k) = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(t)e^{-ikt}\,\mathrm{d}t$$

et si $n \in \mathbf{N}$ et $t \in \mathbf{R}$, la série de Fourier de f

$$(S_n f)(t) = \sum_{k=-n}^{n} \hat{f}(k)e^{ikt}.$$

Rappelons deux résultats classiques sur la convergence des séries de Fourier

— Théorème de Dirichlet : si f est de classe C^1, alors $(S_n f)$ converge uniformément vers f

— Théorème de Parseval : comme $X \subset L^2(0,2\pi)$, on a convergence de $(S_n f)$ vers f dans $L^2(0,2\pi)$.

Il n'est pas vrai que $S_n f$ converge simplement vers f (un théorème très difficile est que $S_n f$ converge vers f presque partout ! Notons qu'il découle du théorème de Parseval qu'une sous-suite de $(S_n f)$ converge presque partout). En effet, on a

Théorème. *L'ensemble*

$$\{f \in X \ : \ \lim_{n\to\infty} (S_n f)(0) = f(0)\}$$

est maigre dans X.

Démonstration. On introduit le noyau de Dirichlet, pour $t \in \mathbf{R}$ (et $t \notin 2\pi\mathbf{Z}$ après la seconde égalité)

$$\begin{aligned} D_n(t) &= \sum_{k=-n}^{n} e^{ikt} \\ &= e^{-int}\frac{1-e^{(2n+1)it}}{1-e^{it}} \\ &= \frac{e^{(n+\frac{1}{2})it} - e^{-(n+\frac{1}{2})it}}{e^{it/2} - e^{-it/2}} \\ &= \frac{\sin(n+\frac{1}{2})t}{\sin\frac{t}{2}} \end{aligned}$$

et on remarque que $S_n f = f \star D_n$, c'est-à-dire

$$S_n f(t) = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(s) D_n(t-s)\,\mathrm{d}s.$$

Définissons une forme linéaire Λ_n sur X par la formule $\Lambda_n(f) = S_n f(0)$.

Lemme. *Pour tout n, $\Lambda_n \in X^*$. De plus, on a $\lim_{n\to\infty} \|\Lambda_n\|_{X^*} = +\infty$.*

Admettons pour l'instant le lemme. Puisque la suite (Λ_n) n'est pas bornée dans X^*, l'ensemble

$$\{f \in X \ : \ \text{la suite } (S_n f(0)) \text{ est bornée}\}$$

est comaigre dans X, et le théorème en découle. □

Il reste à prouver le lemme. Il est immédiate que Λ_n est continue et que

$$\|\Lambda_n\|_{X^*} \leqslant \frac{1}{2\pi} \int_{-\pi}^{\pi} |D_n(t)|\,\mathrm{d}t.$$

Montrons qu'il y a égalité. Pour tout $\varepsilon > 0$, posons $f_\varepsilon(t) = \frac{D_n(t)}{|D_n(t)|+\varepsilon}$ qui est un élément de X de norme $\leqslant 1$. On a donc

$$\|\Lambda_n\|_{X*} \geqslant \Lambda_n(f_\varepsilon) = (S_n f_\varepsilon)(0) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{D_n(t)^2}{|D_n(t)|+\varepsilon}\,\mathrm{d}t$$

et on conclut par convergence dominée en faisant tendre ε vers 0. Finalement, on calcule

$$\begin{aligned} \int_{-\pi}^{\pi} |D_n(t)|\,\mathrm{d}t &\geqslant 4 \int_0^{\pi} |\sin(n+1/2)t|\,\frac{\mathrm{d}t}{t} \\ &= 4 \int_0^{(n+1/2)\pi} |\sin t| \frac{\mathrm{d}t}{t} \end{aligned}$$

qui tend vers l'infini quand n tend vers l'infini.

3.2 Le théorème de l'application ouverte

Soient X, Y deux espaces topologiques. On dit qu'une application $f : X \to Y$ est *ouverte* si pour tout ouvert U de X, l'image $f(U)$ est un ouvert de Y.

Observons que si T est une application linéaire entre espaces de Banach et si $r > 0$, alors pour tout $\lambda > 0$ la condition

$$T(B(0,1)) \supset B(0,r)$$

équivaut à la condition

$$T(B(0,\lambda)) \supset B(0,\lambda r).$$

Théorème (Théorème de l'application ouverte)**.** *Soient X, Y deux espaces de Banach et $T \in \mathcal{L}(X,Y)$ une application linéaire surjective continue. Alors T est ouverte.*

En particulier, il existe $r > 0$ telle que $T(B(0,1)) \supset B(0,r)$.

Réciproquement, une application linéaire ouverte entre espaces normés est surjective puisque son image est un sous-espace ouvert.

Démonstration. Il suffit de prouver le «en particulier». En effet, pour tout $x \in X$ et $\varepsilon > 0$, on a alors
$$B(x,\varepsilon) = x + \varepsilon B(0,1)$$
et, T étant linéaire,
$$T(B(x,\varepsilon)) = T(x) + \varepsilon T(B(0,1)) \supset T(x) + \varepsilon B(0,r) = B(T(x),\varepsilon r).$$

Prouvons le «en particulier». Posons $F_n = \overline{T(B(0,n))}$. Puisque T est surjective, $Y = \bigcup_{n\in\mathbf{N}} F_n$. Par le théorème de BAIRE, il existe donc n tel que F_n est d'intérieur non vide, donc contient une boule $B(y,\varepsilon)$. Puisque F_n est symétrique et convexe, il contient la boule $B(0,\varepsilon)$. Par homogénéité, on a alors pour $\lambda = n/\varepsilon > 0$
$$B(0,1) \subset \overline{T(B(0,\lambda))}.$$
Nous allons prouver que cela implique $B(0,1) \subset T(B(0,2\lambda)))$. Soit $z \in B(0,1)$.

— Il existe $x_1 \in B(0,\lambda)$ tel que $\|z - Tx_1\| < 1/2$.
— Il existe $x_2 \in B(0,\lambda/2)$ tel que $\|z - Tx_1 - Tx_2\| < 1/4$.
— $\vdots$
— Par récurrence, pour tout $n \geqslant 1$, il existe $x_n \in B(0,\lambda/2^{n-1})$ tel que $\|z - Tx_1 - Tx_2 - \cdots - Tx_n\| < 1/2^n$.
— $\vdots$

La série $\sum x_n$ converge normalement et (T étant continue) sa somme x vérifie $z = Tx$. Puisque $\|x\| < 2\lambda$, on a bien montré $B(0,1) \subset T(B(0,2\lambda)))$. Ceci équivaut à $B(0,r) \subset T(B(0,1)))$ pour $r = 1/2\lambda$. □

Corollaire (Théorème d'isomorphisme de BANACH). *Soient X, Y des espaces de BANACH et $T \in \mathcal{L}(X,Y)$ une application linéaire continue bijective. Alors T^{-1} est une application linéaire continue de Y dans X*

Démonstration. La linéarité de T^{-1} est un résultat élémentaire d'algèbre linéaire. Puisque T est surjective, par le théorème de l'application ouverte, elle est ouverte, ce qui revient à dire que T^{-1} est continue. □

Corollaire. *Soit X un espace vectoriel muni de deux normes $\|\cdot\|_1$ et $\|\cdot\|_2$ telles que $(X,\|\cdot\|_1)$ et $(X,\|\cdot\|_2)$ soient complets. S'il existe un réel C tel que $\|\cdot\|_1 \leqslant C\|\cdot\|_2$, alors il existe un réel C' tel que $\|\cdot\|_2 \leqslant C'\|\cdot\|_1$.*

Démonstration. On applique le théorème d'isomorphisme de BANACH à l'application id : $(X,\|\cdot\|_2) \to (X,\|\cdot\|_1)$. □

3.3 Le théorème du graphe fermé

Si $T : X \to Y$ est une application entre des ensembles X et Y, son graphe est $G(T) \subset X \times Y$ est
$$G(T) = \{(x,Tx) \;:\; x \in X\}.$$

On a le résultat suivant.

Proposition. *Soient X et Y des espaces topologiques séparés et $T : X \to Y$ continue. Alors $G(T)$ est fermé (dans $X \times Y$).*

Démonstration. Montrons que $(X \times Y) \setminus G(T)$ est ouvert. Soit $(x, y) \notin G(T)$. Alors $y \neq T(x)$ donc il existe des ouverts disjoints U et V tels que $y \in U$ et $T(x) \in V$. Posons $W = T^{-1}(V)$; c'est un ouvert contenant x. Puisque $T(W) \subset V$, l'ensemble $W \times U$ est un ouvert de $X \times Y$ disjoint de $G(T)$. Nous avons bien montré que $(X \times Y) \setminus G(T)$ est ouvert. □

La réciproque de la proposition n'est pas vraie en général : par exemple, la fonction caractéristique d'un intervalle non trivial est une fonction non continue de $\mathbf{R}$ dans $\mathbf{R}$ dont le graphe est fermé.

Si X et Y sont des espaces vectoriels et si T est linéaire, alors $G(T)$ est un sous-espace vectoriel de $X \times Y$.

Théorème (Théorème du graphe fermé)**.** *Soient X et Y des espaces de* Banach *et $T : X \to Y$ une application linéaire. Alors T est continue si et seulement si $G(T)$ est fermé.*

Fin cours # 3 du 23 janvier

Démonstration. Une partie est couverte par la proposition. Pour l'autre partie, supposons $G(T)$ fermé. Considérons sur X les normes

$$\|x\|_1 = \|x\|_X, \quad \|x\|_2 = \|x\|_X + \|Tx\|_Y$$

L'espace $(X, \|\cdot\|_2)$ est complet. En effet, soit (x_n) une suite de Cauchy pour $\|\cdot\|_2$. Alors les suites de Cauchy (x_n) (pour $\|\cdot\|_X$) et (Tx_n) (pour $\|\cdot\|_Y$) convergent ; notons x et y leurs limites respectives. Puisque (x_n, Tx_n) est une suite de $G(T)$ qui converge vers (x, y), on a $(x, y) \in G(T)$, c'est-à-dire $y = Tx$. Il s'ensuit que (x_n) converge x pour $\|\cdot\|_2$.

Puisque $\|\cdot\|_1 \leqslant \|\cdot\|_2$, par le corollaire du théorème d'isomorphisme de Banach, il existe une constante C telle que $\|\cdot\|_2 \leqslant C\|\cdot\|_1$. On a donc l'inégalité $\|Tx\|_Y \leqslant C\|x\|_X$ qui montre que T est continu. □

Chapitre 4

Les espaces de Banach classiques

Dans l'ensemble du cours, on va considérer des espaces de Banach sur **R** ou **C**. Pour simplifier, on présente les preuves dans le cadre réel. Les résultats s'étendent au cadre complexe mais cela n'est pas forcément immédiat.

4.1 Les espaces de Hilbert

Un espace de Hilbert est un espace de Banach $(X, \|\cdot\|)$ tel qu'il existe un produit scalaire $\langle\cdot,\cdot\rangle$ vérifiant la relation

$$\|x\| = \sqrt{\langle x, x\rangle}$$

pour tout $x \in X$. Remarquons que dans ce cas, le produit scalaire est uniquement déterminé par la norme, via la formule de polarisation

$$4\langle x, y\rangle = \|x+y\|^2 + \|x-y\|^2.$$

Une application linéaire $T : X \to Y$ entre espaces de Banach est isométrique si elle vérifie $\|Tx\| = \|x\|$. Si de plus T est surjective, on peut alors identifier les espaces de Banach X et Y par l'intermédiaire de T. On rappelle le résultat suivant.

Théorème (Théorème de Riesz–Fréchet)**.** *Soit X un espace de Hilbert. L'application de X dans X^* qui à x associe*

$$y \mapsto \langle x, y\rangle$$

est isométrique et surjective.

On peut donc identifier canoniquement un espace de Hilbert à son dual.

4.2 Les espaces $L^p(\mu)$

4.2.1 Compléments de théorie de la mesure : le théorème de Radon–Nikodym

Soit $(\Omega, \mathcal{F})$ un espace mesurable et μ, ν deux mesures sur $(\Omega, \mathcal{F})$. On dit que ν est *absolument continue* par rapport à μ et on écrit $\nu \ll \mu$ si tout ensemble $A \in \mathcal{F}$ tel que $\mu(A) = 0$ vérifie $\nu(A) = 0$.

Soit $f \in L^1(\mu)$ une fonction positive. La formule

$$\nu(A) = \int_A f \,\mathrm{d}\mu \tag{4.1}$$

définit une mesure ν sur $(\Omega, \mathcal{F})$ (pour la σ-additivité, utiliser le théorème de convergence monotone) qui est absolument continue par rapport à μ. Lorsque la relation (4.1) est vérifiée, on écrira « $\mathrm{d}\nu = f\,\mathrm{d}\mu$ ». Cette notation est motivée par la formule suivante : pour toute fonction mesurable positive (resp. ν-intégrable) h, on a

$$\int h\,\mathrm{d}\nu = \int hf\,\mathrm{d}\mu.$$

Si h est étagée, cela se déduit par linéarité de (4.1) ; pour le cas général, utiliser le théorème de convergence monotone (resp. dominée).

Le théorème de Radon–Nikodym affirme que les mesures absolument continues sont toujours de cette forme. Pour simplifier, on l'énonce dans le cas de mesures finies ; il est également vrai pour des mesures σ-finies..

Théorème (Théorème de Radon–Nikodym)**.** *Soit $(\Omega, \mathcal{F})$ un espace mesurable et μ, ν deux mesures finies sur $(\Omega, \mathcal{F})$ telles que $\nu \ll \mu$. Il existe une unique fonction positive $f \in L^1(\mu)$ telle que $\mathrm{d}\nu = f\,\mathrm{d}\mu$.*

On va déduire ce résultat d'un théorème plus général.

Théorème (Théorème de décomposition de Lebesgue)**.** *Soit $(\Omega, \mathcal{F})$ un espace mesuré et μ, ν deux mesures finies sur $(\Omega, \mathcal{F})$. Il existe une fonction positive $f \in L^1(\mu)$ et un ensemble $B \in \mathcal{F}$ avec $\mu(B) = 0$, tels que, pour tout $A \in \mathcal{F}$*

$$\nu(A) = \int_A f\,\mathrm{d}\mu + \nu(A \cap B).$$

Si on applique le théorème décomposition de Lebesgue sous l'hypothèse $\nu \ll \mu$, alors puisque $\mu(B) = 0$ on a $\nu(B) = 0$ et donc $\nu(A \cap B) = 0$ pour tout $A \in \mathcal{F}$. On obtient alors la partie «existence» de la conclusion du théorème de Radon–Nidokym. Pour l'unicité : soit $g \in L^1(\mu)$ telle que $\mathrm{d}\nu = f\mathrm{d}\mu = g\mathrm{d}\mu$. Si on pose $A = \{x \in \Omega \ : \ g(x) > f(x)\}$, alors

$$\int_A (g - f)\,\mathrm{d}\mu = \nu(A) - \nu(A) = 0$$

et donc $g \leqslant f$ μ-p.p. ; on montre de même que $g \geqslant f$ μ-p.p..

Démonstration du théorème de décomposition de Lebesgue. Considérons la mesure $\pi = \mu + \nu$. Définissons une forme linéaire L sur $L^2(\pi)$ par la formule $L(g) = \int g\,\mathrm{d}\nu$. On a par l'inégalité de Cauchy–Schwarz

$$|L(g)| \leqslant \nu(\Omega)^{1/2} \left(\int |g|^2\,\mathrm{d}\nu \right)^{1/2} \leqslant \nu(\Omega)^{1/2} \|g\|_{L^2(\pi)},$$

ce qui montre que L est continue. Par le théorème de Riesz–Fréchet, il existe $h \in L^2(\pi)$ telle que pour tout $g \in L^2(\pi)$ on ait $L(g) = \langle g, h \rangle$, c'est à dire

$$\int_\Omega g\,\mathrm{d}\nu = \int_\Omega gh\,\mathrm{d}\pi = \int_\Omega gh\,\mathrm{d}\mu + \int_\Omega gh\,\mathrm{d}\nu,$$

ce qui se réécrit en

$$\int_\Omega g(1-h)\,\mathrm{d}\nu = \int_\Omega gh\,\mathrm{d}\mu. \tag{4.2}$$

L'idée de la preuve est que si $\mathrm{d}\nu = f\mathrm{d}\mu$, alors $\mathrm{d}\pi = \mathrm{d}\mu + \mathrm{d}\nu = (1+f)\mathrm{d}\mu$ donc $\mathrm{d}\mu = \frac{1}{1+f}\mathrm{d}\pi$ puis $\mathrm{d}\nu = \mathrm{d}\pi - \mathrm{d}\mu = \frac{f}{1+f}\mathrm{d}\pi$: on s'attend donc à ce que f et h soient reliés par la formule $h = \frac{f}{1+f}$. Nous allons vérifier cela rigoureusement.

Lemme. *L'inégalité $h \geqslant 0$ est vraie ν-presque partout et μ-presque partout. L'inégalité $h < 1$ est vraie μ-presque partout.*

Démonstration. En prenant $g = \mathbf{1}_A$ pour $A \in \mathcal{F}$, on a

$$\nu(A) = \int_A h \, \mathrm{d}\mu + \int_A h \, \mathrm{d}\nu.$$

Posons $N = \{h < 0\}$; on a $\nu(N) = \int_N h \, \mathrm{d}\mu + \int_N h \, \mathrm{d}\nu \leqslant 0$, d'où on tire $\nu(N) = 0$ puis $\int_N h \, \mathrm{d}\mu = 0$ donc $\mu(N) = 0$. Posons $B = \{h \geqslant 1\}$; on a $\nu(B) = \int_B h \, \mathrm{d}\mu + \int_B h \, \mathrm{d}\nu \geqslant \mu(B) + \nu(B)$, d'où on dire $\mu(B) = 0$. □

Soit $G_n = \{0 \leqslant h < 1 - 1/n\}$ et posons $G = \bigcup G_n = \Omega \setminus (N \cup B) = \{0 \leqslant h < 1\}$. On définit une fonction mesurable $f : \Omega \to \mathbf{R}$ en posant $f(x) = \frac{h(x)}{1-h(x)}$ si $x \in G$ et $f(x) = 0$ si $x \notin G$. On a, en utilisant (4.2) pour $g = \frac{\mathbf{1}_{A \cap G_n}}{1-h}$ puis le théorème de convergence monotone

$$\nu(A \cap G) = \lim \nu(A \cap G_n) = \lim \int_\Omega \frac{1-h}{1-h} \mathbf{1}_{A \cap G_n} \, \mathrm{d}\nu = \lim \int_\Omega f \mathbf{1}_{A \cap G_n} \, \mathrm{d}\mu = \int_\Omega f \mathbf{1}_A \, \mathrm{d}\mu$$

et le résultat voulu en découle puisque $\nu(A) = \nu(A \cap G) + \nu(A \cap B)$. □

4.2.2 La dualité des espaces $L^p(\mu)$

Soit $(\Omega, \mathcal{F}, \mu)$ un espace mesuré. Pour $1 \leqslant p < +\infty$, on note $\mathcal{L}^p(\mu)$ l'ensemble des fonctions mesurables $f : \Omega \to \mathbf{R}$ telles que $\int_\Omega |f|^p \, \mathrm{d}\mu < +\infty$, et $L^p(\mu)$ l'ensemble des classes d'équivalences de fonctions de $\mathcal{L}^p(\Omega, \mu)$ pour la relation d'équivalence μ-presque partout. Muni de la norme

$$\|f\|_{L^p} = \left(\int_\Omega |f|^p \, \mathrm{d}\mu \right)^{1/p},$$

l'espace $L^p(\mu)$ est un espace de Banach. L'espace $L^2(\mu)$ est un espace de Hilbert.

On définit $L^\infty(\mu)$ comme l'ensemble des classes d'équivalences de fonctions mesurables dont un représentant est borné. Muni de la norme

$$\|f\|_{L^\infty} = \inf\{M > 0 \ : \ |f| \leqslant M \ \mu\text{-p.p.}\},$$

l'espace $L^\infty(\mu)$ est un espace de Banach.

Une fonction *étagée* est une combinaison linéaire de fonctions indicatrices d'éléments de $\mathcal{F}$. Pour tout $p \in [1, \infty]$, les fonctions étagées sont denses dans $L^p(\Omega)$. On dit que deux nombres réels $p, q \in [1, \infty]$ sont *conjugués* si $1/p + 1/q = 1$.

Un espace mesuré $(\Omega, \mathcal{F}, \mu)$ est *σ-fini* si Ω peut s'écrire comme réunion dénombrable de parties mesurables de mesure finie.

Lemme. *Soient $(\Omega, \mathcal{F})$ un espace mesuré σ-fini et p, q des exposants conjugués. Pour toute fonction mesurable $g \in L^q(\mu)$, on a*

$$\|g\|_{L^q} = \sup_{\|h\|_{L^p} \leqslant 1} \int_\Omega gh \, \mathrm{d}\mu.$$

De plus, la borne supérieure ne change pas si on la restreint aux fonctions h étagées telles que $\|h\|_{L^p} \leqslant 1$.

L'hypothèse de σ-finitude peut être allégée mais il faut éviter certains cas très dégénérés : par exemple si $\Omega = \{x\}$ est un singleton et μ est la mesure telle que $\mu(\{x\}) = +\infty$, on a $L^p(\mu) = \{0\}$ si $1 \leqslant p < \infty$ mais $L^\infty(\mu) \neq \{0\}$, donc la conclusion du lemme est fausse quand $p = 1$.

Démonstration. L'inégalité $\geqslant$ découle de l'inégalité de Hölder. Si $1 < q < \infty$, on peut choisir h telle que $gh \geqslant 0$ et $|g|^q = |h|^p$; on a alors $gh = |g|^q = |h|^p$ et il y a égalité dans l'inégalité de Hölder, d'où $\leqslant$. Si $q = 1$ et $p = \infty$, c'est pareil en choisissant $h = \mathrm{signe}(g)$. Si $q = \infty$ et $p = 1$, étant donné $\varepsilon > 0$ on peut choisir une partie $A \in \mathcal{F}$ incluse dans $\{|g| \geqslant \|g\|_{L^\infty} - \varepsilon\}$ telle que $0 < \mu(A) < \infty$ (c'est là où on utilise l'hypothèse de σ-finitude) et poser $h = \mathrm{signe}(g)\mathbf{1}_A$; on obtient que le membre de droite est $\geqslant \|g\|_{L^\infty} - \varepsilon$ et il suffit de faire tendre ε vers 0.

Pour le dernier point, considérer une suite de fonctions étagées h_n telles que $|h_n| \leqslant |h|$ et $\lim h_n = h$, et appliquer le théorème de convergence dominée. □

Soit $g \in L^q(\mu)$. Le lemme dit que la forme linéaire $\ell_g : L^p(\mu) \to \mathbf{R}$ définie par

$$\ell_g(f) = \int_\Omega f g \, \mathrm{d}\mu$$

est continue et de norme $\|g\|_{L^q}$. Ainsi $g \mapsto \ell_g$ est une application linéaire isométrique de $L^q(\mu)$ dans $L^p(\mu)^*$. Nous allons démontrer que si l'espace mesuré $(\Omega, \mathcal{F}, \mu)$ est σ-fini et si $1 \leqslant p < \infty$, cette application est surjective. On peut alors identifier $L^p(\mu)^*$ avec $L^q(\mu)$.

Théorème. *Soient $(\Omega, \mathcal{F}, \mu)$ un espace mesuré σ-fini, $p \in [1, \infty[$ et q son exposant conjugué. Alors l'application $g \mapsto \ell_g$ de $L^q(\mu)$ dans $L^p(\mu)^*$ est surjective.*

Lorsque $1 < p < \infty$, ce résultat est vrai sans l'hypothèse de σ-finitude.

Démonstration. Traitons d'abord le cas où $\mu(\Omega) < \infty$ et d'une forme linéaire $\phi \in L^p(\mu)^*$ qui prend des valeurs positives sur les fonctions positives. Pour $A \in \mathcal{F}$, on pose

$$\nu(A) = \phi(\mathbf{1}_A).$$

Ceci définit une mesure (pour la σ-additivité : si (A_n) est une suite de parties deux à deux disjointes, par le théorème de convergence dominée la série $\sum \mathbf{1}_{A_n}$ converge vers $\mathbf{1}_{\bigcup A_n}$ dans $L^p(\mu)$— c'est là où l'hypothèse $p < \infty$ est utilisée). De plus, ν est absolument continue par rapport à μ. Par le théorème de Radon–Nikodym, il existe une fonction positive $g \in L^1(\mu)$ telle que $\mathrm{d}\nu = g \, \mathrm{d}\mu$. On a

$$\phi(\mathbf{1}_A) = \int_\Omega g \mathbf{1}_A \, \mathrm{d}\mu$$

puis par linéarité, pour toute fonction h étagée

$$\int_\Omega gh \, \mathrm{d}\mu = \phi(h) \leqslant \|\phi\|_{L^p(\mu)^*} \|h\|_{L^p(\mu)}.$$

Cette majoration implique que $g \in L^q(\mu)$ (par exemple : poser $g_n = \min(n, g)$ et utiliser le lemme pour déduire que $\|g_n\|_{L^q} \leqslant \|\phi\|_{L^p(\mu)^*}$, puis utiliser le théorème de convergence monotone pour obtenir $\|g\|_{L^q} \leqslant \|\phi\|_{L^p(\mu)^*}$). Puisque les formes linéaires continues ℓ_g et ϕ coïncident sur les fonctions étagées qui forment une partie dense de $L^p(\mu)$, elles sont égales.

Fin cours #4 du 30 janvier

Toujours dans le cas $\mu(\Omega) < \infty$, soit maintenant $\psi \in L^p(\mu)^*$ quelconque. Pour une fonction positive f de $L^p(\mu)$, on pose

$$\overline{\psi}(f) = \sup\{\psi(h) \ : \ h \text{ mesurable telle que } 0 \leqslant h \leqslant f\} \geqslant 0$$

puis pour une fonction quelconque f de $L^p(\mu)$ décomposée en $f = f^+ - f^-$

$$\phi(f) = \overline{\psi}(f^+) - \overline{\psi}(f^-).$$

Lemme. *L'application ϕ ainsi définie est une forme linéaire continue sur $L^p(\mu)$.*

Admettons le lemme. Puisque ϕ et $\phi - \psi$ sont des formes linéaires continues positives sur $L^p(\mu)$, elle sont de la forme $\phi = \ell_{g_1}$, $\phi - \psi = \ell_{g_2}$ pour g_1, g_2 dans $L^q(\mu)$. On a bien $\psi = \ell_{g_1 - g_2}$.

Preuve du lemme. Soit $f \in L^p(\mu)$. Puisque $(-f)^+ = f^-$ et $(-f)^- = f^+$, on a $\phi(-f) = -\phi(f)$. On vérifie sans difficulté que $\phi(\lambda f) = \lambda\phi(f)$ pour $\lambda > 0$.

Soit f_1, f_2 dans $L^p(\mu)$ positives. Montrons que

$$\overline{\psi}(f_1 + f_2) = \overline{\psi}(f_1) + \overline{\psi}(f_2).$$

$\boxed{\geqslant}$ Soit $\varepsilon > 0$ et h_i mesurables positives telle que $h_i \leqslant f_i$ et $\psi(h_i) \geqslant \overline{\psi}(f_i) - \varepsilon$. Alors $\overline{\psi}(f_1 + f_2) \geqslant \psi(h_1 + h_2) = \psi(h_1) + \psi(h_2) \geqslant \overline{\psi}(f_1) + \overline{\psi}(f_2) - 2\varepsilon$.

$\boxed{\leqslant}$ Soit $\varepsilon > 0$ et h mesurable positive telle que $h \leqslant f_1 + f_2$ et $\psi(h) \geqslant \overline{\psi}(f_1 + f_2) - \varepsilon$. Considérons les fonctions mesurables positives $h_1 = \min(f_1, h)$ et $h_2 = h - h - 1 = \max(h - f_1, 0)$. Alors $\overline{\psi}(f_1 + f_2) \leqslant \psi(h) + \varepsilon = \psi(h_1) + \psi(h_2) + \varepsilon \leqslant \overline{\psi}(f_1) + \overline{\psi}(f_2) + \varepsilon$.

Si f et g sont des fonctions quelconques de $L^p(\mu)$, on a

$$f + g = (f + g)^+ - (f + g)^- = f^+ - f^- + g^+ - g^-$$

donc

$$(f + g)^+ + f^- + g^- = (f + g)^- + f^+ + g^+.$$

On déduit du cas positif que

$$\overline{\psi}((f + g)^+) + \overline{\psi}(f^-) + \overline{\psi}(g^-) = \overline{\psi}((f + g)^-) + \overline{\psi}(f^+) + \overline{\psi}(g^+)$$

et donc $\phi(f + g) = \phi(f) + \phi(g)$.

Il reste enfin à voir que ϕ est continue. Tout d'abord, si $0 \leqslant h \leqslant f$, alors $\|h\|_{L^p} \leqslant \|f\|_{L^p}$; on en déduit que $|\overline{\psi}(f)| \leqslant \|\psi\| \cdot \|f\|_{L^p}$. Pour une fonction $f \in L^p(\Omega)$ quelconque, on écrit

$$|\phi(f)| \leqslant |\overline{\psi}(f^+)| + |\overline{\psi}(f^-)| \leqslant \|\psi\| \, \|f^+\|_{L^p} + \|\psi\| \, \|f^-\|_{L^p} \leqslant 2 \, \|\psi\| \, \|f\|_{L^p},$$

ce qui montre que ϕ est continue. □

Enfin, le cas d'un espace mesuré σ-fini se déduit du cas fini par un changement de mesure (cf. TD). □

4.3 Les espaces ℓ_p

Pour $1 \leqslant p \leqslant \infty$, on note ℓ_p l'espace $L^p(\mu)$ où μ est la mesure de comptage sur $\mathbf{N}$. De manière plus explicite, la norme ℓ_p d'une suite $x = (x_n)$ de réels est

$$\|x\|_p = \left(\sum_{n \in \mathbf{N}} |x_n|^p \right)^{1/p}.$$

$$\|x\|_\infty = \sup_{n \in \mathbf{N}} |x_n|$$

et ℓ_p est l'espace des suites x telles que $\|x\|_p < \infty$.

Comme cas particulier du cas des espaces L^p, lorsque p et q sont conjugués avec $1 \leqslant p < \infty$, on peut identifier les espaces $(\ell_p)^*$ et ℓ_q. On introduit également c_0 comme l'ensemble des suites qui convergent vers 0. C'est un sous-espace fermé de ℓ_∞. De la même manière, on peut identifier les espaces $(c_0)^*$ et ℓ_1.

4.4 Les espaces $C(K)$

Soit K un espace topologique compact et $C(K)$ l'espace des fonctions continues de K dans $\mathbf{R}$. C'est un espace de Banach pour la norme

$$\|f\| = \max_{x \in K} |f(x)|.$$

Si μ est une mesure borélienne sur K, alors l'application $I_\mu : C(K) \to \mathbf{R}$ définie par

$$I_\mu(f) = \int_K f \, \mathrm{d}\mu$$

est une forme linéaire continue sur $C(K)$. Réciproquement, on a le résultat suivant

Théorème (Théorème de représentation de Riesz)**.** *Soit K un espace topologique compact.*

1. *Si $\varphi \in C(K)^*$ prend des valeurs positives sur les fonctions positives, alors il existe une mesure borélienne finie μ sur K telle que $\phi = I_\mu$.*
2. *Si $\varphi \in C(K)^*$, il existe deux mesures boréliennes finies μ et ν sur K telles que $\phi = I_\mu - I_\nu$.*

La preuve est longue (voir le devoir maison) mais la récompense est de taille. En effet, soit I la fonction définie sur $C([0,1])$ par

$$I(f) = \int_0^1 f(x) \, \mathrm{d}x$$

au sens de l'intégration de Riemann. Comme c'est une forme linéaire continue positive sur $C([0,1])$, on a $I = I_\mu$ pour une mesure borélienne μ sur $[0,1]$, qui est nécessairement la mesure de Lebesgue ! Le théorème de représentation de Riesz permet donc de construire la mesure de Lebesgue.

Chapitre 5

Le théorème de Hahn–Banach

5.1 Le théorème de prolongement de Hahn–Banach

Commençons par donner un énoncé général d'algèbre linéaire.

Théorème. *Soit X un $\mathbf{K}$-espace vectoriel, $M \subset X$ un sous-espace vectoriel et $f : M \to \mathbf{K}$ une forme linéaire. Alors il existe une forme linéaire $g : X \to \mathbf{K}$ qui prolonge f, c'est-à-dire telle que $g_{|M} = f$.*

Démonstration. Soit $(e_\alpha)_{\alpha \in A}$ une base de M. Complétons-la en une base $(e_\alpha)_{\alpha \in B}$ de X. Pour tout choix d'éléments $(\lambda_\alpha)_{\alpha \in B \setminus A}$ de $\mathbf{K}$, la formule

$$g(e_\alpha) = \begin{cases} f(e_\alpha) & \text{si } \alpha \in A, \\ \lambda_\alpha & \text{si } \alpha \in B \setminus A \end{cases}$$

définit une unique forme linéaire $g : X \to \mathbf{K}$. Par construction, g prolonge f. □

Si X est un espace vectoriel normé et si f est continue, le prolongement g ainsi construit n'est pas nécessairement continu (par exemple, si $\sup\{|\lambda_\alpha|/\|e_\alpha\| \ : \ \alpha \in B \setminus A\} = +\infty$, alors g n'est pas continue ; mais ce n'est pas la seule obstruction). Dans cette situation, le théorème de Hahn–Banach affirme qu'on peut trouver un prolongement continu et de même norme.

Si M est un sous-espace d'un espace vectoriel normé X et si $f : M \to \mathbf{R}$ est une forme linéaire continue, on notera

$$\|f\|_{M^*} = \sup_{x \in M \ : \ \|x\| \leqslant 1} f(x).$$

Théorème (Théorème de prolongement de Hahn–Banach). *Soit X un espace vectoriel normé, $M \subset X$ un sous-espace vectoriel et $f : M \to \mathbf{R}$ une forme linéaire continue. Alors il existe une forme linéaire continue $g : M \to \mathbf{R}$ qui prolonge f et telle que*

$$\|g\|_{X^*} = \|f\|_{M^*}.$$

Le théorème est facile si X est un espace de Hilbert : il suffit de prendre g nulle sur $M^\perp$. Dans le cas d'un espace de Banach, la preuve va contourner l'absence d'une notion de supplémentaire orthogonal.

Nous allons donner deux preuves de ce théorème : une preuve constructive dans le cas où X est séparable et une preuve basée sur l'axiome du choix dans le cas général. Les deux preuves reposent sur le lemme suivant qui permet de «gagner une dimension» et que l'on va ensuite itérer.

Lemme. *Soit X un espace vectoriel normé, $M \subset X$ un sous-espace et $x \in X$. Si $f : M \to \mathbf{R}$ est une forme linéaire continue, il existe une forme linéaire continue $g : M + \mathbf{R}x \to \mathbf{R}$ qui prolonge f, de même norme que f.*

Démonstration. La clé est la remarque suivante : si $(I_\alpha)_\alpha$ une famille de segments de $\mathbf{R}$ deux à deux non disjoints, alors leur intersection est non vide. (Preuve : écrire $I_\alpha = [x_\alpha, y_\alpha]$. Pour tous α, β on a $x_\beta \leqslant y_\alpha$. La famille (y_α) est donc minorée et sa borne inférieure z vérifie $x_\beta \leqslant z \leqslant y_\beta$ pour tout β).

Montrons le lemme. Si $x \in M$ il n'y a rien à faire. Si $f = 0$, alors $g = 0$ convient. Si $f \neq 0$, on peut supposer par homogénéité que $\|f\| = 1$. La fonction $f : M \to \mathbf{R}$ est 1-lipschitzienne et nous allons la prolonger f en une fonction 1-lipschitzienne $\tilde{f} : M \cup \{x\} \to \mathbf{R}$. On doit pour cela vérifier pour tout $y \in M$ la condition $\tilde{f}(x) \in I_y$ où

$$I_y = \Big[f(y) - \|x - y\|, f(y) + \|x - y\|\Big].$$

Les segments $(I_y)_{y \in M}$ sont deux à deux non disjoints (si $y_1, y_2 \in M$, l'inégalité $f(y_1) - \|x - y_1\| \leqslant f(y_2) + \|x - y_2\|$ se déduit de $f(y_2 - y_1) \leqslant \|y_2 - y_1\| \leqslant \|x - y_2\| + \|x - y_1\|$) et ont donc un point en commun α. On peut donc définir $\tilde{g}$ comme voulu en posant $\tilde{g}(x) = \alpha$.

Soit $g : M + \mathbf{R}x \to \mathbf{R}$ l'unique forme linéaire qui coïncide avec $\tilde{f}$ sur $M \cup \{x\}$. Pour tout $y \in M$ et $\lambda \in \mathbf{R}^*$, on a

$$|g(y + \lambda x)| = |\lambda| \cdot |g(\lambda^{-1} y + x)| \leqslant |\lambda| \cdot \|\lambda^{-1} y + x\| = \|y + \lambda x\|$$

ce qui montre que g est de norme $\leqslant 1$. □

Première preuve du théorème de HAHN–BANACH *: cas séparable.* Supposons X séparable et soit $(x_n)_{n \geqslant 1}$ une suite dense dans X. On considère la famille croissante (M_n) de sous-espaces vectoriels définie en posant $M_0 = M$ et $M_{n+1} = M_n + \mathbf{R}x_n$. Posons aussi $f_0 = f$. À l'aide du lemme, on construit par récurrence, pour tout entier $n \geqslant 1$, une forme linéaire continue $f_n : M_n \to \mathbf{R}$ qui prolonge f_{n-1} et telle que $\|f_n\|_{M_n^*} = \|f\|_{M^*}$. Soit Y le sous-espace $\bigcup_{n \in \mathbf{N}} M_n$. On définit une fonction $g : Y \to \mathbf{R}$ en posant $g(x) = f_n(x)$ lorsque $x \in M_n$. (Cette fonction est bien définie puisque si $x \in M_p$ et $x \in M_q$ alors $f_p(x) = f_q(x)$). La fonction g est une forme linéaire sur Y et vérifie $\|g\|_{Y^*} = \|f\|_M$.

Puisque Y est dense dans X, on peut prolonger de manière unique g en une forme linéaire continue $\tilde{g} : X \to \mathbf{R}$ de même norme. Fin cours #5 du 6 février (Preuve : pour tout $x \in X$, soit (x_p) une suite de Y telle que $\lim x_p = x$. Puisque g est lipschitzienne, $(g(x_p))$ est une suite de CAUCHY, donc convergente, et on pose $\tilde{g}(x) = \lim_{p \to \infty}(g(x_p))$. On vérifie que cette définition ne dépend pas du choix de la suite (x_p) et que la fonction $\tilde{g}$ a toutes les propriétés voulues. □

Avant de donner une autre preuve du théorème de HAHN–BANACH, faisons quelques préliminaires autour du lemme de ZORN.

Soit E un ensemble muni d'une relation d'ordre $\leqslant$ (c'est-à-dire une relation réflexive, transitive et antisymétrique). On dit qu'une partie A de E est *majorée* s'il existe $m \in E$ tel qu'on ait $x \leqslant m$ pour tout x de A. On dit qu'une partie $A \subset E$ est *totalement ordonnée* si pour tous les éléments x, y de A, on a $x \leqslant y$ ou $y \leqslant x$. On dit que E est *inductif* si toute partie totalement ordonnée est majorée. On dit que $x \in E$ est *maximal* si tout $y \in E$ tel que $y \geqslant x$ vérifie $y = x$.

Lemme (Lemme de ZORN)**.** *Tout ensemble ordonné inductif possède (au moins) un élément maximal.*

Le lemme de Zorn est équivalent à l'axiome du choix (ces questions seront évoquées en TD).

Seconde preuve du théorème de Hahn–Banach. Appelons *prolongement partiel* la donnée d'un couple (Y, g) où Y est un sous-espace vectoriel de X tel que $M \subset Y$, et $g : Y \to \mathbf{R}$ est une forme linéaire continue qui prolonge f telle que $\|g\|_Y = \|f\|_M$.

Soit E l'ensemble des prolongements partiels. On le munit de la relation d'ordre $\prec$ définie par

$$(Y, g) \prec (Y', g') \quad \text{si} \quad Y \subset Y' \text{ et } g' \text{ prolonge } g$$

L'ensemble E ainsi ordonné est inductif. En effet, si $A = \{(Y_i, g_i) \ : \ i \in I\}$ est une partie totalement ordonnée de E, un majorant de A est donnée par (Y, g), où Y est le sous-espace $\bigcup_{i \in I} Y_i$ et $g : Y \to \mathbf{R}$ est la forme linéaire définie par $g(y) = g_i(y)$ lorsque $y \in Y_i$. (L'argument qui précède suppose $A \neq \emptyset$; il faut aussi remarquer que (M, f) est un majorant de la partie vide.)

Par le lemme de Zorn, l'ensemble E admet donc un élément maximal (G, g). Il suffit maintenant de montrer que $G = X$. Raisonnons par l'absurde : si $G \subsetneq X$, il existe $x \in X \setminus G$ et le lemme fournit un prolongement $\tilde{g}$ de g tel que $(G + \mathbf{R}x, \tilde{g}) \in E$, contredisant la maximalité. □

Remarque. Une adaptation de cette preuve (avec une version adaptée du lemme permettant de gagner une dimension), laissée en exercice, donne la variante suivante du théorème de prolongement de Hahn–Banach que nous utiliserons. Soit X un $\mathbf{R}$-espace vectoriel et $p : X \to \mathbf{R}^+$ une fonction

1. *positivement homogène*, c'est-à-dire telle que $p(\lambda x) = \lambda p(x)$ si $\lambda \geqslant 0$ et $x \in X$,
2. *sous-additive*, c'est-à-dire telle que $p(x + y) \leqslant p(x) + p(y)$ pour tous x et y dans X.

Soit $M \subset X$ un sous-espace vectoriel et $f : M \to \mathbf{R}$ une forme linéaire telle que $f \leqslant p$. Alors il existe une forme linéaire $g : E \to \mathbf{R}$ prolongeant f et telle que $g \leqslant p$.

Voici des corollaires du théorème de Hahn–Banach.

Corollaire. *Soit X un espace de Banach et $x \in X \setminus \{0\}$. Il existe $f \in X^*$ telle que $\|f\| = 1$ et $f(x) = \|x\|$.*

Démonstration. Prolonger la forme linéaire de norme 1 définie sur $\mathbf{R}x$ par $\lambda x \mapsto \lambda\|x\|$. □

Corollaire (« le dual sépare les points »)**.** *Soit X un espace de Banach et $x_1 \neq x_2$ deux vecteurs de X. Il existe $f \in X^*$ telle que $f(x_1) \neq f(x_2)$*

Démonstration. Appliquer le résultat précédent à $x = x_1 - x_2$. □

Pour terminer, voici un théorème utile pour montrer qu'un sous-espace est dense, qui doit vous rappeler un énoncé analogue pour les espaces de Hilbert.

Théorème. *Soit X un espace de Banach et $M \subset X$ un sous-espace vectoriel. Alors*

$$M \text{ est dense} \iff \text{tout } f \in X^* \text{ vérifiant } f(M) = 0 \text{ vérifie } f = 0.$$

Démonstration. Une forme linéaire continue nulle sur M est nulle sur $\overline{M}$. Ceci montre le sens direct.

Supposons que M n'est pas dense et soit $x \in X \setminus \overline{M}$. Définissons une forme linéaire $g : \overline{M} + \mathbf{R}x \to \mathbf{R}$ par les conditions $g(\overline{M}) = 0$ et $g(x) = 1$.

Posons $r = d(x, \overline{M}) > 0$. Pour $y \in \overline{M}$ et $\lambda \in \mathbf{R}^*$, on a $\|y + \lambda x\| = |\lambda| \cdot \|x - (-\lambda^{-1} y)\| \geqslant |\lambda| r$ et donc

$$g(y + \lambda x) = \lambda \leqslant r^{-1} \|y + \lambda x\|,$$

ce qui montre que g est continue. Par le théorème de prolongement de HAHN–BANACH, il existe une forme linéaire continue $f \in X^*$ qui prolonge g. La forme linéaire f est nulle sur M et non identiquement nulle puisque $f(x) = 1$. □

5.2 Espaces vectoriels topologiques

Définition. On appelle *espace vectoriel topologique* la donnée d'un $\mathbf{R}$-espace vectoriel X muni d'une topologie telle que

1. les applications

$$\begin{aligned} X \times X &\to X, \\ (x, y) &\mapsto x + y, \end{aligned} \qquad \begin{aligned} \mathbf{R} \times X &\to X \\ (\lambda, x) &\mapsto \lambda x \end{aligned}$$

   sont continues. (On munit $X \times X$ et $\mathbf{R} \times X$ de la topologie produit.)

2. $\{0\}$ est fermé

Un espace vectoriel normé, muni de la topologie associée, est un espace vectoriel topologique.

Dans tout espace vectoriel topologique, les translations sont des homéomorphismes. Les voisinages d'un point a sont donc les ensembles de la forme $a + V$ où V est un voisinage de 0.

On utilisera la conséquence suivante de la continuité de l'addition, qui correspond à « couper ε en deux » dans le cas métrique : si V est un voisinage de 0 dans un espace vectoriel topologique, il existe W un voisinage de 0 tel que $W + W \subset V$.

Exercice. Tout espace vectoriel topologique est séparé.

Exercice. Soit X un espace vectoriel topologique et $f : X \to \mathbf{R}$ une forme linéaire. Alors f est continue si et seulement si f est continue en 0.

Lemme (Jauge de MINKOWSKI d'un convexe)**.** *Soit X un espace vectoriel topologique et $W \subset X$ un ouvert convexe contenant* 0*. Pour $x \in X$, on pose*

$$j_W(x) = \inf\{t > 0 \ : \ x \in tW\}.$$

Alors la fonction j_W est à valeurs finies, continue, positivement homogène et sous-additive. De plus,

$$W = \{x \in X \ : \ j_W(x) < 1\}.$$

Démonstration. Voir feuille de TD. □

Si X est un espace vectoriel normé et si $W = B(0, 1)$, alors j_W est la norme.

5.3 Les théorèmes de séparation de HAHN–BANACH

Soient A et B deux parties non vides d'un $\mathbf{R}$-espace vectoriel X, et $g : X \to \mathbf{R}$ une forme linéaire non nulle. On dit que

— la forme linéaire g *sépare A de B au sens large* si pour tout $x \in A$ et $y \in B$ on a $g(x) \leqslant g(y)$; ceci équivaut à dire que

$$\sup_{x \in A} g(x) \leqslant \inf_{y \in B} g(y).$$

— la forme linéaire g *sépare A de B au sens strict* si

$$\sup_{x \in A} g(x) < \inf_{y \in B} g(y).$$

Si on choisit $\alpha \in \mathbf{R}$ telle que $\sup_A g \leqslant \alpha \leqslant \inf_B g$, on dit aussi que l'hyperplan affine $\{g = \alpha\}$ sépare A de B au sens large.

Théorème (Théorème de séparation large de Hahn–Banach)**.** *Soit X un espace vectoriel topologique, $A \subset X$ un convexe non vide et $B \subset X$ un* **ouvert** *convexe non vide, tels que $A \cap B = \emptyset$. Il existe une forme linéaire continue non nulle $g : X \to \mathbf{R}$ qui sépare A de B au sens large.*

Démonstration. Fixons $x_0 \in A$, $y_0 \in B$ et posons $z_0 = y_0 - x_0$. On introduit

$$W = (A - x_0) - (B - y_0) = \{x - x_0 - (y - y_0) \ : \ x \in A, \ y \in B\}.$$

C'est un ouvert (on l'a écrit comme réunion de translatés d'ouverts) convexe contenant 0. Soit j_W sa jauge de Minkowski. Par le lemme précédent, la fonction j_W est continue, positivement homogène, sous-additive et vérifie $W = \{j_W < 1\}$.

Par ailleurs on a $z_0 \notin W$ (puisque A et B sont disjoints) et donc $j_W(z_0) \geqslant 1$. Définissons

$$\begin{aligned} f : \mathbf{R}z_0 &\to \mathbf{R} \\ \lambda z_0 &\mapsto \lambda \end{aligned}$$

C'est une forme linéaire sur $\mathbf{R}z_0$ telle que $f \leqslant j_W$. Par le théorème de prolongement de Hahn–Banach (et plus précisément, par la remarque p. 22), on peut prolonger f en une forme linéaire $g : X \to \mathbf{R}$ telle que $g \leqslant j_W$. Fin cours #6 du 13 février Puisque $|g(z)| \leqslant \max(j_W(z), j_W(-z))$, la forme linéaire g est continue en 0, donc continue. Enfin, pour tout $x \in A$ et $y \in B$, on a

$$g(x) - g(y) + g(z_0) \leqslant g(x - y + z_0) \leqslant j_W(x - y + z_0) < 1 \leqslant g(z_0),$$

d'où on tire comme voulu que $g(x) < g(y)$. □

On dit qu'un espace vectoriel topologique est *localement convexe* si pour tout voisinage V de 0, il existe un ouvert convexe symétrique W tel que $0 \in W \subset V$. Un espace vectoriel normé est localement convexe.

Théorème (Théorème de séparation stricte de Hahn–Banach)**.** *Soit X un espace vectoriel topologique* **localement convexe***, $A \subset X$ un convexe* **fermé** *non vide et $B \subset X$ un convexe* **compact** *non vide, tels que $A \cap B = \emptyset$. Il existe une forme linéaire non nulle continue $f : X \to \mathbf{R}$ qui sépare A de B au sens strict.*

Lemme. *Il existe un voisinage V de 0 tel que $A \cap (B + V) = \emptyset$*

Démonstration. L'ensemble $X \setminus A$ est ouvert. Soit $y \in X \setminus A$. Il existe un ouvert V_y voisinage de 0 tel que $y + V_y + V_y \subset X \setminus A$. L'ensemble $\{y + V_y \ : \ y \in B\}$ est un recouvrement ouvert du compact B ; il existe donc un ensemble fini $F \subset B$ tel que

$$B \subset \bigcup_{y \in F} y + V_y.$$

Soit $V = \bigcap_{y \in F} V_y$; c'est un voisinage de 0 et

$$B + V \subset \bigcup_{y \in F} (y + V_y + V) \subset \bigcup_{y \in F} (y + V_y + V_y) \subset X \setminus A$$

On a ainsi $A \cap (B + V) = \emptyset$. □

Démonstration du théorème de séparation stricte. Puisque X est localement convexe, on peut supposer que le voisinage V produit par le lemme est convexe symétrique. Les ensembles A et $B + V$ sont alors des convexes disjoints non vides, et $B + V$ est ouvert. Par le théorème de séparation large, il existe une forme linéaire non nulle f telle que

$$\sup_{x \in B+V} f(x) \leqslant \inf_{y \in A} f(y).$$

On conclut en remarquant que

$$\sup_{B+V} f = \sup_B f + \sup_V f > \sup_B f.$$

Pour la dernière inégalité : puisque V est symétrique, si f était $\leqslant 0$ sur V, elle serait aussi $\geqslant 0$, donc nulle. Mais $X = \bigcup_{n \in \mathbf{N}} nV$ (en effet, pour x dans X, la suite $(n^{-1}x)$ converge vers 0 par continuité de la multiplication par un scalaire, donc ses termes de rang assez grand sont dans V) et donc f serait nulle. □

5.4 Dual réel vs dual complexe

Sot X un espace de Banach complexe. C'est aussi un espace de Banach réel. On peut définir a priori

$$X^*_{\mathbf{C}} = \{f : X \to \mathbf{C} \ : \ f \text{ continue et } \mathbf{C}\text{-linéaire}\}$$

$$X^*_{\mathbf{R}} = \{f : X \to \mathbf{R} \ : \ f \text{ continue et } \mathbf{R}\text{-linéaire}\}$$

On peut en fait identifier ces deux espaces.

Proposition. *L'application $\varphi : f \mapsto \operatorname{Re} f$ est une bijection $\mathbf{R}$-linéaire et isométrique de $X^*_{\mathbf{C}}$ sur $X^*_{\mathbf{R}}$.*

Démonstration. Il est immédiat que φ est $\mathbf{R}$-linéaire.

— Montrons d'abord que φ est isométrique. Si $f \in X^*_{\mathbf{C}}$, alors $\|\operatorname{Re} f\|_{X^*_{\mathbf{R}}} \leqslant \|f\|_{X^*_{\mathbf{C}}}$ puisque tout nombre complexe λ vérifie $|\operatorname{Re} \lambda\| \leqslant |\lambda|$. Soit $x \in X$ vérifiant $\|x\| \leqslant 1$. Il existe un complexe λ vérifiant $|\lambda| = 1$ et $\lambda f(x) = |f(x)|$. On a alors

$$|f(x)| = \lambda f(x) = f(\lambda x) = \operatorname{Re} f(\lambda x) \leqslant \|\operatorname{Re} f\|_{X^*_{\mathbf{R}}} \|x\|$$

et l'inégalité $\|f\|_{X^*_{\mathbf{C}}} \leqslant \|\operatorname{Re} f\|_{X^*_{\mathbf{R}}}$ s'ensuit en prenant la borne supérieure sur x. Ainsi φ est isométrique et donc injective.

— Montrons que φ est surjective. Si $\ell \in X^*_{\mathbf{R}}$, la formule $f : x \mapsto \ell(x) - i\ell(ix)$ définit une forme $\mathbf{R}$-linéaire continue sur X. On a de plus

$$f(ix) = \ell(ix) - i\ell(-x) = i[\ell(x) - i\ell(ix)] = if(x)$$

et donc f est $\mathbf{C}$-linéaire. Comme $\ell = \operatorname{Re} f = \varphi(f)$, on obtient la surjectivité de φ.
□

A l'aide de cette proposition, il est facile de déduire que le théorème de prolongement de Hahn–Banach (énoncé et démontré dans le cas réel) s'étend avec exactement le même énoncé au cas des espaces vectoriels normés sur $\mathbf{C}$.

Fin du programme pour l'examen partiel

Chapitre 6

Le théorème de Krein–Milman

Soit X un espace vectoriel. Une partie $C \subset X$ est convexe si et seulement si, pour tout $\lambda \in [0,1]$ elle vérifie $\lambda C + (1-\lambda)C = C$. L'inclusion non triviale revient à demander que pour tous x et y dans C, le segment $[x,y] = \{\lambda x + (1-\lambda)y \ : \ 0 \leqslant \lambda \leqslant 1\}$ soit inclus dans C.

Lorsque C est convexe, on montre par récurrence sur n que si $\lambda_1, \ldots, \lambda_n$ sont des réels positifs vérifiant $\lambda_1 + \cdots + \lambda_n = 1$, alors pour tous $x_1, \ldots, x_n$ dans C on a

$$\lambda_1 x_1 + \cdots + \lambda_n x_n \in C.$$

Une telle expression est appelé une combinaison convexe d'éléments de C.

On appelle *enveloppe convexe* de A, et on note $\mathrm{conv}(A)$, le plus petit ensemble convexe qui contient A. On peut le définir par l'une ou l'autre des équations suivantes

$$\begin{aligned}
\mathrm{conv}(A) &= \text{intersection de la famille des convexes contenant } A \\
&= \{\text{combinaisons convexes d'éléments de } A\}, \\
&= \left\{ \sum_{i=1}^{N} \lambda_i x_i,\ N \in \mathbf{N},\ \lambda_i \geqslant 0,\ \sum_{i=1}^{N} \lambda_i = 1,\ x_i \in A \right\}
\end{aligned}$$

Définition. Soit X un espace vectoriel et $C \subset X$ une partie convexe. On dit qu'un point $x \in C$ est un *point extrémal* de C si, dès que $y, z \in C$ et $\lambda \in]0,1[$ vérifient $\lambda y + (1-\lambda)z = x$, alors $y = z = x$.

Autrement dit, x est un point extrémal de C s'il ne peut pas être écrit de manière non triviale comme combinaison convexe d'éléments de C. On peut aussi remarquer que x est extrémal si et seulement si $C \setminus \{x\}$ est convexe.

Exercice. Pour $p = 1, 2, \infty$, déterminer les points extrémaux de la boule unité de $(\mathbf{R}^n, \|\cdot\|_p)$.

Fin cours #7 du 19 février

Dans un espace vectoriel topologique X, l'adhérence $\overline{C}$ d'une partie convexe C est convexe (preuve : pour tout $\lambda \in [0,1]$, l'application $f : X^2 \to X$ donnée par $(x,y) \mapsto \lambda x + (1-\lambda)y$ est continue et vérifie donc (exercice) $f(\overline{C} \times \overline{C}) = f(\overline{C \times C}) \subset \overline{f(C)}$). En particulier, pour toute partie A de X, la partie $\overline{\mathrm{conv}(A)}$ est convexe ; c'est aussi l'intersection de tous les convexes fermés contenant A.

Théorème (Théorème de Krein–Milman)**.** *Soit X un espace vectoriel topologique localement convexe, $K \subset X$ un convexe compact et $\mathcal{E}$ l'ensemble des points extrémaux de K. Alors*

$$K = \overline{\mathrm{conv}(\mathcal{E})}.$$

Pour démontrer le théorème de Krein–Milman, il est utile de disposer du concept suivant. On dit qu'une partie convexe fermée non vide $F \subset K$ est une *face extrémale* si, dès lors que x, y dans K vérifient $\lambda x + (1-\lambda)y \in F$ pour un $0 < \lambda < 1$, alors on a $x \in F$ et $y \in F$. Ainsi, un élément $x \in K$ est un point extrémal de K si et seulement si $\{x\}$ est une face extrémale de K.

Soit $f : X \to \mathbf{R}$ est une forme linéaire continue. Si on pose $m = \sup\{f(x) \ : \ x \in K\}$, alors l'ensemble $\{x \in K \ : \ f(x) = m\}$ est une face extrémale de K.

Preuve du théorème de Krein–Milman. On utilise le lemme suivant.

Lemme. *Toute face extrémale contient un point extrémal.*

Preuve du lemme. Soit F une face extrémale et soit $\mathcal{F}$ l'ensemble des faces extrémales de K incluses dans F. Définissons une relation d'ordre sur $\mathcal{F}$ en posant $F_1 \prec F_2$ si $F_1 \supset F_2$. Montrons que l'ensemble $(\mathcal{F}, \prec)$ est inductif. Soit $A = \{F_i \ : \ i \in I\}$ une partie totalement ordonnée de $\mathcal{F}$; on a donc $F_i \subset F_j$ ou $F_j \subset F_i$ pour tout i, j dans I. Alors la partie de K définie par $\bigcap_{i \in I} F_i$ est un majorant de $(F_i)_{i \in I}$ (elle est convexe fermée comme intersection de convexes fermés et non vide par compacité de K, car l'intersection de toute sous-famille finie est non vide ; enfin, il est facile de vérifier que c'est une face extrémale incluse dans F). Par le lemme de Zorn, il existe un majorant G de $\mathcal{F}$. Montrons que G est un singleton. Si $x \neq y$ sont deux éléments de G, par le théorème de séparation stricte de Hahn–Banach, il existe une forme linéaire continue non nulle $\varphi : X \to \mathbf{R}$ telle que $\varphi(x) \neq \varphi(y)$. Soit M le maximum de φ sur G ; alors $G_0 := G \cap \{\varphi = M\}$ est une face extrémale vérifiant $G_0 \subsetneq G$, d'où contradiction. □

Soit $\mathcal{E}$ l'ensemble des points extrémaux de K (par le lemme, puisque K est une face extrémale, $\mathcal{E}$ est non vide) et $L = \overline{\mathrm{conv}(\mathcal{E})}$; c'est un convexe compact non vide. Montrons que $K = L$. Si $z \in L \setminus K$, par le théorème de séparation stricte de Hahn–Banach, il existe une forme linéaire continue $\psi : X \to \mathbf{R}$ telle que $\sup_L \psi < \psi(z)$. Soit M le maximum de ψ sur K. Alors $K \cap \{\psi = M\}$ est une face extrémale disjointe de L. Par le lemme, elle contient un point extrémal $w \in \mathcal{E} \setminus L$, d'où contradiction. □

Chapitre 7

Dualité ; topologies faible et préfaible

Dans tout ce chapitre, on désigne par X un espace de Banach. Commençons par quelques compléments.

On note X^{**} le dual de X^* ; c'est le *bidual* de X. L'application canonique J_X (ou simplement J) de X dans X^{**} est donnée pour $x \in X$ et $f \in X^*$ par $J_X(x)(f) = f(x)$. Autrement dit, $J_X(x)$ est la forme linéaire $f \mapsto f(x)$. Cette forme linéaire est continue puisque :

$$\|x\|_X = \sup_{f \in X^*,\ \|f\| \leqslant 1} f(x) = \sup_{f \in X^*,\ \|f\| \leqslant 1} J_X(x)(f) = \|J_X(x)\|_{X^{**}},$$

ce qui montre aussi que J_X est isométrique, donc injective.

7.1 La topologie faible sur un espace de Banach

C'est la «topologie sur X la moins fine rendant continues tous les éléments de X^*». De manière plus explicite, on appelle *ouvert faible élémentaire* un ensemble du type [1]

$$V(A, \varepsilon, x_0) = \{x \in X \ : \ \forall f \in A,\ |f(x) - f(x_0)| < \varepsilon\}$$

où A est une partie finie de X^*, x_0 est un point de X et ε un réel > 0 ; on appelle *ouvert faible* une partie de X qui peut s'écrire comme réunion quelconque d'ouverts faibles élémentaires.

Proposition. *L'ensemble $\sigma(X, X^*)$ des ouverts faibles forme une topologie sur X, appelée* topologie faible. *De plus, $(X, \sigma(X, X^*))$ est un espace vectoriel topologique localement convexe.*

Démonstration. Si $y \in V(A, \varepsilon, x_0)$, il existe $\eta > 0$ tel que $V(A, \eta, y) \subset V(A, \varepsilon, x_0)$. Il découle de cette remarque que $\sigma(X, X^*)$ est stable par intersection finie ; les autres axiomes de topologie sont évidents.

La continuité de l'addition et de la multiplication scalaire sont laissés en exercice. Si $x \neq y$ sont deux points de X, par un corollaire de Hahn–Banach, il existe $f \in X^*$ tel que $\varepsilon = |f(x) - f(y)| > 0$. Alors $V(\{f\}, \varepsilon/2, x)$ et $V(\{f\}, \varepsilon/2, y)$ sont des ouverts faibles disjoints contenant respectivement x et y.

Enfin, la locale convexité découle du fait que $V(A, \varepsilon, x_0)$ est convexe. □

1. on a utilisé la notation alternative $V_{x_0, A, \varepsilon}$ lors du cours magistral

Tout ouvert faible est un ouvert fort, c'est-à-dire un ouvert pour la topologie de la norme.

Si X est de dimension finie, la topologie de la norme coïncide avec la topologie faible. En effet, quitte à remplacer la norme par une norme équivalente, on peut considérer le cas de $(\mathbf{R}^n, \|\cdot\|_\infty)$, pour lequel les boules ouvertes sont des ouverts faibles élémentaires.

Si X est de dimension infinie, tout ouvert faible non vide contient un sous-espace de dimension infinie, puisque V_{A,ε,x_0} contient le sous-espace affine $x_0 + \bigcap_{f\in A} \ker(f)$ qui est de codimension finie. Par conséquent, la boule-unité ouverte n'est pas un ouvert faible.

Théorème. *Soit $C \subset X$ une partie convexe. Alors C est fermée si et seulement si C est fermée pour la topologie faible.*

Démonstration. Supposons C fermé et montrons que $X \setminus C$ est un ouvert faible. Soit $x \in X \setminus C$. Par le théorème de séparation stricte de Hahn–Banach, il existe $f \in X^*$ telle que $\varepsilon := f(x) - \sup_C f > 0$. Ainsi, l'ouvert faible élémentaire $V(\{f\}, x, \varepsilon)$ contient x et est disjoint de C. Ceci montre que $X \setminus C$ est un ouvert faible et donc que C est fermé pour la topologie faible. □

Soit $(x_n)_n$ et x dans X. On dit que la suite (x_n) *converge faiblement* vers x si elle converge au sens de la topologie faible. On vérifie (exercice) que c'est équivalent à

$$\forall f \in X^* \quad \lim_{n\to\infty} f(x_n) = f(x).$$

Corollaire (Théorème de Mazur)**.** *Soit (x_n) une suite de X qui converge faiblement vers $x \in X$. Il existe une suite (y_m) d'éléments de* $\operatorname{conv}\{x_n \ : \ n \in \mathbf{N}\}$ *qui converge fortement vers x.*

Démonstration. Cf TD □

Fin cours #8 du 2 avril

7.2 La topologie préfaible sur le dual d'un espace de Banach

C'est la «topologie sur X^* la moins fine rendant continues les applications d'évaluation en des points de X». De manière plus explicite, on appelle *ouvert préfaible élémentaire* un ensemble du type[2]

$$W(B, \varepsilon, f_0) = \{f \in X^* \ : \ \forall x \in B, \ |f(x) - f_0(x)| < \varepsilon\}$$

où B est une partie finie de X, f_0 est un point de X et ε un réel > 0 ; on appelle *ouvert préfaible* une partie de X^* qui peut s'écrire comme réunion quelconque d'ouverts préfaibles élémentaires.

Proposition. *L'ensemble $\sigma(X^*, X)$ des ouverts préfaibles forme une topologie sur X^*, appelée* topologie préfaible *ou* topologie faible-*. *De plus, $(X^*, \sigma(X^*, X))$ est un espace vectoriel topologique localement convexe.*

Démonstration. Similaire au cas de la topologie faible. □

Proposition. *Soit $\varphi \in X^{**}$. Il y a équivalence entre*

1. *φ est continue lorsqu'on munit X^* de la topologie préfaible $\sigma(X^*, X)$,*

2. on a utilisé la notation alternative $W_{f_0,B,\varepsilon}$ lors du cours magistral

2. $\varphi \in J_X(X)$.

On résume cette proposition en disant que «le dual de l'e.v.t. $(X^*, \sigma(X^*, X))$ est X».

Démonstration. Pour tout $x \in X$, la forme linéaire $J_X(x) : X^* \to \mathbf{R}$ est préfaiblement continue puisque l'image réciproque de tout intervalle ouvert est un ouvert préfaible élémentaire.

Réciproquement, supposons φ préfaiblement continue. Puisque $\{|\varphi| < 1\}$ est un ouvert préfaible contenant 0, il contient un ouvert préfaible élémentaire $W(\{x_1, \dots, x_n\}, \varepsilon, 0)$. On a donc, pour tout $f \in X^*$ l'implication

$$\max\{|f(x_1)|, \dots, |f(x_n)|\} \leqslant \varepsilon \Longrightarrow |\varphi(f)| \leqslant 1$$

et donc, par homogénéité

$$f(x_1) = f(x_2) = \dots = f(x_n) = 0 \Longrightarrow \varphi(f) = 0.$$

Le lemme suivant permet de conclure que $f \in \mathrm{Vect}\{J_X(x_1), \dots, J_X(x_n)\}$. □

Lemme. *Soit E un espace vectoriel et $\psi, \psi_1, \dots, \psi_n : E \to \mathbf{R}$ des formes linéaires. Alors*

$$\psi \in \mathrm{Vect}(\psi_1, \dots, \psi_n) \iff \bigcap_{i=1}^{n} \ker \psi_i \subset \ker \psi.$$

Démonstration. Voir TD. □

Théorème (Théoréme de Banach–Alaoglu)**.** *Soit X un espace de Banach. La boule unité fermée B_{X^*} est compacte pour la topologie préfaible.*

Démonstration. Soit $\mathbf{R}^X$ l'ensemble de toutes les fonctions de X dans $\mathbf{R}$, muni de la topologie produit. L'ensemble $L \subset \mathbf{R}^X$ des formes linéaires est fermé comme intersection de fermés, puisqu'on peut écrire

$$L = \bigcap_{x,y \in X} \{f \in \mathbf{R}^X \ : \ f(x+y) = f(x) + f(y)\} \cap \bigcap_{x \in X, \lambda \in \mathbf{R}} \{f \in \mathbf{R}^X \ : \ f(\lambda x) = \lambda f(x)\}.$$

Soit B l'ensemble des fonctions $f \in \mathbf{R}^X$ vérifiant $|f(x)| \leqslant \|x\|$ pour tout x dans X. C'est un espace compact par le théorème de Tychonoff puisque

$$B = \prod_{x \in X} [-\|x\|, \|x\|].$$

Ainsi $L \cap B$, qui s'identifie à B_{X^*}, est une partie compacte de $\mathbf{R}^X$.

Pour conclure la preuve, il suffit d'observer que lorsqu'on identifie X^* à une partie de $\mathbf{R}^X$ muni de la topologie produit, la topologie obtenue coïncide avec la topologie préfaible. □

On peut montrer (cf TD) que les espaces topologiques $(X, \sigma(X, X^*))$ et $(X^*, \sigma(X^*, X))$ ne sont pas métrisables dès que X est de dimension infinie.

Théorème. *Soit X un espace de Banach. Il y a équivalence entre*

1. *X est séparable.*
2. *L'espace topologique compact $(B_{X^*}, \sigma(X^*, X))$ est métrisable.*

Démonstration. Supposons X séparable. Alors B_X contient une suite dense (x_n). On vérifie (exercice, cf TD) que la topologie associée la distance définie sur B_{X^*} par

$$d(f,g) = \sum_{n\in\mathbf{N}} 2^{-n}|f(x_n) - g(x_n)|.$$

est la restriction à B_{X^*} de $\sigma(X^*,X)$.

Réciproquement, supposons $(B_{X^*}, \sigma(X^*,X))$ métrisable. Il existe donc une base dénombrable de voisinages de 0, c'est-à-dire une suite (U_n) d'ouverts préfaibles élémentaires (c'est-à-dire de la forme $U_n = W(B_n, \varepsilon_n, 0)$ avec B_n une partie finie de X) tels que tout voisinage préfaible de 0 contient un des U_n. Comme $\bigcap U_n = \{0\}$, on en déduit qu'une forme linéaire continue qui s'annule sur $\bigcup A_n$ est nulle. Par un corollaire du théorème de Hahn–Banach, cela implique que Vect $(\bigcup A_n)$ est un sous-espace dense, et donc que le $\mathbf{Q}$-espace vectoriel engendré par $\bigcup A_n$ est une partie dénombrable dense de X. □

Corollaire. *Soit X un espace de Banach séparable. Toute suite bornée de X^* admet une sous-suite préfaiblement convergente.*

7.3 Adjoint

Soient X et Y deux espaces de Banach et $T \in \mathcal{L}(X,Y)$. On définit l'application linéaire transposée ou adjointe $T^* : Y^* \to X^*$ par la formule

$$(T^*f)(x) = f(Tx).$$

Puisque $(T^*f)(x) \leqslant \|f\|_{Y^*}\|Tx\|_Y \leqslant \|f\|_{Y^*}\|T\|_{op} \cdot \|x\|_X$, il s'ensuit que T^* est bien définie et continue, et vérifie $\|T^*\|_{op} \leqslant \|T\|_{op}$. Mais on a en fait $\|T^*\|_{op} = \|T\|_{op}$ en prenant la borne supérieure en x dans l'inégalité

$$\|Tx\| = \sup_{f\in Y^*, \|f\|\leqslant 1} f(Tx) = \sup_{f\in Y^*, \|f\|\leqslant 1} (T^*f)(x) \leqslant \|T^*\|_{op}\|x\|$$

7.4 Espaces de Banach réflexifs

Définition. On dit qu'un espace de Banach X est *réflexif* si l'application J_X est surjective.

Dans les cas où on a identifié le dual d'un espace de Banach, on peut déterminer s'il est réflexif.

Proposition. *Si $(\Omega, \mathcal{F}, \mu)$ est un espace mesuré σ-fini, alors pour tout $1 < p < \infty$ l'espace $L^p(\mu)$ est réflexif.*

Démonstration. Soit q l'exposant conjugué de p. L'application $\ell_p : L^p(\mu) \to L^q(\mu)^*$ définie par $\ell_p(f)(g) = \int_X fg \,\mathrm{d}\mu$ est bijective et isométrique. Un calcul direct montre que

$$\ell_q^* \circ J_{L^p(\mu)} = \ell_p$$

et donc l'application $J_{L^p(\mu)}$ est bijective puisque ℓ_p et ℓ_q^* le sont. □

De la même manière, le théorème de Riesz–Fréchet implique que tout espace de Hilbert est réflexif.

Le lemme suivant est utile dans l'étude des espaces de Banach réflexifs.

Lemme (Lemme de GOLDSTINE)**.** *Soit X un espace de BANACH. L'adhérence de $J_X(B_X)$ pour la topologie $\sigma(X^{**}, X^*)$ est $B_{X^{**}}$.*

Démonstration. Notons F l'adhérence de $J_X(B_X)$ pour la topologie $\sigma(X^{**}, X^*)$. On a $F \subset B_{X^{**}}$ puisque cette dernière est préfaiblement compacte. Supposons par l'absurde que l'inclusion est stricte et soit $x \in B_{X^{**}} \setminus F$. Par le théorème de séparation stricte de HAHN–BANACH, il existe une forme linéaire $\phi : X^{**} \to \mathbf{R}$, continue pour la topologie préfaible, telle que $\phi(x) > \sup_F \phi$. Puisque «le dual de $(X^{**}, \sigma(X^{**}, X^*))$ est X^*», il existe $f \in X^*$ telle que $\phi = J_{X^*}(f)$. On a donc $\phi(x) = x(f)$ pour tout $x \in X^{**}$. Il vient alors

$$\|f\| \geqslant x(f) = \phi(x) > \sup_F \phi = \sup_{J_X(B_X)} \phi = \sup_{B_X} f = \|f\|,$$

d'où contradiction. □

Fin cours # 9 du 9 avril

Théorème. *Soit X un espace de BANACH. Alors X est réflexif si et seulement si B_X est faiblement compacte.*

Démonstration. Supposons X réflexif. Par le théorème de BANACH–ALAOGLU, la boule unité $B_{X^{**}}$ est compacte pour la topologie $\sigma(X^{**}, X^*)$. Le résultat en découle en identifiant X et X^{**} via J_X.

Réciproquement, supposons B_X faiblement compacte. L'application J_X est continue de $(X, \sigma(X, X^*))$ vers $(X^{**}, \sigma(X^{**}, X^*))$ (exercice, observer que l'image réciproque d'un ouvert préfaible élémentaire est un ouvert faible élémentaire) et donc $J_X(B_X)$ est compact pour la topologie $\sigma(X^{**}, X^*)$. Par le lemme de GOLDSTINE, on a donc $J_X(B_X) = B_{X^{**}}$, ce qui implique $J_X(X) = X^{**}$. □

Corollaire. *Soit X un espace de BANACH réflexif et $Y \subset X$ un sous-espace fermé. Alors Y est réflexif.*

Démonstration. La topologie $\sigma(Y, Y^*)$ coïncide avec la restriction à Y de $\sigma(X, X^*)$. La boule unité B_Y est un convexe fermé de X, donc faiblement fermé, donc faiblement compact (comme partie fermée d'un compact). □

Corollaire. *Dans un espace de BANACH réflexif, toute suite bornée admet une sous-suite faiblement convergente.*

Démonstration. Soit (y_n) une suite bornée de X et $Y \subset X$ l'espace vectoriel fermé engendré par (y_n). L'espace de BANACH Y est séparable, et réflexif par le corollaire précédent. Le résultat découle alors du corollaire de la page 32 en utilisant le lemme suivant. □

Lemme. *Si Z est un espace de BANACH tel que Z^* est séparable, alors Z est séparable.*

Démonstration. Soit (f_n) une suite dense dans Z^*. Choisissons pour tout n un vecteur $z_n \in Z$ tel que $\|z_n\| \leqslant 1$ et $|f_n(z_n)| \geqslant \frac{1}{2}\|f_n\|$. Le $\mathbf{Q}$-espace vectoriel engendré par $\{z_n : n \in \mathbf{N}\}$ est dénombrable et dense dans $E = \mathrm{Vect}\{z_n : n \in \mathbf{N}\}$. Il suffit donc de montrer que E est dense dans Z ; pour cela on montre que toute forme linéaire continue $f \in Z^*$ nulle sur E est identiquement nulle. En effet, si $f \neq 0$, soit $n \in \mathbf{N}$ tel que $\|f - f_n\| < \frac{1}{4}\|f\|$. Puisque $\|f_n\| \geqslant \frac{3}{4}\|f\|$, on a $\|f - f_n\| \leqslant \frac{1}{3}\|f_n\|$. On en déduit

$$0 = f(x_n) = f_n(x_n) - (f_n - f)(x_n) \geqslant \frac{1}{2}\|f_n\| - \|f - f_n\| \geqslant \frac{1}{2}\|f_n\| - \frac{1}{3}\|f_n\| > 0,$$

d'où contradiction. □

Corollaire. *Soit X un espace de* Banach. *On a l'équivalence*

$$X \text{ réflexif} \iff X^* \text{ réflexif}.$$

Démonstration. Supposons X réflexif. Alors les topologies faible $\sigma(X^*, X^{**})$ et préfaible $\sigma(X^*, X)$ coïncident sur X^*. On déduit du théorème de Banach–Alaoglu que B_{X^*} est (préfaiblement donc) faiblement compacte. Le théorème précédent implique que X^* est réflexif.

Supposons X^* réflexif. Par le point précédent, X^{**} est réflexif et donc son sous-espace fermé $J_X(X)$ est aussi réflexif. Puisque $J_X(X)$ et X sont isométriquement isomorphes, l'espace X est aussi réflexif. $\square$

Chapitre 8

Opérateurs compacts et théorie spectrale

On appelle *opérateur* une application linéaire continue entre espaces de Banach.

Un endomorphisme d'un espace vectoriel de dimension fini est surjectif si et seulement si il est injectif. Voici deux exemples qui montrent que cet énoncé ne s'étend pas aux espaces de dimension infinie.

— L'opérateur de décalage sur l'espace de Hilbert $X = \ell^2(\mathbf{N})$, donné par $e_n \mapsto e_{n-1}$ si $n \geqslant 1$ et $e_n \mapsto 0$, est surjectif mais non injectif (ici $(e_n)_{n\in\mathbf{N}}$ désigne la base hilbertienne usuelle de X).

— Sur l'espace de Hilbert $X = L^2([0,1])$ (pour la mesure de Lesbsgue) l'opérateur de multiplication par t, donné par $f \mapsto (t \mapsto tf(t))$, est injectif mais non surjectif.

Exercice. Déterminer les valeurs propres de ces deux opérateurs. (Solution : dans le premier cas, les valeurs propres sont les scalaires de module < 1 ; dans le second cas, il n'y a aucune valeur propre).

8.1 Opérateurs compacts

Soit X un espace de Banach. Pour une partie A de X, on définit l'*annihilateur* de A comme

$$A^\perp = \{f \in X^* \ : \ f(x) = 0 \ \forall x \in A\}.$$

C'est un sous-espace de X^* qui est fermé pour la topologie préfaible.

Pour une partie B de X^*, on définit le *préannihilateur* de B comme

$${}^\perp B = \{x \in X \ : \ f(x) = 0 \ \forall f \in B\}.$$

C'est un sous-espace de X qui est fermé pour la topologie forte, donc aussi pour la topologie faible (puisque c'est un convexe).

Lemme. *Si A est un sous-espace vectoriel d'un espace de Banach X, alors ${}^\perp(A^\perp) = \overline{A}$.*

Démonstration. L'inclusion $A \subset {}^\perp(A^\perp)$ est immédiate ; puisqu'un préannihilateur est fermé on en déduit que $\overline{A} \subset {}^\perp(A^\perp)$.

Pour montrer l'inclusion réciproque, fixons $x \notin \overline{A}$. Par le théorème de séparation stricte de Hahn–Banach, il existe $f \in X^*$ telle que

$$f(x) > \sup_{\overline{A}} f$$

Comme $f(\overline{A})$ est un sous-espace majoré de $\mathbf{R}$, il est réduit à $\{0\}$. Ceci montre que $f \in A^\perp$ et donc que $x \notin {}^\perp(A^\perp)$, comme voulu. $\square$

Dans la suite du chapitre on suppose que X et Y sont deux espaces de BANACH.

Proposition. *Soient X et Y des espaces de* BANACH *et $T : X \to Y$ une application linéaire continue. Alors*

$$\ker(T^*) = \mathrm{Im}(T)^\perp \quad \text{et} \quad \ker(T) = {}^\perp\mathrm{Im}(T^*).$$

Démonstration. On écrit, pour $f \in Y^*$

$$T^*f = 0 \iff \forall x \in X, (T^*f)(x) = 0 \iff \forall x \in X, f(Tx) = 0 \iff f \in \mathrm{Im}(T)^\perp$$

et pour $x \in X$, en utilisant un corollaire du théorème de HAHN–BANACH pour la première équivalence

$$Tx = 0 \iff \forall f \in Y^*, f(Tx) = 0 \iff \forall f \in Y^*, (T^*f)(x) = 0 \iff x \in {}^\perp\mathrm{Im}(T^*) \quad \square$$

On dit qu'un opérateur $T \in \mathcal{L}(X, Y)$ est compact si $T(B_X)$ est relativement compact pour la topologie forte de Y. On note $\mathcal{K}(X, Y) \subset \mathcal{L}(X, Y)$ l'ensemble des opérateurs compacts.

Remarque. $T \in \mathcal{L}(X, Y)$ est compact si et seulement si, pour toute suite bornée (x_n), la suite (Tx_n) admet une sous-suite convergente.

Proposition. *Soient X et Y des espaces de* BANACH.

1. *Si $T \in \mathcal{L}(X, Y)$ est de rang fini, alors T est compact.*
2. *Si $T \in \mathcal{K}(X, Y)$ est tel que $\mathrm{Im}(T)$ est fermé, alors T est de rang fini.*
3. *$\mathcal{K}(X, Y)$ est un sous-espace fermé de $\mathcal{L}(X, Y)$.*

Démonstration.

1. L'espace vectoriel de dimension finie $\mathrm{Im}(T)$ est fermé dans Y car complet. Puisque $\overline{T(B_X)}$ est une partie fermée bornée d'un espace vectoriel normé de dimension finie, l'opérateur T est compact.
2. Considérons l'opérateur $\tilde{T} : X \to \mathrm{Im}(X)$ définie par $\tilde{T}(x) = T(x)$ pour $x \in X$. Puisque $\mathrm{Im}(X)$ est un sous-espace fermé de Y, c'est un espace de BANACH. L'opérateur $\tilde{T}$ est un opérateur surjectif entre espaces de BANACH. Par le théorème de l'application ouverte, $\tilde{T}(B_X)$ est d'intérieur non vide. L'espace $\mathrm{Im}\, T$ contient une partie compacte d'intérieur non vide, ce qui implique qu'il est de dimension finie (théorème de RIESZ).
3. Il est facile de voir que $0 \in \mathcal{K}(X, Y)$ et que $\mathcal{K}(X, Y)$ est stable par multiplication scalaire. Pour la stabilité par addition, soient S et T dans $\mathcal{K}(X, Y)$. On a

$$(S + T)(B_X) \subset S(B_X) + T(B_X) \subset \overline{S(B_X)} + \overline{T(B_X)}$$

Si K et L sont deux parties compactes d'un espace vectoriel topologique, leur somme $K + L$ est compacte (comme image du compact $K \times L$ par l'application continue $(x, y) \mapsto x + y$). Ainsi $(S + T)(B_X)$ est inclus dans le compact $\overline{S(B_X)} + \overline{T(B_X)}$, donc d'adhérence compacte. On a bien monté $S + T \in \mathcal{K}(X, Y)$.

Il reste à montrer que $\mathcal{K}(X, Y)$ est fermé dans $\mathcal{L}(X, Y)$. Pour cela, soit $T \in \overline{\mathcal{K}(X, Y)}$. Fixons $\varepsilon > 0$ et soit $S \in \mathcal{K}(X, Y)$ tel que $\|S - T\|_{\mathrm{op}} < \frac{\varepsilon}{2}$.

Pour montrer que $T(B_X)$ est d'adhérence compacte, puisque Y est complet, il suffit de montrer qu'il peut pour tout $\varepsilon > 0$ être recouvert par un nombre fini de boules de Y de rayon ε. Soit $S \in \mathcal{K}(X,Y)$ tel que $\|S-T\|_{op} < \frac{\varepsilon}{2}$. Puisque S est un opérateur compact, l'ensemble $S(B_X)$ est d'adhérence compacte. Il existe donc une partie finie $A \subset Y$ telle que

$$S(B_X) \subset \bigcup_{y \in A} B(y, \varepsilon/2).$$

Pour $x \in B_X$, soit $y \in A$ tel que $\|Sx - y\| < \varepsilon/2$. On a

$$\|Tx - y\| \leqslant \|Tx - Sx\| + \|Sx - y\| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$$

et donc

$$T(B_X) \subset \bigcup_{y \in A} B(y, \varepsilon).$$

On a montré que l'ensemble $T(B_X)$ est précompact, c'est-à-dire que pour tout $\varepsilon > 0$ il peut être recouvert par un nombre fini de boules de rayon ε. Puisque Y est complet, cela implique que $T(B_X)$ est d'adhérence compacte, d'où $T \in \mathcal{K}(X,Y)$. □

Fin cours # 10 du 15 avril

Remarquons également que l'ensemble des opérateurs compacts a la *propriété d'idéal* : une composition d'opérateurs est compacte dès lors que l'un des facteurs est compact (exercice).

Théorème (Schauder)**.** *Soit $T \in \mathcal{L}(X,Y)$. Alors T est compact si et seulement si T^* est compact.*

Démonstration. Supposons T compact. Pour montrer que T^* est compact, il suffit de montrer que pour toute suite (f_n) dans B_{Y^*}, la suite $(T^* f_n)$ admet une sous-suite convergente.

Soit $K = \overline{T(B_X)}$; c'est une partie compacte de Y. Notons F_n la restriction de f_n à K. Pour tout n, la fonction $f_n \in C(K)$ est 1-lipschitzienne et vérifie $F_n(0) = 0$. La suite (F_n) est donc bornée et équicontinue dans $C(K)$. Par le théorème d'Ascoli, il existe une sous-suite $(F_{\sigma(n)})$ qui converge uniformément sur K. Pour tous m et n, on a

$$\|T^* f_{\sigma(n)} - T^* f_{\sigma(m)}\|_{X^*} = \sup_{x \in B_X} |T^*(f_{\sigma(n)} - f_{\sigma(m)})(x)| = \sup_{T(B_X)} |f_{\sigma(n)} - f_{\sigma(m)}|,$$

ce qui montre que la suite $(T^* f_{\sigma(n)})$ est de Cauchy, donc converge car X^* est complet.

Supposons maintenant T^* compact. Par ce qu'on vient de montrer, l'opérateur $T^{**} : X^{**} \to Y^{**}$ est compact, ainsi que $T^{**} \circ J_X$ par la propriété d'idéal. Puisque $J_Y \circ T = T^{**} \circ J_X$ (exercice), on obtient que $T(B_X)$ est d'adhérence compacte dans Y^{**}, donc aussi dans Y. □

8.2 Alternative de Fredholm

Le théorème suivant affirme que les perturbations compactes de l'identité se comportent de manières analogues aux endomorphisme d'un espace vectoriel de dimension finie.

Théorème (Alternative de Fredholm)**.** *Soit X un espace de Banach, $T \in \mathcal{K}(X,X)$ et $\lambda \neq 0$. Alors*

1. $\ker(\lambda \mathrm{Id} - T)$ *est de dimension finie*

2. $\mathrm{Im}(\lambda\mathrm{Id} - T)$ *est fermé et égal à* ${}^{\perp}\ker(\lambda\mathrm{Id} - T^*)$*.*

3. On a l'équivalence

$$\lambda\mathrm{Id} - T \text{ injectif } \iff \lambda\mathrm{Id} - T \text{ surjectif}$$

Démonstration. Quitte à remplacer T par $\lambda^{-1}T$, on peut supposer $\lambda = 1$.

1. Soit $E = \ker(\mathrm{Id} - T)$; c'est un sous-espace fermé L'opérateur $T_{|E} : E \to X$ est compact car T est compact. Mais il coïncide avec l'identité sur E, et donc la boule unité de E est compacte, ce qui implique que E est de dimension finie.
2. Soit (y_n) une suite de $\mathrm{Im}(\mathrm{Id} - T)$ qui converge vers y. On considère

$$E_n = \{x \in X \ : \ (\mathrm{Id} - T)x = y_n\}.$$

C'est un sous-espace affine fermé. Soit $d_n = \inf\{\|x\| \ : \ x \in E_n\}$ et $x_n \in E_n$ tel que $\|x_n\| \leqslant 2d_n$.[1] On a donc $y_n = x_n - Tx_n$. Remarquons que pour tout $z \in \ker(\mathrm{Id} - T)$, on a $x_n - d_n z \in E_n$ et donc $\|x_n/d_n - z\| \geqslant 1$.

Montrons d'abord que la suite (x_n) (ou de manière équivalente la suite (d_n)) est bornée. Si elle ne l'était pas, en utilisant la compacité de T, on pourrait trouver une sous-suite telle que $d_{\sigma(n)} \to \infty$ et $T(\frac{x_{\sigma(n)}}{d_{\sigma(n)}}) \to z$. En passant à la limite dans l'équation $y_n = x_n - Tx_n$, on en déduit que $\lim \frac{x_{\sigma(n)}}{d_{\sigma}(n)} = z$ et donc $Tz = z$, c'est-à-dire $z \in \ker(\mathrm{Id} - T)$. On a donc $\|x_n/d_n - z\| \geqslant 1$, contradiction.

Puisque (x_n) est bornée et T compact, il existe une sous-suite $(Tx_{\sigma(n)})$ qui converge convergente vers z. L'équation

$$y_{\sigma(n)} = x_{\sigma(n)} - Tx_{\sigma(n)}$$

montre que $(x_{\sigma(n)})$ converge vers $y + z$. Si on pose $x = y + z$, on a $Tx = z$ puisque T est continue, et donc $y = (\mathrm{Id} - T)(x)$ ce qui montre que l'image de $\mathrm{Id} - T$ est fermée.
3. Supposons $\mathrm{Id} - T$ injectif et (par l'absurde) non surjectif. Posons $X_n = \mathrm{Im}(\mathrm{Id} - T)^n$. La suite (X_n) est une suite décroissante de sous-espaces fermés (en effet il découle de la propriété d'idéal que l'on opérateur T_n défini par la relation $(\mathrm{Id} - T)^n = \mathrm{Id} - T_n$ est compact ; par conséquent $(\mathrm{Id} - T)^n$ est d'image fermée). Ils sont tous distincts : puisqu'il existe $x_0 \in X \setminus \mathrm{Im}(\mathrm{Id} - T)$, il vient que $(\mathrm{Id} - T)^n x_0 \in X_n \setminus X_{n-1}$ (injectivité...). Par le lemme de Riesz, on peut trouver $e_n \in X_n$ vérifiant $\|e_n\| = 1$ et $d(e_n, X_{n+1}) \geqslant 1/2$. Pour $n > m$ on écrit alors

$$Te_m - Te_n = e_m + (\mathrm{Id} - T)(e_n - e_m) - e_n \in e_m + X_{m+1}$$

et donc $\|Te_m - Te_n\| \geqslant 1/2$, ce qui contredit la compacité de T.

Pour la réciproque, on utilise les formules précédemment démontrées

$$\ker(T^* - \lambda\mathrm{Id}) = \mathrm{Im}(T - \lambda\mathrm{Id})^{\perp} \tag{8.1}$$

$$\ker(T - \lambda\mathrm{Id}) = {}^{\perp}\,\mathrm{Im}(T^* - \lambda\mathrm{Id}) \tag{8.2}$$

Si $\mathrm{Id} - T$ est surjectif, alors $\mathrm{Id} - T^*$ est injectif d'après (8.1). Puisque T^* est compact (théorème de Schauder), par le paragraphe précédent, $\mathrm{Id} - T^*$ est surjectif. Mais alors (8.2) implique que $\mathrm{Id} - T$ est injectif.

1. En réalité, la borne inférieure est atteinte car E_n est de dimension finie par le point 1 du théorème, donc localement compact

$\square$

Rappelons le lemme de Riesz qui a été utilisé dans la preuve.

Lemme (Lemme de RIESZ)**.** *Soit X un espace de BANACH et $E \subset X$ un sous-espace fermé. Il existe un vecteur $x \in X \setminus E$ tel que $\|x\| = 1$ et $d(x, E) \geqslant 1/2$.*

Démonstration. Soit $x_0 \in X \setminus E$ arbitraire. Comme E est fermé, on a $d(x_0, E) > 0$. Soit $y_0 \in E$ tel que $\|x_0 - y_0\| \leqslant 2d(x_0, E)$. Posons $x = \frac{x_0 - y_0}{\|x_0 - y_0\|}$. Ce choix convient : pour tout $y \in E$,

$$\|x - y\| = \left\| \frac{x_0 - y_0}{\|x_0 - y_0\|} - y \right\| \geqslant \frac{d(x_0, E)}{\|x_0 - y_0\|} \geqslant \frac{1}{2}$$

$\square$

8.3 Spectre d'un opérateur

Dans cette partie, les espaces de BANACH considérés sont sur le corps des nombres complexes.

Soit X un espace de BANACH. Rappelons qu'un opérateur $T \in \mathcal{L}(X)$ est *inversible* s'il existe $S \in \mathcal{L}(X)$ tel que $ST = TS = \mathrm{Id}_X$. Il découle du théorème d'isomorphisme de BANACH qu'un opérateur est inversible si et seulement si il est injectif et surjectif.

Soit $T \in \mathcal{L}(X)$ on appelle *spectre* de T l'ensemble

$$\sigma(T) = \{\lambda \in \mathbf{C} \ : \ \lambda \mathrm{Id} - T \text{ n'est pas inversible}\}.$$

Si X est de dimension finie, le spectre coïncide avec l'ensemble des valeurs propres. En dimension infinie, si λ est une valeur propre de T, alors $\lambda\mathrm{Id} - T$ n'est pas injectif et donc $\lambda \in \sigma(T)$; mais il existe des éléments du spectre qui ne sont pas des valeurs propres.

Théorème. *Soit X un espace de BANACH de dimension infinie et $T \in \mathcal{K}(X, X)$. Alors*

1. $0 \in \sigma(T)$,
2. $\sigma(T) \setminus \{0\}$ *est l'ensemble des valeurs propres non nulles de T, qui sont toutes de multiplicité finie.*
3. $\sigma(T) \setminus \{0\}$ *est soit un ensemble fini, soit un ensemble infini dénombrable dont le seul point d'accumulation est 0.*

Fin cours # 11 du 16 avril

Démonstration.
1. Un opérateur compact T ne peut pas être inversible, sinon $\mathrm{Id} = TT^{-1}$ serait compact par propriété d'idéal, et donc la boule unité serait une partie compacte, ce qui est impossible en dimension infinie.
2. Si $\lambda \in \sigma(T)$, alors $\lambda\mathrm{Id} - T$ n'est pas inversible, donc non injectif par l'alternative de FREDHOLM, c'est-à-dire que λ est une valeur propre de T.
3. Sinon, il existerait une suite (λ_n) de nombres complexes deux à deux distincts vérifiant $|\lambda_n| \geqslant r$ pour un $r > 0$. Soit pour tout n un vecteur $x_n \neq 0$ tel que $Tx_n = \lambda_n x_n$. La famille (x_n) est libre. Soit $E_n = \mathrm{vect}(x_1, \ldots, x_n)$. Par le lemme de RIESZ, il existerait pour tout n un vecteur $y_n \in E_n$ tel que $\|y_n\| = 1$ et $d(y_n, E_{n-1}) \geqslant 1/2$. Remarquons que $Ty_n - \lambda_n y_n \in E_{n-1}$.

   La suite (Ty_n/λ_n) est incluse dans la partie compacte $r^{-1}T(B_X)$, donc admet une sous-suite convergente. Nous allons montrer que c'est absurde car aucune sous-suite n'est de CAUCHY : en effet, pour $n > m$, on a

$$\frac{Ty_n}{\lambda_n} - \frac{Ty_m}{\lambda_m} = y_n + \frac{Ty_n - \lambda_n y_n}{\lambda_n} - \frac{Ty_m}{\lambda_m} \in y_n + E_{n-1}$$

et donc $\left\|\frac{Ty_n}{\lambda_n} - \frac{Ty_m}{\lambda_m}\right\| \geqslant 1/2$.

$\square$

Lemme. *Soit* $A \in \mathcal{L}(X)$ *un opérateur tel que* $\alpha = \|A\|_{op} < 1$. *Alors* $\mathrm{Id} - A$ *est inversible et* $\|(\mathrm{Id} - A)^{-1} - \mathrm{Id}\|_{op} \leqslant \frac{\alpha}{1-\alpha}$.

Démonstration. Posons $\alpha = \|A\|_{op}$. Puisque $\|A^n\|_{op} \leqslant \alpha^n$, la série $\sum A^n$ converge normalement dans $\mathcal{L}(X)$; soit S sa somme. On a pour tout N

$$(\mathrm{Id} - A)\left(\sum_{n=0}^{N} A^n\right) = \left(\sum_{n=0}^{N} A^n\right)(\mathrm{Id} - A) = \mathrm{Id} - A^{N+1}$$

En faisant $N \to \infty$, par continuité de la composition dans $\mathcal{L}(X)$, on a $(\mathrm{Id} - A)S = S(\mathrm{Id} - A) = \mathrm{Id}$. La dernière inégalité découle de $\|S - \mathrm{Id}\| \leqslant \sum_{n=1}^{\infty} \alpha^n$. $\square$

Proposition. *Soit* $G \subset \mathcal{L}(X)$ *l'ensemble des opérateurs inversibles.*

1. *L'ensemble* G *est ouvert.*
2. *L'application* $T \mapsto T^{-1}$ *est continue de* G *dans* G
3. *(dans le cas complexe) Pour tous* $S, T \in \mathcal{L}(X)$ *et* $\phi \in \mathcal{L}(X)^*$, *l'application*

$$\lambda \in \mathbf{C} \mapsto \phi((S + \lambda T)^{-1})$$

est holomorphe sur son domaine de définition

Démonstration.

1. Remarquons que G est un groupe pour la composition. Si $T \in G$ et $\|H\| < \|T^{-1}\|^{-1}$, alors $\|T^{-1}H\| < 1$. Donc $\mathrm{Id} - T^{-1}H = T^{-1}(T - H)$ est inversible, puis $T - H$ aussi. Ceci montre que G est ouvert.
2. On écrit
$$\|(T + H)^{-1} - T^{-1}\| = \|T^{-1}\| \cdot \|(\mathrm{Id} + HT^{-1})^{-1} - \mathrm{Id}\|$$
qui tend vers 0 quand $\|H\| \to 0$ grâce au lemme.
3. Soit $A \subset \mathbf{C}$ le domaine de définition, qui est ouvert par le 1. Pour λ, μ dans A, on a
$$(S + \lambda T)^{-1} - (S + \mu T)^{-1} = (\mu - \lambda)(S + \lambda T)^{-1}T(S + \mu T)^{-1},$$
ce qui montre que la limite
$$\lim_{\mu \to \lambda} \frac{\phi((S + \lambda T)^{-1}) - \phi(S + \mu T)^{-1})}{\lambda - \mu}$$
existe, d'où le côté holomorphe.

$\square$

C'est pour le théorème suivant qu'il est important que X soit complexe.

Théorème. *Soit* $T \in \mathcal{L}(X)$. *Alors* $\sigma(T)$ *est un compact non vide de* $\mathbf{C}$, *contenu dans le disque fermé de centre* 0 *et de rayon* $\|T\|$.

Démonstration. Le fait que $\sigma(T)$ est fermé découle du fait que G est ouvert. Si $\lambda > \|T\|$, alors $\mathrm{Id} - \lambda^{-1}T$ est inversible par le lemme, donc $\lambda \notin \sigma(T)$.

Il reste à montrer que $\sigma(T)$ est non vide. Si on avait $\sigma(T) = \emptyset$, alors T serait inversible. Soit $\phi \in \mathcal{L}(X)^*$ telle que $\phi(T^{-1}) \neq 0$, donnée par un corollaire de Hahn–Banach. La fonction

$$\lambda \mapsto \phi((\lambda\mathrm{Id} - T)^{-1})$$

est une fonction entière vérifiant $\phi(0) \neq 0$ et de limite 0 en l'infini. Ceci contredit le théorème de Liouville. $\square$

8.4 Formule de Gelfand

On appelle *rayon spectral* d'un élément $T \in \mathcal{L}(X)$ la quantité

$$\rho(T) = \sup_{\lambda \in \sigma(T)} |\lambda|.$$

Lemme. *Soit $P \in \mathbf{C}[X]$ et $T \in \mathcal{L}(X)$. Alors*

$$\sigma(P(T)) = \{P(\lambda) \ : \ \lambda \in \sigma(T)\}$$

Démonstration. Si $P = 0$ c'est évident ; sinon on écrit $P(X) - \lambda = \alpha \prod_i (X - \lambda_i)$ avec $\alpha \neq 0$. On a alors $P(T) - \lambda \mathrm{Id} = \alpha \prod_i (T - \lambda_i \mathrm{Id})$. Puisque les facteurs de ce produit comment, le produit est inversible si et seulement si chaque facteur est inversible (exercice). On a donc

$$\lambda \in \sigma(P(T)) \iff \exists i \ : \ \lambda_i \in \sigma(T) \iff \exists \alpha \in \mathbf{C} \ : \ P(\alpha) = \lambda,$$

d'où le résultat. □

Théorème (Formule de Gelfand)**.** *Pour tout $T \in \mathcal{L}(X)$, on a*

$$\rho(T) = \lim_{n \to \infty} \|T^n\|^{1/n}.$$

Démonstration. Il découle du lemme précédent que $\rho(T^n) = \rho(T)^n$. On a donc

$$\rho(T) = \rho(T^n)^{1/n} \leqslant \|T^n\|^{1/n}$$

et donc

$$\rho(T) \leqslant \liminf_{n \to \infty} \|T^n\|^{1/n}.$$

Si $z \in \mathbf{C}$ est tel que $z\|T\| < 1$, alors

$$(\mathrm{Id} - zT)^{-1} = \sum_{n \geqslant 0} z^n T^n.$$

Pour tout $\phi \in \mathcal{L}(X)^*$, on a donc

$$\phi((\mathrm{Id} - zT)^{-1}) = \sum_{n \geqslant 0} z^n \phi(T^n).$$

Le terme de gauche définit une fonction holomorphe sur l'ouvert $\{z \in \mathbf{C} \ : \ 1/z \notin \sigma(T)\}$ qui contient le disque ouvert de rayon $1/\rho(T)$. Le rayon de convergence R de la série entière du terme de droite vérifie donc $R \geqslant 1/\rho(T)$. Or on a $\frac{1}{R} = \limsup |\phi(T^n)|^{1/n}$, c'est-à-dire

$$\limsup_{n \to \infty} |\phi(T^n)|^{1/n} \leqslant \rho(T).$$

Soit $\varepsilon > 0$ et posons $S = \frac{T}{\rho(T) + \varepsilon}$. Soit J l'injection canonique de $\mathcal{L}(X)$ dans son bidual. Pour tout $\phi \in \mathcal{L}(X)^{**}$, on a $\lim J(S^n)(\phi) = \lim \phi(S^n) = 0 =$ et donc la famille $\{J(S^n) \ : \ n \in \mathbf{N}\}$ est ponctuellement bornée dans $\mathcal{L}(X)^{**}$. Par le théorème de Banach–Steinhaus, elle est bornée en norme : il existe une constante C telle que $\|T^n\| \leqslant C(\rho(T) + \varepsilon)^n$. Il s'ensuit que $\limsup \|T^n\|^{1/n} \leqslant \rho(T) + \varepsilon$ et il suffit de faire tendre ε vers 0. □

Fin cours # 12 du 29 avril

8.5 Opérateurs auto-adjoints compacts

Théorème. *Soit $T \in \mathcal{K}(H)$ un opérateur compact et auto-adjoint. Il existe une base hilbertienne de H formée de vecteurs propres de T.*

Dans la fin du cours on suppose que H est un espace de Hilbert (dans le cas complexe, le produit scalaire d'un espace de Hilbert vérifie la propriété de sesquilinéarité $\langle x, \alpha y\rangle = \alpha\langle x, y\rangle$ et $\langle \alpha x, y\rangle = \overline{\alpha}\langle x, y\rangle$ pour x et y dans H, et α dans $\mathbf{C}$). On dit que $T \in \mathcal{L}(H)$ est auto-adjoint s'il vérifie la relation

$$\langle x, Ty\rangle = \langle Tx, y\rangle$$

pour tous x et y dans H.

Lemme. *Soit $T \in \mathcal{L}(H)$ un opérateur auto-adjoint. Alors $\sigma(T) \subset \mathbf{R}$.*

Démonstration. Montrons d'abord que les valeurs propres de T sont réelles : si $Tx = \lambda x$ pour $\lambda \in \mathbf{C}$ et $x \in H$ non nul, alors

$$\overline{\lambda}\|x\|^2 = \langle Tx, x\rangle = \langle x, Tx\rangle = \lambda\|x\|^2$$

donc $\lambda \in \mathbf{R}$.

Soit $\lambda \in \mathbf{C} \setminus \mathbf{R}$, écrit $\lambda = \alpha + i\beta$ avec $\alpha \in \mathbf{R}$ et $\beta \in \mathbf{R}^*$. Montrons que $\lambda\mathrm{Id} - T$ est inversible. On sait déjà qu'il est injectif car $\lambda \notin \mathbf{R}$. Soit E l'image de $\lambda\mathrm{Id} - T$. On calcule pour tout $x \in H$

$$\langle x, (\lambda\mathrm{Id} - T)x\rangle - \langle(\lambda\mathrm{Id} - T)x, x\rangle = (\lambda - \overline{\lambda})\|x\|^2 = 2\,\mathrm{Im}(\lambda)\|x\|^2$$

et

$$|\langle x, (\lambda\mathrm{Id} - T)x\rangle - \langle(\lambda\mathrm{Id} - T)x, x\rangle| \leqslant 2\|(\lambda\mathrm{Id} - T)x\| \cdot \|x\|$$

d'où on déduit l'identité $\|(\lambda\mathrm{Id} - T)x\| \geqslant \|x\|$. Il s'ensuit que E est fermée : si une suite $y_n = (\lambda\mathrm{Id} - T)x_n$ converge, alors (x_n) est de Cauchy donc converge vers x, et (y_n) converge vers $(\lambda\mathrm{Id} - T)x \in E$. Enfin, si $z \perp E$, on a pour tout $x \in \mathbf{R}$

$$0 = \langle z, (\lambda\mathrm{Id} - T)x\rangle = \langle(\overline{\lambda}\mathrm{Id} - T)z, x\rangle$$

d'où on déduit $(\overline{\lambda}\mathrm{Id} - T)z = 0$ puis $z = 0$ car $\overline{\lambda}\mathrm{Id} - T$ est injectif. Ainsi $E = H$, l'opérateur $\lambda\mathrm{Id} - T$ est bijectif et donc $\lambda \notin \sigma(T)$ □

Lemme. *Si T est auto-adjoint, alors $\rho(T) = \|T\|$.*

Démonstration. On a $\|T^2\| \leqslant \|T\|^2$, mais également pour tout $x \in H$,

$$\|Tx\|^2 = \langle Tx, Tx\rangle = \langle T^2x, x\rangle \leqslant \|T^2x\| \cdot \|x\| \leqslant \|T^2\| \cdot \|x\|^2,$$

d'où $\|T\|^2 \leqslant \|T^2\|$ en prenant la borne supérieure pour x dans la boule unité. On a donc $\|T^2\| = \|T\|^2$. Il s'ensuit par récurrence sur k (et parce que les puissances d'un opérateur auto-adjoint sont auto-adjointes) que $\|T^{2^k}\| = \|T\|^{2^k}$. Par la formule de Gelfand,

$$\rho(T) = \lim_{k\to\infty} \|T^{2^k}\|^{1/2^k} = \|T\|.$$ □

Théorème. *Soit $T \in \mathcal{K}(H)$ un opérateur compact et auto-adjoint. Il existe une base hilbertienne de H formée de vecteurs propres de T.*

Démonstration. On sait déjà que l'on peut écrire

$$\sigma(T) = \{\lambda_n \ : \ n \in E\} \subset \mathbf{R}$$

avec $I = \mathbf{N}$ ou $I = \{0, \ldots, N\}$, avec de plus $\lambda_0 = 0$. Notons $E_n = \ker(\lambda_n \mathrm{Id} - T)$. Montons que les sous-espaces (E_n) sont deux à deux orthogonaux : si $x \in E_m$ et $y \in E_n$ pour $m \neq n$, alors

$$\lambda_m \langle x, y\rangle = \langle Tx, y\rangle = \langle x, Ty\rangle = \lambda_n \langle x, y\rangle.$$

Comme $\lambda_m \neq \lambda_n$, cela implique $\langle x, y\rangle = 0$.

Soit $F = \bigoplus_{n \in I \setminus \{0\}} E_n$. On a $E_0 \subset F^\perp$, montrons que l'on a égalité. Puisque F est stable par T qui est auto-adjoint, c'est aussi le cas de $F^\perp$. L'opérateur $S : F^\perp \to F^\perp$ induit par T et un opérateur compact auto-adjoint dont la seule valeur propre est 0. On a donc $\rho(S) = 0$ puis $S = 0$ par le lemme précédent. Ceci montre bien que $F^\perp \subset E_0$.

Pour terminer, il suffit de prendre pour chaque n une base hilbertienne de E_n ; la réunion de ces bases est une base hilbertienne de H formée de vecteurs propres de T. $\square$

Fin du cours

8.6 Calcul fonctionnel continu

Proposition (Calcul fonctionnel continu)**.** *Soit T un opérateur auto-adjoint et $K = \sigma(T) \subset \mathbf{R}$. Il existe une unique application linéaire $\tau : C(K) \to \mathcal{L}(H)$ telle que*

1. $\tau(1_K) = \mathrm{Id}$
2. $\tau(x \mapsto x) = T$
3. $\tau(fg) = \tau(f)\tau(g)$
4. *τ est une isométrie*