Analyse Fonctionnelle 2
M1 Mathématiques Générales

Guillaume Aubrun

Année universitaire 2025–2026

Chapitre 1

Le théorème de BAIRE

1.1 Le théorème de BAIRE

On considère un espace métrique complet (X, d). Étant donnés $x \in X$ et $r > 0$, on note $B_f(x, r)$ la boule fermée de centre x et de rayon r. On note également $B(x, r)$ la boule ouverte de mêmes centre et rayon.

Théorème (Théorème de BAIRE)**.** *Soit $(F_n)_n$ une suite de fermés d'intérieur vide d'un espace métrique complet (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 $x_n \in B_n$ et observer que la suite (x_n) est de CAUCHY donc converge vers $x \in \bigcap B_n$). L'ensemble non vide $\bigcap B_n$ est un singleton puisque son diamètre vaut 0.

Exercice. Montrer qu'un espace métrique qui vérifie la propriété des fermés emboîtés est complet.

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

Théorème (Théorème de BAIRE)**.** *Soit $(O_n)_n$ une suite d'ouverts denses d'un espace métrique complet (X, d). Alors $\bigcap O_n$ est dense.*

Montrons-le sous cette forme, équivalente à la précédente en considérant $F_n = X \setminus O_n$. Une partie A d'un espace métrique X est dense si et seulement si tout ouvert non vide de X intersecte A. Cette remarque implique, dans tout espace métrique, 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 par récurrence 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, donc contient une boule $B_0 = B_f(x_0, r_0)$ avec $0 < r_0 \leqslant 1$.
1. Puisque O_1 est dense et $B'_0 \cap \Omega$ est un ouvert non vide, l'ouvert $B'_0 \cap \Omega \cap O_1$ est non vide, donc contient une boule $B_1 = B_f(x_1, r_1)$ avec $0 < r_1 \leqslant 1/2$.

$\vdots$

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

$\vdots$

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 $\bigcap O_n$. $\square$

Fixons quelques points de terminologie

- Étant donnée une partie non vide $A \subset X$, on note $d(x, A) = \inf_{y \in A} d(x, y)$. La fonction $x \mapsto d(x, A)$ est une fonction continue (et même 1-lipschitzienne) de X dans $\mathbf{R}$, qui s'annule exactement sur $\overline{A}$.
- On dit qu'une partie $A \subset X$ est *maigre* si elle est incluse dans une réunion dénombrable de fermés d'intérieur vide.
- Dans un espace métrique complet, 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.

Remarquons que tout fermé F est un G_δ puisqu'on peut écrire $F = \bigcap O_n$ avec $O_n = \{x \in X \ : \ d(x, F) < 2^{-n}\}$.

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.

1.2 Fonctions de première classe

On dit qu'une fonction $f : \mathbf{R} \to \mathbf{R}$ est *de première classe* si elle est 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 par la formule $F = \bigcap_{k \in \mathbf{N}} O_k$ 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 F 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 $d(f_{n_k}(x), F) < 2^{-k}$. En prenant la limite $k \to \infty$, on obtient $d(f(x), F) = 0$ et donc $f(x) \in F$.

□

Preuve du théorème. Soit $(V_n)_{n\in\mathbf{N}}$ 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

$$\mathbf{R} \setminus C = \bigcup_{n\in\mathbf{N}} f^{-1}(V_n) \setminus \mathrm{int}[f^{-1}(V_n)].$$

Par le lemme, $f^{-1}(V_n)$ est un F_σ. Remarquons que si A est un F_σ, alors $A \setminus \mathrm{int}(A)$ est un F_σ d'intérieur vide, donc est maigre. Ainsi, on a écrit $\mathbf{R}\setminus C$ comme réunion dénombrable de parties maigres. Par le théorème de Baire, l'ensemble $\mathbf{R} \setminus C$ est maigre et donc l'ensemble C est comaigre. □

Chapitre 2

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. Dans le cours tous les espaces de BANACH seront sur le corps $\mathbf{K} = \mathbf{R}$ ou $\mathbf{K} = \mathbf{C}$ (on rappelle qu'un espace de BANACH est un espace vectoriel normé complet). Un cas particulier important est le cas des espaces de HILBERT : ce sont les espaces de BANACH pour lesquels la norme peut s'écrire à partir d'un produit scalaire.

2.1 Le théorème de BANACH–STEINHAUS

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 $Y = \mathbf{K}$, on trouve l'espace dual de X, noté X^*. La norme sur $X^* = \mathcal{L}(X, \mathbf{K})$ 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 propriétés 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 $B(x_0, \varepsilon)$, étant symétrique et convexe il contiendrait $B(0, \varepsilon)$, d'où pour tout $T \in \Phi$ l'implication $\|x\|_X \leqslant \varepsilon \Longrightarrow \|Tx\|_Y \leqslant n$ 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. □

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.

Il existe donc (une partie comaigre de X n'étant pas vide) des fonctions continues dont la série de Fourier ne converge pas en 0.

Démonstration. Pour $f \in X$, posons $\Lambda_n(f) = S_n f(0)$. Nous allons montrer le résultat suivant

Lemme. *Pour tout n, $\Lambda_n : X \to \mathbf{R}$ est une forme linéaire continue. 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^*, le théorème de Banach–Steinhaus permet de conclure que l'ensemble

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

est comaigre dans X. Le théorème en découle puisque cet ensemble contient toute fonction

$$\{f \in X \ : \ \lim \Lambda_n(f) = f(0)\}$$

(car si la suite $(\Lambda_n(f))$ converge, elle est bornée).

Pour démontrer le lemme, on introduit le noyau de Dirichlet, pour $t \in \mathbf{R}$ (et $t \notin 2\pi\mathbf{Z}$ après la seconde égalité)

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

La fonction D_n est paire. Par ailleurs, on a pour tout $t \in \mathbf{R}$

$$\begin{aligned} (S_n f)(t) &= \sum_{k=-n}^{n} \hat{f}(k)e^{ikt} \\ &= \sum_{k=-n}^{n} \int_{-\pi}^{\pi} f(s)e^{ik(t-s)}\,\mathrm{d}s \\ &= \frac{1}{2\pi}\int_{-\pi}^{\pi} f(s)D_n(t-s)\,\mathrm{d}s \end{aligned}$$

On en déduit l'inégalité

$$|\Lambda_n(0)| \leqslant \|f\|_\infty \frac{1}{2\pi}\int_{-\pi}^{\pi} |D_n(s)|\,\mathrm{d}s$$

qui montre que la forme linéaire Λ_n est continue et que

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

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 $+\infty$ quand n tend vers l'infini (exercice). $\square$

2.2 Le théorème de l'application ouverte

Soient X, Y deux espaces métriques. 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.

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 vectoriel ouvert.

Fin cours # 1 du 13 janvier

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

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

De même, la condition

$$\overline{T(B(0, 1))} \supset B(0, r)$$

équivaut à la condition

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

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 $r = 1/2\lambda$. $\square$

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 : $(X, \|\cdot\|_2) \to (X, \|\cdot\|_1)$. □

2.3 Le théorème du graphe fermé

Si $T : X \to Y$ est une application entre des ensembles X et Y, son graphe $G(T) \subset X \times Y$ est défini par

$$G(T) = \{(x, Tx) \ : \ x \in X\}.$$

On a le résultat suivant.

Proposition. *Soient X et Y des espaces métriques 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 $f : \mathbf{R} \to \mathbf{R}$ définie par $f(x) = 1/x$ si $x \neq 0$ et $f(0) = 0$ a un graphe fermé bien qu'elle ne soit pas continue.

Si X et Y sont des espaces vectoriels et si $T : X \to Y$ 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 dans $X \times Y$.*

Démonstration. Une implication est couverte par la proposition. Pour l'autre implication, 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 continue. □

Chapitre 3

Les espaces de BANACH classiques

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

3.1 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

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

On peut montrer (c'est un exercice plutôt difficile) qu'un espace de BANACH X est un espace de HILBERT si et seulement si il vérifie l'identité du parallélogramme : pour tous x, y dans X on a

$$\|x+y\|^2 + \|x-y\|^2 = 2\left(\|x\|^2 + \|y\|^2\right)$$

On dit qu'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 qui permet d'identifier canoniquement un espace de HILBERT à son dual.

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.

Un autre point important est qu'il existe toujours une application isométrique surjective entre deux espaces de HILBERT séparables de dimension infinie. On utilise le fait que tout espace de HILBERT séparable de dimension infinie admet une base hilbertienne $(x_n)_{n\in\mathbf{N}}$.

Théorème. *Soit H un espace de* HILBERT *admettant une base hilbertienne (x_n) et $\ell^2(\mathbf{N})$ l'espace de Hilbert formé des suites de carré sommable. L'application linéaire de H dans $\ell^2(\mathbf{N})$ qui à $x \in H$ associe la suite $(\langle x, x_n\rangle)_{n\in\mathbf{N}}$ est isométrique et surjective.*

Démonstration. Cette application est isométrique par l'égalité de Parseval. Montrons qu'elle est surjective. Soit une suite $(\lambda_n)_n \in \ell^2(\mathbf{N})$. Puisque pour $p \leqslant q$

$$\left\| \sum_{n=p}^{q} \lambda_n x_n \right\| = \left(\sum_{n=p}^{q} \lambda_n^2 \right)^{1/2},$$

la suite des sommes partielles de la série $\sum \lambda_n x_n$ est de Cauchy, donc converge. On peut donc définir un élément $x \in H$ comme $x = \sum \lambda_n x_n$ (on prendra garde au fait qu'il n'y a pas forcément convergence normale). Par continuité de $y \mapsto \langle y, x \rangle$, on a $\langle x, x_n \rangle = \lambda_n$. On a obtenu la surjectivité voulue. □

La notion d'orthogonalité entre vecteurs a un sens dans un espace de Hilbert et permet de définir la notion de projection orthogonale. Si $Y \neq \{0\}$ est un sous-espace fermé d'un espace de Hilbert, il existe une projection d'image Y qui est continue et de norme 1 : la projection orthogonale sur Y. Ce n'est pas le cas dans un espace de Banach général.

3.2 Les 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 $\mathcal{L}^p(\mu)$ pour la relation d'équivalence donnée par l'égalité 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'équivalence 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.

3.2.1 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

$$\nu(A) = \int_A f \, \mathrm{d}\mu \tag{3.1}$$

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 (3.1) 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

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

Si h est étagée, cela se déduit par linéarité de (3.1) ; 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 (un espace mesuré $(\Omega, \mathcal{F}, \mu)$ est dit σ-fini si l'espace Ω peut s'écrire comme réunion dénombrable de parties mesurables de mesure finie).

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 mesurable 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 de 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

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

ce qui se réécrit en

$$\int_\Omega g(1-h)\,\mathrm{d}\nu = \int_\Omega gh\,\mathrm{d}\mu. \tag{3.2}$$

On cherche f telle que $\mathrm{d}\nu = f\mathrm{d}\mu$. En injectant cela dans (3.2) cela suggère que $(1 - h)f = h$ et donc $f = \frac{h}{1-h}$. 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

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

Posons $N = \{h < 0\}$. On a $0 \leqslant \nu(N) = \int_N h \,\mathrm{d}\pi \leqslant 0$, d'où on tire $\pi(N) = 0$ donc $\mu(N) = \nu(N) = 0$. 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 (3.2) 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 $\nu(A) = \nu(A \cap G) + \nu(A \cap B)$. □

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

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 $1/p + 1/q = 1$.

Lemme. *Soient $(\Omega, \mathcal{F}, \mu)$ un espace mesuré σ-fini et p, q des exposants conjugués. Pour toute fonction $g \in L^q(\mu)$, 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 en demandant que toute partie mesurable de mesure infinie contient un sous-ensemble mesurable de mesure finie non nulle (ce sera clair dans la preuve ci-dessous). 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. On peut supposer $\|g\|_{L^q} = 1$. L'inégalité $\geqslant$ découle de l'inégalité de Hölder. Pour montrer l'autre inégalité, On distingue plusieurs cas

1. 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 donc

$$\int gh \,\mathrm{d}\mu = \|g\|_{L^q} = \|h\|_{L^p} = 1$$

   et ce choix de h montre l'égalité voulue.
2. Si $q = 1$ et $p = \infty$, c'est pareil en choisissant $h = \mathrm{signe}(g)$.
3. Le cas $q = \infty$ et $p = 1$ est un peu plus compliqué. Soit $\varepsilon > 0$. Par définition de la norme $\|\cdot\|_\infty$, l'ensemble $\{|g| \geqslant 1 - \varepsilon\}$ est de mesure non nulle et contient donc un sous-ensemble $A \in \mathcal{F}$ tel que $0 < \mu(A) < \infty$—c'est là où on utilise l'hypothèse de σ-finitude. Soit $h = \frac{\mathrm{signe}(g)\mathbf{1}_A}{\mu(A)}$; on a alors $\|h\|_{L^1} = 1$ et $\int gh \,\mathrm{d}\mu \geqslant 1 - \varepsilon$, donc le membre de droite est $\geqslant 1 - \varepsilon$. 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. $\square$

Fin cours #2 du 23 janvier

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.*

Pour $p = 2$, on retrouve le théorème de Risez–Fréchet. 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.

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). □

3.3 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.

3.4 Les espaces $C(K)$

Soit K un espace métrique 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 finie 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 métrique 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 $\varphi = I_\mu - I_\nu$.*

La preuve est longue 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 4

Le théorème de Hahn–Banach

4.1 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.

Énonçons d'abord une forme plus générale.

Théorème (Théorème de prolongement de Hahn–Banach, cas général)**.** *Soit X un espace vectoriel, $p : X \to \mathbf{R}$ une fonction convexe, $M \subset X$ un sous-espace vectoriel et $f : M \to \mathbf{R}$ une forme linéaire vérifiant $f \leqslant p$ sur M. Alors il existe une forme linéaire $g : X \to \mathbf{R}$ qui prolonge f et vérifie $g \leqslant p$ sur X.*

On applique souvent le résultat au cas où X est un espace vectoriel normé et p est un multiple de la norme. Soit $t > 0$. Remarquons qu'une forme linéaire $f : X \to \mathbf{R}$ vérifie l'inégalité $f \leqslant t\|\cdot\|$ si et seulement elle est continue et de norme $\leqslant t$. On a donc le corollaire suivant.

Corollaire (Théorème de prolongement de Hahn–Banach, cas continu)**.** *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 $g : X \to \mathbf{R}$ qui prolonge f et telle que*

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

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

Nous allons donner deux preuves : une preuve du corollaire dans le cas séparable et une preuve du théorème dans le cas général basée sur l'axiome du choix. 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 réel, $p : X \to \mathbf{R}$ une fonction convexe, $M \subset X$ un sous-espace vectoriel et $z \in X$. Si $f : M \to \mathbf{R}$ est une forme linéaire vérifiant $f \leqslant p$ sur M, il existe une forme linéaire $g : M + \mathbf{R}z \to \mathbf{R}$ prolongeant f et vérifiant $g \leqslant p$ sur $M + \mathbf{R}z$.*

Démonstration. Si $z \in M$, il suffit de considérer $g = f$. Supposons donc $z \notin M$. Pour tout choix de $\alpha \in \mathbf{R}$, il existe une unique forme linéaire $g_\alpha : M + \mathbf{R}z$ prolongeant f et vérifiant $g_\alpha(z) = \alpha$. Pour $x \in M$ et $t \in \mathbf{R}$, elle est donnée par

$$g_\alpha(x + tz) = f(x) + \alpha t$$

L'inégalité $g_\alpha \leqslant p$ sur $M + \mathbf{R}z$ équivaut donc à

$$\forall x \in M, \forall t \in \mathbf{R} \quad f(x) + \alpha t \leqslant p(x + tz)$$

Cette inégalité est vérifiée si $t = 0$. Traitant séparément les cas $t > 0$ et $t = -s < 0$, cette condition équivaut à

$$\sup_{y \in M, s>0} \frac{f(y) - p(y - sz)}{s} \leqslant \alpha \leqslant \inf_{x \in M, t>0} \frac{p(x + tz) - f(x)}{t}$$

Pour voir qu'un choix de α vérifiant ces conditions est possible, il suffit de montrer pour tous $x, y \in M$ et $s, t > 0$ l'inégalité

$$\frac{f(y) - p(y - sz)}{s} \leqslant \frac{p(x + tz) - f(x)}{t} \tag{4.1}$$

On peut alors écrire

$$\begin{aligned}(4.1) \iff & tf(y) - tp(y - sz) \leqslant sp(x + tz) - sf(x) \\ \iff & f\left(\frac{s}{s+t}x + \frac{t}{s+t}y\right) \leqslant \frac{s}{s+t}p(x + tz) + \frac{t}{s+t}p(y - sz)\end{aligned}$$

Cette égalité est vraie : elle découle de la convexité de p puisque

$$f\left(\frac{s}{s+t}x + \frac{t}{s+t}y\right) \leqslant p\left(\frac{s}{s+t}x + \frac{t}{s+t}y\right) = p\left(\frac{s}{s+t}(x + tz) + \frac{t}{s+t}(y - sz)\right) \quad \square$$

Dans le cas d'un espace vectoriel normé, le lemme implique que toute forme linéaire continue sur M peut être prolongée en une forme linéaire continue de même norme sur $M + \mathbf{R}z$. On peut alors donner une preuve «constructive» du corollaire dans le cas séparable.

Preuve du corollaire dans le cas 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 $M_{n+1} = M_n + \mathbf{R}x_{n+1}$. 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. (Preuve : pour tout $x \in X$, soit (y_p) une suite de Y telle que $\lim x_p = x$. Puisque g est lipschitzienne, $(g(y_p))$ est une suite de Cauchy, donc convergente, et on pose $\tilde{g}(x) = \lim_{p\to\infty}(g(y_p))$. On vérifie que cette définition ne dépend pas du choix de la suite (y_p) et que la fonction $\tilde{g}$ a toutes les propriétés voulues. □

Avant de donner la preuve du théorème de Hahn–Banach dans le cas général, faisons quelques préliminaires autour du lemme de Zorn.

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

Preuve du théorème de prolongement 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 qui prolonge f et telle que $g \leqslant p$ sur M.

Soit $\mathscr{E}$ 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 $\mathscr{E}$ ainsi ordonné est inductif. En effet, si $\mathscr{A} = \{(Y_i, g_i) \;:\; i \in I\}$ est une partie totalement ordonnée de E, un majorant de $\mathscr{A}$ 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 $\mathscr{A} \neq \emptyset$; il faut aussi remarquer que (M, f) est un majorant de la partie vide.)

Par le lemme de Zorn, l'ensemble $\mathscr{E}$ 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 $x \in X \backslash G$ et le lemme fournit un prolongement $\tilde{g}$ de g tel que $(G + \mathbf{R}x, \tilde{g}) \in \mathscr{E}$, contredisant la maximalité. □

Fin cours # 3 du 30 janvier

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 $\lambda x \mapsto \lambda\|x\|$. □

Exercice. Si X est un espace de Hilbert, montrer que pour tout $x \in X$ non nul il existe un unique $f \in X^*$ vérifiant $\|f\| = 1$ et $f(x) = 1$. Montrer que ce n'est pas le cas si $X = \ell_1$.

Il est instructif de comparer les deux formules suivantes, pour $f \in X^*$ et $x \in X$

$$\|f\|_{X^*} = \sup_{x \in B_X} f(x) \quad \|x\|_X = \max_{f \in B_{X^*}} f(x)$$

La formule de gauche est la définition de la norme sur X^* ; la borne supérieure n'y est pas nécessairement atteinte. La formule de droite et le fait que la borne supérieure y est atteinte se déduisent du corollaire précédant.

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. Si X est un espace de Banach, on définit l'annihilateur d'une partie $M \subset X$ comme

$$M^\perp = \{f \in X^* \ : \ \forall x \in M, \quad f(x) = 0\}$$

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

$$M \text{ est dense } \iff M^\perp = \{0\}.$$

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 appartient à $M^\perp$ et n'est pas identiquement nulle puisque $f(x) = 1$. □

4.2 Espaces vectoriels topologiques

Le cadre des espaces vectoriels normés est trop restreint : il va être nécessaire de considérer des topologies sur un espace vectoriel qui ne sont pas induites par une norme, ni même par une distance. Faisons tout d'abord des rappels de topologie générale (non métrique).

4.2.1 Espaces topologiques

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

(a) l'ensemble $\mathcal{T}$ est stable par union quelconque,

(b) 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*. Il découle des axiomes que $\mathcal{T}$ contient nécessairement les éléments $\emptyset$ (en considérant dans (a) la réunion de la famille vide) et X (en considérant dans (b) l'intersection de la famille vide).

Exemple. Soit (X, d) un espace métrique. 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 dès que X contient au moins deux points.

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.
— Un espace topologique est dit *compact* s'il est séparé et si tout recouvrement ouvert admet un sous-recouvrement fini.

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.

On prendra garde au fait que les notions d'adhérence et de compacité ne sont pas définies à l'aide de limites de (sous-)suites comme dans le cas métrique.

Si X est un espace topologique séparé et $K \subset X$ une partie compacte, alors K est un fermé dans X. (Preuve. Soit $y \in X \setminus K$. Pour tout $x \in K$, il existe des ouverts disjoints U_x et V_x contenant respectivement x et y. Par compacité, il existe un sous-ensemble fini $F \subset K$ tel que $K \subset \bigcup_{x \in F} U_x$. L'ouvert $\bigcap_{x \in F} V_x$ est un voisinage de y disjoint de K. Cet argument montre que $X \setminus K$ est ouvert, donc K est fermé.)

4.2.2 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 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

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

4.2.3 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 de somme et de multiplication scalaire

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

   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 $W + W \subset V$. [Preuve : puisque $s^{-1}(V)$ est un voisinage de $(0, 0)$ dans $X \times X$ par continuité de s en $(0, 0)$, il existe un pavé ouvert $O_1 \times O_2$ tel que $(0, 0) \in O_1 \times O_2 \subset s^{-1}(V)$, et $W = O_1 \cap O_2$ convient.]

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 et convexe. De plus,

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

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

Démonstration. Soit $x \in X$. La fonction $f : \mathbf{R} \to X$ donnée par $f(\lambda) = \lambda x$ est continue, donc $f^{-1}(W)$ est un ouvert de $\mathbf{R}$ contenant 0 et il existe $\varepsilon > 0$ tel que $\varepsilon x \in W$, ce qui montre que j_W est à valeurs finies.

La convexité de W (ou plus précisément le fait que W soit étoilé par rapport à 0) implique que l'on a $sW \subset tW$ pour tous $0 < s < t$. En particulier, si $x \in X$ et $t > 0$ sont tels que $j_W(x) < t$, alors $x \in tW$.

Montrer que pour tous $x, y \in X$ et $\lambda > 0$ on a

1. $j_W(\lambda x) = \lambda j_W(x)$,
2. $j_W(x+y) \leqslant j_W(x) + j_W(y)$,

ces deux propriété impliquant ensemble la convexité de la fonction j_W. La première relation est élémentaire ($x \in tW$ équivaut à $\lambda x \in \lambda tW$). Pour montrer la seconde, soient x, y dans X. Pour $s > j_W(x)$ et $t > j_W(y)$, on a

$$x + y \in sW + tW = (s+t)W$$

(la dernière égalité découlant de la convexité de W). On a donc $j_W(x+y) \leqslant s + t$, et l'inégalité 2. en découle en prenant la borne inférieure sur s et t.

Si $j_W(x) < 1$, il existe $t < 1$ tel que $x \in tW$, ce qui implique $x \in W$. Réciproquement, si $x \in W$, alors par continuité de la multiplication scalaire, on a $\lambda x \in W$ pour λ dans un voisinage de 1 et donc $j_W(x) < 1$.

Pour $\varepsilon > 0$, on a donc $\{j_W < \varepsilon\} = \varepsilon W$, ce qui montre que j_W est continue en 0. Puisque $|j_W(x) - j_W(x+h)| \leqslant \max(j_W(h), j_W(-h))$, il s'ensuit que j_W est continue en tout $x \in X$. □

4.3 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, convexe 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, on peut prolonger f en une forme linéaire $g : X \to \mathbf{R}$ telle que $g \leqslant j_W$. 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 voisinage ouvert V_y 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. □

Fin cours #4 du 6 février

4.4 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 $\mathbf{C}$.

4.5 Le théorème de Krein–Milman

Soit X un espace vectoriel réel et C une partie convexe de X. On montre par récurrence sur n que si $\lambda_1, \dots, \lambda_n$ sont des réels positifs vérifiant $\lambda_1 + \cdots + \lambda_n = 1$, alors pour tous $x_1, \dots, x_n$ 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 réel 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 $(\mathbf{R}^n, \|\cdot\|_p)$.

Lemme. *Dans un espace vectoriel topologique X, l'adhérence $\overline{C}$ d'une partie convexe C est convexe.*

Démonstration. On a d'abord $\overline{C \times C} = \overline{C} \times \overline{C}$ (exercice). Soit $\lambda \in [0,1]$ et $f_\lambda : X^2 \to X$ l'application continue $(x,y) \mapsto \lambda x + (1-\lambda) y$. Il faut montrer que pour tous x, y dans $\overline{C}$, on a $\lambda x + (1-\lambda) y \in \overline{C}$. On doit donc vérifier que tout voisinage ouvert V de $\lambda x + (1-\lambda) y$ intersecte C. Mais $f_\lambda^{-1}(V)$ est un ouvert de $X \times X$ contenant $(x,y) \in \overline{C \times C}$: il existe donc $(x', y') \in (C \times C) \cap f_\lambda^{-1}(V)$. Le point $\lambda x' + (1-\lambda) y'$ est dans $C \cap V$, comme recherché. □

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 $y \in F$. Un élément $x \in K$ est un point extrémal de K si et seulement si $\{x\}$ est une face extrémale de K.

Il est facile de vérifier que l'intersection d'une famille quelconque de faces extrémales de K, dès lors qu'elle est non vide, est encore une face extrémale de K.

Soit $f : X \to \mathbf{R}$ une forme linéaire continue et $m = \sup\{f(x) \ :\ x \in K\}$. 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 élément maximal 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 fermé 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. □

En dimension finie, on peut remplacer l'utilisation du lemme de Zorn par une récurrence sur la dimension, et montrer le résultat plus précis suivant : si $K \subset \mathbf{R}^n$ est un convexe compact et si $\mathcal{E}$ est l'ensemble de ses points extrémaux, alors $K = \mathrm{conv}(\mathcal{E})$.

Chapitre 5

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.

5.1 La topologie faible sur un espace de Banach

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

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

où A est une partie finie de X^*, x est un point de X (le «centre» de $V(x, \varepsilon, A)$) 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 $y \in V(x, \varepsilon, A)$, il existe $\eta > 0$ tel que $V(y, \eta, A) \subset V(x, \varepsilon, A)$. En effet, pour tout $f \in A$, puisque $f(y) \in]f(x)-\varepsilon, f(x)+\varepsilon[$, il existe $\eta_f > 0$ tel que $]f(y)-\eta_f, f(y)+\eta_f[\subset]f(x) - \varepsilon, f(x) + \varepsilon[$ et il suffit de considérer alors $\eta = \min\{\eta_f \ :\ f \in A\}$.

Il découle de cette remarque que $\sigma(X, X^*)$ est stable par intersection finie. La stabilité par réunion quelconque est évidente.

La continuité de l'addition et de la multiplication scalaire découlent du fait que

$$V(x, \varepsilon/2, A) + V(y, \varepsilon/2, A) \subset V(x + y, \varepsilon, A)$$

et si $\lambda \neq 0$

$$\lambda V(x, \varepsilon/|\lambda|, A) = V(\lambda x, \varepsilon, A).$$

Si $x \in X \setminus \{0\}$, par un corollaire de HAHN–BANACH, il existe $f \in X^*$ tel que $\varepsilon = f(x) > 0$. Alors $V(x, \varepsilon, \{f\})$ est un ouvert faible contenant x et non 0, ce qui montre que $X \setminus \{0\}$ est un ouvert faible.

On a ainsi montré que $(X, \sigma(X, X^*)$ est un espace vectoriel topologique. Sa locale convexité découle du fait que $V(x, \varepsilon, A)$ est convexe. □

On utilise les adjectifs «faible» et «fort» en relation respectivement avec la topologie faible et la topologie (dite forte) de la norme. Tout ouvert faible est un ouvert fort, c'est-à-dire un ouvert pour la topologie de la norme : il suffit de le vérifier pour les ouverts faibles élémentaires, pour lesquels c'est immédiat.

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 $V(x, \varepsilon, A)$ contient le sous-espace affine $x + \bigcap_{f \in A} \ker(f)$ 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. On a déjà observé que toute partie faiblement fermée est fortement fermée. 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 $V(x, \varepsilon, \{f\})$ 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, c'est-à-dire que pour tout voisinage faible V de x on a $x_n \in V$ pour n assez grand. On vérifie (exercice) que c'est équivalent à

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

Corollaire (Lemme 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. On pose $C = \operatorname{conv}\{x_n\}$. Son adhérence forte $\overline{C}$ est convexe, donc faiblement fermée par le théorème précédent. On a donc $x \in \overline{C}$, donc il existe une suite (y_m) qui converge vers x. □

Il est instructif de considérer un cas concret. Soit $(e_n)_{n \geqslant 1}$ une base hilbertienne d'un espace de Hilbert séparable X. La suite (e_n) converge faiblement vers 0 : par le théorème de RIESZ–FRÉCHET, toute forme linéaire continue sur X est de la forme $f : x \mapsto \langle x, y \rangle$ pour $y \in X$; puisque $\|y\|^2 = \sum_n \langle e_n, y \rangle^2 = \sum_n f(e_n)^2$, on a bien $\lim_{n \to \infty} f(e_n) = 0$. La suite (e_n) ne converge pas fortement vers 0. Pour illustrer le lemme de MAZUR, on peut remarquer que la suite de ses moyennes de CESARÓ converge fortement vers 0 puisque

$$\left\| \frac{e_1 + \cdots + e_n}{n} \right\| = \frac{1}{\sqrt{n}}$$

5.2 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

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

où B est une partie finie de X, f 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. □

Fin cours #5 du 13 février