4. Les théorèmes fondamentaux

4.1. Le théorème de Baire et ses conséquences

Soit X un espace topologique ; un ouvert U de X est dense dans X si et seulement si le fermé complémentaire U^c est d'intérieur vide. Si on a un nombre fini d'ouverts denses, on vérifie facilement de proche en proche que $U_1 \cap \ldots \cap U_n$ est encore un ouvert dense. Le théorème de Baire donne un cas où cette propriété triviale d'intersection finie peut s'étendre aux *suites* d'ouverts denses.

Théorème 4.1.1 : théorème de Baire. *Soit* X *un espace métrique complet ; si* $(U_n)_{n\ge 0}$ *est une suite de parties ouvertes et* denses *dans* X, *l'intersection* $\bigcap_{n\ge 0} U_n$ *est dense dans l'espace* X.

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

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

Notons maintenant B_n la boule *fermée* de centre x_n et de rayon r_n. On a

$$B_{n+1} \subset B(x_{n+1}, 2r_{n+1}) \subset B(x_n, r_n) \subset B_n.$$

Comme l'espace X est complet, que les ensembles B_n sont fermés, décroissants, non vides et que leur diamètre tend vers 0, on a $\bigcap_{n\ge 0} B_n \ne \emptyset$; or, par construction, $\bigcap_{n\ge 0} B_n \subset V \cap \bigcap_{n\ge 0} U_n$, ce qui montre que cette dernière intersection est non vide.

//

Corollaire 4.1.2. *Soient* X *un espace métrique complet* non vide *et* $(F_n)_{n\ge 0}$ *une suite de parties fermées de* X *telle que* $\bigcup_{n\ge 0} F_n = X$ *; alors l'un des fermés* F_n *a un intérieur non vide ; en réalité, on peut même dire que* $\bigcup_{n\ge 0} \mathring{F}_n$ *est dense dans* X.

Démonstration. Soit V un ouvert non vide de X ; dans l'espace métrique complet $Y = \overline{V}$, considérons les ouverts (relatifs) $U_n = \overline{V} \setminus F_n$. Puisque $\bigcup_{n\ge 0} F_n = X$, l'intersection $\bigcap_{n\ge 0} U_n$ est vide, ce qui entraîne par le théorème 1 appliqué à Y que l'un au moins des ouverts U_n n'est pas dense dans Y. Il existe donc n_0 et un ouvert non vide U de Y qui soit disjoint de U_{n_0}. Cet ouvert U doit rencontrer V ; on peut donc trouver $x \in V \cap U$ et $r > 0$ tels que $B(x, r) \subset V$ ne rencontre pas U_{n_0}, c'est à dire $B(x,r) \subset F_{n_0}$ et en particulier $x \in \mathring{F}_{n_0}$. On a ainsi montré que la réunion des intérieurs des (F_n) rencontre tout ouvert non vide V donné.

//

Les résultats obtenus sur l'espace métrique complet X sont de nature purement topologique ; l'hypothèse devrait être formulée ainsi : si la topologie de l'espace topologique X peut être définie par une distance d qui rend (X, d) complet, alors le théorème de Baire est valable. Par exemple, on peut montrer facilement que tout ouvert U d'un espace métrique complet (X, d) peut être muni d'une nouvelle métrique δ qui ne change pas la topologie, et qui fait de (U, δ) un espace métrique complet (exercice) ; le théorème de Baire est donc valable pour tout ouvert d'un espace métrique complet, ce que l'on voit aussi facilement en adaptant la démonstration du théorème 1.

On dit qu'un espace topologique X est un *espace de Baire* s'il vérifie le théorème de Baire : toute intersection d'une suite d'ouverts denses dans X est dense dans X.

Exercice 4.1.1. Montrer qu'il n'existe pas de norme sur l'espace vectoriel $\mathbb{R}[X]$ des polynômes qui rende cet espace complet.

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

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

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

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

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

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

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

Lemme 4.1.4. *On suppose que* E *est complet,* $f : \mathrm{E} \to \mathrm{F}$ *linéaire continue, et que* A *est un sous-ensemble borné de* F *tel que* $\mathrm{A} \subset f(\mathrm{B_E}) + \varepsilon \mathrm{A}$, *avec* $0 < \varepsilon < 1$. *Alors*

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

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

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

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

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

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

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

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

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

Remarque 4.1.2. Le théorème précédent a peu de chose à voir avec les normes. Il reste vrai si E et F sont deux espaces vectoriels munis de métriques vérifiant les propriétés (*i*), (*ii*) et (*iii*) de la remarque 1.1.8, et si E et F sont **complets** pour ces métriques. En particulier le théorème des isomorphismes, et ses conséquences immédiates que nous allons énoncer dans ce qui suit, sont valables pour les espaces de Fréchet.

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

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

Théorème 4.1.5 : théorème de l'application ouverte. *Soient* E *et* F *deux espaces de Banach ; toute application linéaire, continue, surjective* f *de* E *sur* F *est ouverte.*

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

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

donnée par la proposition 1.3.3 ; la première flèche π est la projection canonique de E sur le quotient par le noyau de f. Par des arguments algébriques, la deuxième flèche g est bijective, et elle est continue d'après la proposition 1.3.3. C'est donc un isomorphisme, et il en résulte que f est ouverte parce que π et g sont ouvertes.

//

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

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

Démonstration. Soit f une application linéaire de E dans F dont le graphe $\mathrm{G} \subset \mathrm{E} \times \mathrm{F}$ est fermé ; alors G est un espace de Banach. Tout point z du graphe G est de la forme $z = (x, f(x))$ pour un certain $x \in \mathrm{E}$ unique ; notons $p : \mathrm{G} \to \mathrm{E}$ l'application définie par $p(z) = p(x, f(x)) = x \in \mathrm{E}$. Il est clair que p est linéaire, continue et bijective (l'inverse –algébrique– étant l'application $x \to (x, f(x))$ de E dans G). D'après le théorème des isomorphismes, cet inverse $x \to (x, f(x))$ est continu de E dans G ; il en résulte que $x \to f(x)$ est continue de E dans F.

//

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

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

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

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

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

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

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

//

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

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

//

Corollaire 4.1.9. *Soient* E *un espace de Banach (ou bien un espace vectoriel* (E, d) *complet comme dans la proposition 7),* Y *un espace normé et* (f_n) *une suite d'applications linéaires continues de* E *dans* Y *; on suppose que, pour tout* $x \in \mathrm{E}$*, la suite* $(f_n(x))$ *converge dans* Y *; notons* $f(x)$ *sa limite. Alors* f *est linéaire et continue.*

Démonstration. D'abord, il est évident que la limite f est linéaire. Soit $x \in \mathrm{E}$; comme la suite $(f_n(x))$ est convergente, elle est bornée ; par le théorème 8, la suite $(\|f_n\|)$ est alors bornée. Il existe alors un nombre $\mathrm{M} \geq 0$ tel que, pour tout $x \in \mathrm{E}$ et tout entier $n \geq 0$ on ait $\|f_n(x)\| \leq \mathrm{M}\,\|x\|$. Passant à la limite on trouve $\|f(x)\| \leq \mathrm{M}\,\|x\|$, pour tout $x \in \mathrm{E}$.

//

Corollaire 4.1.10. *Soient* E *un espace de Banach (ou bien un* (E, d) *complet comme dans la proposition 7),* Y *un espace normé et* (u_k) *une suite d'applications linéaires continues de* E *dans* Y *; on suppose que, pour tout* $x \in \mathrm{E}$*, la série* $\sum_k u_k(x)$ *converge dans* Y *; notons* $\mathrm{T}(x)$ *sa somme. Alors* T *est linéaire et continue.*

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

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

4.2. Théorème de Hahn-Banach

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

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

Exemples 4.2.1. Les semi-normes sont des fonctions sous-linéaires. Une forme linéaire (réelle) est une fonction sous-linéaire. Une fonction sous-linéaire sur $\mathbb{R}$ est linéaire par morceaux : elle vaut $q(t) = at$ pour $t < 0$ et $q(t) = bt$ pour $t \geq 0$, avec $a \leq b$ (en effet, on doit avoir $-a + b = q(-1) + q(1) \geq q(0) = 0$).

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

et il est alors clair que cette propriété est vraie :

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

Le lemme est donc établi.

Le lemme de Zorn

Le lemme de Zorn est assez directement équivalent à un axiome de la théorie des ensembles, *l'axiome du choix*. Il permet de valider certains types de raisonnements où on cherche à garantir l'existence d'objets maximaux.

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

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

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

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

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

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

Dans le cas normé séparable, on peut donner une démonstration de Hahn-Banach qui n'utilise pas l'axiome du choix. Supposons X normé réel séparable de dimension infinie. On suppose y^* donnée sur le sous-espace vectoriel Y, majorée par une fonction sous-linéaire continue q. On remarque d'abord qu'on peut prolonger y^* par continuité à l'adhérence Y_0 de Y, en gardant $y_0^* \leq q$ sur Y_0 pour le prolongement y_0^*. Soit $(x_n)_{n \geq 0}$ une suite dense dans X ; si $Y_0 = X$, c'est fini. Sinon, soit k_0 le plus petit entier $k \geq 0$ tel que $x_k \notin Y_0$ (s'il n'en existait pas, Y_0 fermé contiendrait la suite dense (x_n) toute entière, donc on aurait $Y_0 = X$, ce qui a été exclu ici) ; on pose $Y_1 = Y_0 + \mathbb{R}x_{k_0}$ et on prolonge y_0^* de Y_0 à une forme y_1^* sur Y_1, vérifiant toujours $y_1^* \leq q$ sur Y_1, en utilisant le lemme 2. On remarque que Y_1 contient tous les vecteurs x_k pour $0 \leq k \leq k_0$. Si $Y_1 = X$ c'est fini. Sinon, soit k_1 le plus petit entier $k > k_0$ tel que $x_k \notin Y_1$ (s'il n'en existait pas, on aurait $Y_1 = X$ qui est exclu ici). On pose $Y_2 = Y_1 + \mathbb{R}x_{k_1}$ et on prolonge y_1^* en y_2^* sur Y_2, vérifiant $y_2^* \leq q$ sur Y_2. On remarque que Y_2 contient tous les vecteurs x_k pour $0 \leq k \leq k_1$.

On continue ainsi, et on construit une suite croissante (Y_n) de sous-espaces vectoriels de X (ou alors il existe N tel que $X = Y_N$) avec sur Y_{n+1} une forme linéaire y_{n+1}^* qui prolonge y_n^* et qui vérifie $y_{n+1}^* \leq q$. On construit aussi une suite strictement croissante (k_n) telle que Y_{n+1} contienne tous les x_k pour $0 \leq k \leq k_n$. Comme la suite d'espaces est croissante, la réunion $Y_\infty = \bigcup_n Y_n$ est un sous-espace vectoriel, et comme les formes (y_n^*) se prolongent les unes les autres, on peut définir y_∞^* sur Y_∞ qui les prolonge toutes, et $y_\infty^* \leq q$. Enfin, Y_∞ est dense dans X parce qu'il contient toute la suite (x_n). Pour terminer, il suffit de prolonger par continuité y_∞^* en une forme linéaire continue x^* sur l'espace X.

Théorème 4.2.3 : théorème de séparation de Hahn-Banach. *Soient* X *un espace normé réel,* A *un convexe ouvert non vide et* B *un convexe non vide tels que* A *et* B *soient disjoints. Il existe alors une forme linéaire continue f sur* X *telle que*

$$f(a) < \inf f(B)$$

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

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

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

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

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

Puisque A et B sont disjoints, ce convexe est disjoint de $\mathrm{D} = \{0_{\mathrm{X}}\}$. D'après l'étape préliminaire, on peut trouver une forme linéaire continue m sur X telle que $m(c) < m(d_0)$ pour tout $c \in \mathrm{C}$, où $d_0 = 0_{\mathrm{X}}$ est l'unique élément de D. Cela signifie que $m(a - b) < m(d_0) = 0$ pour tous $a \in \mathrm{A}$, $b \in \mathrm{B}$, soit encore $m(a) < m(b)$ pour tous $a \in \mathrm{A}$, $b \in \mathrm{B}$, donc $m(a) \leq \inf m(\mathrm{B})$. Puisque m n'est pas nulle, on peut choisir $u \in \mathrm{X}$ tel que $m(u) > 0$, puis $\varepsilon > 0$ tel que $a + \varepsilon u \in \mathrm{A}$ puisque A est ouvert ; on termine en écrivant $m(a) < m(a + \varepsilon u) \leq \inf m(\mathrm{B})$, et en prenant $f = m$.

//

Rappelons que X^* désigne le dual topologique d'un espace normé X.

Exercices 4.2.2.

Si C est fermé, K compact, C et K convexes non vides et disjoints d'un espace normé X, montrer qu'il existe une forme linéaire continue $x^* \in \mathrm{X}^*$ telle que $\sup x^*(\mathrm{C}) < \min x^*(\mathrm{K})$.

Soient X un espace normé et q une fonction sous-linéaire sur X ; montrer que q est continue sur X si et seulement s'il existe une constante M telle que $|q(x)| \leq \mathrm{M}\,\|x\|$ pour tout $x \in \mathrm{X}$.

Corollaire 4.2.4. *Si* C *est un sous-ensemble convexe fermé non vide d'un espace normé réel* X*, alors* C *est l'intersection de demi-espaces affines fermés.*

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

Pour tout $x \notin \mathrm{C}$, on peut trouver une boule ouverte $\mathrm{A} = \mathrm{B}(x, r)$ disjointe de C ; d'après le théorème de séparation il existe une forme linéaire continue x^* telle que $x^*(a) < \inf x^*(\mathrm{C})$ pour tout $a \in \mathrm{A}$, et en particulier $x^*(x) < \inf x^*(\mathrm{C})$. On voit donc que si on pose $d = \inf x^*(\mathrm{C})$ et

$$\mathrm{D}_x = \{y \in \mathrm{X} : x^*(y) \geq d\}$$

on aura $\mathrm{C} \subset \mathrm{D}_x$ mais $x \notin \mathrm{D}_x$.

//

Théorème 4.2.5 : théorème de Hahn-Banach. *Soient* X *un espace normé (réel ou complexe) et* Y *un sous-espace vectoriel de* X *; pour tout* $\ell \in \mathrm{Y}^*$*, il existe* $m \in \mathrm{X}^*$ *dont la restriction à* Y *soit* ℓ *et telle que* $\|m\| = \|\ell\|$.

Démonstration. Considérons d'abord le cas réel. Ici la fonction sous-linéaire q de l'énoncé du théorème 1 sera un multiple convenable de la norme N de X. Par définition de la norme de la forme linéaire ℓ, on a $\ell \le \|\ell\|\,\mathrm{N} = q$ sur le sous-espace vectoriel Y. On peut donc trouver un prolongement m tel que $m \le q$ sur X, ce qui donne le résultat : on a en effet $m(x) \le \|\ell\|\,\|x\|$ pour tout $x \in \mathrm{X}$, d'où aussi $|m(x)| \le \|\ell\|\,\|x\|$ en appliquant à x et $-x$; tout ceci montre que m est continue et $\|m\| \le \|\ell\|$, mais $\|\ell\| \le \|m\|$ puisque m prolonge ℓ.

Si X est un espace vectoriel complexe, on commence par le considérer comme un espace vectoriel réel, et on considère sur Y la forme linéaire réelle $\ell_1 = \mathrm{Re}\,\ell$. On trouve alors une forme linéaire réelle m_1 sur X telle que m_1 prolonge la forme linéaire réelle ℓ_1 et $\|m_1\| = \|\ell_1\|$. Par la proposition 1.6.1, on sait que m_1 est la partie réelle d'une forme linéaire complexe m sur X, et de plus $\|m\| = \|m_1\| \le \|\ell_1\| = \|\ell\|$; d'autre part m prolonge ℓ (ici Y est un sous-espace vectoriel complexe ; si $y \in \mathrm{Y}$ on a aussi $iy \in \mathrm{Y}$ ce qui permet d'écrire $m(y) = m_1(y) - im_1(iy)$, et alors $m(y) = \ell_1(y) - i\ell_1(iy) = \ell(y)$).

//

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

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

//

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

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

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

//

Remarque 4.2.3. Le dual de X/Y est identifiable isométriquement au sous-espace de X^* formé des x^* dont la restriction à Y est nulle.

Notons π la projection de X sur X/Y ; à $z^* \in (\mathrm{X}/\mathrm{Y})^*$ on associe $x^* = z^* \circ \pi \in \mathrm{X}^*$, qui est nulle sur Y. Inversement, si $x^* \in \mathrm{X}^*$ est nulle sur Y, on peut l'écrire $x^* = z^* \circ \pi$ d'après la proposition 1.3.3, et $\|z^*\| = \|x^*\|$.

Proposition 4.2.8. *Soit* X *un espace normé ;*

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

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

Démonstration. Si X est de dimension finie, toute forme linéaire sur X est continue, donc le dual topologique est égal au dual algébrique pour lequel on connaît l'égalité des dimensions. Si X^* est de dimension finie, X^{**} est de dimension finie par ce qui précède, donc X est de dimension finie par la proposition 3.1.

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

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

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

//

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

Proposition 4.2.9. *Soit* X *un espace normé ; si le dual* X^* *est séparable, alors* X *est séparable.*

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

//

Hahn-Banach et transposition

On a déjà montré que la transposée ${}^t\mathrm{T}$ d'une application linéaire continue T a une norme inférieure ou égale à celle de T. Grâce au théorème de Hahn-Banach on peut compléter ce résultat.

Proposition 4.2.10. *Soient* X *et* Y *deux espaces normés ; pour tout* $\mathrm{T} \in \mathcal{L}(\mathrm{X}, \mathrm{Y})$*, on a* $\|{}^t\mathrm{T}\| = \|\mathrm{T}\|$.

Démonstration. On sait déjà que $\|{}^t\mathrm{T}\| \le \|\mathrm{T}\|$ par la proposition 1.6.2, nous allons montrer l'égalité. Pour tout $\varepsilon > 0$, on peut trouver un vecteur $x \in \mathrm{X}$ tel que $\|x\| \le 1$ et tel que $\|\mathrm{T}(x)\| > \|\mathrm{T}\| - \varepsilon$, puis une forme linéaire $y^* \in \mathrm{F}^*$ telle que $\|y^*\| \le 1$ et $y^*(\mathrm{T}(x)) = \|\mathrm{T}(x)\|$. Alors

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

//

Exemple 4.2.4. La transposée de l'injection de $\mathrm{Y} \subset \mathrm{X}$ dans X est l'application de restriction de X^* sur Y^* (surjective par Hahn-Banach). Ces deux applications sont de norme 1 lorsque $\mathrm{Y} \ne \{0\}$.

Pour $\mathrm{T} \in \mathcal{L}(\mathrm{X}, \mathrm{Y})$ on notera im(T) le sous-espace de Y *image de l'application* T, noté aussi T(X),

$$\mathrm{im}(\mathrm{T}) = \mathrm{T}(\mathrm{X}) = \{y \in \mathrm{Y} : \exists x \in \mathrm{X},\ y = \mathrm{T}(x)\}.$$

Lemme 4.2.11. *Soient* X, Y *deux espaces normés et* $\mathrm{T} \in \mathcal{L}(\mathrm{X}, \mathrm{Y})$ *; l'application* ${}^t\mathrm{T}$ *est injective si et seulement si* im(T) *est dense dans* Y. *De plus, si* $\mathrm{im}({}^t\mathrm{T})$ *est dense dans* X^**, l'application* T *est injective.*

Démonstration. Si im(T) n'est pas dense, son adhérence Z est un sous-espace vectoriel fermé de Y, distinct de Y. D'après le corollaire 6, il existe une forme linéaire y^* non nulle sur Y, mais dont la restriction à Z est nulle ; en particulier, $y^*(\mathrm{T}(x)) = 0$ pour tout $x \in \mathrm{X}$ puisque Z contient l'image de T. On a donc ${}^t\mathrm{T}(y^*)(x) = 0$ pour tout $x \in \mathrm{X}$, ce qui signifie que ${}^t\mathrm{T}(y^*) = 0$, donc ${}^t\mathrm{T}$ n'est pas injective.

Si ${}^t\mathrm{T}$ n'est pas injective, il existe $y^* \in \mathrm{Y}^*$ non nulle telle que ${}^t\mathrm{T}(y^*) = 0$, ce qui signifie que $y^*(\mathrm{T}(x)) = 0$ pour tout $x \in \mathrm{X}$. On voit alors que l'image de T est contenue dans le noyau de y^*, qui est un sous-espace fermé de Y, distinct de Y. Il en résulte que im(T) n'est pas dense dans Y.

Si $\mathrm{T}(x) = 0$, on a ${}^t\mathrm{T}(y^*)(x) = y^*(\mathrm{T}(x)) = 0$ pour tout $y^* \in \mathrm{Y}^*$, ce qui montre que $x^*(x) = 0$ pour tout $x^* = {}^t\mathrm{T}(y^*) \in \mathrm{im}({}^t\mathrm{T})$; si $\mathrm{im}({}^t\mathrm{T})$ est dense dans X^*, on en déduit par continuité que $x^*(x) = 0$ pour tout $x^* \in \mathrm{X}^*$, donc $x = 0$ par Hahn-Banach ; il en résulte que T est injective.

//

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

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

$$\forall x^* \in X^*, \quad J_X(x^*) = x^*(x).$$

Pour tout $x^* \in X^*$, on a $|J_X(x)(x^*)| = |x^*(x)| \le \|x^*\| \, \|x\|$, donc $J_X(x) \in X^{**}$ et $\|J_X(x)\| \le \|x\|$. On dit que $J_X \in \mathcal{L}(X, X^{**})$ est *l'application canonique* de X dans son bidual.

Exemple 4.3.1. L'injection de c_0 dans ℓ_∞ correspond à l'injection J_{c_0} de c_0 dans son bidual, modulo les identifications habituelles entre duaux et espaces de suites.

Proposition 4.3.1. *L'application canonique* $J_X : X \to X^{**}$ *est isométrique.*

Démonstration. Soit $x \in X$; par le corollaire 2.7, il existe $x^* \in X^*$ tel que $\|x^*\| \le 1$ et $x^*(x) = \|x\|$. Alors

$$\|x\| = |x^*(x)| = |J_X(x)(x^*)| \le \|x^*\| \, \|J_X(x)\| \le \|J_X(x)\|,$$

vu que $\|x^*\| \le 1$; donc $\|J_X(x)\| = \|x\|$.

//

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

Si X est un espace normé l'application J_X est injective par la proposition 1. Remarquons que si J_X est bijective alors J_X est une isométrie de X sur un espace de Banach (par la proposition 1.2.5). Il s'ensuit que si J_X est bijective, alors nécessairement X est un espace de Banach. Ceci explique que nous restreindrons la définition qui suit aux espaces de Banach.

Définition 4.3.2. Un espace de Banach E est dit *réflexif* si l'application canonique $J_E : E \to E^{**}$ est bijective.

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

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

Proposition 4.3.2. *Tout espace de Hilbert est réflexif.*

Démonstration. Soient H un espace de Hilbert et $x^{**} \in H^{**}$; pour tout vecteur $y \in H$ soit $\ell_y \in H^*$ la forme linéaire sur H définie par $\ell_y(x) = \langle x, y \rangle$; l'application $y \to \overline{x^{**}(\ell_y)}$ est une forme linéaire et continue sur H. Par la proposition 2.3.5, il existe $x \in H$ tel que, pour tout $y \in H$ on ait $\overline{x^{**}(\ell_y)} = \langle y, x \rangle$. D'après la proposition 2.3.5, toute $f \in H^*$ est de la forme $f = \ell_y$ pour un certain $y \in H$, donc on a $x^{**}(f) = x^{**}(\ell_y) = \langle x, y \rangle = f(x)$, c'est à dire que x^{**} est l'image de x par l'application canonique de H dans H^{**}, qui est donc surjective.

//

Exemple 4.3.3. Les espaces ℓ_p, $\mathrm{L}_p(\Omega,\mu)$, sont réflexifs lorsque $1 < p < +\infty$; on pourrait dire un peu vite : le dual de L_p est L_q, et celui de L_q est L_p, donc ça marche ; c'est un peu trop rapide, parce que le dual de L_p n'**est** pas L_q, mais *s'identifie* à L_q au moyen d'une certaine bijection. Il faut donc prendre la peine, au moins une fois, de vérifier que tout colle bien.

Expliquons le cas de $\mathrm{X} = \mathrm{L}_p$; soit j_q l'application isométrique de L_q sur le dual X^* de L_p. Si x^{**} est une forme linéaire continue sur $\mathrm{X}^* = (\mathrm{L}_p)^*$, la composée $x^{**} \circ j_q$ est une forme linéaire continue sur L_q ; il existe donc une fonction $f \in \mathrm{L}_p = \mathrm{X}$ telle que

$$\forall g \in \mathrm{L}_q, \quad x^{**}(j_q(g)) = \int_\Omega f\, g\, d\mu.$$

Soit $x^* \in \mathrm{X}^*$; il existe $g \in \mathrm{L}_q$ tel que $x^* = j_q(g)$, et alors $x^*(f) = \int_\Omega f\, g\, d\mu$. La ligne précédente signifie donc bien que l'on a trouvé un vecteur $f \in \mathrm{X} = \mathrm{L}_p$ tel que

$$\forall x^* \in \mathrm{X}^* = (\mathrm{L}_p)^*, \quad x^{**}(x^*) = x^*(f).$$

En revanche, les espaces c_0, ℓ_1 et ℓ_∞ sont des espaces de Banach non réflexifs, comme on va le voir. D'après les résultats généraux qui suivent, il suffit de voir que ℓ_1 n'est pas réflexif.

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

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

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

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

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

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

Espaces normés isomorphes

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

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

Lemme 4.3.3. *Si* X *est réflexif et si* Y *est isomorphe à* X*, alors* Y *est réflexif.*

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

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

ce qu'il fallait démontrer.

//

Proposition 4.3.4. *Si* X *est réflexif, alors* X^* *est réflexif.*

Démonstration. Posons $\mathrm{Z} = \mathrm{X}^*$. Si $z^{**} = x^{***}$ est une forme linéaire sur le dual $\mathrm{Z}^* = \mathrm{X}^{**}$ de $\mathrm{Z} = \mathrm{X}^*$, elle définit une forme linéaire continue $z = x^* = x^{***} \circ \mathrm{J_X}$ sur X. Il reste seulement à vérifier que z définit la forme z^{**}, au sens précédent. Soit $z^* \in \mathrm{Z}^* = \mathrm{X}^{**}$; puisque X est réflexif il existe $x \in \mathrm{X}$ tel que $z^* = \mathrm{J_X}(x)$. Alors

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

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

//

Proposition 4.3.5. *Si* X *est réflexif, tout sous-espace fermé* Y *de* X *est réflexif.*

Démonstration. Soit π l'application de restriction définie de X^* sur Y^* (surjective par le théorème de Hahn-Banach). Soit y^{**} une forme linéaire continue sur Y^*. Alors $x^{**} = y^{**} \circ \pi$ est une forme linéaire continue sur X^*, donc il existe $x \in \mathrm{X}$ tel que $x^{**}(x^*) = x^*(x)$ pour tout x^* in X^*. Il suffit de voir que $x \in \mathrm{Y}$ pour pouvoir conclure assez facilement ; si on avait $x \notin \mathrm{Y}$, on pourrait trouver d'après le corollaire 2.6 une forme linéaire $x^* \in \mathrm{X}^*$ telle que $x^*(x) = 1$ mais $x^*(y) = 0$ pour tout $y \in \mathrm{Y}$. On aurait alors $\pi(x^*) = 0$, donc $x^{**}(x^*) = y^{**}(\pi(x^*)) = 0$, ce qui contredit $x^{**}(x^*) = x^*(x) = 1$.

//

Corollaire 4.3.6. *Si* E *est un espace de Banach et si* E^* *est réflexif, alors* E *est réflexif.* En effet E^{**} est alors réflexif et E est isomorphe à un sous-espace fermé de E^{**}.

Pourquoi s'intéresser aux espaces réflexifs ?

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

4.4. Théorème de Riesz

Lemme 4.4.1. *Soit* Z *un espace normé de dimension* n *; pour tout* $\varepsilon \in]0,1[$*, on peut trouver dans la boule unité de* Z *une famille* A *d'au moins* ε^{-n} *points dont les distances mutuelles sont* $\ge \varepsilon$ *: si* $x, y \in \mathrm{A}$ *et* $x \ne y$*, alors* $\|x-y\| \ge \varepsilon$*, et* $\operatorname{card} \mathrm{A} \ge \varepsilon^{-n}$.

Démonstration. Soit A une famille maximale de points de la boule unité $\mathrm{B_Z}$ de Z dont les distances mutuelles soient $\ge \varepsilon$; alors les boules de rayon ε centrées aux points de A recouvrent $\mathrm{B_Z}$: en effet, si $x \in \mathrm{B_Z}$ et $x \notin \mathrm{A}$, on ne peut pas, d'après la maximalité de A, ajouter le point x à la famille A pour former une nouvelle famille A' de points à distances mutuelles $\ge \varepsilon$; cela signifie qu'il existe un point $y \in \mathrm{A}$ tel que $d(y,x) < \varepsilon$, donc x est bien contenu dans une boule de rayon ε centrée en un point y de A. Soit V le volume de $\mathrm{B_Z}$; puisque Z est de dimension n, les boules de rayon ε ont un volume égal à $\varepsilon^n \mathrm{V}$ (dans le cas des scalaires réels), et puisque les boules de ce rayon centrées aux points de A recouvrent $\mathrm{B_Z}$, on a $(\operatorname{card} \mathrm{A})\, \varepsilon^n \mathrm{V} \ge \mathrm{V}$, d'où le résultat.

//

Théorème 4.4.2. *Si la boule unité d'un espace normé* X *est compacte, alors* X *est de dimension finie.*

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

//

Une autre démonstration est fondée sur le résultat suivant : *si* F *et* G *sont deux sous-espaces fermés d'un espace de Banach* E*, et si on a* $\mathrm{B_G} \subset \mathrm{F} + \varepsilon \mathrm{B_G}$ *pour un certain* $\varepsilon < 1$*, alors* $\mathrm{G} = \mathrm{F}$. On démontre ce résultat par itération, comme dans le lemme 1.4. Pour l'application au théorème de Riesz, on procède de la façon suivante : si la boule unité de E était compacte, on pourrait la recouvrir par une famille finie de boules de rayon 1/4, centrées en des points $x_1, \ldots, x_\mathrm{N}$. Appliquons le résultat précédent avec $\mathrm{G} = \mathrm{E}$ et en prenant pour F le sous-espace de dimension finie engendré par $x_1, \ldots, x_\mathrm{N}$. On aura bien que $\mathrm{B_G} \subset \mathrm{F} + \frac{1}{4}\mathrm{B_G}$, ce qui entraîne que $\mathrm{E} = \mathrm{G} = \mathrm{F}$ est de dimension finie.

On dit qu'un opérateur borné T d'un espace de Banach E dans un espace de Banach F est *compact* si l'adhérence dans F de l'image de la boule unité de E est compacte dans l'espace F.

Corollaire 4.4.3. *Soit* E *un espace de Banach, réel ou complexe ; si* $\mathrm{T} \in \mathcal{L}(\mathrm{E})$ *est compact et si* $\lambda \ne 0$*, le sous-espace* $\mathrm{F}_\lambda = \ker(\mathrm{T} - \lambda \operatorname{Id_E}) = \{y \in \mathrm{E} : \mathrm{T}(y) = \lambda y\}$ *est de dimension finie.*

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

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

//

5. Topologies faibles

Dans ce chapitre on introduira la notion de *suite faiblement convergente*, et on verra deux résultats importants :

– si E est un espace de Banach réflexif, toute suite bornée $(x_n) \subset \mathrm{E}$ admet des sous-suites faiblement convergentes ;

– si f est une fonction convexe continue sur un espace normé X et si $(x_n) \subset \mathrm{X}$ converge faiblement vers $x \in \mathrm{X}$, alors $f(x) \leq \liminf_n f(x_n)$.

Ces deux résultats permettent de minimiser certaines fonctions convexes définies sur des espaces réflexifs.

5.1. Topologies initiales

Une façon courante de définir une topologie sur un ensemble X est de partir d'une famille $\mathcal{C}$ de parties élémentaires, dont on veut imposer qu'elles soient des ouverts de la topologie $\mathcal{T}$ à construire. Supposons que cette famille $\mathcal{C}$ vérifie

— tout point $x \in \mathrm{X}$ est contenu dans au moins une partie $\mathrm{C} \in \mathcal{C}$;

— pour tous $\mathrm{C}_1, \mathrm{C}_2 \in \mathcal{C}$ et tout $x \in \mathrm{C}_1 \cap \mathrm{C}_2$, il existe un ensemble $\mathrm{C} \in \mathcal{C}$ tel que $x \in \mathrm{C} \subset \mathrm{C}_1 \cap \mathrm{C}_2$.

Si ces deux conditions sont vérifiées, on peut définir une topologie $\mathcal{T}$ sur X en disant que Ω est un ouvert de $\mathcal{T}$ si pour tout $x \in \Omega$, il existe $\mathrm{C} \in \mathcal{C}$ tel que $x \in \mathrm{C} \subset \Omega$. Cela revient exactement à dire que les ouverts de la topologie $\mathcal{T}$ sont les réunions quelconques de familles d'éléments de $\mathcal{C}$ (y compris la famille vide dont la réunion est l'ensemble vide).

C'est ce que l'on fait pour les espaces métriques, en partant de la famille $\mathcal{C}$ de toutes les boules ouvertes $\mathrm{B}(x, r)$, $x \in \mathrm{X}$ et $r > 0$. Mais la deuxième condition ci-dessus, qui est vérifiée pour les espaces métriques, est en général peu naturelle. On préfère souvent partir d'une famille $\mathcal{B}$ de parties de X vérifiant une propriété plus forte

— tout point $x \in \mathrm{X}$ est contenu dans au moins une partie $\mathrm{B} \in \mathcal{B}$;

— pour tous $\mathrm{B}_1, \mathrm{B}_2 \in \mathcal{B}$, l'intersection $\mathrm{B}_1 \cap \mathrm{B}_2$ est un ensemble de $\mathcal{B}$.

La deuxième condition dit que $\mathcal{B}$ est stable par intersection finie. On définit à nouveau une topologie $\mathcal{T}$ sur X dont les ouverts sont les réunions quelconques de familles d'ensembles de $\mathcal{B}$. Un cas important est le suivant : donnons-nous une famille $(\mathrm{Y}_i)_{i \in \mathrm{I}}$ d'espaces topologiques et, pour tout $i \in \mathrm{I}$, une application $f_i : \mathrm{X} \to \mathrm{Y}_i$. On veut construire une topologie $\mathcal{T}$ sur X qui rende continue toutes les f_i, mais en y mettant *le moins possible* d'ouverts.

Proposition 5.1.1. *Donnons-nous deux ensembles* I *et* X, *une famille* $(\mathrm{Y}_i)_{i \in \mathrm{I}}$ *d'espaces topologiques et, pour tout* $i \in \mathrm{I}$, *une application* $f_i : \mathrm{X} \to \mathrm{Y}_i$. *Il existe une topologie* $\mathcal{O}$ *sur* X *caractérisée par la propriété suivante :*

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

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

La topologie initiale n'est en général pas une topologie associée à une distance – et les suites ne jouent donc pas un rôle aussi important que dans le cas métrique. Indiquons cependant qu'une suite (x_n) de points de X tend vers $x \in \mathrm{X}$ si et seulement si : pour tout $i \in \mathrm{I}$ la suite $f_i(x_n)$ converge vers $f_i(x)$ (appliquer la proposition à l'application g de $\mathrm{Z} = \overline{\mathbb{N}}$ dans X définie par $g(n) = x_n$ pour tout $n \in \mathbb{N}$ et $g(\infty) = x$).

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

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

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

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

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

Autrement dit, cette topologie initiale est la topologie sur X obtenue à partir de la famille $\mathcal{B}$, stable par intersection finie, formée de tous les ensembles $\mathrm{W}(\mathrm{J}, \mathrm{V}_\mathrm{J})$,

$$\mathrm{W}(\mathrm{J}, \mathrm{V}_\mathrm{J}) = \bigcap_{j \in \mathrm{J}} f_j^{-1}(\mathrm{V}_j)$$

où J est un sous-ensemble fini quelconque de I et $\mathrm{V}_\mathrm{J} = (\mathrm{V}_j)_{j \in \mathrm{J}}$ une famille telle que pour chaque $j \in \mathrm{J}$, V_j soit un ouvert quelconque de Y_j. Puisque $f_i^{-1}(\mathrm{V}_i)$ est dans $\mathcal{B}$ pour tout $i \in \mathrm{I}$ et tout ouvert V_i de Y_i, tous ces ensembles sont des ouverts de X et il en résulte que toutes les (f_i) sont continues ; si g est continue de Z dans X, toutes les $f_i \circ g$ sont donc continues. Mais inversement, si $f_i \circ g$ est continue pour tout i, $g^{-1}(f_i^{-1}(\mathrm{V}_i)) = (f_i \circ g)^{-1}(\mathrm{V}_i)$ est ouvert dans Z pour tout i et tout ouvert V_i de Y_i, donc par intersection finie $g^{-1}(\mathrm{W}(\mathrm{J}, \mathrm{V}_\mathrm{J}))$ est ouvert dans Z pour tout ensemble fini $\mathrm{J} \subset \mathrm{I}$ et toute famille V_J, et enfin $g^{-1}(\mathrm{U})$ est ouvert pour tout U ouvert de X, puisqu'un tel ouvert U est réunion d'une famille d'ensembles $\mathrm{W}(\mathrm{J}, \mathrm{V}_\mathrm{J})$.

Proposition 5.1.2. *Donnons-nous un ensemble* I*, un espace vectoriel* X*, une famille* $(\mathrm{Y}_i)_{i \in \mathrm{I}}$ *d'espaces vectoriels topologiques (cf. définition 1.1.3) et, pour tout* $i \in \mathrm{I}$*, une application linéaire* $f_i : \mathrm{X} \to \mathrm{Y}_i$*. Muni de la topologie initiale associée aux applications* f_i*, l'espace* X *est un espace vectoriel topologique.*

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

//

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

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

Exemple 5.1.1. Soient X, Y des espaces vectoriels et $\mathrm{B} : \mathrm{X} \times \mathrm{Y} \to \mathbb{K}$ une forme bilinéaire ; pour $y \in \mathrm{Y}$ notons $p_y : \mathrm{X} \to \mathbb{R}_+$ l'application $x \to |\mathrm{B}(x, y)|$. C'est une semi-norme sur X. La topologie sur X associée à la famille des semi-normes p_y s'appelle la topologie *faible* associée à B, ou à Y s'il n'y a pas d'ambiguïté sur la forme B ; cette topologie se note $\sigma(\mathrm{X}, \mathrm{Y})$. Cette topologie est la topologie initiale associée aux applications $f_y : \mathrm{X} \to \mathbb{K}$, où, pour $y \in \mathrm{Y}$ on a noté f_y l'application $x \to \mathrm{B}(x, y)$. En d'autres termes, la topologie faible est la topologie la plus faible pour laquelle les applications f_y sont continues.

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

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

Terminons ce paragraphe par une définition.

Définition 5.1.2. Soient X et Y deux espaces normés ; chaque élément x de X définit une semi-norme N_x sur $\mathcal{L}(\mathrm{X}, \mathrm{Y})$ donnée par $\mathrm{N}_x(f) = \|f(x)\|$. On appelle *topologie forte* sur $\mathcal{L}(\mathrm{X}, \mathrm{Y})$ la topologie associée à la famille de semi-normes $(\mathrm{N}_x)_{x \in \mathrm{X}}$.

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

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

5.2. Topologie faible sur un espace normé

Soit X un espace vectoriel normé ; considérons la forme bilinéaire $(x^*, x) \to x^*(x)$ sur $\mathrm{X}^* \times \mathrm{X}$. Cette forme bilinéaire définit des topologies faibles $\sigma(\mathrm{X}, \mathrm{X}^*)$ sur X et $\sigma(\mathrm{X}^*, \mathrm{X})$ sur X^*, comme indiqué dans l'exemple du paragraphe précédent.

La topologie $\sigma(\mathrm{X}, \mathrm{X}^*)$ sur X s'appelle aussi la *topologie faible* sur X. C'est la topologie la moins fine rendant continues toutes les applications $x \in \mathrm{X} \to x^*(x)$, où x^* décrit l'ensemble de toutes les formes linéaires continues (en norme) sur X ; bien entendu la topologie de la norme rend déjà continues toutes ces applications, donc la topologie faible $\sigma(\mathrm{X}, \mathrm{X}^*)$ est plus faible que la topologie de la norme. D'après le théorème de Hahn-Banach, la topologie faible est séparée : si $x_1 \neq x_2$, on peut trouver $x^* \in \mathrm{X}^*$ telle que $x^*(x_1 - x_2) \neq 0$.

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

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

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

On peut aussi considérer la boule unité fermée B(X) de X, et la munir de la topologie induite par la topologie faible ; même dans ce cadre, la topologie de $(B(X), \sigma(X, X^*))$ est strictement plus faible en général que celle de B(X) munie de la topologie de la norme : soit $X = \ell^2$, l'espace hilbertien des suites de carré sommable ; notons (e_n) sa base hilbertienne canonique. On vérifie que la suite (e_n) tend faiblement vers 0 (voir lemme 3.1.b). Cependant la suite (e_n) ne peut converger en norme, vu que $\|e_n - e_m\|^2 = 2$ (pour $m \neq n$). En fait, pour tout espace normé de dimension infinie la topologie de $(B(X), \sigma(X, X^*))$ est strictement plus faible que la topologie de la norme, pour la raison évoquée plus haut : si X est de dimension infinie, un voisinage faible de 0_X dans $(B(X), \sigma(X, X^*))$ contient toujours des points de la sphère unité de X.

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

Proposition 5.2.1. *Soient* X *et* Y *deux espaces normés et* $T \in \mathcal{L}(X, Y)$ *; alors* T *est continue de* X *muni de la topologie* $\sigma(X, X^*)$ *dans* Y *muni de la topologie* $\sigma(Y, Y^*)$.

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

//

Théorème 5.2.2. *Soient* X *un espace normé et* C *un sous-ensemble convexe de* X *; l'ensemble* C *est fermé en norme si et seulement s'il est faiblement fermé.*

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

Démonstration. Comme la topologie de la norme est plus fine que la topologie faible, toute partie fermée pour la topologie faible est fermée pour la topologie de la norme. Démontrons l'inverse ; si X est réel, les demi-espaces affines fermés de la forme $\{x \in X : x^*(x) \geq c\}$, où $x^* \in X^*$, sont faiblement fermés puisque x^* est faiblement continue ; dans le cas complexe la forme $\mathbb{R}$-linéaire $\mathrm{Re}\, x^*$ est une application faiblement continue de X dans $\mathbb{R}$; il résulte de ces considérations et du corollaire 4.2.4 que tout convexe fermé est faiblement fermé.

//

Soit X un espace vectoriel normé ; on va maintenant s'intéresser à une topologie faible sur le dual X^*, la topologie $\sigma(X^*, X)$ ou *topologie $*$-faible* sur X^*. C'est la topologie la moins fine sur X^* rendant continues toutes les applications $x^* \in X^* \to x^*(x)$, où x décrit X ; bien entendu, puisque la topologie de la norme de X^* rend continues toutes ces applications, la topologie $*$-faible est plus faible que la topologie de la norme sur X^*. La topologie $*$-faible est la *topologie de la convergence simple sur* X, topologie sur l'ensemble d'applications X^*, applications de X dans $\mathbb{K}$.

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

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

Il en résulte que si X est de dimension infinie, un voisinage $*$-faible de x_0^* n'est jamais borné. On peut aussi considérer la boule unité fermée $B(X^*)$ de X^*, et la munir de la topologie induite par la topologie $*$-faible ; pour tout espace normé de dimension infinie la topologie de $(B(X^*), \sigma(X^*, X))$ est strictement plus faible que la topologie de la norme, toujours pour les mêmes raisons : si X est de dimension infinie, un voisinage faible de 0_{X^*} dans $(B(X^*), \sigma(X^*, X))$ contient toujours des points de la sphère unité de X^*. On verra aussi une autre raison avec le théorème qui suit : l'espace topologique $(B(X^*), \sigma(X^*, X))$ est toujours compact, alors que $B(X^*)$ n'est jamais compacte en norme si la dimension est infinie (théorème 4.4.2).

Théorème 5.2.3. *Muni de la topologie $\sigma(X^*, X)$ la boule unité de X^* est compacte.*

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

Théorème 5.2.4 : théorème de Tykhonov. *Tout produit d'espaces compacts (muni de la topologie produit) est compact.*

Rappelons que la topologie produit sur $\prod_{i \in I} X_i$ est la topologie la moins fine qui rende continues toutes les projections $\prod_{i \in I} X_i \to X_j$, $j \in I$.

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

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

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

Toutes ces conditions définissent des fermés, donc $B = F_1 \cap F_2$ est un fermé de $\mathbb{K}^X$. Pour $r \geq 0$, posons $D_r = \{\lambda \in \mathbb{K} : |\lambda| \leq r\}$. On a $B \subset \prod_{x \in X} D_{\|x\|}$, qui est compact par le théorème 4 ; étant fermé dans un compact, B est compact.

//

Corollaire 5.2.5. *Si* E *est réflexif, les convexes fermés bornés de* E *sont faiblement compacts.*

5.3. Suites faiblement convergentes

Il est intéressant de revoir certaines de ces propriétés de compacité “à la main”, et avec des *suites*. Rappelons qu'une suite $(x_n^*) \subset \mathrm{X}^*$ est $*$-faiblement convergente vers un vecteur x^* si $\lim_n x_n^*(x) = x^*(x)$ pour tout $x \in \mathrm{X}$; une suite $(x_n) \subset \mathrm{X}$ est faiblement convergente vers $x \in \mathrm{X}$ si $x^*(x) = \lim_n x^*(x_n)$ pour tout $x^* \in \mathrm{X}^*$.

Exemples 5.3.1.

1. Lorsque H est un espace de Hilbert, toute forme linéaire continue sur H est de la forme $x \to \langle x, y\rangle$ pour un certain vecteur $y \in \mathrm{H}$; il en résulte qu'une suite $(x_n) \subset \mathrm{H}$ converge faiblement vers $x \in \mathrm{H}$ si et seulement si

$$\forall y \in \mathrm{H}, \quad \langle x, y\rangle = \lim_n \langle x_n, y\rangle.$$

2. Soit $(e_n)_{n\geq 0}$ une suite orthonormée dans un espace de Hilbert ; alors e_n converge faiblement vers 0.

En effet, si E est un Hilbert et (e_n) une suite orthonormée dans E, on a

$$\sum_{n\geq 0} |\langle e_n, x\rangle|^2 \leq \|x\|^2$$

pour tout $x \in \mathrm{E}$ (inégalité de Bessel - théorème 6.3.2) ; la suite $(\langle e_n, x\rangle)$ est de carré sommable donc tend vers 0.

3. La suite canonique $(\mathbf{e}_n)_{n\geq 0}$ tend faiblement vers 0 dans c_0, et dans l'espace ℓ_p si $1 < p < +\infty$, mais **pas** dans ℓ_1 ou ℓ_∞.

Lemme 5.3.1. *Dans un espace normé toute suite faiblement convergente est bornée. Si* X *est complet, toute suite $*$-faiblement convergente dans* X^* *est bornée.*

Démonstration. Soit (x_n) une suite faiblement convergente ; en plongeant isométriquement X dans X^{**} on peut considérer (x_n) comme une suite d'applications linéaires de l'espace de Banach X^* dans $\mathbb{K}$ qui converge en tout point $x^* \in \mathrm{X}^*$; il résulte alors du théorème de Banach-Steinhaus (corollaire 4.1.9) que $\{\|x_n\| : n \geq 0\}$ est borné. Le deuxième cas est une application directe du même corollaire 4.1.9 : si la suite d'applications (x_n^*) converge en tout point de l'espace de Banach X, elle est bornée en norme.

//

Proposition 5.3.2. *Si f est une fonction réelle convexe définie sur un convexe fermé* C *d'un espace normé* X, *et si $(x_n) \subset \mathrm{C}$ converge faiblement vers x, on a $f(x) \leq \liminf_n f(x_n)$.*

Démonstration. Posons $\ell = \liminf_n f(x_n)$. Si $\ell = +\infty$ l'inégalité à démontrer est évidente. Sinon, soit $m > \ell$. Considérons l'ensemble $\mathrm{D}_m = \{y \in \mathrm{C} : f(y) \leq m\}$. C'est un convexe fermé, donc faiblement fermé, et puisque $\liminf_n f(x_n) < m$ il existe une sous-suite $(x_{n_k})_{k\geq 0}$ contenue dans D_m, donc sa limite faible x reste dans D_m. On a donc $f(x) \leq m$ pour tout $m > \ell$, ce qui montre que $f(x) \leq \liminf_n f(x_n)$.

//

Lemme 5.3.3. *Soit* (x_n^*) *une suite* **bornée** *dans le dual d'un espace normé* X *; pour que cette suite soit* $*$*-faiblement convergente vers* x^**, il suffit que* $x^*(d) = \lim_n x_n^*(d)$ *pour tout* d *d'un ensemble* D *total dans* X*. Pour que cette suite soit* $*$*-faiblement convergente vers une limite dans* X^**, il suffit que* $\lim_n x_n^*(d)$ *existe pour tout* d *d'un ensemble* D *total dans* X*.*

Démonstration. Montrons la deuxième variante de l'énoncé. Supposons $\|x_n^*\| \leq 1$ pour tout $n \geq 0$. Si la limite existe pour tout d d'un ensemble total T, elle existe aussi, par linéarité, pour tout d de l'ensemble dense $\mathrm{D} = \mathrm{Vect}(\mathrm{T})$. Montrons que $(x_n^*(x))_{n \geq 0}$ converge pour tout $x \in \mathrm{X}$. Il suffit de montrer que cette suite de scalaires est de Cauchy. Choisissons $\varepsilon > 0$ et $d \in \mathrm{D}$ tels que $\|x - d\| < \varepsilon/3$. Pour tout entier n, on a $|x_n^*(x) - x_n^*(d)| < \varepsilon/3$, et la suite $(x_n^*(d))_{n \geq 0}$ converge vers une limite ℓ ; il en résulte que pour n assez grand, on aura $|x_n^*(x) - \ell| < \varepsilon/2$, et si n, m sont assez grands, on aura $|x_n^*(x) - x_m^*(x)| < \varepsilon$. La suite $(x_n^*(x))$ est donc de Cauchy, donc convergente. Il est alors clair que la formule

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

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

//

Exemples 5.3.2.

1. La suite canonique $(\mathbf{e}_k)_{k \geq 0}$ est totale dans $\mathrm{X} = \ell_p$, $1 \leq p < \infty$. Soit $q \in]1, +\infty]$ le conjugué de p ; pour qu'une suite $(y^{(n)})$, bornée dans $\mathrm{X}^* \simeq \ell_q$ soit $*$-faiblement convergente (vers un certain $y \in \ell_q$) il suffit que $\lim_n y_k^{(n)}$ existe pour tout k : pour une suite bornée, la convergence $*$-faible est identique à la convergence pour la topologie produit de $\mathbb{K}^{\mathbb{N}}$ (convergence simple, coordonnée par coordonnée). En particulier, la suite canonique $(\mathbf{e}_k)$ tend $*$-faiblement vers 0 dans ℓ_q, considéré comme dual de ℓ_p, pour tout $q \in]1, +\infty]$; elle tend aussi $*$-faiblement vers 0 dans ℓ_1, considéré comme dual de c_0.

2. Définissons une suite (f_n) de fonctions dans $\mathrm{L}_\infty(0,1)$ par $f_n(t) = (-1)^{[nt]}$ (partie entière). Cette suite est formée de fonctions de module un, donc de norme un dans L_∞. Elle tend vers 0 dans $\sigma(\mathrm{L}_\infty, \mathrm{L}_1)$:

prenons $\mathrm{T} = \mathrm{C}([0,1])$, qui est dense dans $\mathrm{X} = \mathrm{L}_1$. On aura pour g continue

$$(*) \qquad (f_n, g) = \int_0^1 f_n(t) g(t)\, dt = \sum_{k=0}^{n-1} (-1)^k \int_{k/n}^{(k+1)/n} g(t)\, dt.$$

Utilisons la continuité uniforme de g. Etant donné $\varepsilon > 0$, il existe $\delta > 0$ tel que $|g(t) - g(s)| < \varepsilon$ si $|t - s| < \delta$. On prend n_0 tel que $n_0 > \delta^{-1}$. On voit que si $n \geq n_0$

$$\Bigl| \int_{k/n}^{(k+1)/n} g(t)\, dt - \int_{(k+1)/n}^{(k+2)/n} g(t)\, dt \Bigr| \leq \frac{\varepsilon}{n}$$

et on en déduit en regroupant dans $(*)$ les morceaux d'intégrale deux par deux que $|(f_n, g)| \leq \varepsilon + \|g\|_\infty / n$.

Proposition 5.3.4. *Si* X *est un espace normé séparable, toute suite bornée de* X^* *admet des sous-suites* $*$*-faiblement convergentes.*

Démonstration. Pour exprimer la démonstration, il est utile d'introduire une petite convention de notation. Si $M = \{n_0 < \ldots < n_j < \ldots\}$ est un sous-ensemble infini de $\mathbb{N}$, convenons de noter la sous-suite (x_{n_j}) par $(x_n)_{n\in M}$. Soit donc (y_k) une suite dense dans X, et (x_n^*) une suite bornée dans X^*, telle que par exemple $\|x_n^*\| \le 1$ pour tout entier $n \ge 0$. La suite de scalaires $(x_n^*(y_0))$ est bornée, donc elle admet une sous-suite convergente $(x_n^*(y_0))_{n\in M_0}$. La suite $(x_n^*(y_1))_{n\in M_0}$ est encore bornée, donc on peut trouver un nouvel ensemble infini $M_1 \subset M_0$ tel que la sous-suite $(x_n^*(y_1))_{n\in M_1}$ soit convergente. En continuant ainsi, on construit une suite décroissante $M_0 \supset M_1 \supset \ldots \supset M_j \supset \ldots$ telle que $(x_n^*(y_j))_{n\in M_j}$ soit convergente pour tout $j \ge 0$.

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

//

Exemple 5.3.3. Si (μ_n) est une suite de probabilités sur le compact $[0,1]$, il existe une sous-suite (μ_{n_j}) et une probabilité μ sur $[0,1]$ telles que $\int f\,d\mu = \lim_j \int f\,d\mu_{n_j}$ pour toute fonction continue f sur $[0,1]$; on dit que la sous-suite (μ_{n_j}) converge *vaguement* vers μ. Le résultat provient du fait que l'espace des mesures sur $[0,1]$ est le dual de l'espace *séparable* $C([0,1])$.

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

Théorème 5.3.5. *Si* X *est réflexif, toute suite bornée dans* X *admet des sous-suites faiblement convergentes.*

Démonstration. Soit (x_n) une suite bornée dans X ; le sous-espace fermé Y engendré par la suite (x_n) est un espace réflexif séparable, donc son dual Y^* est séparable et réflexif. On peut donc appliquer le théorème précédent à $Y^{**} \simeq Y$; en considérant (x_n) comme une suite bornée dans Y^{**}, on peut trouver une sous-suite (x_{n_j}) qui soit $*$-faiblement convergente dans Y^{**} vers un élément y^{**}, c'est à dire telle que

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

Puisque Y est réflexif, il existe un vecteur $y \in Y$ tel que $y^{**} = J_Y(y)$, et la relation ci-dessus nous dit que $y^*(y) = y^{**}(y^*) = \lim_j y^*(x_{n_j})$ pour tout $y^* \in Y^*$, ce qui signifie que la sous-suite (x_{n_j}) converge faiblement vers y dans Y. Si x^* est une forme linéaire continue sur X, elle n'agira sur les (x_n) et sur $y \in Y$ que par sa restriction $y^* \in Y^*$ à l'espace Y, et on aura encore $x^*(y) = y^*(y) = \lim_j y^*(x_{n_j}) = \lim_j x^*(x_{n_j})$. On a donc montré que la sous-suite (x_{n_j}) converge faiblement dans X vers le vecteur y.

//

Théorème 5.3.6. *Si* C *est un convexe fermé borné non vide d'un espace réflexif et si f est une fonction convexe continue sur* C*, elle atteint son minimum sur* C.

Démonstration. On peut trouver une suite $(x_n) \subset \mathrm{C}$ telle que $(f(x_n))$ converge en décroissant vers $\inf f(\mathrm{C})$ (peut-être $-\infty$) ; la suite (x_n) est bornée puisque C est borné ; quitte à passer à une sous-suite on peut supposer que (x_n) converge faiblement vers $x \in \mathrm{E}$; comme C est faiblement fermé, on sait que $x \in \mathrm{C}$. Fixons momentanément un entier $m \geq 0$ et considérons l'ensemble $\mathrm{D}_m = \{y \in \mathrm{C} : f(y) \leq f(x_m)\}$. C'est un convexe fermé, donc faiblement fermé, et la suite $(x_n)_{n \geq m}$ est contenue dans D_m, donc sa limite faible x reste dans D_m. On a donc $f(x) \leq f(x_m)$ pour tout m, ce qui montre que f atteint son minimum au point x.

//

Corollaire 5.3.7. *Si* C *est un convexe fermé borné non vide d'un espace réflexif réel* E *et si $x^* \in \mathrm{E}^*$, la forme linéaire x^* atteint son maximum sur* C.

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

Exercice.

A. Montrer que la forme linéaire continue ℓ sur $\mathrm{L}_1(0,1)$, définie par

$$\forall f \in \mathrm{L}_1(0,1), \quad \ell(f) = \int_0^1 t f(t)\, dt$$

n'atteint pas son sup sur la boule unité de $\mathrm{L}_1(0,1)$.

B. Sur $\mathrm{E} = \mathrm{C}([0,1])$, on considère la forme linéaire continue

$$\ell(f) = \int_0^{1/2} f(t)\, dt - \int_{1/2}^1 f(t)\, dt.$$

Montrer que ℓ n'atteint pas son sup sur la boule unité fermée de $\mathrm{C}([0,1])$.

Corollaire 5.3.8. *Si f est convexe continue sur un convexe fermé non vide d'un espace réflexif* E *et si $f(x)$ tend vers $+\infty$ lorsque $\|x\| \to +\infty$, la fonction f atteint son minimum sur* E.

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

//

Remarque 5.3.4. C'est ce type de résultat qui permet de minimiser les fonctionnelles telles que celle qui a été évoquée dans l'introduction.

5.4. Appendice : démonstration du Théorème de Tykhonov

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

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

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

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

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

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

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

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

Remarque 5.4.1. On a vu que la famille maximale $\mathcal{U}$ est stable par intersection finie ; on voit aussi par maximalité que si $\mathrm{U} \in \mathcal{U}$ et $\mathrm{U} \subset \mathrm{Y}$, alors $\mathrm{Y} \in \mathcal{U}$. Il en résulte que pour tout $\mathrm{Y} \subset \mathrm{X}$, ou bien $\mathrm{Y} \in \mathcal{U}$, ou bien $\mathrm{Y} \notin \mathcal{U}$. Une telle famille $\mathcal{U}$ de parties de X est appelée *ultrafiltre* sur X.