4. Les théorèmes fondamentaux

4.1. Le théorème de Baire et ses conséquences

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

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

Démonstration. Soit $(\mathrm{U}_n)_{n\geq 0}$ une suite d'ouverts denses de X ; soit V une partie ouverte non vide de X ; on doit montrer que $\bigcap_{n\geq 0} \mathrm{U}_n$ rencontre V. Comme U_0 est dense, U_0 rencontre V et on peut choisir un point $x_0 \in \mathrm{V} \cap \mathrm{U}_0$. Comme $\mathrm{V} \cap \mathrm{U}_0$ est ouvert, il existe un nombre $r_0 > 0$, que l'on peut choisir ≤ 1, tel que la boule ouverte $\mathrm{B}(x_0, 2r_0)$ de centre x_0 et de rayon $2r_0$ soit contenue dans $\mathrm{V} \cap \mathrm{U}_0$.

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

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

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

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

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

Démonstration. Soit V un ouvert non vide de X ; dans l'espace métrique complet $\mathrm{Y} = \overline{\mathrm{V}}$, considérons les ouverts (relatifs) $\mathrm{U}_n = \overline{\mathrm{V}} \setminus \mathrm{F}_n$. Puisque $\bigcup_{n\geq 0} \mathrm{F}_n = \mathrm{X}$, l'intersection $\bigcap_{n\geq 0} \mathrm{U}_n$ 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 $x \in \mathrm{V} \cap \mathrm{U}$ et $r > 0$ tels que $\mathrm{B}(x, r) \subset \mathrm{V}$ ne rencontre pas U_{n_0}, c'est à dire $\mathrm{B}(x,r) \subset \mathrm{F}_{n_0}$ et en particulier $x \in \mathring{\mathrm{F}}_{n_0}$. On a ainsi montré que la réunion des intérieurs des (F_n) rencontre tout ouvert non vide V donné.

//

Exercice 4.1.1. Montrer qu'il n'existe pas de norme sur l'espace vectoriel $\mathbb{R}[\mathrm{X}]$ 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.*

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. Notons 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.

//

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 $\varphi \in \mathcal{S}$.

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 $-a + b = q(-1) + q(1) \geq q(0) = 0$).

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 $m(t) = c\,t$ 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* $g(z) \leq q(z)$ *pour tout* $z \in \mathrm{Z}$ *; soit* $x \in \mathrm{X}$ *tel que* $x \notin \mathrm{Z}$ *; il existe une forme linéaire* $\widetilde{g}$ *sur* $\mathrm{Z} + \mathbb{R}x$ *telle que* $\widetilde{g}$ *prolonge* g *et* $\widetilde{g} \leq q$ *sur* $\mathrm{Z} + \mathbb{R}x$.

Démonstration du lemme. Bien entendu, prolonger g à $\mathrm{Z} + \mathbb{R}x$ demande seulement de définir $\gamma = \widetilde{g}(x)$. 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 $t\neq 0$:

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

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

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

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*

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

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

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

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 $\ell \leq \|\ell\|\,\mathrm{N} = q$ sur le sous-espace vectoriel Y. On peut donc trouver un prolongement m tel que $m \leq q$ 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 $\ell_1 = \operatorname{Re} \ell$. 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* $r = \operatorname{dist}(x, \mathrm{Y}) > 0$ *; 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 $\|x^*\| = 1$.

//

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 $x \ne 0_{\mathrm{X}}$, 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.

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 $n \geq 0$; 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 $\|x_{n_0}^*\| \geq \|x^*\| - \|x_{n_0}^* - x^*\| \geq 3/4$, ce qui est contradictoire avec $x_{n_0}^*(x_{n_0}) \geq \|x_{n_0}^*\|/2 \geq 3/8 > 1/4$.

//

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 $\|{}^t\mathrm{T}\| \leq \|\mathrm{T}\|$ 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 $\|x\| \leq 1$ et tel que $\|\mathrm{T}(x)\| > \|\mathrm{T}\| - \varepsilon$, puis une forme linéaire $y^* \in \mathrm{F}^*$ telle que $\|y^*\| \leq 1$ et $y^*(\mathrm{T}(x)) = \|\mathrm{T}(x)\|$. Alors

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

//

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 $\mathrm{Y} \neq \{0\}$.

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

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

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* $\mathrm{im}(\mathrm{T})$ *est dense dans* Y. *De plus, si* $\mathrm{im}({}^t\mathrm{T})$ *est dense dans* X^**, l'application* T *est injective.*

Démonstration. Si $\mathrm{im}(\mathrm{T})$ 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 $\mathrm{im}(\mathrm{T})$ 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 $|J_X(x)(x^*)| = |x^*(x)| \leq \|x^*\| \, \|x\|$, donc $J_X(x) \in X^{**}$ et $\|J_X(x)\| \leq \|x\|$. 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 $\|x^*\| \leq 1$ et $x^*(x) = \|x\|$. Alors

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

vu que $\|x^*\| \leq 1$; 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.

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.

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 $z = x^* = x^{***} \circ \mathrm{J}_\mathrm{X}$ 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 $z^* = \mathrm{J}_\mathrm{X}(x)$. Alors

$$z^{**}(z^*) = x^{***}(\mathrm{J}_\mathrm{X}(x)) = x^*(x) = \mathrm{J}_\mathrm{X}(x)(x^*) = z^*(z),$$

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 $x \in X$ tel que $x^{**}(x^*) = x^*(x)$ pour tout x^* in X^*. Il suffit de voir que $x \in Y$ pour pouvoir conclure assez facilement ; si on avait $x \notin Y$, on pourrait trouver d'après le corollaire 2.6 une forme linéaire $x^* \in X^*$ telle que $x^*(x) = 1$ mais $x^*(y) = 0$ pour tout $y \in Y$. 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 $\bigcap_n C_n$ 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* $\geq \varepsilon$ *: si* $x, y \in A$ *et* $x \neq y$, *alors* $\|x - y\| \geq \varepsilon$, *et* $\operatorname{card} A \geq \varepsilon^{-n}$.

Démonstration. Soit A une famille maximale de points de la boule unité B_Z de Z dont les distances mutuelles soient $\geq \varepsilon$; alors les boules de rayon ε centrées aux points de A recouvrent B_Z : en effet, si $x \in B_Z$ et $x \notin A$, on ne peut pas, d'après la maximalité de A, ajouter le point x à la famille A pour former une nouvelle famille A' de points à distances mutuelles $\geq \varepsilon$; cela signifie qu'il existe un point $y \in A$ 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 B_Z ; puisque Z est de dimension n, les boules de rayon ε ont un volume égal à $\varepsilon^n V$ (dans le cas des scalaires réels), et puisque les boules de ce rayon centrées aux points de A recouvrent B_Z, on a $(\operatorname{card} A)\, \varepsilon^n V \geq V$, 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 $Z \subset X$ d'une dimension finie n telle que $2^n > N$; il existerait alors dans la boule unité de Z une famille d'au moins 2^n points tels que $\|z_i - z_j\| \geq 1/2$. Mais alors chacune des boules B_α contiendrait *au plus un* des points (z_i), donc $\mathrm{N} \geq 2^n$, contradiction.

//

On dit qu'un opérateur 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* $\lambda \neq 0$, *le sous-espace* $\mathrm{F}_\lambda = \ker(\mathrm{T} - \lambda \,\mathrm{Id}_\mathrm{E}) = \{y \in \mathrm{E} : \mathrm{T}(y) = \lambda y\}$ *est de dimension finie.*

Démonstration. Désignons par K le compact de E égal à l'adhérence de $\mathrm{T}(\mathrm{B}_\mathrm{E})$. 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 $\mathrm{B}_\mathrm{F} \subset |\lambda|^{-1}\mathrm{K}$. Soit $y \in \mathrm{B}_\mathrm{F}$; on a

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

//

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

Voir la version longue du poly.

5.2. Topologie faible sur un espace normé

On essaie de définir une topologie $\mathcal{T}$ sur un espace normé X, en y mettant *le moins possible* d'ouverts, mais de façon que toutes les applications $x \to x^*(x)$, pour $x^* \in \mathrm{X}^*$, restent continues pour $\mathcal{T}$. On supposera X réel pour l'instant, pour simplifier la description. Désignons par $\mathcal{O}$ la famille des ouverts de la topologie normée usuelle de X.

Si $\mathcal{T}$ vérifie cette condition et si $x^* \in \mathrm{X}^*$, $a, b \in \mathbb{R}$, alors l'ensemble

$$\{x \in \mathrm{X} : a < x^*(x) < b\}$$

doit être un ouvert de $\mathcal{T}$. C'est aussi un ouvert de $\mathcal{O}$. Puisque toute intersection finie d'ouverts est un ouvert, la famille d'ouverts $\mathcal{T}$ doit contenir tout ensemble B de la forme

$$\mathrm{B} = \bigcap_{i=1}^{k} \{x \in \mathrm{X} : a_i < x_i^*(x) < b_i\}$$

pour tout entier $k \geq 1$, toutes suites finies $x_1^*, \ldots, x_k^*$ dans X^* et $a_1, b_1, \ldots, a_k, b_k$ dans $\mathbb{R}$. Notons encore que tout ensemble B est dans $\mathcal{O}$.

Pour finir $\mathcal{T}$ doit contenir toutes les réunions $\bigcup_{i \in \mathrm{I}} \mathrm{B}_i$ d'ensembles B_i de la forme précédente. Mais la famille de ces réunions est une topologie sur X ; c'est la topologie voulue, *la topologie la moins fine qui rend continues* toutes les applications $x \to x^*(x)$, quand x^* décrit X^*. Puisque tous les ensembles de la forme B_i sont dans $\mathcal{O}$, il en résulte que tout ouvert de $\mathcal{T}$ est un ouvert de $\mathcal{O}$.

La topologie $\mathcal{T}$ sur X s'appelle la *topologie faible* sur X. La topologie faible est plus faible que la topologie de la norme, puisqu'on a vérifié que tous les ouverts faibles sont des ouverts de la topologie de la norme.

Soit X un espace vectoriel normé, réel ou complexe ; pour que $\mathrm{W} \subset \mathrm{X}$ soit un voisinage du point $x_0 \in \mathrm{X}$ pour la topologie faible (on dira que W est un *voisinage faible* de x_0), il faut et il suffit qu'il existe un nombre fini de formes linéaires continues $x_1^*, \ldots, x_n^* \in \mathrm{X}^*$ et un nombre $\varepsilon > 0$ tels que

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

La topologie faible sur X 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 $\sigma(X, X^*)$ 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 $x^* \in X^*$ telle que $x^*(x_1 - x_2) \neq 0$; les deux ensembles $U_j = \{x \in X : |x^*(x - x_j)| < |x^*(x_1 - x_2)|/2\}$, $j = 1, 2$, sont deux ouverts faibles disjoints contenant x_1 et x_2.

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* $\sigma(Y, Y^*)$.

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 $\operatorname{Re} x^*$ 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 $\mathbb{K}$.

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.

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 $\mathrm{B} = \mathrm{F}_1 \cap \mathrm{F}_2$ est un fermé de $\mathbb{K}^{\mathrm{X}}$. Pour $r \geq 0$, posons $\mathrm{D}_r = \{\lambda \in \mathbb{K} : |\lambda| \leq r\}$. On a $\mathrm{B} \subset \prod_{x \in \mathrm{X}} \mathrm{D}_{\|x\|}$, 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 $f(x) \le \liminf_n f(x_n)$.*

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

//

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 $\|x_n^*\| \le 1$ pour tout $n \ge 0$. 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 $(x_n^*(x))_{n \ge 0}$ 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 $(x_n^*(d))_{n \ge 0}$ 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 $\|x^*\| \le 1$, et la suite $(x_n^*)_{n \ge 0}$ converge $*$-faiblement vers x^*.

//

Exemples 5.3.2.

1. La suite canonique $(\mathbf{e}_k)_{k \ge 0}$ est totale dans $\mathrm{X} = \ell_p$, $1 \le p < \infty$. 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 $n \geq n_0$

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

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

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

//

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

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 $\mathrm{Y}^{**} \simeq \mathrm{Y}$; 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

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

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

//

Théorème 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.

//