Cours de topologie et d'analyse fonctionnelle
Master première année

Raphaël Danchin

7 janvier 2013

Table des matières

Chapitre 1

Un peu de topologie générale

But recherché : Formaliser les concepts "intuitifs" de proximité de deux objets mathématiques de "même nature", et de continuité des applications agissant sur ces objets.

1.1 Définitions

En licence, nous avons vu que l'introduction d'une norme permettait d'atteindre ce but. Dans ce cours, nous serons parfois amenés à manipuler des objets pour lesquels l'introduction d'une norme ne permet pas de mesurer la proximité de façon adéquate, et nous aurons besoin de définir une notion encore plus générale.

Cela peut être fait par le biais de la définition suivante.

Définition. Soit X un ensemble. On appelle **topologie** sur X la donnée d'un ensemble $\mathcal{O}$ de parties de X possédant les propriétés suivantes :

(i) $\mathcal{O}$ contient $\emptyset$ et X,

(ii) la réunion quelconque d'éléments de $\mathcal{O}$ est encore dans $\mathcal{O}$,

(iii) l'intersection finie d'éléments de $\mathcal{O}$ est encore dans $\mathcal{O}$.

Les éléments de $\mathcal{O}$ sont appelés **ouverts** et les complémentaires des ouverts sont appelés **fermés**. Le couple $(X, \mathcal{O})$ est appelé **espace topologique**.

Exemples.

1. Soit X un ensemble. Alors $\{\emptyset, X\}$ est une topologie sur X, parfois appelée **topologie grossière**.
2. Si l'on prend pour $\mathcal{O}$ l'ensemble de toutes les parties de X, on obtient également une topologie, appelée **topologie discrète**.
3. Sur $\mathbb{R}$, l'ensemble des réunions arbitraires d'intervalles ouverts $]a, b[$ est une topologie. Sauf mention explicite, on munit toujours $\mathbb{R}$ de cette topologie.

Définition. Soit X un ensemble et $\mathcal{O}_1$, $\mathcal{O}_2$ deux topologies sur X. On dit que $\mathcal{O}_1$ est **plus fine** (ou plus forte) que $\mathcal{O}_2$ si $\mathcal{O}_2 \subset \mathcal{O}_1$.

Ainsi, la topologie discrète est la plus fine et la topologie grossière, la moins fine de toutes les topologies. Dans le cas de $\mathbb{R}$, la topologie "usuelle" se situe entre les deux.

La donnée d'une topologie $\mathcal{O}$ sur l'ensemble X permet de définir les notions de voisinage, adhérence, intérieur, frontière, etc.

Définition. Soit $(X, \mathcal{O})$ un espace topologique.

– Soit $x \in X$ et $V \subset X$. On dit que V est un **voisinage** de x si V contient un ouvert contenant x. L'ensemble des voisinages de x est noté $\mathcal{V}_X(x)$.

– On dit qu'un sous-ensemble $\mathcal{V}$ de $\mathcal{V}_X(x)$ est une **base de voisinages** de x si tout élément de $\mathcal{V}_X(x)$ contient un élément de $\mathcal{V}$.
– On dit que x est un **point intérieur** à V si V est un voisinage de x.
– **L'adhérence** $\overline{A}$ d'une partie A de X est le plus petit fermé contenant A. On dit que A est **dense** dans X si $\overline{A} = X$.
– **L'intérieur** $\mathring{A}$ d'une partie A de X est le plus grand ouvert contenu dans A.
– L'ensemble $\overline{A}\backslash\mathring{A}$ est appelé **frontière** de A.

Remarque. On retiendra que l'ensemble des ouverts contenant un point x donné est une base de voisinages de x, et qu'*une partie est ouverte si et seulement si elle est voisinage de tous ses points.*

Attention : La définition de boule (vue en licence dans le cadre des espaces vectoriels normés) n'a pas d'équivalent dans un espace topologique "général".

Exercice : Soit $(X, \mathcal{O})$ un espace topologique et A une partie de X. Montrer les propriétés suivantes :

(i) $A = \overline{A}$ si et seulement si A est fermé.

(ii) $A = \mathring{A}$ si et seulement si A est ouvert.

(iii) A est ouvert si et seulement si A est un voisinage de tous ses points.

(iv) $X\backslash\overline{A} = \overbrace{X\backslash A}^{\circ}$ et $X\backslash\mathring{A} = \overline{X\backslash A}$.

(v) La frontière de A est égale à $X \setminus \big(\mathring{A} \cup \overbrace{X\backslash A}^{\circ}\big)$.

Retenons la caractérisation suivante de l'adhérence et de la densité :

Proposition. *Soit A une partie d'un espace topologique $(X, \mathcal{O})$.*

(i) Un point x de X est dans $\overline{A}$ si et seulement si tout voisinage de x rencontre A.

(ii) L'ensemble A est dense dans X si et seulement si A rencontre tout ouvert non vide de X.

Preuve : Tout d'abord, si $x \notin \overline{A}$ alors x appartient à l'ouvert $X \setminus \overline{A}$ qui est un voisinage de x ne rencontrant pas A.

Réciproquement, s'il existe un ouvert Ω contenant x et disjoint de A alors $X \setminus \Omega$ est un fermé contenant A, donc $\overline{A}$. En conséquence, x n'est pas dans $\overline{A}$.

Prouvons (ii). Soit A dense, Ω un ouvert non vide de X et x un point de Ω. Alors Ω est voisinage de $x \in \overline{A}$ donc rencontre A. Réciproquement, supposons que tout ouvert non vide de X rencontre A. Soit $x \in X$ quelconque et V un voisinage de x (que l'on peut toujours supposer ouvert). Alors V rencontre A donc $x \in \overline{A}$ puis $\overline{A} = X$. ■

La notion de "séparation des points" est fondamentale dans les espaces topologiques.

Définition. On dit qu'un espace topologique $(X, \mathcal{O})$ est **séparé** si pour tout couple (x, y) de points distincts de X il existe un voisinage V de x et un voisinage V' de y tels que $V \cap V' = \emptyset$.

Exemples. 1. Soit X un ensemble contenant au moins deux éléments et muni de la topologie grossière $\mathcal{O} = \{\emptyset, X\}$. Alors $(X, \mathcal{O})$ n'est pas séparé.

2. Un ensemble muni de la topologie discrète est toujours séparé. En effet, pour cette topologie, tout singleton est voisinage du point qu'il contient.

3. L'ensemble $\mathbb{R}$ muni de sa topologie usuelle est séparé.

La preuve de la proposition suivante constitue un excellent exercice.

Proposition. *Soit* $(X, \mathcal{O})$ *un espace topologique et* $A \subset X$. *L'ensemble* $\mathcal{O}_A$ *des parties* Ω *de* A *telles qu'il existe* $\Omega' \in \mathcal{O}$ *vérifiant* $\Omega = \Omega' \cap A$ *est une topologie sur* A. *On l'appelle* **topologie induite** *par* $\mathcal{O}$ *sur* A.

Remarque. Sauf mention contraire, on munit toujours les parties d'un ensemble topologique de la topologie induite.

Deux cas simples à retenir : si A est un ouvert de X alors tout ouvert de A est un ouvert de X, si A est un fermé de X alors tout fermé de A est un fermé de X.

1.2 Continuité

Il est maintenant temps d'aborder la deuxième partie du but recherché : généraliser la notion de continuité aux fonctions définies sur des espaces topologiques. Sans distance ou norme, il n'est plus question de donner une définition à l'aide d'ϵ et de η. L'utilisation de *voisinages* est le bon substitut. Commençons par définir la notion de limite.

Définition. Soit $(X, \mathcal{O})$ et $(X', \mathcal{O}')$ deux espaces topologiques. Soit A une partie de X, $f \in \mathcal{F}(A; X')$, $x_0 \in \overline{A}$ et $\ell \in X'$. On dit que f a pour **limite** ℓ en x_0 si

$$\forall V' \in \mathcal{V}_{X'}(\ell), \exists V \in \mathcal{V}_X(x_0), \big(x \in A \cap V \Rightarrow f(x) \in V'\big).$$

Remarque. Dans la définition ci-dessus, on peut bien sûr remplacer $\mathcal{V}_{X'}(\ell)$ (resp. $\mathcal{V}_X(x_0)$) par une base de voisinages de ℓ (resp x_0).

Exercice : Montrer l'unicité de la limite dans le cas où $(X', \mathcal{O}')$ est séparé. Que peut-on dire dans le cas général ?

On peut maintenant définir la continuité :

Définition. Sous les hypothèses précédentes, si de plus $x_0 \in A$ et $f(x_0) = \ell$, on dit que f est **continue** en x_0. Si f est continue en tout point de A, on dit que f est continue sur A.

Pour montrer la continuité en tout point, nul n'est besoin de recourir à la définition grâce au théorème suivant.

Théorème. *Soit* $(X, \mathcal{O})$ *et* $(X', \mathcal{O}')$ *deux espaces topologiques et* $f \in \mathcal{F}(X; X')$. *Les trois propriétés suivantes sont équivalentes :*

(i) la fonction f *est continue sur* X,

(ii) l'image réciproque par f *de tout ouvert de* X' *pour la topologie* $\mathcal{O}'$ *est un ouvert de* X *pour la topologie* $\mathcal{O}$,

(iii) l'image réciproque par f *de tout fermé de* X' *pour la topologie* $\mathcal{O}'$ *est un fermé de* X *pour la topologie* $\mathcal{O}$.

Preuve : Montrons juste l'équivalence entre les deux premières assertions.

Soit donc f continue sur X et Ω' un ouvert de X'. Considérons $x \in f^{-1}(\Omega')$ et posons $x' = f(x)$. L'ouvert Ω' est voisinage de x'. Par continuité de f en x, il existe donc un voisinage V de x tel que $f(V) \subset \Omega'$. Cela signifie que $V \subset f^{-1}(\Omega')$. Donc $f^{-1}(\Omega')$ est bien ouvert.

Réciproquement, supposons (ii) vérifiée, et fixons $x \in X$ et V' un voisinage de $f(x)$ (que l'on peut toujours supposer ouvert quitte à le diminuer). Alors $f^{-1}(V')$ contient x et est ouvert grâce à (ii). C'est donc un voisinage de x dont l'image est incluse dans V'. Cela assure la continuité en x. ■

Théorème (de composition)**.** *Soit* $(X, \mathcal{O})$, $(X', \mathcal{O}')$ *et* $(X'', \mathcal{O}'')$ *trois espaces topologiques. La composée d'une fonction continue* $f\ : X \to X'$ *et d'une fonction continue* $f'\ : X' \to X''$ *est une fonction continue* $f' \circ f\ : X \to X''$.

Preuve : Grâce au théorème précédent, la démonstration est très facile. En effet, si Ω'' est un ouvert de X'' alors, par continuité de f', l'ensemble ${f'}^{-1}(\Omega'')$ est un ouvert de X', puis, par continuité de f, l'ensemble $f^{-1}\big({f'}^{-1}(\Omega'')\big)$ est un ouvert de Ω. Il ne reste plus qu'à remarquer que $(f' \circ f)^{-1}(\Omega'') = f^{-1}\big({f'}^{-1}(\Omega'')\big)$. ■

1.3 Topologie produit (hors-programme)

Il arrive souvent que l'on dispose d'une famille d'espaces topologiques $(X_i, \mathcal{O}_i)_{i \in I}$ et que l'on cherche une topologie "naturelle" à mettre sur le produit cartésien $\prod_{i \in I} X_i$. En pratique, il est souhaitable que la topologie choisie sur l'espace produit rende continues toutes les **projections canoniques** p_j définies par

$$p_j : \left\{ \begin{array}{lcl} \prod_{i \in I} X_i & \longrightarrow & X_j \\ (x_i)_{i \in I} & \longmapsto & x_j \end{array} \right.$$

Pour simplifier, examinons d'abord le cas de deux espaces topologiques $(X_1, \mathcal{O}_1)$ et $(X_2, \mathcal{O}_2)$. Les deux projections canoniques associées à $X_1 \times X_2$ sont

$$p_1 : \left\{ \begin{array}{lcl} X_1 \times X_2 & \longrightarrow & X_1 \\ (x_1, x_2) & \longmapsto & x_1 \end{array} \right. \quad \text{et} \quad p_2 : \left\{ \begin{array}{lcl} X_1 \times X_2 & \longrightarrow & X_2 \\ (x_1, x_2) & \longmapsto & x_2. \end{array} \right.$$

Remarquons que pour tout $A_1 \subset X_1$ et $A_2 \subset X_2$ on a

$$p_1^{-1}(A_1) = A_1 \times X_2 \quad \text{et} \quad p_2^{-1}(A_2) = X_1 \times A_2.$$

Une condition nécessaire et suffisante de continuité pour p_1 et p_2 est donc que pour tout ouvert O_1 de X_1 et tout ouvert O_2 de X_2, les rectangles $O_1 \times X_2$ et $X_1 \times O_2$ soient des ouverts pour la topologie de $X_1 \times X_2$.

Comme une topologie doit être stable par intersection finie et réunion quelconque, celle que nous souhaitons construire doit en particulier contenir tous les **ouverts élémentaires** $O_1 \times O_2$ avec O_1 ouvert de X_1 et O_2 ouvert de X_2. On en déduit la définition suivante de topologie produit sur $X_1 \times X_2$:

Définition. La **topologie produit** sur $X_1 \times X_2$ est la topologie engendrée par les rectangles $O_1 \times X_2$ et $X_1 \times O_2$. C'est la plus petite topologie contenant tous les ouverts élémentaires.

Nous laissons au lecteur le soin de vérifier que la topologie produit est l'ensemble des réunions arbitraires d'ouverts élémentaires, et que c'est la topologie la moins fine rendant continues les projections canoniques p_1 et p_2.

Exemple. Si l'on prend $X_1 = X_2 = \mathbb{R}$ muni de la topologie usuelle, la topologie produit sur $\mathbb{R}^2 = X_1 \times X_2$ est l'ensemble des réunions arbitraires de rectangles ouverts.

Examinons maintenant le cas d'une famille arbitraire $(X_i, \mathcal{O}_i)_{i \in I}$ d'espaces topologiques. Les considérations ci-dessus nous amènent à la définition suivante de la topologie produit.

Définition. On appelle topologie produit sur $\prod_{i \in I} X_i$ la plus petite topologie contenant tous les rectangles élémentaires

$$\prod_{i \in I} O_i \quad \text{avec} \quad O_i \ \text{ ouvert de } \ X_i.$$

En reprenant les considérations du cas de deux espaces topologiques, on conclut aisément au résultat suivant.

Proposition. *La topologie produit est la topologie la moins fine rendant continues toutes les projections canoniques. C'est l'ensemble des réunions arbitraires d'ouverts élémentaires.*

Terminons cette section par un résultat fort utile.

Proposition. *Si tous les $(X_i, \mathcal{O}_i)$ sont séparés alors $\prod_{i\in I} X_i$ muni de la topologie produit aussi.*

Preuve : Soit a et b deux éléments distincts de $\prod_{i\in I} X_i$. Alors il existe $i_0 \in I$ tel que $a_{i_0} \neq b_{i_0}$. Comme X_{i_0} est séparé, il existe deux ouverts disjoints U_{i_0} et V_{i_0} de X_{i_0} contenant respectivement a_{i_0} et b_{i_0}. Il est alors clair que

$$U_{i_0} \times \prod_{i\neq i_0} X_i \quad \text{et} \quad V_{i_0} \times \prod_{i\neq i_0} X_i$$

sont deux ouverts disjoints de l'espace produit contenant respectivement a et b. ∎

1.4 Suites des espaces topologiques

Il est souvent commode de pouvoir exprimer les notions de topologie abstraite à l'aide de suites. Pour cela, il faut avant tout donner un sens à la notion de *convergence*.

Définition. Soit $(X, \mathcal{O})$ un espace topologique, $\ell \in X$ et $(x_n)_{n\in\mathbb{N}}$ une suite d'éléments de X. On dit que la suite $(x_n)_{n\in\mathbb{N}}$ converge vers ℓ si pour tout voisinage V de ℓ il existe $N \in \mathbb{N}$ tel que $x_n \in V$ pour tout $n \geq N$.

Exercice : Dans le cas d'un espace topologique séparé, montrer l'unicité de la limite d'une suite. Que dire pour un espace topologique non séparé ?

Proposition 1. *Soit F un fermé de $(X, \mathcal{O})$ et $(x_n)_{n\in\mathbb{N}}$ une suite convergente de points de F. Alors la limite de cette suite est encore dans F. On dit que F est* **séquentiellement fermé.**

Preuve : Supposons par l'absurde que la suite $(x_n)_{n\in\mathbb{N}}$ admette une limite x dans le complémentaire Ω de F. Comme Ω est un voisinage (ouvert) de x, les termes de la suite doivent tous se trouver dans Ω à partir d'un certain rang, ce qui est contraire aux hypothèses. ∎

Remarque. Nous verrons plus loin que la réciproque est vraie dans un espace métrique (ou plus généralement dans tout espace topologique admettant des bases de voisinages dénombrables en chaque point).

On laisse au lecteur le soin d'établir le résultat suivant :

Proposition 2. *Soit $(X, \mathcal{O})$ et $(X', \mathcal{O}')$ deux espaces topologiques, $x \in X$ et $f : X \to X'$ continue en x. Alors pour toute suite $(x_n)_{n\in\mathbb{N}}$ d'éléments de X convergeant vers x, on a*

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

La propriété exhibée ci-dessus est appelée **continuité séquentielle**. D'après la proposition 2, la continuité séquentielle est donc une propriété *plus faible* que la continuité. Il existe des espaces topologiques pour lesquels la notion de continuité séquentielle est strictement plus faible que celle de continuité.

Exercice : Donner un exemple d'espaces topologiques $(X, \mathcal{O})$ et $(X', \mathcal{O}')$ et de fonction $f : X \to X'$ continue séquentiellement mais non continue.

Terminons ces généralités par la définition de *valeur d'adhérence* (à comparer à celle de limite).

Définition. Soit $(X, \mathcal{O})$ espace topologique et $(x_n)_{n\in\mathbb{N}}$ suite d'éléments de X. On dit que $a \in X$ est **valeur d'adhérence** de $(x_n)_{n\in\mathbb{N}}$ si pour tout voisinage V de a et pour tout $N \in \mathbb{N}$, il existe un $n \geq N$ tel que $x_n \in V$.

Remarque. S'il existe une **extraction**[1] φ telle que la suite $(x_{\varphi(n)})_{n\in\mathbb{N}}$ (appelée **suite extraite** ou **sous-suite** de $(x_n)_{n\in\mathbb{N}}$) converge vers a alors a est valeur d'adhérence de $(x_n)_{n\in\mathbb{N}}$.

Exercice : Soit $(x_n)_{n\in\mathbb{N}}$ une suite de l'espace topologique $(X, \mathcal{O})$ et A l'ensemble des valeurs d'adhérence de $(x_n)_{n\in\mathbb{N}}$. On note $S_n = \{x_p \,/\, p \geq n\}$. Montrer que

$$A = \bigcap_{n\in\mathbb{N}} \overline{S}_n. \tag{1.1}$$

1.5 Compacité

La notion de compacité est d'une importance fondamentale car c'est elle qui permet par exemple de montrer que des suites admettent des sous-suites convergentes sous des hypothèses favorables portant principalement sur l'espace topologique auquel appartient la suite. Dans cette section, nous définissons la notion de compacité et en donnons les propriétés les plus classiques dans le cadre des espaces topologiques généraux. Nous reverrons la notion de compacité plus loin, dans le cadre des espaces métriques puis des espaces vectoriels normés.

Définition. Un espace topologique X est **compact** s'il est séparé et si tout recouvrement de X par des ouverts admet un sous-recouvrement fini.
Une partie A d'un espace topologique quelconque X est **compacte** si, munie de la topologie induite, c'est un espace topologique compact.

Remarque. On laisse au lecteur le soin de vérifier qu'un ensemble topologique séparé est compact si et seulement si de toute famille de fermés d'intersection vide, on peut extraire une sous-famille finie d'intersection vide.

Donnons une caractérisation des parties compactes bien plus maniable que la définition.

Proposition. *Soit $(X, \mathcal{O})$ un espace topologique séparé et A une partie de X. Alors A est compacte si et seulement si de toute famille d'ouverts de X recouvrant A on peut extraire un sous-recouvrement fini.*

Preuve : Supposons d'abord A compacte et considérons un recouvrement $(\Omega_i)_{i\in I}$ de A par des ouverts de X. On a alors aussi

$$A \subset \bigcup_{i\in I} (\Omega_i \cap A).$$

Comme chaque $\Omega_i \cap A$ est ouvert pour la topologie induite, on peut extraire un sous-recouvrement fini $(\Omega_{i_k} \cap A)_{1\leq k\leq n}$ de A. On a bien sûr

$$A \subset \bigcup_{k=1}^{n} \Omega_{i_k}$$

ce qui est la propriété souhaitée.

1. c'est-à-dire une application strictement croissante de $\mathbb{N}$ dans $\mathbb{N}$

Réciproquement supposons que de toute famille d'ouverts de X recouvrant A on puisse extraire un sous-recouvrement fini. Soit $(\Omega'_i)_{i\in I}$ un recouvrement de A pour la topologie induite. Par définition de la topologie induite, pour chaque $i \in I$, il existe un ouvert Ω_i de X tel que $\Omega'_i = \Omega_i \cap A$. Extrayons de $(\Omega_i)_{i\in I}$ un sous-recouvrement fini $(\Omega_{i_k})_{1\le k\le n}$. Alors $(\Omega'_{i_k})_{1\le k\le n}$ est un sous-recouvrement fini de A. Donc A est bien une partie compacte. ■

On retiendra donc qu'une partie A d'un espace topologique séparé $(X, \mathcal{O})$ est compacte si et seulement si pour toute famille $(\Omega_i)_{i\in I}$ d'ouverts de X telle que $A \subset \bigcup_{i\in I} \Omega_i$ on peut trouver un nombre fini d'indices $i_1, \cdots, i_n$ tels que $A \subset \bigcup_{k=1}^{n} \Omega_{i_k}$.

Exemples.

1. Tous les sous-ensembles finis d'un espace topologique séparé sont compacts.
2. Dans un espace topologique séparé, l'ensemble constitué par une suite convergente et sa limite est compact.
3. L'ensemble $\mathbb{R}$ muni de sa topologie usuelle n'est pas compact, mais $[a, b]$ l'est.

Théorème. *Soit $(X, \mathcal{O})$ un espace topologique séparé et A une partie de X.*

(i) Si A est compacte alors A est fermée.

(ii) Réciproquement si A est fermée et si de plus X est compact alors A est compacte.

Preuve : Pour établir la première propriété, nous allons démontrer que $X \setminus A$ est ouvert. Soit donc $x \in X \setminus A$. Comme $x \notin A$ et la topologie est séparée, pour tout $a \in A$ il existe deux ouverts disjoints V_a et W_a tels que

$$x \in V_a,\ a \in W_a \quad \text{et} \quad V_a \cap W_a = \emptyset.$$

La famille $(W_a)_{a\in A}$ constitue un recouvrement du compact A par des ouverts, dont un peut extraire un sous-recouvrement fini $(W_{a_i})_{1\le i\le n}$. Il est alors clair que $\bigcap_{i=1}^{n} V_{a_i}$ est un voisinage de x ne rencontrant pas A. Le point x étant arbitraire dans $X \setminus A$, on peut conclure que $X \setminus A$ est ouvert.

Pour établir la propriété (ii), considérons une famille $(F_i)_{i\in I}$ de fermés de A d'intersection vide. Comme A est fermé, chaque F_i est aussi un fermé du compact X. D'après la remarque 1.5, on en déduit que la famille de fermés considérée admet une sous-famille finie d'intersection vide. Par conséquent, la partie A est compacte. ■

Proposition 3. *Dans un compact, toute suite a au moins une valeur d'adhérence. De plus, si cette valeur d'adhérence est unique alors la suite converge.*

La preuve de la première partie de la proposition repose sur le lemme suivant que nous laissons au lecteur en guise d'exercice.

Lemme. *Soit $(X, \mathcal{O})$ un espace topologique compact et $(F_n)_{n\in\mathbb{N}}$ une suite décroissante (au sens de l'inclusion) de fermés non vides de X. Alors $\bigcap_{n\in\mathbb{N}} F_n$ n'est pas vide.*

Revenons à la preuve de la proposition. Soit X un compact et $(x_n)_{n\in\mathbb{N}} \in X^{\mathbb{N}}$. Pour $n \in \mathbb{N}$, notons $S_n = \{x_p \,/\, p \ge n\}$ et A l'ensemble des valeurs d'adhérence de la suite. La famille $(\overline{S}_n)_{n\in\mathbb{N}}$ est une suite décroissante de fermés non vides du compact X donc, en vertu du lemme ci-dessus, son intersection (qui n'est autre que A d'après (1.1)) est non vide.

Supposons maintenant que l'ensemble S soit réduit à un point ℓ mais (par l'absurde) que la suite ne converge pas. Il existe donc un voisinage V de ℓ que l'on peut supposer ouvert, et une sous-suite $(x_{\varphi(n)})_{n\in\mathbb{N}}$ tels que $x_{\varphi(n)}$ ne soit jamais dans V. L'ensemble $X \setminus V$ est un fermé du compact X, donc est compact. On en déduit que la sous-suite considérée admet une valeur d'adhérence dans $X \setminus V$, nécessairement distincte de ℓ. Cela contredit l'hypothèse faite.

Proposition. *Soit $(X, \mathcal{O})$ et $(X', \mathcal{O}')$ deux espaces topologiques, le premier étant compact et le deuxième, séparé. Alors l'image de X par toute application continue de X dans X' est compacte.*

Preuve : Soit $f : X \to X'$ continue et $(\Omega_i)_{i \in I}$ un recouvrement de $f(X)$ par des ouverts de X'. On vérifie facilement que $\big(f^{-1}(\Omega_i)\big)_{i \in I}$ est un recouvrement de X. De plus, par continuité de f, chaque ensemble $f^{-1}(\Omega_i)$ est ouvert. On peut donc trouver un nombre fini d'indices $i_1, \cdots, i_N$ tels que $X \subset \bigcup_{k=1}^{N} f^{-1}(\Omega_{i_k})$, d'où l'on tire $f(X) \subset \bigcup_{k=1}^{N} \Omega_{i_k}$. Comme X' est séparé, cela assure la compacité de $f(X)$. $\blacksquare$

Remarque. En d'autres termes, dans des espaces topologiques séparés, l'image d'un compact par une application continue, est compacte.

Chapitre 2

Espaces métriques

Dans ce chapitre, on donne une présentation détaillée des espaces métriques. Une importance particulière est accordée aux notions de compacité et de complétude dans ce cadre.

2.1 Définitions

Définition. Soit X un ensemble. On dit qu'une application $d : X \times X \to \mathbb{R}$ est une **distance** sur X si les trois propriétés suivantes sont vérifiées :

(i) symétrie : pour tout $(x, y) \in X^2$, on a $d(x, y) = d(y, x)$,

(ii) positivité : pour tout $(x, y) \in X^2$, on a $d(x, y) \geq 0$ et $d(x, y) = 0$ si et seulement si $x = y$,

(iii) inégalité triangulaire : pour tout $(x, y, z) \in X^3$, on a $d(x, z) \leq d(x, y) + d(y, z)$.

Le couple (X, d) est alors appelé **espace métrique**.

De la définition d'une distance, on peut déduire la **deuxième inégalité triangulaire** valable pour tout $(x, y, z) \in X^3$:

$$\boxed{\big|d(x, y) - d(y, z)\big| \leq d(x, z).}$$

Exemples.

1. La fonction d définie par $d(x, y) = |x - y|$ pour tout $(x, y) \in \mathbb{R}^2$ est une distance sur $\mathbb{R}$. C'est la **distance usuelle** sur $\mathbb{R}$.
2. La fonction $\overline{d}$ définie par $\overline{d}(x, y) = |\arctan x - \arctan y|$ pour tout $(x, y) \in \mathbb{R}^2$ est aussi une distance sur $\mathbb{R}$.
3. N'importe quel ensemble non vide X peut être muni de la **distance triviale** δ définie pour tout $(x, y) \in X^2$ par
$$\delta(x, y) = \begin{cases} 1 & \text{si} \quad x \neq y, \\ 0 & \text{si} \quad x = y. \end{cases}$$

Définition. Supposons l'ensemble X muni de deux distances d et d'. On dit que d et d' sont **équivalentes** s'il existe $C \in [1, +\infty[$ tel que

$$\forall (x, y) \in X^2,\ C^{-1} d(x, y) \leq d'(x, y) \leq C d(x, y).$$

Définition. On dit qu'une partie A d'un espace métrique (X, d) est **bornée** s'il existe $x \in X$ et un réel positif M tels que

$$\forall a \in A,\ d(a, x) \leq M.$$

Remarque. On établit aisément que A est une partie bornée si et seulement si la borne supérieure $\delta(A)$ de l'ensemble $\{d(x, y) \,/\, (x, y) \in A^2\}$ est finie. On dit que $\delta(A)$ est le **diamètre** de A.

Exemples. 1. Les parties bornées de $\mathbb{R}$ muni de la distance d définie ci-dessus sont les parties bornées "habituelles" (par exemple $[a,b[$, $[-3,1]\cup]1,1000000[$, etc.).

2. Toutes les parties de $\mathbb{R}$ muni de la distance $\overline{d}$ sont bornées. En fait $\mathbb{R}$ lui-même est de diamètre π.
3. Sur X muni de la distance triviale, toutes les parties ayant au moins deux éléments sont bornées et de diamètre 1. (Quant aux singletons, ils sont de diamètre nul comme dans n'importe quel espace métrique.)

Exercice : Vérifier que si X est muni de deux distances équivalentes d et d' alors les parties bornées de (X,d) sont les parties bornées de (X,d'). En déduire que sur $\mathbb{R}$, la distance $\overline{d}(x,y) = |\arctan x - \arctan y|$ n'est pas équivalente à la distance usuelle.

Définition. Soit A et B deux parties non vides de l'espace métrique (X,d) et $x \in X$.
- Le réel positif
$$d(x,A) \overset{\text{déf}}{=} \inf_{a\in A} d(x,a)$$
est appelé **distance de x à l'ensemble A**.
- Le réel positif
$$d(A,B) \overset{\text{déf}}{=} \inf\{d(a,b)\,/\,(a,b)\in A\times B\}$$
est appelé **distance de A à B**.

Proposition. *Soit (X_1,d_1) et (X_2,d_2) deux espaces métriques. Soit δ la fonction définie pour tout $(x_1,x_2)\in X_1\times X_2$ et $(x_1',x_2')\in X_1\times X_2$ par*
$$\delta\big((x_1,x_2),(x_1',x_2')\big) = \max\big(d_1(x_1,x_1'),d_2(x_2,x_2')\big).$$
Alors δ est une distance sur $X_1\times X_2$ appelée **distance produit**.

Remarque. La définition ci-dessus se généralise aisément au cas du produit d'un nombre fini d'espaces métriques. Par ailleurs, on peut aussi définir la distance produit entre deux espaces métriques par
$$\delta\big((x_1,x_2),(x_1',x_2')\big) = d_1(x_1,x_1')+d_2(x_2,x_2') \quad\text{ou}\quad \delta\big((x_1,x_2),(x_1',x_2')\big) = \sqrt{d_1^2(x_1,x_1')+d_2^2(x_2,x_2')}.$$
Ce choix n'a guère d'importance car les différentes distances obtenues sont équivalentes.

Définition. Soit (X,d) un espace métrique et Y une partie de X. La restriction de la fonction d à l'ensemble $Y\times Y$ est une distance sur Y appelée **distance induite**.

2.2 Topologie des espaces métriques

Dans tout espace métrique, la distance permet de définir des boules.

Définition. Soit (X,d) un espace métrique, x_0 un élément de X et r un réel positif.
- L'ensemble $B_X(x_0,r) \overset{\text{déf}}{=} \{x\in X,\ d(x_0,x)<r\}$ (noté aussi $B(x_0,r)$) est appelé **boule ouverte** de centre x_0 et de rayon r.
- L'ensemble $\overline{B}_X(x_0,r) \overset{\text{déf}}{=} \{x\in X,\ d(x_0,x)\le r\}$ (noté aussi $\overline{B}(x_0,r)$) est appelé **boule fermée** de centre x_0 et de rayon r.
- L'ensemble $S_X(x_0,r) \overset{\text{déf}}{=} \{x\in X,\ d(x_0,x)=r\}$ (noté aussi $S(x_0,r)$) est appelé **sphère** de centre x_0 et de rayon r.

Les boules sont les pièces élémentaires permettant de définir une topologie sur les espaces métriques :

Définition. Un sous-ensemble Ω d'un espace métrique (X,d) est dit **ouvert** si pour tout $x\in\Omega$ il existe $r>0$ tel que $B_X(x,r)\subset\Omega$.

Un espace métrique est donc un espace topologique particulier...

Remarque. L'ensemble des boules ouvertes (ou fermées) ayant pour centre un point x donné constitue une base de voisinages de x. En fait, chaque point possède une base *dénombrable* de voisinages, par exemple la suite des boules ouvertes de centre x et de rayon $1/n$.

Remarque. Nous laissons au lecteur le soin d'établir que la topologie associée à un espace métrique n'est autre que la topologie engendrée par les boules ouvertes, c'est-à-dire la plus petite topologie contenant toutes les boules ouvertes, ou encore l'ensemble des réunions quelconques d'intersections finies de boules ouvertes.

Exemples.

1. Pour tout ensemble, la topologie associée à la distance triviale coïncide avec la topologie discrète.
2. La topologie sur $\mathbb{R}$ associée à la distance usuelle n'est autre que la topologie usuelle introduite dans le chapitre précédent.

Remarque. Désormais, si A est une partie non vide d'un espace métrique (X,d), nous avons donc *deux* façons naturelles de munir A d'une topologie :
- on peut d'abord considérer la topologie $\mathcal{O}$ sur X associée à la distance d, puis munir A de la topologie $\mathcal{O}_A$ induite par $\mathcal{O}$,
- on peut au contraire munir A de la distance d_A induite par d, puis considérer la topologie $\mathcal{O}'_A$ sur A associée à d_A.

Heureusement, les deux topologies obtenues sont les mêmes ! (Le vérifier constitue un excellent exercice).

Retenons aussi le fait suivant qui est fort utile :

Proposition 4. *Deux distances équivalentes engendrent la même topologie.*

Preuve : Supposons l'ensemble X muni de deux métriques d et d'. Soit C une constante telle que
$$\forall(x,y)\in X^2,\ C^{-1}d(x,y)\leq d'(x,y)\leq Cd(x,y). \tag{2.1}$$
Notons $\mathcal{O}$ et $\mathcal{O}'$ les topologies associées. Soit Ω un ouvert pour la topologie $\mathcal{O}$ et x un point de Ω. Nous allons montrer que Ω est un voisinage de x pour la topologie $\mathcal{O}'$, ce qui assurera que $\mathcal{O}\subset\mathcal{O}'$. Comme Ω est voisinage de x pour la topologie $\mathcal{O}$, il existe $r>0$ tel que tout point $y\in X$ tel que $d(x,y)<r$ soit dans Ω. En vertu de (2.1), on en déduit que tout point $y\in X$ tel que $d'(x,y)<C^{-1}r$ est dans Ω. En conséquence, Ω est bien un voisinage de x pour la topologie $\mathcal{O}'$. Le point x étant arbitraire dans Ω, on peut alors conclure que Ω est un ouvert de $\mathcal{O}'$, puis que $\mathcal{O}\subset\mathcal{O}'$.

L'inclusion $\mathcal{O}'\subset\mathcal{O}$ se démontre de la même façon. Il suffit d'échanger les rôles de Ω et de Ω'. ■

Exercice : Montrer que la réciproque de la proposition ci-dessus est fausse. On pourra montrer que les topologies associées aux distances d et $\overline{d}$ sur $\mathbb{R}$ définies respectivement par
$$d(x,y)=|x-y|\quad\text{et}\quad\overline{d}(x,y)=\bigl|\arctan x-\arctan y\bigr|$$
sont les mêmes bien que ces deux distances ne soient pas équivalentes.

Nous allons voir que la présence d'une distance confère à (X,d) de nombreuses propriétés que les espaces topologiques généraux n'ont pas.

La première propriété est celle de séparation :

Proposition. *Tout espace métrique est un espace topologique séparé.*

Preuve : Soit x et y deux points distincts de l'espace métrique (X, d). Il est clair que les boules ouvertes de centre respectifs x et y, et de rayon $d(x, y)/2$ sont disjointes. Par ailleurs, ce sont des voisinages de x et de y respectivement. ■

Dans un espace métrique, la propriété de fermeture peut être caractérisée à l'aide de suites (dans un espace topologique général, seule une implication est vraie, voir proposition 1 page 9) :

Proposition. *Dans un espace métrique un ensemble est fermé si et seulement si il est séquentiellement fermé.*

Preuve : Seule la réciproque est à justifier. Raisonnons par contraposition. Soit donc A une partie non fermée de X, et $a \in \overline{A} \setminus A$. Pour tout $n \in \mathbb{N}$, on choisit un point x_n de l'ensemble (non vide) $A \cap B(a, 2^{-n})$. La suite obtenue est une suite de points de A qui converge vers a. Mais a n'est pas dans A donc A n'est pas séquentiellement fermé. ■

Remarque : Dans un espace métrique, l'adhérence $\overline{A}$ d'une partie A est donc l'ensemble des limites des suites convergentes d'éléments de A.

Exercice : Soit A une partie non vide d'un espace métrique (X, d). Montrer que $x \in \overline{A}$ si et seulement si $d(x, \overline{A}) = 0$.

Dans les espaces métriques, on dispose d'une caractérisation des valeurs d'adhérence fort commode :

Proposition. *Soit $(x_n)_{n\in\mathbb{N}}$ une suite de points de l'espace métrique (X, d). Alors a est valeur d'adhérence de $(x_n)_{n\in\mathbb{N}}$ si et seulement si il existe une suite extraite $(x_{\varphi(n)})_{n\in\mathbb{N}}$ qui converge vers a.*

Preuve : Seule l'implication directe est à justifier (l'implication réciproque découle de la définition de valeur d'adhérence et reste vraie dans n'importe quel espace topologique). Supposons donc a valeur d'adhérence de la suite. En se limitant aux voisinages de a constitués par la famille de boules ouvertes $B(a, 2^{-k})$ dans la définition de la valeur d'adhérence, on voit que

$$\forall k \in \mathbb{N}, \forall N \in \mathbb{N}, \exists n \geq N, d(a, x_n) < 2^{-k}.$$

On construit alors une suite extraite qui converge vers a par récurrence : on définit $\varphi(0)$ comme étant le plus petit indice tel que $d(a, x_{\varphi(0)}) < 1$, puis $\varphi(1)$ comme le plus petit indice strictement supérieur à $\varphi(0)$ vérifiant $d(a, x_{\varphi(1)}) < 1/2$, etc. ■

Exemple important : Soit une suite $(x_n)_{n\in\mathbb{N}}$ d'éléments de $\mathbb{R}$ muni de la distance usuelle. Rappelons que, par définition,

$$\liminf x_n = \lim_{n\to+\infty} \inf_{p\geq n} x_p \quad \text{et} \quad \limsup x_n = \lim_{n\to+\infty} \sup_{p\geq n} x_p.$$

Si la suite $(x_n)_{n\in\mathbb{N}}$ est minorée, l'ensemble de ses valeurs d'adhérence (lorsqu'il n'est pas vide) l'est aussi, et la borne inférieure de toutes les valeurs d'adhérence est finie et coïncide avec $\liminf x_n$. On vérifie alors qu'il existe une suite extraite de $(x_n)_{n\in\mathbb{N}}$ qui converge vers $\liminf x_n$.

De même, si la suite est majorée et l'ensemble des valeurs d'adhérence n'est pas vide, la borne supérieure des valeurs d'adhérence. est finie et coïncide avec $\limsup x_n$.

Bien sûr, il y a convergence de la suite si et seulement si $\liminf x_n$ et $\limsup x_n$ sont finies et égales et, dans ce cas, la limite de la suite est égale à la valeur commune de $\liminf x_n$ et $\limsup x_n$.

Rappelons finalement au lecteur que $\liminf x_n$ et $\limsup x_n$ gardent un sens dans $\mathbb{R} \cup \{-\infty, +\infty\}$ pour n'importe quelle suite réelle. Par exemple si $(x_n)_{n\in\mathbb{N}}$ n'est pas majorée, on a $\limsup x_n = +\infty$.

2.3 Continuité dans les espaces métriques

Dans un espace métrique, l'ensemble des boules ouvertes centrées en un point constitue une base de voisinages de ce point.

En conséquence, une fonction f définie sur une partie A d'un espace métrique (X_1, d_1) et à valeurs dans un espace métrique (X_2, d_2) est continue en $x \in A$ si et seulement si on a :

$$\forall \varepsilon > 0, \exists \delta > 0, \forall y \in X_1 \cap A,\ d_1(x, y) < \delta \Longrightarrow d_2(f(x), f(y)) < \varepsilon.$$

Remarque : Si au lieu de prendre des boules ouvertes, on prend les boules fermées centrées en x dans la définition de la continuité, on obtient des inégalités larges au lieu d'inégalités strictes. Cela peut toujours servir...

La caractérisation ci-dessus de la continuité va nous permettre d'établir la propriété importante suivante (qui, elle non plus, n'est pas vraie dans les espaces topologiques "généraux").

Proposition. *Une application entre deux espaces métriques est continue en un point si et seulement si elle est séquentiellement continue en ce point.*

Preuve : Seule la réciproque est à montrer, l'implication directe étant vraie dans n'importe quel espace topologique (voir proposition 2 page 9). Raisonnons par contraposition. Soit donc une fonction f définie sur une partie A d'un espace métrique (X_1, d_1) et à valeurs dans un espace métrique (X_2, d_2). Soit $a \in A$. Supposons que f ne soit pas continue en a. Alors il existe un $\varepsilon > 0$ tel que pour tout voisinage V de a il existe un $x \in V \cap A$ vérifiant $d_2(f(x), f(a)) \geq \varepsilon$. En prenant pour V les boules ouvertes $B(a, 2^{-n})$, on construit une suite $(x_n)_{n \in \mathbb{N}}$ telle que

$$\forall n \in \mathbb{N},\ d_2(f(x_n), f(a)) \geq \varepsilon \quad \text{et} \quad d_1(x_n, a) < 2^{-n}.$$

En conséquence, $(f(x_n))_{n \in \mathbb{N}}$ ne converge pas vers $f(a)$ bien que $(x_n)_{n \in \mathbb{N}}$ tende vers a. ■

Nous avons vu que pour une application $f : (X_1, d_1) \to (X_2, d_2)$ entre deux espaces métriques, la continuité en tout point pouvait être caractérisée de la façon suivante :

$$\forall \varepsilon > 0, \boxed{\forall x \in X_1,\ \exists \delta > 0,} \forall y \in X_1,\ d_1(x, y) < \delta \Longrightarrow d_2(f(x), f(y)) < \varepsilon.$$

On peut définir une notion plus forte de continuité appelée continuité uniforme :

Définition. On dit qu'une application $f : X_1 \to X_2$ entre deux espaces métriques (X_1, d_1) et (X_2, d_2) est **uniformément continue** si l'on a :

$$\forall \varepsilon > 0, \boxed{\exists \delta > 0,\ \forall x \in X_1,} \forall y \in X_1,\ d_1(x, y) < \delta \Longrightarrow d_2(f(x), f(y)) < \varepsilon.$$

Bien sûr continuité uniforme entraîne continuité. La réciproque est fausse en général mais vraie dans le cas d'un espace métrique *compact* (voir plus loin).

Définition. Soit k un réel positif. Une application f entre deux espaces métriques (X_1, d_1) et (X_2, d_2) est dite lipschitzienne de rapport k si

$$\forall (x, y) \in X_1 \times X_1,\ d_2(f(x), f(y)) \leq k d_1(x, y).$$

Exemple. Dans $\mathbb{R}^+$ muni de la distance associée à la valeur absolue, la fonction $x \mapsto 2x + 3$ est lipschitzienne de rapport 2, la fonction $x \mapsto \sqrt{x}$ est uniformément continue sans être lipschitzienne, et la fonction $x \mapsto x^2$ est continue sans être ni uniformément continue ni lipschitzienne.

Exercice : Montrer que lipschitzienne entraîne uniformément continue, et qu'uniformément continue entraîne continue.

Voici un exemple particulier d'application lipschitzienne.

Définition. Soit (X_1, d_1) et (X_2, d_2) deux espaces métriques. On dit que $f : X_1 \to X_2$ est une **isométrie** si f est bijective et conserve la distance :

$$\forall (x, y) \in X_1^2,\ d_2(f(x), f(y)) = d_1(x, y).$$

Remarque : Les propriétés de continuité, d'uniforme continuité et lipschitziennes sont invariantes par changement de distance en une distance équivalente.

2.4 Compacité dans les espaces métriques

En général, l'uniforme continuité est une notion strictement plus forte que la continuité. Le théorème suivant établit que dans un espace métrique compact, les deux notions se rejoignent.

Théorème (de Heine)**.** *Soit (X, d_X) et (Y, d_Y) deux espaces métriques. Supposons (X, d_X) compact. Alors toute fonction continue de X vers Y est uniformément continue sur X.*

Preuve : Fixons $\varepsilon > 0$. Par définition de la continuité, pour chaque $x \in X$, il existe $\eta_x > 0$ tel que

$$d_X(x, y) < \eta_x \Longrightarrow d_Y(f(x), f(y)) < \frac{\varepsilon}{2}.$$

Par compacité de X, on peut trouver un nombre fini de points $x_1, \cdots, x_N$ tels que

$$X \subset \bigcup_{k=1}^{N} B_X\Big(x_k, \frac{\eta_{x_k}}{2}\Big).$$

Posons $\eta = \min_{1 \le k \le N} \eta_{x_k}$. Pour tout couple (x, y) tel que $d_X(x, y) < \frac{\eta}{2}$, on peut trouver un indice k tel que x et y soient dans $B_X(x_k, \eta_{x_k})$. On a alors

$$d_Y(f(x), f(y)) \le d_Y(f(x), f(x_k)) + d_Y(f(x_k), f(y)) < \varepsilon,$$

d'où l'uniforme continuité. ■

Dans les espaces métriques, la compacité peut se caractériser à l'aide de suites. Ce fait fondamental est l'objet du théorème ci-dessous.

Théorème (de Bolzano-Weierstrass)**.** *Un espace métrique (X, d) est compact si et seulement si toute suite d'éléments de X admet une sous-suite convergente.*

Preuve : D'après la proposition 3 page 11, toute suite d'un compact admet une valeur d'adhérence. Comme par ailleurs, dans un espace métrique, il y a équivalence entre avoir des sous-suites convergentes et posséder des valeurs d'adhérence, l'implication directe est établie.

Concentrons-nous maintenant sur la démonstration de l'implication réciproque. Soit donc X un espace métrique pour lequel toute suite a une valeur d'adhérence. Soit $(\Omega_i)_{i \in I}$ un recouvrement de X par des ouverts. Il s'agit d'extraire de cette famille un sous-recouvrement fini de X.

1. On se ramène à un recouvrement par des boules ouvertes de même rayon.
   Plus précisément, on montre qu'il existe un $\varepsilon > 0$ tel que toute boule ouverte de rayon ε soit contenue dans un des Ω_i.
   On procède par l'absurde en supposant qu'un tel ε n'existe pas. Alors pour tout $n \in \mathbb{N}$, il existe $x_n \in X$ tel que $B(x_n, 2^{-n})$ soit contenue dans aucun Ω_i. Cette suite admet une valeur d'adhérence x qui, fatalement, appartient à un des Ω_i. Notons i_x l'indice correspondant et $r > 0$ tel que $B(x, r) \subset \Omega_{i_x}$. Comme x est valeur d'adhérence de la suite, il existe $n \in \mathbb{N}$ tel que $d(x, x_n) < r/2$ et $2^{-n} \le r/2$. On a donc $B(x_n, 2^{-n}) \subset \Omega_{i_x}$, ce qui contredit la définition de x_n.

2. On montre que l'on peut recouvrir X par un nombre fini de boules de même rayon ε. Pour cela, on part d'un point x_0 quelconque. Si $X \subset B(x_0, \varepsilon)$, il n'y a rien a faire, la construction est terminée. Sinon, on choisit un x_1 tel que $x_1 \notin B(x_0, \varepsilon)$. Si $B(x_0, \varepsilon) \cup B(x_1, \varepsilon)$ recouvre X, la construction est terminée, sinon on choisit un point x_2 tel que $x_2 \notin B(x_0, \varepsilon) \cup B(x_1, \varepsilon)$.
   Par récurrence, on construit ainsi une famille $(x_0, \cdots, x_n)$ telle que $x_n \notin B(x_0, \varepsilon) \cup \cdots \cup B(x_{n-1}, \varepsilon)$. Le procédé de construction s'arrête nécessairement au bout d'un nombre fini d'étapes sinon on obtiendrait une suite $(x_n)_{n \in \mathbb{N}}$ telle que $d(x_n, x_m) \geq \varepsilon$ pour $n \neq m$. Une telle suite ne saurait avoir de sous-suite convergente.
3. Conclusion.
   D'après la première étape, il existe $\varepsilon > 0$ tel que pour chaque $x \in X$ il existe un indice $i_x \in I$ tel que $B(x, \varepsilon) \subset \Omega_{i_x}$. D'après la deuxième étape, il existe une famille finie $(x_1, \cdots, x_N)$ d'éléments de X telle que $X \subset \bigcup_{k=1}^N B(x_k, \varepsilon)$. On a a fortiori $X \subset \bigcup_{k=1}^N \Omega_{i_{x_k}}$.

■

Donnons plusieurs conséquences importantes du théorème de Bolzano-Weierstrass.

Définition. On dit qu'une partie A d'un espace topologique séparé X est relativement compacte si $\overline{A}$ est compacte.

Corollaire 1. *Soit (X, d) un espace métrique. Une partie A de X est relativement compacte si et seulement si de toute suite d'éléments de A, on peut extraire une sous-suite qui converge dans X.*

Preuve : Supposons A relativement compacte. Alors toute suite d'éléments de A est aussi une suite d'éléments du compact $\overline{A}$. En conséquence, elle admet au moins une valeur d'adhérence dans $\overline{A}$, et donc une sous-suite convergente dans X.

Réciproquement, supposons que de toute suite de A on puisse extraire une sous-suite convergente et considérons une suite $(x_n)_{n \in \mathbb{N}}$ d'éléments de $\overline{A}$. Il existe alors une suite $(y_n)_{n \in \mathbb{N}}$ d'éléments de A telle que $d(x_n, y_n) \leq 2^{-n}$ pour tout $n \in \mathbb{N}$. Cette suite admet une sous-suite convergente $(y_{\varphi(n)})_{n \in \mathbb{N}}$ et il est clair que $(x_{\varphi(n)})_{n \in \mathbb{N}}$ converge vers la même limite. Cette limite se trouve nécessairement dans le fermé $\overline{A}$. En conséquence A est bien relativement compact. ■

Corollaire 2. *Dans un espace métrique, toute partie relativement compacte est bornée et toute partie compacte est fermée bornée.*

Preuve : Pour démontrer la première partie du corollaire, raisonnons par contraposition. Soit donc A une partie non bornée de l'espace métrique (X, d). Alors on peut construire (par récurrence) une suite $(x_n)_{n \in \mathbb{N}}$ telle que $d(x_n, x_m) \geq 1$ pour tout $(n, m) \in \mathbb{N}^2$ tel que $n \neq m$. Une telle suite ne peut avoir de sous-suite convergente.

La deuxième partie du corollaire devient alors évidente. En effet, tout compact est relativement compact (donc borné), et fermé. ■

Corollaire 3. *Le produit cartésien d'un nombre fini d'espaces métriques compacts est un espace métrique compact.*

Preuve : Pour simplifier, limitons nous au produit de deux espaces métriques compacts (X_1, d_1) et (X_2, d_2). Soit donc $(x_n, y_n)_{n \in \mathbb{N}}$ une suite d'éléments de $X_1 \times X_2$ (muni de la métrique produit). Par compacité de X_1, la suite $(x_n)_{n \in \mathbb{N}}$ admet une sous-suite convergente $(x_{\varphi(n)})_{n \in \mathbb{N}}$. Par compacité de X_2, la suite $(y_{\varphi(n)})_{n \in \mathbb{N}}$ admet une sous-suite convergente $(y_{\varphi(\psi(n))})_{n \in \mathbb{N}}$. Il est clair que la suite $(x_{\varphi \circ \psi(n)}, y_{\varphi \circ \psi(n)})_{n \in \mathbb{N}}$ est convergente. La réciproque du théorème de Bolzano-Weierstrass permet alors de conclure. ■

Remarque. En utilisant le *procédé diagonal de Cantor*, on peut établir que le produit d'une famille dénombrable d'espaces métriques compacts est un espace métrique compact. Cela reste vrai pour une famille *quelconque* si l'on fait appel à l'axiome du choix.

2.5 Complétude

Rappelons tout d'abord la définition de suite de Cauchy.

Définition. Soit (X,d) un espace métrique. On dit qu'une suite $(x_n)_{n\in\mathbb{N}}$ de $X^{\mathbb{N}}$ est une **suite de Cauchy** si elle vérifie :

$$\forall \varepsilon > 0, \exists N \in \mathbb{N},\ (p \geq N \ \text{ et } \ n \geq N) \Longrightarrow d(x_n, x_p) < \varepsilon.$$

À titre d'exercice, nous laissons le soin au lecteur d'établir que toute suite convergente est de Cauchy.

La réciproque est fausse, en général. En effet, considérons l'ensemble $\mathbb{Q}$ muni de la distance de la valeur absolue. Alors la suite définie par

$$x_0 = 1 \quad \text{et} \quad x_{n+1} = \frac{1}{1+x_n} \quad \text{pour} \quad n \geq 1$$

est une suite de Cauchy de $\mathbb{Q}$ (car elle converge dans $\mathbb{R}$ pour la même distance). Cependant, sa limite $\frac{1}{2}(\sqrt{5}-1)$ est irrationnelle et donc cette suite ne converge pas dans $\mathbb{Q}$. Cet exemple motive la définition suivante :

Définition. L'espace métrique (X,d) est dit **complet** si toute suite de Cauchy d'éléments de X converge dans X.
Une partie A d'un espace métrique est dite complète si, munie de la distance induite, c'est un espace métrique complet.

Exemple. L'ensemble des réels muni de la distance usuelle est, par construction, complet. En revanche, $\mathbb{Q}$ muni de la même distance ne l'est pas.

Exercice : Vérifier que dans tout espace métrique une suite de Cauchy ayant une valeur d'adhérence est convergente.

Proposition. *Soit (X,d) et (X',d') deux espaces métriques. Supposons qu'il existe une application bijective $f : X \to X'$ telle que f et f^{-1} soient uniformément continues. Alors (X,d) est complet si et seulement si (X',d') est complet.*

Preuve : Supposons (X,d) complet. Soit $(x'_n)_{n\in\mathbb{N}}$ une suite de Cauchy de (X',d'). Comme f^{-1} est uniformément continue, on vérifie (en revenant à la définition) que $(f^{-1}(x'_n))_{n\in\mathbb{N}}$ est une suite de Cauchy de (X,d). Mais comme l'espace (X,d) est complet, cette suite converge vers un élément de x de X. Enfin, comme f est continue, on conclut alors que $(f(x_n))_{n\in\mathbb{N}}$ converge vers $f(x)$.
La réciproque se traite en échangeant les rôles de X (resp. f) et de X' (resp. f^{-1}). ∎

Attention : Contrairement aux propriétés topologiques vues jusqu'à présent, la complétude n'est pas préservée par homéomorphisme [1].

Proposition. *Toute partie complète est fermée.*

1. Rappelons qu'un **homéomorphisme** est une application bijective continue et d'application réciproque continue entre deux espaces topologiques.

Preuve : Soit A une partie complète de l'espace métrique (X,d). Si $(x_n)_{n\in\mathbb{N}}$ est une suite convergente d'éléments de A alors c'est une suite de Cauchy. Comme A est complète, la limite de cette suite est dans A. En conséquence A est fermée. ∎

Proposition. *Toute partie fermée d'un espace métrique complet est complète.*

Preuve : Soit A une partie fermée d'un espace métrique complet (X,d) et $(x_n)_{n\in\mathbb{N}}$ une suite de Cauchy de A. Alors $(x_n)_{n\in\mathbb{N}}$ est aussi une suite de Cauchy de X donc converge dans X. Comme A est fermé, la limite de cette suite est dans A. ∎

Nous laissons au lecteur le soin de montrer le résultat suivant :

Proposition. *Le produit cartésien d'un nombre fini d'espaces métriques complets est un espace métrique complet.*

La proposition suivante permet de comparer les notions de complétude et de compacité.

Proposition. *Soit (X,d) un espace métrique. Les deux propriétés suivantes sont équivalentes :*

(i) *(X,d) est compact,*

(ii) *(X,d) est complet et, pour tout $\varepsilon>0$, X peut être recouvert par un nombre fini de boules de rayon ε.*

Preuve : $(i)\Rightarrow(ii)$: soit $(x_n)_{n\in\mathbb{N}}$ une suite de Cauchy de X. Comme (X,d) est un espace métrique compact, la suite $(x_n)_{n\in\mathbb{N}}$ a une valeur d'adhérence. Mais comme elle est de Cauchy, elle converge. En conséquence (X,d) est complet. Par ailleurs, la compacité assure aussi que X peut être recouvert par un nombre fini de boules ouvertes de n'importe quel rayon donné.

$(ii)\Rightarrow(i)$: Supposons que (X,d) soit complet et que, pour tout $\varepsilon>0$, X puisse être recouvert par un nombre fini de boules ouvertes de rayon ε.

Soit $(x_n)_{n\in\mathbb{N}}$ une suite de $X^{\mathbb{N}}$. Il s'agit de montrer que cette suite admet une sous-suite convergente. En fait, on va plutôt établir qu'il existe une sous-suite $(x_{\psi(n)})_{n\in\mathbb{N}}$ de $(x_n)_{n\in\mathbb{N}}$, et une suite $(y_n)_{n\in\mathbb{N}}$ de $X^{\mathbb{N}}$ telles que pour tout $n\in\mathbb{N}$ et $p\in\mathbb{N}$ on ait $x_{\psi(n+p)}\in B(y_n,2^{-n})$. Cela assurera que $(x_{\psi(n)})_{n\in\mathbb{N}}$ est de Cauchy et donc converge puisque (X,d) est complet.

Afin de construire les $(y_n)_{n\in\mathbb{N}}$ et les extractions successives, on commence par remarquer que comme X est recouvert par un nombre *fini* de boules de rayon 1, il existe $y_0\in X$ tel que $B(y_0,1)$ contienne une infinité de termes de $(x_n)_{n\in\mathbb{N}}$. Cela permet de définir une suite extraite $(x_{\varphi_0(n)})_{n\in\mathbb{N}}\in B(y_0,1)^{\mathbb{N}}$. De même, X peut-être recouvert par un nombre fini de boules de rayon $1/2$ donc il existe $y_1\in X$ tel que la boule $B(y_1,1/2)$ contienne une infinité de termes de $(x_{\varphi_0(n)})_{n\in\mathbb{N}}$. Cela assure l'existence d'une deuxième extraction φ_1 telle que $B(y_1,1/2)$ contienne tous les termes de la suite $(x_{\varphi_0(\varphi_1(n))})_{n\in\mathbb{N}}$. Par une récurrence élémentaire, on obtient ainsi une suite $(y_n)_{n\in\mathbb{N}}$ de points de X et des extractions $\varphi_0,\cdots,\varphi_k,\cdots$ telles que $B(y_k,2^{-k})$ contienne tous les termes de $(x_{\varphi_0\circ\cdots\circ\varphi_k(n)})_{n\in\mathbb{N}}$. On conclut à l'aide du *procédé diagonal* qui consiste à poser $\psi(n)=\varphi_0\circ\cdots\circ\varphi_n(n)$. La suite extraite ainsi construite vérifie la propriété souhaitée. ∎

Corollaire 1. *Soit (X,d) un espace métrique complet et A une partie de X. Il y a équivalence entre les deux énoncés suivants :*

(i) *la partie A est relativement compacte,*

(ii) *pour tout $\varepsilon>0$, la partie A peut être recouverte par un nombre fini de boules centrées en des points de A.*

Preuve : $(i) \Rightarrow (ii)$: Il suffit de remarquer que

$$\overline{A} \subset \bigcup_{x \in A} B(x, \varepsilon).$$

Comme $\overline{A}$ est compact, on peut extraire du recouvrement ci-dessus un sous-recouvrement fini de $\overline{A}$ donc (a fortiori) de A.

$(ii) \Rightarrow (i)$: on remarque que la complétude de (X, d) assure que $(\overline{A}, d)$ est complet. Il est clair que si pour tout $\varepsilon > 0$, la partie A peut être recouverte par un nombre fini de boules centrées en des points de A, il en est de même pour $\overline{A}$. En effet

$$A \subset \bigcup_{i=1}^{n} B\Big(x_i, \frac{\varepsilon}{2}\Big) \Longrightarrow \overline{A} \subset \bigcup_{i=1}^{n} B(x_i, \varepsilon).$$

La proposition précédente permet de conclure que $(\overline{A}, d)$ est compact. ■

Corollaire 2. *Les compacts de $\mathbb{R}$ sont les ensembles fermés bornés.*

Preuve : On sait déjà que dans un espace métrique, les compacts sont fermés bornés. Réciproquement, sachant que $\mathbb{R}$ est complet et que toute partie bornée de $\mathbb{R}$ peut être recouverte par un nombre fini de boules de rayon ε arbitrairement fixé, le corollaire précédent assure que toute partie bornée de $\mathbb{R}$ est relativement compacte. ■

Attention : Il existe des espaces métriques complets qui ne sont pas compacts : $\mathbb{R}$ par exemple. Dans un espace complet, on dispose du résultat de prolongement suivant :

Théorème. *Soit (X, d_X) et (Y, d_Y) deux espaces métriques. Supposons (Y, d_Y) complet. Soit A une partie dense de X et $f : A \to Y$ uniformément continue.*

Alors il existe une unique application $\widetilde{f} : X \to Y$ continue sur X qui prolonge[2] f sur X. De plus ce prolongement est uniformément continu sur X.

Preuve : Démontrons d'abord l'unicité. Soit f_1 et f_2 deux prolongements continus de f, et $x \in X$ arbitraire. Fixons $(x_n)_{n\in\mathbb{N}} \in A^{\mathbb{N}}$ convergeant vers x. Comme chaque x_n est dans A, on a $f_1(x_n) = f_2(x_n)$. En passant à la limite, on conclut que $f_1(x) = f_2(x)$.

Passons maintenant à l'existence du prolongement. Soit $x \in X$ arbitraire et $(x_n)_{n\in\mathbb{N}} \in A^{\mathbb{N}}$ convergeant vers x. La suite $(x_n)_{n\in\mathbb{N}}$ est donc de Cauchy dans (X, d_X). En utilisant l'uniforme continuité de f, on vérifie aisément que $(f(x_n))_{n\in\mathbb{N}}$ est de Cauchy dans l'espace métrique complet (Y, d_Y) donc converge vers une certaine limite ℓ. On laisse au lecteur le soin de vérifier que la valeur de ℓ ne dépend pas du choix de la suite $(x_n)_{n\in\mathbb{N}}$ convergeant vers x. On pose alors $\widetilde{f}(x) = \ell$.

En passant à la limite dans la définition de l'uniforme continuité de f, on conclut à l'uniforme continuité de $\widetilde{f}$ (en fait, à ε fixé, tout η convenant à f convient aussi à $\widetilde{f}$). ■

Définition. Soit (X, d) un espace métrique. On dit d'une application f de X dans X qu'elle est **contractante** s'il existe $k \in [0, 1[$ tel que

$$\forall (x, y) \in X^2,\ d(f(x), f(y)) \leq k d(x, y).$$

Théorème (du point fixe)**.** *Soit (X, d) un espace métrique complet et f une application contractante de X dans X. Alors f admet un unique point fixe*[a].

a. On appelle **point fixe** de f tout élément x de X tel que $f(x) = x$.

2. c'est-à-dire $\widetilde{f}(x) = f(x)$ pour tout $x \in A$.

Preuve : La preuve est au moins aussi intéressante que l'énoncé du théorème. Fixons un réel $k \in [0,1[$ tel que

$$d(f(x), f(y)) \leq kd(x,y) \quad \text{pour tout} \quad (x,y) \in X^2. \tag{2.2}$$

L'unicité du point fixe est évidente. En effet si $f(x) = x$ et $f(y) = y$ alors (2.2) assure que $d(x,y) \leq kd(x,y)$, d'où $(1-k)d(x,y) \leq 0$ puis $d(x,y) = 0$.

Pour montrer l'existence du point fixe, partons de $x_0 \in X$ quelconque et définissons la suite $(x_n)_{n\in\mathbb{N}}$ par la relation de récurrence $x_{n+1} = f(x_n)$. En exploitant (2.2), on voit que

$$\forall n \in \mathbb{N},\ d(x_{n+1}, x_n) \leq k^n d(x_1, x_0)$$

puis que

$$\forall (n,p) \in \mathbb{N}^2,\ d(x_{n+p}, x_n) \leq \frac{k^n}{1-k} d(x_1, x_0)\cdot$$

En conséquence, la suite $(x_n)_{n\in\mathbb{N}}$ est de Cauchy. Comme (X,d) est complet, elle converge vers un point x de X qui, clairement, vérifie $f(x) = x$. ■

Chapitre 3

Espaces vectoriels normés

3.1 Définitions

Définition. Soit E un espace vectoriel sur $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$. Une fonction $\| \cdot \| : E \to \mathbb{R}^+$ est appelée **norme** sur E si elle vérifie les conditions suivantes :
- (*i*) $\|x\| = 0 \iff x = 0$,
- (*ii*) $\forall \lambda \in \mathbb{K}, \forall x \in E,\ \|\lambda x\| = |\lambda| \|x\|$,
- (*iii*) $\forall (x, y) \in E^2,\ \|x + y\| \leq \|x\| + \|y\|$.

Le couple $(E, \| \cdot \|)$ est alors appelé **espace vectoriel normé**.

Remarque. Sur $E \times E$, on définit la fonction d par :

$$\forall (x, y) \in E^2,\ d(x, y) = \|x - y\|.$$

On vérifie aisément que d est une distance sur E. On l'appelle **distance associée à la norme**. En plus des trois propriétés définissant les distances, la fonction d est invariante par translation :

$$\forall (x, y, z) \in X^3,\ d(x + z, y + z) = d(x, y)$$

et homogène de degré un :

$$\forall (x, y, \lambda) \in X \times X \times \mathbb{K},\ d(\lambda x, \lambda y) = |\lambda| d(x, y).$$

La **deuxième inégalité triangulaire** pour d peut s'écrire en termes de norme :

$$\forall (x, y) \in E^2,\ \big| \|x\| - \|y\| \big| \leq \|x - y\|.$$

Dans tout ce qui suit, on munit l'e.v.n. E de la topologie associée à la distance d. Notons que les boules et les sphères peuvent se définir en termes de norme :
- la **boule ouverte** $B_E(x_0, r)$ de centre x_0 et de rayon r est égale à $\{x \in E \mid \|x - x_0\| < r\}$,
- la **boule fermée** $\overline{B}_E(x_0, r)$ de centre x_0 et de rayon r est égale à $\{x \in E \mid \|x - x_0\| \leq r\}$,
- la **sphère** $S_E(x_0, r)$ de centre x_0 et de rayon r est égale à $\{x \in E \mid \|x - x_0\| = r\}$.

Exercice : Soit E un e.v.n, $r > 0$ et $x_0 \in E$. Montrer que l'adhérence de $B_E(x_0, r)$ est égale à $\overline{B}_E(x_0, r)$ et que l'intérieur de $\overline{B}_E(x_0, r)$ vaut $B_E(x_0, r)$. En déduire que la frontière de $B_E(x_0, r)$ est la sphère $S_E(x_0, r)$. Que reste-t-il de ces résultats dans un espace métrique "quelconque" ?

Exemple 1. Soit $E = \mathbb{R}^n$. Il est facile d'établir que la fonction $x \mapsto \|x\|_\infty = \sup_{i=1}^n |x_i|$ est une norme sur $\mathbb{R}^n$.

Fixons maintenant un réel $p \geq 1$. Pour $x \in \mathbb{R}^n$, on pose

$$\|x\|_p = \left(\sum_{i=1}^n |x_i|^p\right)^{\frac{1}{p}}.$$

Alors $\|\cdot\|_p$ est une norme sur $\mathbb{R}^n$. L'inégalité triangulaire

$$\|x+y\|_p \leq \|x\|_p + \|y\|_p$$

qui n'est pas triviale si $p > 1$ est appelée **inégalité de Minkowski**.

L'inégalité de Minkowski se montre à l'aide de l'**inégalité de Hölder** valable sur $\mathbb{R}^n$ ou $\mathbb{C}^n$:

$$\left|\sum_{i=1}^n x_i y_i\right| \leq \|x\|_p \|y\|_q$$

où $1 \leq p \leq +\infty$ et q est **l'exposant conjugué** de p défini par la relation $1/p + 1/q = 1$ (avec la convention $1/\infty = 0$).

Le cas $p = q = 2$ n'est autre que l'**inégalité de Cauchy-Schwarz**. Le cas général repose sur l'**inégalité de Young**[1] :

$$\forall (a,b) \in \mathbb{R}^+ \times \mathbb{R}^+,\ ab \leq \frac{a^p}{p} + \frac{b^q}{q}\cdot$$

En effet, si l'on écarte les cas triviaux où x ou y est nul, on peut diviser par $\|x\|_p\|y\|_q$ et l'on a, d'après l'inégalité de Young :

$$\forall i \in \{1, \cdots, n\},\ \frac{|x_i|}{\|x\|_p}\frac{|y_i|}{\|y\|_p} \leq \frac{1}{p}\frac{|x_i|^p}{\|x\|_p^p} + \frac{1}{q}\frac{|y_i|^q}{\|y\|_q^q}.$$

Une simple sommation sur i donne l'inégalité de Hölder.

Revenons à la preuve de l'inégalité de Minkowski. En appliquant deux fois l'inégalité de Hölder, on obtient :

$$\begin{aligned}
\|x+y\|_p^p = \sum_{i=1}^n |x_i+y_i||x_i+y_i|^{p-1} &\leq \sum_{i=1}^n |x_i||x_i+y_i|^{p-1} + \sum_{i=1}^n |y_i||x_i+y_i|^{p-1},\\
&\leq \|x\|_p\left(\sum_{i=1}^n |x_i+y_i|^p\right)^{\frac{p-1}{p}} + \|y\|_p\left(\sum_{i=1}^n |x_i+y_i|^p\right)^{\frac{p-1}{p}},\\
&= (\|x\|_p + \|y\|_p)\|x+y\|_p^{p-1}.
\end{aligned}$$

Exemple 2. Pour tout $p \in [1, +\infty]$, on peut munir l'ensemble $\ell^p(\mathbb{K})$ des suites de $\mathbb{K}$ de puissance p-ième sommable d'une structure d'espace vectoriel normé en posant :

$$\|x\|_{\ell^p} = \left(\sum_{n=0}^{+\infty} |x_n|^p\right)^{\frac{1}{p}}.$$

L'inégalité triangulaire se démontre comme dans l'exemple précédent en faisant tendre n vers l'infini.

L'ensemble $\ell^\infty(\mathbb{K})$ des suites bornées de $\mathbb{K}$ peut être muni de la norme $\|x\|_{\ell^\infty} = \sup_{n\in\mathbb{N}} |x_n|$.

1. qui se démontre dans le cas $a > 0$ et $b > 0$ en utilisant la concavité de la fonction logarithme.

Exemple 3. L'ensemble $\mathcal{C}([0,1];\mathbb{R})$ peut être muni d'une norme en posant

$$\|f\|_{L^\infty} \stackrel{\text{déf}}{=} \sup_{t\in[0,1]} |f(t)|.$$

La norme $\|\cdot\|_{L^\infty}$ ainsi définie est appelée **norme uniforme**.

D'autres choix sont possibles. Par exemple

$$\|f\|_{L^2} \stackrel{\text{déf}}{=} \sqrt{\int_0^1 |f(t)|^2\,dt}, \qquad \|f\|_{L^1} \stackrel{\text{déf}}{=} \int_0^1 |f(t)|\,dt \quad \text{ou même} \quad \|f\|_{L^p} \stackrel{\text{déf}}{=} \left(\int_0^1 |f(t)|^p\,dt\right)^{\frac{1}{p}}$$

avec $p \in [1,+\infty[$.

Exemple 4. L'ensemble $\mathcal{M}_n(\mathbb{R})$ des matrices carrées de taille n à coefficients réels peut être muni des normes

$$\|A\|_\infty = \max_{1\leq i,j\leq n} |a_{ij}| \quad \text{ou} \quad \|A\|_p = \left(\sum_{i=1}^n \sum_{j=1}^n |a_{ij}|^p\right)^{\frac{1}{p}} \quad \text{avec} \quad p\in[1,+\infty[.$$

Définition. Soit E un e.v. muni de deux normes $\|\cdot\|_1$ et $\|\cdot\|_2$.
On dit que $\|\cdot\|_2$ est plus forte que $\|\cdot\|_1$ s'il existe une constante $c>0$ telle que

$$\forall x\in E,\ \|x\|_1 \leq c\|x\|_2.$$

On dit que $\|\cdot\|_1$ et $\|\cdot\|_2$ sont équivalentes s'il existe $c>0$ telle que

$$\forall x\in E,\ c^{-1}\|x\|_2 \leq \|x\|_1 \leq c\|x\|_2.$$

Il est clair que les distances associées à deux normes équivalentes, sont équivalentes. En conséquence, d'après la proposition 4 page 15, on a :

Proposition. *Les topologies engendrées par deux normes équivalentes sont les mêmes.*

À titre d'exercice, le lecteur pourra vérifier la proposition suivante :

Proposition. *Soit $\|\cdot\|_1$ et $\|\cdot\|_2$ deux normes sur E.*

Alors $\|\cdot\|_2$ est plus forte que $\|\cdot\|_1$. si et seulement si tout ouvert de E pour la topologie associée à $\|\cdot\|_1$ est ouvert pour la topologie de $\|\cdot\|_2$.

On retiendra que plus la norme est forte plus la topologie associée est fine (c'est-à-dire plus elle a d'ouverts).

Exemple 1. Dans $\mathcal{M}_n(\mathbb{R})$, les normes $\|\cdot\|_1$ et $\|\cdot\|_\infty$ sont équivalentes.

Exemple 2. Dans $\mathcal{C}([0,1];\mathbb{R})$, la norme $\|\cdot\|_{L^\infty}$ est (strictement) plus forte que $\|\cdot\|_{L^1}$.

Remarque : Si $(E,\|\cdot\|_E)$ et $(F,\|\cdot\|_F)$ sont deux espaces vectoriels normés, on peut munir $E\times F$ d'une **norme produit** en posant pour tout $(u,v)\in E\times F$,

$$\|(u,v)\|_1 = \|u\|_E + \|v\|_F, \quad \|(u,v)\|_2 = \sqrt{\|u\|_E^2+\|v\|_F^2} \quad \text{ou} \quad \|(u,v)\|_\infty = \max(\|u\|_E,\|v\|_F).$$

Les trois normes ci-dessus sont équivalentes. Les topologies associées sont donc les mêmes.

Nous laissons au lecteur le soin de définir une norme pour le produit d'un nombre fini d'e.v.n.

3.2 Applications linéaires

Notation. Soit E et F deux espaces vectoriels normés sur le même corps $\mathbb{K}$. On note $L(E;F)$ l'ensemble des applications linéaires de E dans F.

L'équivalence entre les deux premières assertions de la proposition suivante sera fondamentale pour la suite du cours.

Proposition 5. *Soit $u \in L(E;F)$. Les cinq propriétés suivantes sont équivalentes.*

(i) $\exists M \geq 0, \forall x \in E, \|u(x)\|_F \leq M\|x\|_E$,

(ii) *u est continue sur E,*

(iii) *u est continue en 0,*

(iv) *u est bornée sur la boule unité fermée,*

(v) *u est bornée sur la sphère unité.*

Preuve : Il suffit de prouver $(i) \Rightarrow (ii) \Rightarrow (iii) \Rightarrow (iv) \Rightarrow (v) \Rightarrow (i)$.

$(i) \Rightarrow (ii)$ Par linéarité de u, on a $u(y) - u(x) = u(y-x)$. Donc,

$$\forall (x,y) \in E^2,\ \|u(y) - u(x)\|_F \leq M\|y - x\|_E.$$

Donc u est lipschitzienne donc continue.

$(ii) \Rightarrow (iii)$ Trivial.

$(iii) \Rightarrow (iv)$ De la continuité en 0 de u, on déduit l'existence d'un $\eta > 0$ tel que $\|u(x)\|_F \leq 1$ dès que $x \in \overline{B}_E(0,\eta)$. Or $x \in \overline{B}_E(0,\eta)$ si et seulement si $\eta^{-1}x \in \overline{B}_E(0,1)$. Par linéarité de u, on conclut que $\|u(x)\|_F \leq \eta^{-1}$ pour tout $x \in \overline{B}_E(0,1)$. Donc u est bornée sur la boule unité.

$(iv) \Rightarrow (v)$ Trivial.

$(v) \Rightarrow (i)$ Notons M une borne de u restreinte à la sphère unité. Si $x \neq 0$, le point $x/\|x\|_E$ appartient à la sphère unité. En utilisant la linéarité de u et le fait que u est bornée par M sur la sphère unité, on obtient donc

$$\|u(x)\|_F = \|x\|_E \left\| u\Big(\frac{x}{\|x\|_E}\Big) \right\|_F \leq M\|x\|_E.$$

■

Exercice : À l'aide de la proposition ci-dessus, démontrer que l'application

$$\begin{cases} E \times E & \longrightarrow & E \\ (x,y) & \longmapsto & x + y \end{cases}$$

est continue.

Définition. Soit $\mathcal{L}(E,F)$ l'ensemble des applications linéaires continues de E dans F. Pour $f \in \mathcal{L}(E,F)$, on note $\|f\|_{\mathcal{L}(E,F)}$ la plus petite constante M telle que

$$\forall x \in E,\ \|f(x)\|_F \leq M\|x\|_E.$$

On a donc

$$\forall x \in E,\ \|f(x)\|_F \leq \|f\|_{\mathcal{L}(E;F)}\|x\|_E.$$

Le lecteur vérifiera facilement que l'on a aussi

$$\|f\|_{\mathcal{L}(E;F)} = \sup_{x \in E\setminus\{0\}} \frac{\|f(x)\|_F}{\|x\|_E} = \sup_{\substack{x \in E \\ \|x\|_E = 1}} \|f(x)\|_F. \tag{3.1}$$

Proposition. *L'espace* $\big(\mathcal{L}(E,F);\|\cdot\|_{\mathcal{L}(E,F)}\big)$ *est un espace vectoriel normé.*

Preuve : Vérifions rapidement que $\|\cdot\|_{\mathcal{L}(E;F)}$ est une norme. La proposition 5 assure que pour tout $f\in\mathcal{L}(E,F)$, la quantité $\|f\|_{\mathcal{L}(E;F)}$ est finie (et positive). Il est de plus immédiat que $\|\lambda f\|_{\mathcal{L}(E;F)}=|\lambda|\|f\|_{\mathcal{L}(E;F)}$ pour tout $\lambda\in\mathbb{K}$ et que $\|f\|_{\mathcal{L}(E;F)}=0$ si et seulement si $f=0$.

Si f et g sont deux éléments de $\mathcal{L}(E,F)$, on a pour tout $x\in E$,

$$\|(f+g)(x)\|_F=\|f(x)+g(x)\|_F\leq\|f(x)\|_F+\|g(x)\|_F\leq\big(\|f\|_{\mathcal{L}(E;F)}+\|g\|_{\mathcal{L}(E;F)}\big)\|x\|_E.$$

Donc $f+g$ est linéaire continue et vérifie $\|f+g\|_{\mathcal{L}(E;F)}\leq\|f\|_{\mathcal{L}(E;F)}+\|g\|_{\mathcal{L}(E;F)}$. Donc $\|\cdot\|_{\mathcal{L}(E;F)}$ est bien une norme. ∎

Attention : La définition de $\mathcal{L}(E;F)$ dépend du choix des normes sur E et F.

Pour illustrer ce fait, prenons $E=\mathcal{C}([0,1];\mathbb{R})$ et $F=\mathbb{R}$. On munit F de la norme donnée par la valeur absolue. Considérons la forme linéaire L définie sur E par $L(f)=f(0)$.

Il est immédiat que L est continue si l'on munit E de la norme $\|\cdot\|_{L^\infty}$. En revanche, L n'est pas continue si l'on munit E de la norme $\|\cdot\|_{L^1}$. Pour s'en persuader, on pourra considérer la suite de fonctions $(f_n)_{n\in\mathbb{N}}$ définie par

$$f_n(x)=\begin{cases} n-n^2x & \text{si}\quad x\in[0,\frac{1}{n}],\\ 0 & \text{si}\quad x\in[\frac{1}{n},1].\end{cases}$$

Remarque. Cependant, changer la norme sur E et la norme sur F en des normes équivalentes ne modifie pas $\mathcal{L}(E;F)$, et change $\|\cdot\|_{\mathcal{L}(E;F)}$ en une norme équivalente.

Proposition. *Soit E, F et G trois e.v.n, $f\in\mathcal{L}(E;F)$ et $g\in\mathcal{L}(F;G)$. Alors $g\circ f\in\mathcal{L}(E;G)$ et l'on a l'inégalité suivante :*

$$\|g\circ f\|_{\mathcal{L}(E;G)}\leq\|f\|_{\mathcal{L}(E;F)}\,\|g\|_{\mathcal{L}(F;G)}.$$

Preuve : Tout d'abord, le théorème de composition des applications continues assure que $g\circ f$ est continue. La linéarité de $g\circ f$ est évidente. Soit $x\in E$. Par définition de $\|g\|_{\mathcal{L}(F;G)}$, on a

$$\|g\circ f(x)\|_G\leq\|g\|_{\mathcal{L}(F;G)}\|f(x)\|_F,$$

puis par définition de $\|f\|_{\mathcal{L}(E;F)}$,

$$\|g\circ f(x)\|_G\leq\|g\|_{\mathcal{L}(F;G)}\|f\|_{\mathcal{L}(E;F)}\|x\|_E.$$

En conséquence, on a bien $\|g\circ f\|_{\mathcal{L}(E;G)}\leq\|f\|_{\mathcal{L}(E;F)}\,\|g\|_{\mathcal{L}(F;G)}$. ∎

Remarque. La proposition ci-dessus assure que l'ensemble $\mathcal{L}(E)$ des applications linéaires continues de E dans E est tel que :

$$\forall(f,g)\in\mathcal{L}(E),\ \|g\circ f\|_{\mathcal{L}(E)}\leq\|f\|_{\mathcal{L}(E)}\|g\|_{\mathcal{L}(E)}.$$

On vérifie aisément que $\|\mathrm{Id}_E\|_{\mathcal{L}(E)}=1$.

Ces deux propriétés supplémentaires confèrent à l'e.v.n. $\big(\mathcal{L}(E);\|\cdot\|_{\mathcal{L}(E)}\big)$ une structure d'**algèbre normée**. On dit que $\|\cdot\|_{\mathcal{L}(E)}$ est une **norme d'algèbre**.

Pour les applications multilinéaires, on peut établir une caractérisation de la continuité similaire à celle que l'on a pour les applications linéaires :

Proposition. *Soit E_1, $\cdots$, E_k et F des espaces vectoriels sur $\mathbb{K}$, munis de normes $\|\cdot\|_{E_1}$, $\cdots$, $\|\cdot\|_{E_k}$ et $\|\cdot\|_F$. Soit u une application k-linéaire de $E_1\times\cdots\times E_k$ dans F.*

Les quatre propriétés suivantes sont équivalentes :

(i) *u est continue,*

(ii) *u est continue en* $(0,\cdots,0)$*,*

(iii) *u est bornée sur* $\overline{B}_{E_1}(0,1)\times\cdots\times\overline{B}_{E_k}(0,1)$*,*

(iv) *il existe* $M\in\mathbb{R}^+$ *tel que*

$$\forall(x_1,\cdots,x_k)\in E_1\times\cdots\times E_k,\ \|u(x_1,\cdots,x_k)\|_F\leq M\|x_1\|_{E_1}\cdots\|x_k\|_{E_k}. \tag{3.2}$$

Preuve : Pour simplifier la présentation, on suppose que $k=2$ et l'on munit $E_1\times E_2$ de la norme produit

$$\|(x_1,x_2)\|_{E_1\times E_2}=\max\bigl(\|x_1\|_{E_1},\|x_2\|_{E_2}\bigr).$$

Il suffit de prouver $(i)\Rightarrow(ii)\Rightarrow(iii)\Rightarrow(iv)\Rightarrow(i)$.

$(i)\Rightarrow(ii)$ Évident.

$(ii)\Rightarrow(iii)$ De la continuité en $(0,0)$, on déduit l'existence d'un $\eta>0$ tel que $\|x_1\|_{E_1}\leq\eta$ et $\|x_2\|_{E_2}\leq\eta$ entraîne $\|u(x_1,x_2)\|_F\leq 1$. On en déduit que

$$\|(x_1,x_2)\|_{E_1\times E_2}\leq 1\Longrightarrow\|u(x_1,x_2)\|_F\leq\eta^{-2}.$$

$(iii)\Rightarrow(iv)$ Soit M une borne de u sur $\overline{B}_{E_1}(0,1)\times\overline{B}_{E_2}(0,1)$. Soit $(x_1,x_2)\in E_1\times E_2$ tel que $x_1\neq 0$ et $x_2\neq 0$. On a alors $\bigl(\frac{x_1}{\|x_1\|_{E_1}},\frac{x_2}{\|x_2\|_{E_2}}\bigr)\in\overline{B}_{E_1}(0,1)\times\overline{B}_{E_2}(0,1)$ donc

$$\|u(x_1,x_2)\|_{E_1\times E_2}=\|x_1\|_{E_1}\|x_2\|_{E_2}\left\|u\Bigl(\frac{x_1}{\|x_1\|_{E_1}},\frac{x_2}{\|x_2\|_{E_2}}\Bigr)\right\|_F\leq M\|x_1\|_{E_1}\|x_2\|_{E_2}.$$

$(iv)\Rightarrow(i)$ Soit (x_1,x_2) et (y_1,y_2) deux éléments de $E_1\times E_2$. Par bilinéarité de u, on a

$$u(y_1,y_2)-u(x_1,x_2)=u(y_1-x_1,y_2)+u(x_1,y_2-x_2),$$

donc

$$\|u(y_1,y_2)-u(x_1,x_2)\|_F\leq M\bigl(\|y_1-x_1\|_{E_1}\|y_2\|_{E_2}+\|x_1\|_{E_1}\|y_2-x_2\|_{E_2}\bigr).$$

Cela assure visiblement la continuité en (x_1,y_1).

■

Exercice : À l'aide de la proposition ci-dessus, montrer que l'application $\left\{\begin{array}{ccc}\mathbb{K}\times E&\longrightarrow&E\\(\lambda,x)&\longmapsto&\lambda x\end{array}\right.$ est continue.

Notations : On note $\mathcal{L}^k(E_1\times\cdots\times E_k;F)$ l'ensemble des applications k-linéaires continues de $E_1\times\cdots\times E_k$ vers F.

Pour $u\in\mathcal{L}^k(E_1\times\cdots\times E_k;F)$, on note $\|\cdot\|_{\mathcal{L}^k(E_1\times\cdots\times E_k;F)}$ la borne inférieure de toutes les constantes M vérifiant (3.2). On peut vérifier que c'est une norme sur $\mathcal{L}^k(E_1\times\cdots\times E_k;F)$.

3.3 Sous-espaces vectoriels

Proposition. *Soit* $(E,\|\cdot\|_E)$ *un e.v.n. et* F *un s.e.v. de* E*. Alors l'adhérence de* F *dans* E *est un s.e.v. de* E*.*

Preuve : C'est immédiat en utilisant la caractérisation de l'adhérence par les suites. ■

Voici un résultat fort utile de prolongement des applications linéaires continues :

Théorème. *Soit* $(E, \|\cdot\|_E)$ *un e.v.n,* $(G, \|\cdot\|_G)$ *un e.v.n. complet et* F *un s.e.v. dense de* E. *Soit* $L \in \mathcal{L}(F;G)$.

Il existe une unique application $\widetilde{L} \in \mathcal{L}(E;G)$ *qui prolonge* L *sur* E. *De plus, on a*

$$\|\widetilde{L}\|_{\mathcal{L}(E;G)} = \|L\|_{\mathcal{L}(F;G)}.$$

Preuve : Sachant que toute application linéaire continue est en fait lipschitzienne et donc uniformément continue, la proposition de la page 22 assure l'existence et l'unicité d'un prolongement continu $\widetilde{L}$ défini sur E. Vérifions la linéarité de $\widetilde{L}$. Soit donc $(\lambda,\mu) \in \mathbb{K}^2$, $(x,y) \in E^2$, $(x_n)_{n\in\mathbb{N}} \in F^{\mathbb{N}}$ une suite tendant vers x et $(y_n)_{n\in\mathbb{N}} \in F^{\mathbb{N}}$ une suite tendant vers y. En passant à la limite dans l'égalité $L(\lambda x_n + \mu y_n) = \lambda L(x_n) + \mu L(y_n)$, on en déduit que $L(\lambda x + \mu y) = \lambda L(x) + \mu L(y)$.

Enfin, sachant que

$$\|\widetilde{L}\|_{\mathcal{L}(E;G)} = \sup_{x\in E\setminus\{0\}} \frac{\|\widetilde{L}(x)\|_G}{\|x\|_E} \quad \text{et} \quad \|L\|_{\mathcal{L}(F;G)} = \sup_{x\in F\setminus\{0\}} \frac{\|L(x)\|_G}{\|x\|_E},$$

et que $L(x) = \widetilde{L}(x)$ pour $x \in F$, il est clair que $\|\widetilde{L}\|_{\mathcal{L}(E;G)} \geq \|L\|_{\mathcal{L}(F;G)}$. Mais si $x \in E$ et $(x_n)_{n\in\mathbb{N}} \in F^{\mathbb{N}}$ tend vers x alors on a

$$\forall n \in \mathbb{N},\ \|\widetilde{L}(x_n)\|_G \leq \|L\|_{\mathcal{L}(F;G)} \|x_n\|_E,$$

donc

$$\|\widetilde{L}(x)\|_G \leq \|L\|_{\mathcal{L}(F;G)} \|x\|_E.$$

Cela assure que $\|\widetilde{L}\|_{\mathcal{L}(E;G)} \leq \|L\|_{\mathcal{L}(F;G)}$. ■

Théorème. *Soit* f *une forme linéaire sur* $(E, \|\cdot\|_E)$. *Alors* f *est continue si et seulement si* $\ker f$ *est fermé.*

Preuve : Si f est continue alors l'ensemble $\ker f$ est fermé car image réciproque du fermé $\{0\}$ de $\mathbb{K}$. Cela établit l'implication directe.

Pour établir la réciproque, supposons que f ne soit pas continue. Alors il existe une suite $(x_n)_{n\in\mathbb{N}}$ tendant vers 0 et telle que $f(x_n) = 1$ pour tout $n \in \mathbb{N}$.

En effet, d'après la proposition 5, la forme linéaire f n'est pas continue en 0. Donc il existe $\varepsilon > 0$ et une suite $(y_n)_{n\in\mathbb{N}}$ tendant vers 0 telle que $|f(y_n)| \geq \varepsilon$ pour tout $n \in \mathbb{N}$. La suite $(x_n)_{n\in\mathbb{N}}$ définie par $x_n = y_n/f(y_n)$ vérifie les propriétés souhaitées.

Soit maintenant $z \in E$ n'appartenant pas à $\ker f$ (c'est possible car f n'est pas nulle car non continue). On constate que la suite de terme général $z - f(z)x_n$ est dans $(\ker f)^{\mathbb{N}}$ et converge vers z. ■

De la démonstration ci-dessus découle en fait un résultat bien plus précis, à savoir :

Corollaire. *Une forme linéaire non nulle sur un e.v.n.* E *n'est pas continue si et seulement si son noyau est dense dans* E.

Théorème (de Riesz)**.** *Soit* $(E, \|\cdot\|_E)$ *un e.v.n. et* M *un s.e.v. fermé de* E *distinct de* E. *Alors pour tout* $\varepsilon > 0$, *il existe* $x_\varepsilon \in E$ *tel que*

$$\|x_\varepsilon\|_E = 1 \quad \textit{et} \quad \inf_{m\in M} \|x_\varepsilon - m\| \geq 1 - \varepsilon.$$

Preuve : Soit $y \in E \setminus M$. Posons $\alpha \stackrel{\text{déf}}{=} d(y, M)$. Comme M est fermé et $y \notin M$, on a $\alpha > 0$. Pour $\varepsilon \in]0,1[$ fixé, on choisit alors $m_\varepsilon \in M$ tel que $\|y - m_\varepsilon\|_E \leq \alpha/(1-\varepsilon)$.

Un calcul facile montre que le point $x_\varepsilon \stackrel{\text{déf}}{=} \dfrac{y - m_\varepsilon}{\|y - m_\varepsilon\|_E}$ répond à la question. ■

Remarque : Dans un espace euclidien, en faisant appel à l'orthogonalité, on peut facilement construire un $x \in E$ correspondant à $\varepsilon = 0$. Mais pour un e.v.n. "général", on ne peut pas faire mieux que le théorème de Riesz.

3.4 Espaces de Banach

Définition. Un espace vectoriel normé complet est appelé **espace de Banach**.

Exemples. 1) Nous verrons dans le chapitre 5 que l'espace vectoriel normé $\mathcal{C}([0,1];\mathbb{R})$ muni de la norme $\|\cdot\|_{L^\infty}$ est complet.
2) En revanche l'espace $\mathcal{C}([0,1];\mathbb{R})$ muni de la norme $\|\cdot\|_{L^1}$ n'est pas complet. Pour le voir, on peut par exemple considérer la suite $(f_n)_{n\geq 2}$ définie par

$$f_n(t) = \begin{cases} 1 & \text{si} \quad t \in [0, 1/2], \\ n(\frac{1}{2}+\frac{1}{n} - t) & \text{si} \quad t \in [\frac{1}{2}, \frac{1}{2}+\frac{1}{n}], \\ 0 & \text{si} \quad t \in [\frac{1}{2}+\frac{1}{n}, 1]. \end{cases}$$

Elle est de Cauchy au sens de la norme $\|\cdot\|_{L^1}$ mais ne converge pas dans $\mathcal{C}([0,1];\mathbb{R})$.
3) Nous verrons dans la section suivante que tous les e.v.n. de dimension finie sont complets.

Rappelons qu'une série $\sum u_n$ d'éléments de E est dite **absolument convergente** si $\sum \|u_k\|$ converge.

Proposition. *Dans un espace de Banach, toute série absolument convergente est convergente.*

Preuve : En vertu de la propriété de complétude, il suffit d'établir que la suite de terme général $\sum_{k=0}^n u_k$ est de Cauchy. Cela découle du fait que pour tout $m > n$, on a

$$\left\| \sum_{k=0}^m u_k - \sum_{k=0}^n u_k \right\| \leq \sum_{k=n+1}^m \|u_k\|,$$

et que $\sum \|u_k\|$ converge. ∎

Exercice : Montrer que, réciproquement, si E est un e.v.n. dans lequel toute série absolument convergente, est convergente, alors E est complet.

Donnons un résultat de complétude fort utile :

Proposition. *Soit $(E, \|\cdot\|_E)$ un e.v.n. et $(F, \|\cdot\|_F)$ un Banach. Alors $\big(\mathcal{L}(E;F); \|\cdot\|_{\mathcal{L}(E;F)}\big)$ est un Banach.*

Preuve : Nous donnons juste la structure de la preuve et laissons au lecteur le plaisir d'écrire les détails. Soit donc $(L_n)_{n\in\mathbb{N}}$ une suite de Cauchy d'éléments de $\mathcal{L}(E;F)$. Il s'agit de montrer la convergence vers un élément L de $\mathcal{L}(E;F)$.

1. *Convergence simple :* On montre que pour tout $x \in E$ la suite $(L_n(x))_{n\in\mathbb{N}}$ est une suite de Cauchy d'éléments de F. Comme F est complet, elle converge vers un élément $L(x)$ de F.
2. *Étude de la limite :* On vérifie d'abord que $L \in L(E;F)$ puis que L est continue.
3. *Convergence en norme :* On vérifie que $(L_n)_{n\in\mathbb{N}}$ converge vers L dans $\mathcal{L}(E;F)$.

∎

Corollaire. *Soit E un e.v.n. (quelconque). L'ensemble des formes linéaires continues sur E est un espace de Banach.*

Notation. L'ensemble des formes linéaires continues sur E est appelé **dual topologique de** E, et noté E'. La norme $\|f\|_{E'}$ d'un élément de E' est donc définie par

$$\|f\|_{E'} = \sup_{x\neq 0} \frac{|f(x)|}{\|x\|_E}.$$

3.5 Le cas de la dimension finie

Le résultat suivant est fondamental :

Théorème. *Dans un e.v. de dimension finie, les compacts sont les fermés bornés et toutes les normes sont équivalentes.*

Preuve : Nous avons déjà vu qu'un compact est fermé borné. Il suffit donc d'établir qu'en dimension finie tout fermé borné est compact et que toutes les normes sont équivalentes.
Soit donc E un e.v. de dimension finie p et $(e_1, \cdots, e_p)$ une base de E. Dans tout ce qui suit, on note

$$\|x\|_E^\infty = \max_{1 \le i \le p} |x_i| \quad \text{pour} \quad x = \sum_{i=1}^p x_i e_i.$$

Première étape : On montre que dans $(E, \|\cdot\|_E^\infty)$ les fermés bornés sont compacts.
Soit donc A une partie fermée bornée de E au sens de la norme $\|\cdot\|_E^\infty$. Supposons pour simplifier que $\mathbb{K} = \mathbb{R}$ (l'adaptation au cas $\mathbb{K} = \mathbb{C}$ est laissée au lecteur). Soit

$$\varphi : \begin{cases} (\mathbb{R}^p, \|\cdot\|_\infty) \longrightarrow (E, \|\cdot\|_E^\infty) \\ (x_1, \cdots, x_p) \longmapsto \sum_{i=1}^p x_i e_i. \end{cases}$$

Alors φ est un isomorphisme isométrique entre $(\mathbb{R}^p, \|\cdot\|_\infty)$ et $(E, \|\cdot\|_E^\infty)$. En particulier, φ^{-1} est continue, donc $\varphi^{-1}(A)$ est un fermé borné de $(\mathbb{R}^p, \|\cdot\|_\infty)$. Soit $M \ge 0$ tel que $\varphi^{-1}(A) \subset [-M, M]^p$. L'ensemble $[-M, M]^p$ est compact car produit cartésien de compacts. Comme $\varphi^{-1}(A)$ est fermé, on conclut que $\varphi^{-1}(A)$ est compact, puis que $A = \varphi(\varphi^{-1}(A))$ est également compact.

Deuxième étape : Équivalence des normes.
Soit maintenant $\|\cdot\|_E$ une norme quelconque sur E. Définissons :

$$\Psi : \begin{cases} (E, \|\cdot\|_E^\infty) & \longrightarrow & (\mathbb{R}, |\cdot|) \\ x & \longmapsto & \|x\|_E. \end{cases}$$

Pour $(x, y) \in E^2$, on a (avec des notations évidentes) grâce à la deuxième inégalité triangulaire,

$$\begin{aligned} |\Psi(y) - \Psi(x)| &\le \|y - x\|_E, \\ &\le \left\| \sum_{i=1}^p (y_i - x_i) e_i \right\|_E, \\ &\le \left(\sum_{i=1}^p \|e_i\|_E \right) \|y - x\|_E^\infty. \end{aligned}$$

En conséquence, l'application Ψ est continue sur E. Notons au passage que le calcul précédent montre que la norme $\|\cdot\|_E^\infty$ est plus forte que $\|\cdot\|_E$ (prendre $y = 0$ pour le voir).
Sachant que d'après la première étape la sphère unité de E pour la norme $\|\cdot\|_E^\infty$ est compacte, on en déduit que Ψ restreinte à S est bornée et atteint ses bornes. Cela signifie en particulier qu'il existe un x_0 tel que

$$\|x_0\|_E^\infty = 1 \quad \text{et} \quad \|x\|_E \ge \|x_0\|_E \quad \text{pour tout} \quad x \in E \quad \text{tel que} \quad \|x\|_E^\infty = 1.$$

Par homogénéité, on conclut que

$$\forall x \in E,\ \|x\|_E \ge \|x_0\|_E \|x\|_E^\infty.$$

Enfin, il est clair que x_0 ne peut pas être nul. Donc $\|x_0\|_E > 0$. En conséquence, la norme $\|\cdot\|_E$ est plus forte que $\|\cdot\|_\infty$.
Finalement, les deux normes sont donc équivalentes.

Troisième étape : Fin de la preuve de la compacité.

Sachant que toute norme de E est équivalente à $\|\cdot\|_E^\infty$, la première étape permet maintenant conclure que tout fermé borné de E est compact.

■

Corollaire. *Tous les e.v.n. de dimension finie sont complets.*

Preuve : Soit E un e.v.n. de dimension finie. Comme toutes les normes sur E sont équivalentes on peut, sans nuire à la généralité, supposer que E est muni de la norme $\|\cdot\|_E^\infty$. Avec ce choix de norme, il apparaît clairement qu'une suite $(x_n)_{n\in\mathbb{N}}$ est de Cauchy dans E si et seulement si les suites de ses composantes $(x_n^i)_{n\in\mathbb{N}}$ pour $i\in\{1,\cdots,p\}$ sont de Cauchy dans $\mathbb{K}$. Comme $\mathbb{K}$ est complet, on en déduit que les suites des composantes convergent, puis que $(x_n)_{n\in\mathbb{N}}$ converge aussi. ■

En dimension infinie, les sous-espaces vectoriels ne sont pas nécessairement fermés. Par exemple, si l'on considère l'espace vectoriel c_0 des suites réelles tendant vers 0 à l'infini, muni de la norme $\|x\|_\infty=\sup_{n\in\mathbb{N}}|x_n|$, et le sous-espace vectoriel F de c_0 constitué par les suites réelles nulles à partir d'un certain rang, alors F est un s.e.v. dense de c_0 qui n'est pas fermé.

Nous disposons cependant du résultat suivant :

Théorème. *Soit $(E,\|\cdot\|_E)$ un e.v.n. quelconque et F un s.e.v. de E de dimension finie. Alors F est fermé dans E.*

Preuve : Fixons une base $(e^1,\cdots,e^p)$ de F. Comme toutes les normes sont équivalentes en dimension finie, la norme de E restreinte à F est équivalente à la norme $\|\cdot\|_\infty$ sur F associée à $(e^1,\cdots,e^p)$.

Soit $(x_n)_{n\in\mathbb{N}}$ une suite convergente de F. Elle est donc de Cauchy au sens de la norme $\|\cdot\|_\infty$ introduite ci-dessus et il est alors immédiat que toutes les composantes de la suite par rapport à $(e^1,\cdots,e^p)$ sont des suites de Cauchy de $\mathbb{K}$, donc convergent. ■

Théorème. *La boule unité fermée de l'e.v.n. E est compacte si et seulement si E est de dimension finie.*

Preuve : Seule l'implication directe reste à prouver. On raisonne par contraposition : supposons E de dimension infinie. On peut alors construire par récurrence une suite $(e_k)_{k\in\mathbb{N}}$ de vecteurs linéairement indépendants. La suite $(V_k)_{k\in\mathbb{N}}$ des sous-espaces vectoriels définis par $V_k\stackrel{\text{déf}}{=}\operatorname{Vect}(e_0,\cdots,e_k)$ est une suite strictement croissante de s.e.v. fermés distincts de E. En appliquant le théorème de Riesz, on peut alors construire par récurrence une suite $(x_n)_{n\in\mathbb{N}}$ de vecteurs unitaires telle que $d(x_n,V_{n-1})\geq 1/2$ et $x_n\in V_n$. On a visiblement $\|x_n-x_m\|\geq 1/2$ pour $n\neq m$. En conséquence la suite $(x_n)_{n\in\mathbb{N}}$ n'a pas de valeur d'adhérence et la boule $\overline{B}_E(0,1)$ n'est donc pas compacte. ■

Corollaire. *Soit E un e.v.n. de dimension finie. Alors*

(i) toute suite bornée de E admet une valeur d'adhérence,

(ii) toute suite bornée ayant une seule valeur d'adhérence converge.

Exemple. Considérons l'espace $E=\mathcal{C}([0,\pi];\mathbb{R})$ muni de la norme

$$\|f\|_{L^2}\stackrel{\text{déf}}{=}\sqrt{\frac{1}{\pi}\int_0^\pi|f(t)|^2\,dt}\,.$$

Pour $n\in\mathbb{N}^*$, posons $f_n(x)=\sin nx$. Il est clair que $\|f_n\|_{L^2}^2=1/2$ et que $\|f_n-f_m\|_{L^2}=1$ pour $n\neq m$. Donc la suite $(f_n)_{n\in\mathbb{N}}$ est une suite de $\overline{B}_E(0,1)$ qui n'a pas de valeur d'adhérence. On en conclut que E est de dimension infinie.

Remarque : Dans un e.v.n. $(E, \|\cdot\|_E)$ de dimension infinie, tous les compacts sont d'intérieur vide. En effet si l'ensemble K n'est pas d'intérieur vide alors il contient une boule fermée $\overline{B}_E(x_0, r)$ avec $r > 0$. Si K était compact alors la boule fermée $\overline{B}_E(x_0, r)$ serait aussi compacte et donc E, de dimension finie.

Théorème. *Soit $(E, \|\cdot\|_E)$ et $(F, \|\cdot\|_F)$ deux e.v.n, et $f \in L(E; F)$ une application linéaire. Si E est de dimension finie alors f est nécessairement continue.*

Preuve : Fixons une base $(e^1, \cdots, e^p)$ de E. Puisque sur E toutes les normes sont équivalentes, on peut supposer que $\|x\|_E = \max_{1 \le i \le p} |x^i|$ où $(x^1, \cdots, x^p)$ sont les coordonnées de x par rapport à $(e^1, \cdots, e^p)$.

On a donc, d'après l'inégalité triangulaire et la linéarité de f,

$$\|f(x)\|_F = \|\sum_{i=1}^{p} x^i f(e^i)\|_F \le \Big(\sum_{i=1}^{p} \|f(e^i)\|_F\Big) \|x\|_E,$$

d'où la continuité de f. ∎

Le lecteur pourra vérifier que plus généralement on a le résultat suivant :

Théorème. *Supposons que $E_1, \cdots, E_p$ soient <u>de dimension finie</u>. Alors toute application p-linéaire de $E_1 \times \cdots \times E_p$ dans F est continue.*

Chapitre 4

Connexité et convexité

4.1 Connexité

La notion de connexité est très importante en topologie. Si l'on cherche à s'en faire une représentation intuitive, on retiendra qu'une partie connexe ne comporte "qu'un seul morceau".

4.1.1 Cadre général

La définition mathématique de connexité que nous donnons ci-dessous peut sembler de prime abord assez éloignée de la représentation intuitive que nous avons tenté d'imposer au lecteur :

Proposition. *Soit $(X, \mathcal{O})$ un espace topologique et A une partie de X. Les trois propriétés suivantes sont équivalentes*[1] *:*

(i) A et $\emptyset$ sont les seules parties ouvertes et fermées de A,

(ii) il n'existe pas de couple d'ouverts non vides de A, disjoints et de réunion égale à A,

(iii) il n'existe pas de couple de fermés de A non vides, disjoints et de réunion égale à A.

Si l'une de ces trois propriétés est vérifiée, on dit que A est une partie connexe de X.

Preuve : $(i) \Rightarrow (ii)$: Soit Ω_1 et Ω_2 deux ouverts disjoints et de réunion égale à A. Alors $\Omega_2 = A \setminus \Omega_1$ donc Ω_2 est à la fois ouvert et fermé. On en déduit que Ω_2 vaut $\emptyset$ ou A.

$(ii) \Rightarrow (i)$: Soit Ω une partie ouverte et fermée de A. Alors il en est de même de $A \setminus \Omega$, donc l'un des deux ensembles Ω et $A \setminus \Omega$ doit être vide.

$(ii) \iff (iii)$: Il suffit de passer au complémentaire. ■

Exemples. (i) Soit $(X, \mathcal{O})$ un espace topologique séparé et A une partie finie de X ayant au moins deux éléments. Il est facile de voir que tous les singletons de A sont à la fois ouverts et fermés. En conséquence A n'est pas connexe.

(ii) Tout espace vectoriel normé est connexe.

Proposition. *L'image d'une partie connexe par une application continue est connexe.*

Preuve : Soit $(X, \mathcal{O})$ et $(X', \mathcal{O}')$ deux espaces topologiques, A une partie connexe de X et $f \in \mathcal{C}(A; X')$. Notons $B = f(A)$. Soit U une partie ouverte et fermée de B (pour la topologie induite). Par continuité de f, la partie $f^{-1}(U)$ est ouverte et fermée dans A. Puisque A est connexe, on a donc $f^{-1}(U) = \emptyset$ (auquel cas $U = \emptyset$) ou bien $f^{-1}(U) = A$ (et alors $U = B$). ■

1. Ci-dessous, quand on parle d'ouverts ou de fermés, c'est au sens de la topologie induite sur A.

Corollaire 1. *Une partie A de l'espace topologique $(X, \mathcal{O})$ est connexe si et seulement si toute application continue de A dans $\{0,1\}$ muni de la topologie discrète est constante.*

Preuve : Soit A connexe et f une application continue de A dans $\{0,1\}$. Alors $f(A)$ doit être une partie connexe de $\{0,1\}$. Sachant que $\{0,1\}$ n'est pas connexe, cela signifie que $f(A)$ ne doit comporter qu'un seul élément.

Pour montrer la réciproque, procédons par contraposition. Supposons donc que A n'est pas connexe. Alors il existe deux fermés B et C non vides, disjoints et tels que $A = B \cup C$. On définit la fonction f sur A par $f(x) = 0$ si $x \in B$ et $f(x) = 1$ si $x \in C$. Par construction de f, on vérifie facilement que pour tout fermé F de $\{0,1\}$, l'ensemble $f^{-1}(F)$ est un fermé de A. Donc f est continue et non constante. ■

Corollaire 2. *La réunion d'une famille de parties connexes d'intersection non vide est connexe.*

Preuve : Soit $(A_i)_{i \in I}$ une famille d'ensembles connexes d'intersection B non vide et f une application continue de $\bigcup_{i \in I} A_i$ dans $\{0,1\}$. Par connexité de A_i, f est constante sur chaque A_i. Si l'on note a_i la valeur de cette constante, on en déduit que f restreinte à B (qui n'est pas vide) doit être constante et égale à a_i pour tout i. En conséquence, tous les a_i sont égaux entre eux, et f est donc constante sur $\bigcup_{i \in I} A_i$. On conclut grâce au corollaire précédent. ■

Attention : La réunion d'une famille *quelconque* de connexes n'est pas connexe en général. De même, l'intersection de deux connexes n'est pas toujours connexe (dans $\mathbb{R}^2$, considérer par exemple l'intersection d'un anneau et d'un rectangle).

Corollaire 3. *Soit A une partie connexe de $(X, \mathcal{O})$. Alors toute partie B de X telle que $A \subset B \subset \overline{A}$ est connexe.*

Preuve : Soit f une application continue de B à valeurs dans $\{0,1\}$ et b un point de B. Alors f restreinte à A est constante. Supposons par exemple que f vaille 0 sur A. Comme $b \in \overline{A}$, on a

$$f(b) = \lim_{\substack{a \to b \\ a \in A}} f(a) = 0.$$

On conclut que f est nulle sur B. Donc B est connexe. ■

4.1.2 Parties connexes de $\mathbb{R}$

Théorème. *Les parties connexes de $\mathbb{R}$ sont les intervalles.*

Preuve : Considérons d'abord le cas d'un intervalle fermé borné de $\mathbb{R}$: $I = [a,b]$. Soit F_1 et F_2 deux fermés disjoints de $[a,b]$. Comme $[a,b]$ est fermé dans $\mathbb{R}$, F_1 et F_2 sont en fait deux fermés de $\mathbb{R}$, et comme ils sont bornés, ils sont compacts. Supposons par l'absurde que ces deux compacts soient non vides. Comme ils sont disjoints, on a $d(F_1, F_2) > 0$, et il existe $x_1 \in F_1$ et $x_2 \in F_2$ tels que $d(F_1, F_2) = |x_1 - x_2|$ (exercice : le prouver). Le point $(x_1 + x_2)/2$ appartient aussi à l'intervalle $[a,b]$ et doit donc appartenir à l'un des deux fermés, par exemple F_1. Mais

$$d\Big(\frac{x_1 + x_2}{2}, x_2\Big) = \frac{|x_1 - x_2|}{2} = \frac{d(F_1, F_2)}{2},$$

ce qui est absurde. Donc F_1 ou F_2 est vide et $[a,b]$ est bien connexe.

Comme tout intervalle de $\mathbb{R}$ est réunion croissante d'intervalles fermés bornés, le corollaire 2 de la page 38 permet de conclure que tout intervalle de $\mathbb{R}$ est connexe.

Si $A \subset \mathbb{R}$ n'est pas un intervalle, il existe deux points x et y de A tels que $x < y$ et $[x, y] \not\subset A$. En prenant $y_0 \in [x, y] \setminus A$, on constate que $A \cap]-\infty, y_0]$ et $A \cap [y_0, +\infty[$ sont deux fermés non vides et disjoints de A dont la réunion vaut A. Donc A n'est pas connexe. ∎

Théorème (des valeurs intermédiaires)**.** *Soit A une partie connexe d'un espace topologique $(X, \mathcal{O})$, et f une application continue de A dans $\mathbb{R}$. Alors $f(A)$ est un intervalle de $\mathbb{R}$.*

Preuve : On sait que $f(A)$ est un ensemble connexe de $\mathbb{R}$. Il ne reste plus qu'à appliquer le théorème précédent. ∎

Corollaire. *Soit f une fonction continue de A dans $\mathbb{R}$ avec A à la fois compact et connexe. Alors il existe $(a, b) \in \mathbb{R}^2$ avec $a \leq b$ tel que $f(A) = [a, b]$.*

Preuve : D'après le théorème des valeurs intermédiaires, l'ensemble $f(A)$ est un intervalle. Par ailleurs, par compacité de A, l'ensemble $f(A)$ doit être un compact de $\mathbb{R}$. En conséquence, c'est un intervalle fermé borné. ∎

4.1.3 La connexité par arcs

Définition. Soit $(X, \mathcal{O})$ un espace topologique, et A une partie non vide de X. Soit (a, b) un couple de points de A. On dit qu'une application φ définie sur $[0, 1]$ est un **chemin** de A allant de a vers b si elle est continue de $[0, 1]$ dans A, et vérifie $\varphi(0) = a$ et $\varphi(1) = b$.

Remarque. L'image de $[0, 1]$ par φ est une courbe continue d'extrémités a et b.

Exemple. Soit E un espace vectoriel muni d'une topologie pour laquelle l'addition vectorielle et la multiplication par un scalaire sont des opérations continues (on parle **d'espace vectoriel topologique**). À tout **segment** fermé $[a, b] = \{(1-t)a + tb,\ t \in [0, 1]\}$ de E, on peut associer un chemin allant de a vers b. Il suffit de considérer $\varphi : t \mapsto (1-t)a + tb$.

Définition. On dit que la partie A de X est **connexe par arcs** si pour tout couple (a, b) de points de A il existe un chemin de A allant de a vers b.

Proposition. *Toute partie connexe par arcs est connexe.*

Preuve : On va utiliser la caractérisation de la connexité donnée par le corollaire 1 de la page 38. Soit donc A connexe par arcs et $f \in \mathcal{C}(A; \{0, 1\})$. Soit a et b deux points quelconques de A. Il existe un chemin $\varphi \in \mathcal{C}([0, 1]; A)$ tel que $\varphi(0) = a$, $\varphi(1) = b$. L'application $f \circ \varphi$ est continue de $[0, 1]$ (partie connexe de $\mathbb{R}$) dans $\{0, 1\}$ et est donc constante. En particulier

$$f(a) = f \circ \varphi(0) = f \circ \varphi(1) = f(b).$$

Donc f est constante. ∎

Attention : La réciproque est, en général, fausse.

On a cependant le résultat suivant :

Théorème. *Pour les parties <u>ouvertes</u> des espaces vectoriels normés, la connexité est équivalente à la connexité par arcs.*

Preuve : Soit A une partie connexe non vide et ouverte de E. Fixons $a \in A$ et considérons l'ensemble

$$B \stackrel{\text{déf}}{=} \{b \in A \mid \text{il existe un chemin de } A \text{ joignant } a \text{ et } b\}.$$

Par définition même, l'ensemble B est connexe par arcs. Reste à montrer que $B = A$.

- B n'est pas vide car contient le point a.
- L'ensemble B est ouvert.
En effet, si $b \in B$, il existe un chemin de A allant de a vers b et, puisque A est ouvert, une boule non vide $B(b,r)$ incluse dans A. Pour tout point c de $B(b,r)$, le segment $[b,c]$ est inclus dans $B(b,r)$ donc dans A. Il existe donc un chemin de A allant de b vers c. Enfin, la réunion d'un chemin allant de a à b avec un chemin allant de b vers c est un chemin allant de a vers c. On conclut donc que $c \in B$, puis que $B(b,r) \subset B$.
- L'ensemble B est fermé.
Soit $b \in \overline{B} \cap A$. Choisissons $r > 0$ tel que $B(b,r) \subset A$. Comme $b \in \overline{B}$, l'ensemble $B(b,r) \cap B$ n'est pas vide et contient donc un point c. Par définition de l'ensemble B, il existe un chemin de A joignant a à c. Par ailleurs, le segment $[b,c]$ appartient à $B(b,r)$ donc est inclus dans A, et l'on en déduit finalement que $b \in B$.

Comme l'ensemble A est connexe, on conclut que $B = A$. ■

4.2 Un peu de convexité

Définition. Une partie A d'un e.v. E est dite **convexe** si

$$\forall (x,y) \in A \times A,\ [x,y] \subset A.$$

Remarque : Les notions de connexité et de connexité par arc sont des notions topologiques (i.e. elles ne dépendent que de la topologie choisie). Dans le cas d'un e.v.n, elles sont donc invariantes par changement de norme en une norme équivalente.
La notion de convexité est une *notion algébrique* et est donc indépendante de la topologie choisie.

Proposition. *Dans un espace vectoriel topologique, tout ensemble convexe est connexe par arc.*

Preuve : Soit A une partie convexe de E. Soit $(x,y) \in A^2$. Par convexité, l'image de l'application continue

$$\begin{cases} [0,1] & \longrightarrow & E \\ t & \longmapsto & (1-t)x + ty \end{cases}$$

est incluse dans A.

Donc A est bien connexe par arc. ■

Proposition 1. *Soit A convexe. Alors pour tout n-uplet $(x_1, \cdots, x_n)$ d'éléments de A et tout n-uplet $(\alpha_1, \cdots, \alpha_n) \in [0,1]^n$ vérifiant $\alpha_1 + \cdots + \alpha_n = 1$, on a*

$$\sum_{i=1}^{n} \alpha_i x_i \in A.$$

Preuve : On raisonne par récurrence sur n. Le cas $n = 1$ est évident. Supposons le résultat établi pour tout $(x_1, \cdots, x_n) \in E^n$ et $(\alpha_1, \cdots, \alpha_n) \in [0,1]^n$ vérifiant $\alpha_1 + \cdots + \alpha_n = 1$.
Soit $(x_1, \cdots, x_{n+1}) \in A^{n+1}$ et $(\beta_1, \cdots, \beta_{n+1}) \in [0,1]^{n+1}$ tel que $\beta_1 + \cdots + \beta_{n+1} = 1$. On peut de plus supposer que tous les coefficients β_i sont dans $]0,1[$ (sinon le résultat est contenu dans l'hypothèse de récurrence).
Soit $x = \sum_{i=1}^{n+1} \beta_i x_i$. On constate que

$$x = \beta_{n+1} x_{n+1} + (1 - \beta_{n+1}) y \quad \text{avec} \quad y \stackrel{\text{déf}}{=} \sum_{i=1}^{n} \underbrace{(1-\beta_{n+1})^{-1} \beta_i}_{\alpha_i}\, x_i.$$

Comme $\sum_{i=1}^{n} \alpha_i = 1$, l'hypothèse de récurrence assure que $y \in A$. Par convexité de A, on a donc $x \in A$ comme souhaité. ■

Nous laissons au lecteur le soin de vérifier les propriétés suivantes :

1. L'image d'une partie convexe par une application linéaire ou affine est convexe.
2. L'image réciproque d'un convexe par une application linéaire est convexe.
3. L'intersection d'une famille quelconque de convexes est convexe.
4. Soit A et B deux convexes, et $(\alpha, \beta) \in \mathbb{R}^2$. Alors $\alpha A + \beta B$ est convexe.

Définition. Soit A une partie quelconque de E. On appelle **enveloppe convexe** de A le plus petit ensemble convexe contenant A.

Proposition. *Soit A une partie de E. L'enveloppe convexe de A est l'ensemble des combinaisons linéaires finies à coefficients positifs dont la somme vaut 1, d'éléments de A.*

Preuve : Notons B l'ensemble des combinaisons linéaires finies à coefficients positifs dont la somme vaut 1, d'éléments de A. Il est clair que cet ensemble est convexe. Par ailleurs, si C est un autre ensemble convexe contenant A, il doit en particulier contenir toutes les combinaisons convexes d'éléments de A, donc B. ■

Proposition. *Dans un e.v.n, l'adhérence et l'intérieur d'un ensemble convexe sont convexes.*

Preuve : Soit A un convexe non vide, $(x, y) \in \overline{A}^2$ et $t \in [0,1]$. Il existe alors deux suites $(x_n)_{n\in\mathbb{N}}$ et $(y_n)_{n\in\mathbb{N}}$ d'éléments de A telles que

$$\lim_{n\to+\infty} x_n = x \quad \text{et} \quad \lim_{n\to+\infty} y_n = y.$$

Par convexité de A, on a $tx_n + (1-t)y_n \in A$ pour tout $n \in \mathbb{N}$. De plus il est clair que

$$\lim_{n\to+\infty} tx_n + (1-t)y_n = x + (1-t)y$$

donc $tx + (1-t)y \in \overline{A}$. Ceci montre que $\overline{A}$ est convexe.
Montrons maintenant que $\mathring{A}$ est aussi convexe. On écarte le cas $\mathring{A} = \emptyset$ qui est trivial. Soit $(x, y) \in \mathring{A}^2$. Il existe alors $r > 0$ tel que les boules ouvertes $B(x, r)$ et $B(y, r)$ soient incluses dans A. Par convexité de A, on a donc $tB(x,r) + (1-t)B(y,r) \subset A$ pour tout $t \in [0,1]$. On établit facilement que

$$tB(x,r) + (1-t)B(y,r) = B(tx + (1-t)y, r).$$

Donc $tx + (1-t)y \in \mathring{A}$, et $\mathring{A}$ est bien convexe. ■

Définition. Soit A une partie convexe de E et $f : A \to \mathbb{R}$. On dit que f est **convexe** si

$$\forall \alpha \in [0,1],\ \forall (x,y) \in E^2,\ f(\alpha x + (1-\alpha)y) \le \alpha f(x) + (1-\alpha) f(y).$$

On dit que f est **concave** si $-f$ est convexe.

Proposition. *Soit $f : A \to \mathbb{R}$ une fonction convexe. Alors pour tout réel c les ensembles $f^{-1}(]-\infty, c])$ et $f^{-1}(]-\infty, c[)$ sont convexes.*

Preuve : Supposons $f^{-1}(]-\infty, c])$ non vide et donnons nous x et y deux éléments de $f^{-1}(]-\infty, c])$ et $t \in [0,1]$. Par convexité de f, on a

$$f(tx + (1-t)y) \le tf(x) + (1-t)f(y) \le tc + (1-t)c = c,$$

d'où le résultat. La preuve de la convexité de $f^{-1}(]-\infty, c[)$ est similaire. ■

Exemple. Soit $f : E \to \mathbb{R}$ une fonction affine, et $a \in \mathbb{R}$. Alors les ensembles $f^{-1}([a, +\infty[)$, $f^{-1}(]a, +\infty[)$, $f^{-1}(] - \infty, a[)$ et $f^{-1}(] - \infty, a])$ sont convexes.

En utilisant la proposition 1 de la page 40, et en raisonnant par récurrence, on obtient le résultat suivant :

Proposition. *Si $f : A \to \mathbb{R}$ est une fonction convexe alors pour tout $(\alpha_1, \cdots, \alpha_n) \in [0,1]^n$ tel que $\sum_{i=1}^n \alpha_i = 1$, on a*

$$f\left(\sum_{i=1}^n \alpha_i x_i\right) \leq \sum_{i=1}^n \alpha_i f(x_i).$$

Théorème. *Soit E un e.v.n. de dimension finie et U un ouvert convexe de E. Alors toute fonction convexe définie sur U est continue.*

Preuve : Soit $f : U \to \mathbb{R}$ une fonction convexe. On fixe une base $(e_1, \cdots, e_n)$ de E et l'on munit E de la norme $\|x\| = \sum_{i=1} |x_i|$ pour $x = \sum_{i=1}^n x_i e_i$ (cela n'influe en rien sur les propriétés de continuité de f puisqu'en dimension finie toutes les normes sont équivalentes). Fixons un $b \in U$. Il s'agit de montrer la continuité de f en b.

1. Réduction au cas $b = 0$, $f(0) = 0$ et $\overline{B}(0,1) \subset U$.

Quitte à considérer la fonction $g : x \mapsto f(\lambda x + b) - f(b)$ avec $\lambda > 0$ tel que $\overline{B}(b, \lambda) \subset U$, on peut se ramener à étudier la continuité en 0 pour une fonction convexe définie sur un ouvert contenant la boule unité fermée $\overline{B}(0,1)$ et s'annulant en 0. On remarquera en effet que f et g sont simultanément convexes et que f est continue en b si et seulement si g est continue en 0.

2. Majoration de g sur $\overline{B}(0,1)$.

Notons a_0 l'origine, $a_i^+ = e_i$ et $a_i^- = -e_i$ pour $i = 1, \cdots, n$. On remarque que tout point $x = \sum_{i=1}^n x_i e_i$ de $\overline{B}(0,1)$ se décompose en

$$x = (1 - \|x\|)a_0 + \sum_{i=1}^n |x_i| a_i^{\varepsilon_i} \quad \text{avec} \quad \varepsilon_i = \text{“}+\text{”} \ \text{si} \ x_i \geq 0 \ \text{et} \ \varepsilon_i = \text{“}-\text{”} \ \text{sinon.}$$

Il est clair que $1 - \|x\| \in [0,1]$, $|x_i| \in [0,1]$ pour $i \in \{1, \cdots, n\}$ et $1 - \|x\| + \sum_{i=1}^n |x_i| = 1$. Par convexité de g et comme $g(a_0) = 0$, on a donc, d'après la proposition précédente,

$$g(x) \leq M\|x\| \quad \text{avec} \quad M = \max(g(a_1^+), \cdots, g(a_n^+), g(a_1^-), \cdots, g(a_n^-)). \tag{4.1}$$

3. Fin de la preuve.

En écrivant que $0 = \frac{x}{2} + \frac{(-x)}{2}$ et en utilisant $g(0) = 0$, on obtient

$$g(x) \geq -g(-x) \geq -M\|x\|.$$

On conclut que

$$\forall x \in \overline{B}(0,1), \ |g(x)| \leq M\|x\|.$$

Donc g est continue en 0, et f, en b.

■

Attention : Le résultat ci-dessus est faux en général si l'on ne suppose pas que U est ouvert ou si l'on se place en dimension infinie.

Chapitre 5

Espaces d'applications continues

5.1 Généralités

Dans tout ce chapitre, (X, d_X) et (Y, d_Y) désignent deux espaces métriques et l'on note $\mathcal{C}(X;Y)$ l'ensemble des applications continues de X vers Y. Comme d'habitude, $\mathcal{F}(X;Y)$ est l'ensemble des fonctions de X vers Y.

Considérons une suite $(f_n)_{n\in\mathbb{N}}$ de fonctions de $\mathcal{F}(X;Y)$. Nous souhaitons exprimer le fait que cette suite converge vers une fonction f de $\mathcal{F}(X;Y)$. Si l'on se réfère au cas $X = Y = \mathbb{R}$, il y a au moins deux types de convergence à considérer : la **convergence simple** (ou ponctuelle) et la **convergence uniforme**. Cela motive la définition suivante :

Définition. Soit $(f_n)_{n\in\mathbb{N}}$ une suite de fonctions de $\mathcal{F}(X;Y)$, et f une fonction de $\mathcal{F}(X;Y)$. On dit que

- $(f_n)_{n\in\mathbb{N}}$ **converge simplement** vers f si $(f_n(x))_{n\in\mathbb{N}}$ tend vers $f(x)$ pour tout $x \in X$,
- $(f_n)_{n\in\mathbb{N}}$ **converge uniformément** vers f si $\sup_{x\in X} d_Y(f_n(x), f(x))$ tend vers 0 quand n tend vers $+\infty$.

Dans le cas $X = [0,1]$ et $Y = \mathbb{R}$, il est bien connu que la convergence uniforme peut être exprimée en termes de norme : dire qu'une suite de fonctions $(f_n)_{n\in\mathbb{N}}$ définies sur $[0,1]$ converge uniformément vers f signifie que

$$\lim_{n\to+\infty} \|f_n - f\|_{L^\infty} = 0.$$

En conséquence, la norme $\|\cdot\|_{L^\infty}$ est associée à la notion de convergence uniforme. La topologie correspondante est souvent appelée *topologie de la convergence uniforme.*

Cela est en fait vrai dans un cadre bien plus général :

Proposition. *Supposons (X, d_X) compact. Alors la fonction δ définie sur $\mathcal{C}(X;Y) \times \mathcal{C}(X;Y)$ par*

$$\delta(f,g) = \sup_{x\in X} d_Y(f(x), g(x))$$

est une distance sur $\mathcal{C}(X;Y)$.

Preuve : La seule chose à vérifier est que la fonction δ est bien à valeurs finies. Soit donc (f,g) un couple de fonctions de $\mathcal{C}(X;Y)$. Sachant que X est compact, les ensembles $f(X)$ et $g(X)$ sont aussi compacts, et donc bornés. En conséquence, l'ensemble $\{d_Y(f(x), g(x)), x \in X\}$ est un borné de $\mathbb{R}^+$. Des vérifications de routine permettent alors de conclure que δ est bien une distance. ■

Théorème. *Supposons (X, d_X) compact et (Y, d_Y) complet. Alors l'ensemble $\mathcal{C}(X;Y)$ muni de la distance δ est un espace métrique complet.*

Preuve : Soit $(f_n)_{n\in\mathbb{N}}$ une suite de Cauchy de fonctions de $\mathcal{C}(X;Y)$. Il s'agit de montrer que $(f_n)_{n\in\mathbb{N}}$ converge vers un élément f de $\mathcal{C}(X;Y)$ au sens de la distance δ.

1. *Convergence simple :*

En combinant les définitions de suite de Cauchy et de δ, on voit que

$$\forall\varepsilon>0,\ \exists N\in\mathbb{N},\ \big(n\geq N\ \text{ et }\ p\geq N\big)\Longrightarrow\big(\forall x\in X,\ d_Y(f_n(x),f_p(x))\leq\varepsilon\big). \tag{5.1}$$

Cela assure en particulier que pour $x\in X$ fixé, la suite $(f_n(x))_{n\in\mathbb{N}}$ est de Cauchy dans Y. Comme Y est complet, cette suite converge donc vers un élément $f(x)$ de Y.

2. *Convergence uniforme :*

Fixons $\varepsilon>0$ et $n\geq N$ dans (5.1). En faisant tendre p vers $+\infty$, on obtient

$$\forall x\in X,\ d_Y(f_n(x),f(x))\leq\varepsilon.$$

En conséquence la suite $(f_n)_{n\in\mathbb{N}}$ converge uniformément vers f.

3. *Continuité de la fonction limite :*

Pour achever la démonstration du théorème, il ne reste plus qu'à établir que f est continue sur X. Soit donc $\varepsilon>0$ et $x_0\in X$. Pour tout $x\in X$ et $n\in\mathbb{N}$, on peut écrire :

$$d_Y(f(x),f(x_0))\leq d_Y(f(x),f_n(x))+d_Y(f_n(x),f_n(x_0))+d_Y(f_n(x_0),f(x_0)).$$

Choisissons n de telle sorte que $\sup_{x\in X}d_Y(f(x),f_n(x))\leq\frac{\varepsilon}{3}$ (c'est possible d'après l'étape précédente). On obtient

$$\forall x\in X,\ d_Y(f(x),f(x_0))\leq\frac{2}{3}\varepsilon+d_Y(f_n(x),f_n(x_0)). \tag{5.2}$$

Sachant que f_n est continue en x_0, il existe $\eta>0$ tel que

$$\forall x\in X,\ d_X(x,x_0)\leq\eta\Longrightarrow d_Y(f_n(x),f_n(x_0))\leq\frac{\varepsilon}{3}.$$

D'après (5.2), on a donc

$$\forall x\in X,\ d_X(x,x_0)\leq\eta\Longrightarrow d_Y(f(x),f(x_0))\leq\varepsilon,$$

d'où la continuité en x_0. ■

5.2 Equicontinuité

Définition. Soit $\mathcal{F}$ une partie de $\mathcal{C}(X;Y)$ et x_0 un point de X. On dit que $\mathcal{F}$ est **équicontinue** en x_0 si

$$\forall\varepsilon>0,\exists\alpha>0,\ \forall f\in\mathcal{F},\ \forall y\in X,\ d_X(x_0,y)<\alpha\Rightarrow d_Y(f(x_0),f(y))<\varepsilon.$$

On dit que $\mathcal{F}$ est équicontinue sur X si $\mathcal{F}$ est équicontinue en tout point de X.

Si l'on a la condition plus forte :

$$\forall\varepsilon>0,\exists\alpha>0,\ \forall f\in\mathcal{F},\forall(x,y)\in X^2,\ d_X(x,y)<\alpha\Rightarrow d_Y(f(x),f(y))<\varepsilon, \tag{5.3}$$

on dit que $\mathcal{F}$ est **uniformément équicontinue**.

Rappelons que sur un compact, continuité entraîne uniforme continuité (c'est le théorème de Heine). Nous disposons d'un résultat analogue pour l'équicontinuité :

Lemme 1. *Soit $\mathcal{F}$ une partie équicontinue de $\mathcal{C}(X;Y)$. Supposons (X, d_X) compact. Alors $\mathcal{F}$ est uniformément équicontinue sur X.*

Preuve : Il s'agit d'une adaptation facile de la démonstration du théorème de Heine. Les détails sont laissés en exercice. ■

Exemples.
1. Une partie $\mathcal{F}$ de $\mathcal{C}(X;Y)$ comportant un nombre fini d'éléments est toujours équicontinue. Si de plus les éléments de $\mathcal{F}$ sont uniformément continus alors $\mathcal{F}$ est uniformément équicontinue.
2. La réunion de deux parties équicontinues de $\mathcal{C}(X;Y)$ est une partie équicontinue de $\mathcal{C}(X;Y)$.
3. Fixons $M \in \mathbb{R}^+$. L'ensemble $\mathcal{F}$ des fonctions $f : X \longrightarrow Y$ lipschitziennes de rapport M est uniformément équicontinu sur X : à ε fixé, il suffit de prendre $\alpha = \varepsilon/M$ dans (5.3).
4. On prend $X = [0,1]$ et $Y = \mathbb{R}$ munis tous les deux de la distance $d(x,y) = |x-y|$. Soit
$$\mathcal{F} = \Big\{ f \in \mathcal{C}^1([0,1];\mathbb{R}) \;/\; \int_0^1 |f'(t)|^2\, dt \leq 1 \Big\}.$$
Alors $\mathcal{F}$ est équicontinue sur $[0,1]$.
En effet, si x et y sont deux points de $[0,1]$ tels que $x \leq y$, on a d'après l'inégalité de Cauchy-Schwarz,
$$|f(y) - f(x)| \leq \sqrt{y-x}\,\sqrt{\int_x^y |f'(t)|^2\, dt} \leq \sqrt{y-x}.$$
5. Prenons maintenant
$$\mathcal{F} = \{ f_n : x \mapsto e^{-nx},\, n \in \mathbb{N} \}.$$
Alors $\mathcal{F}$ est équicontinue en tout point de $]0,1]$ mais pas en 0. Le défaut d'équicontinuité en 0 est dû au fait que la suite $(f_n'(0))_{n\in\mathbb{N}}$ n'est pas bornée.
6. Considérons enfin
$$\mathcal{F} = \{ f_n : x \mapsto \sin(nx),\, n \in \mathbb{N} \}.$$
Alors $\mathcal{F}$ n'est équicontinue en aucun point de $[0,1]$.

En général la propriété de convergence uniforme est strictement plus forte que celle de convergence simple. Cependant, si l'on se restreint à des suites de fonctions appartenant à une partie *uniformément équicontinue* de $\mathcal{C}(X;Y)$, on a le résultat remarquable suivant :

Lemme 2. *Supposons (X, d_X) compact. Soit $\mathcal{F}$ une partie uniformément équicontinue de $\mathcal{C}(X;Y)$ et $(f_n)_{n\in\mathbb{N}}$ une suite de fonctions de $\mathcal{F}$. On a l'équivalence suivante :*
$$(f_n)_{n\in\mathbb{N}} \text{ converge simplement vers } f \iff (f_n)_{n\in\mathbb{N}} \text{ converge uniformément vers } f.$$

Preuve : Il suffit de justifier l'implication directe. Soit ε un réel strictement positif arbitraire. Par hypothèse, il existe un réel α strictement positif tel que pour tout $n \in \mathbb{N}$, on ait
$$d_X(x,x') < \alpha \implies d_Y(f_n(x), f_n(x')) < \frac{\varepsilon}{3}.$$
Par passage à la limite, on constate que l'inégalité ci-dessus est aussi vérifiée par f. (En particulier f est donc uniformément continue.) On recouvre ensuite le compact X par une famille finie de boules $(B_X(x_j,\alpha))_{1\leq j\leq N_\varepsilon}$. Soit $x \in X$ et x_j tel que $x \in B_X(x_j,\alpha)$. On a
$$\begin{aligned} d_Y(f_n(x), f(x)) &\leq d_Y(f_n(x), f_n(x_j)) + d_Y(f_n(x_j), f(x_j)) + d_Y(f(x_j), f(x)), \\ &\leq \frac{2}{3}\varepsilon + \max_{1\leq i\leq N_\varepsilon} d_Y(f_n(x_i), f(x_i)). \end{aligned}$$

En vertu de l'hypothèse de convergence simple, il existe n_0 tel que

$$\forall n \geq n_0\,,\ \max_{1\leq i\leq N_\varepsilon} d_Y(f_n(x_i), f(x_i)) < \frac{\varepsilon}{3}.$$

D'où le lemme. ■

5.3 Le théorème d'Ascoli

Il s'agit d'un résultat fondamental d'analyse fonctionnelle qui a des applications dans de nombreux domaines des mathématiques (comme nous le verrons dans la suite du cours ainsi que dans le module équations aux dérivées partielles).

Pour le démontrer, nous ferons appel aux lemmes 1 et 2, ainsi qu'au résultat suivant :

Lemme 3. *Tout espace métrique compact admet une partie dénombrable dense*[1].

Preuve : Pour tout $p \in \mathbb{N}$, l'ensemble X peut être recouvert par un nombre fini N_p de boules $B_X(x_i^p, 2^{-p})$. Par construction, l'ensemble $\{x_i^p \,/\, p \in \mathbb{N},\ 1 \leq i \leq N_p\}$ est dénombrable et dense dans X. ■

Nous pouvons maintenant énoncer le théorème d'Ascoli :

Théorème (d'Ascoli)**.** *Soit (X, d_X) un espace métrique compact, (Y, d_Y) un espace métrique complet et $\mathcal{F}$ une partie de $\mathcal{C}(X;Y)$. On suppose que*

(i) $\mathcal{F}$ est équicontinue sur X,

(ii) pour tout $x \in X$ l'ensemble $\{f(x) \,/\, f \in \mathcal{F}\}$ est relativement compact dans Y.

Alors $\mathcal{F}$ est relativement compacte dans $\mathcal{C}(X;Y)$.

Preuve : Nous allons donner deux démonstrations du théorème d'Ascoli.

La première démonstration consiste à établir que toute suite $(f_n)_{n\in\mathbb{N}}$ d'éléments de $\mathcal{F}$ admet une sous-suite convergente dans $\mathcal{C}(X;Y)$. Pour construire cette sous-suite, on part d'une suite $(x_p)_{p\in\mathbb{N}}$ dense dans X (dont l'existence est assurée par le lemme 3). Par hypothèse, l'ensemble

$$\{f(x_0) \,|\, f \in \mathcal{F}\}$$

est d'adhérence compacte. Donc il existe une fonction φ_0 strictement croissante de $\mathbb{N}$ dans $\mathbb{N}$ et un élément $g(x_0)$ de Y tels que

$$\lim_{n\to\infty} f_{\varphi_0(n)}(x_0) = g(x_0).$$

De même, il existe un point $g(x_1)$ de Y et une fonction φ_1 strictement croissante de $\mathbb{N}$ dans $\mathbb{N}$ telle que

$$\lim_{n\to\infty} f_{\varphi_0\circ\varphi_1(n)}(x_1) = g(x_1).$$

On définit ainsi par récurrence une sous-suite $(\varphi_p)_{p\in\mathbb{N}}$ de fonctions strictement croissantes de $\mathbb{N}$ dans $\mathbb{N}$ telle que

$$\forall p \in \mathbb{N}\,,\ \forall k \leq p\,,\ \lim_{n\to\infty} f_{\varphi_0\circ\cdots\circ\varphi_p(n)}(x_k) = g(x_k).$$

On souhaite maintenant obtenir une sous-suite de $(f_n)_{n\in\mathbb{N}}$ qui converge simplement pour tout point x_p. Cela peut se faire à l'aide du procédé diagonal de Cantor : on définit la fonction ψ de $\mathbb{N}$ dans $\mathbb{N}$ par

$$\psi(n) = \varphi_0 \circ \cdots \circ \varphi_n(n).$$

1. Un ensemble topologique jouissant de cette propriété est dit **séparable**.

C'est clairement une fonction strictement croissante de $\mathbb{N}$ dans $\mathbb{N}$ qui, par construction, vérifie

$$\forall p \in \mathbb{N}\,,\ \lim_{n\to\infty} f_{\psi(n)}(x_p) = g(x_p). \tag{5.4}$$

L'ensemble X étant compact, le lemme 1 assure que $\mathcal{F}$ est uniformément équicontinue. Démontrons maintenant que la fonction g est uniformément continue sur $\{x_p \mid p \in \mathbb{N}\}$. Soit ε un réel strictement positif et α vérifiant (5.3). Pour tout couple d'entiers (p,q) tel que $d_X(x_p,x_q) < \alpha$, on a

$$\begin{aligned} d_Y(g(x_p),g(x_q)) &\leq d_Y(g(x_p),f_{\psi(n)}(x_p)) + d_Y(f_{\psi(n)}(x_p),f_{\psi(n)}(x_q)) + d_Y(f_{\psi(n)}(x_q),g(x_q)),\\ &\leq \varepsilon + d_Y(g(x_p),f_{\psi(n)}(x_p)) + d_Y(f_{\psi(n)}(x_q),g(x_q)). \end{aligned}$$

L'inégalité ci-dessus étant vraie pour tout n, on obtient, en passant à la limite $n \to \infty$,

$$d_X(x_p,x_q) < \alpha \Rightarrow d_Y(g(x_p),g(x_q)) \leq \varepsilon.$$

La fonction g est donc uniformément continue sur $\{x_p \mid p \in \mathbb{N}\}$ qui est une partie dense de X. Comme l'espace métrique (Y,d_Y) est complet, le théorème de la page 22 garantit que g peut être prolongée de manière unique en une application uniformément continue définie sur X entier.

Notons encore g ce prolongement. Pour conclure la preuve du théorème d'Ascoli, il suffit de démontrer que

$$\forall x \in X\,,\ \lim_{n\to\infty} f_{\psi(n)}(x) = g(x).$$

En effet, $\mathcal{F}$ est une partie uniformément équicontinue. En vertu du lemme 2, il suffit donc d'établir la convergence simple. Soit ε un réel strictement positif arbitraire et x un élément de X. Il existe un réel $\alpha > 0$ tel que

$$\forall n \in \mathbb{N}\,,\ d_X(x,x') < \alpha \Rightarrow \sup_{n\in\mathbb{N}} d_Y(f_{\psi(n)}(x),f_{\psi(n)}(x')) + d_Y(g(x),g(x')) < \frac{\varepsilon}{2}.$$

Il existe un entier p tel que $d_X(x,x_p) < \alpha$. Il en résulte que

$$\begin{aligned} d_Y(f_{\psi(n)}(x),g(x)) &\leq d_Y(f_{\psi(n)}(x),f_{\psi(n)}(x_p)) + d_Y(f_{\psi(n)}(x_p),g(x_p)) + d_Y(g(x_p),g(x))\\ &\leq \frac{\varepsilon}{2} + d_Y(f_{\psi(n)}(x_p),g(x_p)). \end{aligned}$$

La relation (5.4) permet de conclure à la convergence simple. Ceci achève la preuve du théorème d'Ascoli.

Donnons maintenant une autre démonstration du théorème d'Ascoli. reposant sur la caractérisation des ensembles relativement compacts des espaces métriques complets donnée par le corollaire 1 de la page 21. Comme l'ensemble $\mathcal{C}(X;Y)$ muni de la distance δ est bien complet, il suffit finalement de montrer que pour tout $\varepsilon > 0$ l'ensemble $\mathcal{F}$ peut être recouvert par un nombre fini de boules $B_{\mathcal{C}(\mathcal{X};\mathcal{Y})}(f_i,\varepsilon)$ avec $f_i \in \mathcal{F}$.

Fixons donc un $\varepsilon > 0$. Soit $\eta > 0$ tel que

$$\forall f \in \mathcal{F}, \forall (x,y) \in X^2, d_X(x,y) < \eta \Longrightarrow d_Y(f(x),f(y)) < \varepsilon/3. \tag{5.5}$$

Comme X est compact, il existe une famille finie $(x_1,\cdots,x_n)$ de points de X telle que

$$X \subset \bigcup_{i=1}^{n} B_X(x_i,\eta).$$

Par hypothèse, les n ensembles $\{f(x_i) \,/\, f \in \mathcal{F}\}$ sont relativement compacts dans Y, donc l'ensemble

$$\{(f(x_1), \cdots, f(x_n)) \,/\, f \in \mathcal{F}\}$$

est relativement compact dans Y^n muni de la distance produit.
L'espace Y^n est complet car produit d'espaces métriques complets. Donc en vertu du corollaire 1 de la page 21, on peut le recouvrir par un nombre fini p de boules de type $B_{Y^n}\big((f_j(x_1), \cdots, f_j(x_n)), \frac{\varepsilon}{3}\big)$ avec $f_j \in \mathcal{F}$. Montrons que

$$\mathcal{F} \subset \bigcup_{j=1}^{p} B_{\mathcal{C}(\mathcal{X};\mathcal{Y})}(f_j, \varepsilon),$$

ou, de manière équivalente, que pour tout $f \in \mathcal{F}$, il existe $j \in \{1, \cdots, p\}$ tel que

$$\forall x \in X, \ d_Y(f(x), f_j(x)) < \varepsilon.$$

Soit donc $f \in \mathcal{F}$ et $x \in X$. Choisissons $i \in \{1, \cdots, n\}$ tel que $d_X(x, x_i) < \eta$ et $j \in \{1, \cdots, p\}$ tel que $(f(x_1), \cdots, f(x_n)) \in B_{Y^n}\big((f_j(x_1), \cdots, f_j(x_n)), \frac{\varepsilon}{3}\big)$.
Écrivons que

$$d_Y(f(x), f_j(x)) \leq d_Y(f(x), f(x_i)) + d_Y(f(x_i), f_j(x_i)) + d_Y(f_j(x_i), f_j(x)).$$

Le choix de j assure que le deuxième terme du membre de droite est strictement plus petit que $\varepsilon/3$. Il en est de même des deux autres termes d'après (5.5). ■

Corollaire. *Soit K un compact de $\mathbb{R}^n$. Munissons l'ensemble $\mathcal{C}(K;\mathbb{R})$ de la norme*

$$\|f\|_\infty = \sup_{x \in K} |f(x)|.$$

Soit $\mathcal{F}$ une partie équicontinue de $\mathcal{C}(K;\mathbb{R})$ telle que l'ensemble $\{f(x) \,|\, f \in \mathcal{F}\}$ soit borné pour tout $x \in K$. Alors $\mathcal{F}$ est relativement compacte dans $\mathcal{C}(K;\mathbb{R})$.

Exemple. Soit

$$\mathcal{F} = \Big\{ f \in \mathcal{C}^1([0,1];\mathbb{R}) \,/\, \int_0^1 |f(t)|^2 \, dt + \int_0^1 |f'(t)|^2 \, dt \leq 1 \Big\}.$$

On a déjà vu dans un exemple précédent que $\mathcal{F}$ était une partie uniformément équicontinue de $\mathcal{C}([0,1];\mathbb{R})$. Par ailleurs, si $f \in \mathcal{F}$ alors la fonction $|f|$ qui est aussi continue sur $[0,1]$ atteint son minimum en un point $x_0 \in [0,1]$. D'après l'inégalité de Cauchy-Schwarz, on a pour tout $x \in [0,1]$,

$$\big(|f(x)| - |f(x_0)|\big)^2 \leq |f(x) - f(x_0)|^2 \leq \left(\int_{x_0}^x |f'(t)| \, dt\right)^2 \leq |x - x_0| \left|\int_{x_0}^x |f'(t)|^2 \, dt\right| \leq \int_0^1 |f'(t)|^2 \, dt,$$

et, clairement,

$$|f(x_0)|^2 \leq \int_0^1 |f(t)|^2 \, dt.$$

En conséquence, pour tout $x \in [0,1]$, on a

$$|f(x)| = |f(x_0)| + |f(x)| - |f(x_0)| \leq \sqrt{\int_0^1 |f(t)|^2 \, dt} + \sqrt{\int_0^1 |f'(t)|^2 \, dt} \leq 2.$$

Donc toutes les hypothèses du corollaire précédent sont vérifiées par $\mathcal{F}$. En conséquence, $\mathcal{F}$ est une partie relativement compacte de $\mathcal{C}([0,1];\mathbb{R})$.

Autrement dit de toute suite $(f_n)_{n\in\mathbb{N}}$ de fonctions de $\mathcal{C}^1([0,1];\mathbb{R})$ telles que

$$\int_0^1 |f_n(t)|^2\,dt + \int_0^1 |f_n'(t)|^2\,dt \le 1,$$

on peut extraire une sous-suite convergente dans $\mathcal{C}([0,1];\mathbb{R})$.

5.4 Le théorème de Stone-Weierstrass (hors-programme)

Dans cette partie, on cherche à établir un critère simple permettant de déterminer si une partie $\mathcal{F}$ de $\mathcal{C}(X;\mathbb{K})$ (avec $\mathbb{K}=\mathbb{R}$ ou $\mathbb{C}$) est dense.

On suppose dans un premier temps que $\mathbb{K}=\mathbb{R}$.

Lemme 4. *Pour tout $a>0$ il existe une suite de fonctions polynômes qui converge uniformément vers $f : x\mapsto |x|$ sur $[-a,a]$.*

Preuve : Pour $y\in[0,1]$, posons $g(y)=\sqrt{1+y}$. D'après la formule de Taylor-Lagrange, pour tout $y\in[0,1]$ et $n\in\mathbb{N}$, il existe $\theta\in]0,1[$ tel que

$$g(y)=\sum_{k=0}^{n}\frac{y^k}{k!}g^{(k)}(0)+\frac{y^{n+1}}{(n+1)!}g^{(n+1)}(\theta y).$$

La somme P_n apparaissant au membre de droite est un polynôme de degré n. Un calcul facile montre que pour tout $z\in[0,1]$,

$$g^{(k)}(z)=(1+z)^{\frac12-k}\prod_{j=0}^{k}\Big(\frac12-j\Big).$$

Donc, pour tout $y\in[0,1]$, la valeur absolue du terme de reste dans le développement de Taylor-Lagrange est majorée par

$$u_n\stackrel{\text{déf}}{=}\frac12\prod_{j=1}^{n}\Big(1+\frac{1}{2j}\Big).$$

On constate que $u_{n+1}/u_n=1-\frac{1}{2n}+O(\frac{1}{n^2})$. En conséquence, la suite $(u_n)_{n\in\mathbb{N}}$ converge vers 0 (on peut montrer que u_n est un $O(n^{-\frac12})$). Cela assure la convergence uniforme de la suite de polynômes $(P_n)_{n\in\mathbb{N}}$ vers g sur $[0,1]$.

En remarquant que $f(x)=|x|=g(x^2-1)$, et en posant $Q_n(x)=P_n(x^2-1)$, on en déduit que $(Q_n)_{n\in\mathbb{N}}$ converge uniformément vers f sur $[-1,1]$.

Le cas général $a\neq1$ s'en déduit par dilatation : on vérifie que la suite de polynômes $(aQ_n(a^{-1}\cdot))_{n\in\mathbb{N}}$ converge vers f sur $[-a,a]$. ■

Théorème (de Stone-Weierstrass)**.** *Soit (X,d) un espace métrique compact et $\mathcal{F}$ une partie de $\mathcal{C}(X;\mathbb{R})$. On suppose que $\mathcal{F}$ est stable par combinaison linéaire et multiplication*[a]*, contient toutes les fonctions constantes sur X, et* **sépare les points**[b]. *Alors $\mathcal{F}$ est dense dans $\mathcal{C}(X;\mathbb{R})$.*

a. Autrement dit, $\mathcal{F}$ est une **sous-algèbre** de $\mathcal{C}(X;\mathbb{R})$

b. ce qui signifie que pour tout $(x_1,x_2)\in X^2$ avec $x_1\neq x_2$, et $(y_1,y_2)\in\mathbb{R}^2$, il existe une fonction f de $\mathcal{F}$ telle que $f(x_1)=y_1$ et $f(x_2)=y_2$.

Preuve : Nous allons en fait démontrer que $\overline{\mathcal{F}}$ est *dense* dans $\mathcal{C}(X;\mathbb{R})$. Comme $\overline{\mathcal{F}}$ est fermé, cela entraînera que $\overline{\mathcal{F}} = \mathcal{C}(X;\mathbb{R})$. Nous laissons au lecteur le soin de vérifier que $\overline{\mathcal{F}}$ est stable par combinaison linéaire et multiplication, contient toutes les fonctions constantes et sépare les points.

1. Stabilité de $\overline{\mathcal{F}}$ par passage à la valeur absolue.
   Soit $f \in \overline{\mathcal{F}}$ et $M > 0$ une borne de f sur X (dont l'existence est garantie par la compacité de X). Le lemme 5 assure l'existence d'une suite de polynômes $(P_n)_{n\in\mathbb{N}}$ qui converge uniformément vers $x \mapsto |x|$ sur $[-M, M]$. Posons $f_n = P_n(f)$. Comme $\overline{\mathcal{F}}$ contient les fonctions constantes, et est stable par addition et multiplication, il est clair que $f_n \in \overline{\mathcal{F}}$. Par ailleurs, on constate aisément que $(f_n)_{n\in\mathbb{N}}$ converge uniformément vers $|f|$ sur X.
2. Stabilité par passage au sup ou à l'inf .
   Soit f et g deux fonctions de $\overline{\mathcal{F}}$. En remarquant que
   $$\inf(f,g) = \frac{1}{2}\bigl(f+g-|f-g|\bigr) \quad \text{et} \quad \sup(f,g) = \frac{1}{2}\bigl(f+g+|f-g|\bigr),$$
   et en utilisant la stabilité de $\overline{\mathcal{F}}$ par combinaison linéaire et passage à la valeur absolue, on en déduit que
   $$\inf(f,g) \in \overline{\mathcal{F}} \quad \text{et} \quad \sup(f,g) \in \overline{\mathcal{F}}. \tag{5.6}$$
3. Démonstration proprement dite.
   Fixons une fonction f de $\mathcal{C}(X;\mathbb{R})$ et $\varepsilon > 0$. Il s'agit de montrer l'existence d'une fonction g de $\overline{\mathcal{F}}$ telle que $\sup_{x\in X} |f(x) - g(x)| \leq \varepsilon$.
   Pour commencer, fixons un point x de X. Pour tout $y \in X \setminus \{x\}$, donnons-nous une fonction f_y de $\mathcal{F}$ telle que $f_y(x) = f(x)$ et $f_y(y) = f(y)$. L'ensemble
   $$O_y \stackrel{\text{déf}}{=} \bigl\{x' \in X \,/\, f_y(x') > f(x') - \varepsilon\bigr\}$$
   est un ouvert de X contenant x et y donc $(O_y)_{y\neq x}$ est un recouvrement de X par des ouverts. Par compacité, il existe donc $y_1, \cdots, y_p$ tels que
   $$X = \bigcup_{i=1}^{p} O_{y_i}.$$
   Posons alors $g_x \stackrel{\text{déf}}{=} \sup(f_{y_1}, \cdots, f_{y_p})$. D'après (5.6), la fonction g_x est dans $\overline{\mathcal{F}}$ et il est facile de voir que $g_x(x) = f(x)$ et
   $$\forall x' \in X,\ g_x(x') > f(x') - \varepsilon.$$
   On considère maintenant les ensembles
   $$\Omega_x \stackrel{\text{déf}}{=} \bigl\{x' \in X \,/\, g_x(x') < f(x') + \varepsilon\bigr\}.$$
   Comme précédemment, ces ensembles sont des ouverts non vides dont la réunion recouvre X. On peut donc trouver un nombre fini de points $x_1, \cdots, x_q$ tels que
   $$X = \bigcup_{i=1}^{q} \Omega_{x_i}.$$
   La fonction $g \stackrel{\text{déf}}{=} \inf(g_{x_1}, \cdots, g_{x_q})$ est dans $\overline{\mathcal{F}}$ et vérifie
   $$\forall x' \in X,\ f(x') - \varepsilon < g(x') < f(x') + \varepsilon,$$
   donc répond à la question.

■

Exemples. 1. Soit Lip l'ensemble des fonctions lipschitziennes de l'espace métrique compact (X, d) dans $\mathbb{R}$. Il est clair que l'ensemble Lip est stable par combinaison linéaire et multiplication, et contient toutes les fonctions constantes. Il est de plus séparant. En effet, soit $(x_1, x_2) \in X^2$ tel que $x_1 \neq x_2$, et $(\lambda_1, \lambda_2) \in \mathbb{R}^2$. Alors la fonction

$$f : x \longmapsto \lambda_1 + \frac{d(x_1, x)}{d(x_1, x_2)}(\lambda_2 - \lambda_1)$$

est lipschitzienne, égale à λ_1 en x_1, et à λ_2 en x_2.
Le théorème de Stone-Weierstrass assure donc que l'ensemble des fonctions lipschitziennes sur X est dense dans $\mathcal{C}(X; \mathbb{R})$.

2. Pour tout $a < b$, l'ensemble des fonctions polynômes de $[a, b]$ dans $\mathbb{R}$ est dense dans $\mathcal{C}([a, b]; \mathbb{R})$. En effet, cet ensemble contient les constantes, sépare les points, et est bien sûr stable par combinaison linéaire et multiplication.
3. À titre d'exercice, nous laissons au lecteur le soin de vérifier que l'ensemble des polynômes trigonométriques du type
$$\sum_{k=0}^{n} a_k \cos(kx) + b_k \sin(kx)$$
est dense dans l'ensemble des fonctions continues 2π-périodiques de $\mathbb{R}$ dans $\mathbb{R}$.

Le théorème de Stone-Weierstrass se généralise aux fonctions à valeurs complexes comme suit :

Théorème. *Soit (X, d) un espace métrique compact et $\mathcal{F}$ une partie de $\mathcal{C}(X; \mathbb{C})$. On suppose que $\mathcal{F}$ est* stable par conjugaison [2]*, combinaison linéaire et multiplication, contient les fonctions constantes et sépare les points. Alors $\mathcal{F}$ est dense dans $\mathcal{C}(X; \mathbb{C})$.*

Preuve : Notons $\mathcal{F}_{\mathcal{R}e} = \big\{\mathcal{R}e\, f \,/\, f \in \mathcal{F}\big\}$ et $\mathcal{F}_{\mathcal{I}m} = \big\{\mathcal{I}m\, f \,/\, f \in \mathcal{F}\big\}$.
La stabilité par conjugaison et le fait que
$$\mathcal{R}e\, f = \frac{f + \overline{f}}{2} \quad \text{et} \quad \mathcal{I}m\, f = \frac{f - \overline{f}}{2i}$$
assure que ces deux ensembles vérifient les hypothèses du théorème de Stone-Weirstrass pour les fonctions à valeurs réelles. En conséquence, si $f \in \mathcal{F}(X; \mathbb{C})$ et $\varepsilon > 0$, il existe $g \in \mathcal{F}_{\mathcal{R}e}$ et $h \in \mathcal{F}_{\mathcal{I}m}$ tels que
$$\delta(\mathcal{R}e\, f, g) \leq \varepsilon/2 \;\text{ et }\; \delta(\mathcal{I}m\, f, h) \leq \varepsilon/2.$$
Donc $\delta(f, g + ih) \leq \varepsilon$. ■

Remarque : Comme le montre l'exemple suivant, l'hypothèse de stabilité par conjugaison est indispensable. Considérons en effet l'espace métrique $X = \big\{z \in \mathbb{C} \,/\, |z| = 1\big\}$ et l'ensemble $\mathcal{F}$ constitué par la restriction des polynômes à coefficients complexes, à l'ensemble X. Alors $\mathcal{F}$ est une sous-algèbre de $\mathcal{C}(X; \mathbb{C})$ qui contient les constantes et sépare les points. Cependant, la fonction continue $\varphi : z \mapsto 1/z$ n'est pas dans l'adhérence de $\mathcal{F}$ car pour tout $P \in \mathcal{F}$, on a
$$\int_0^{2\pi} e^{i\theta} P(e^{i\theta})\, d\theta = 0$$
alors que
$$\int_0^{2\pi} e^{i\theta} \varphi(e^{i\theta})\, d\theta = 2\pi.$$

2. c'est-à-dire que $f \in \mathcal{F}$ entraîne $\overline{f} \in \mathcal{F}$

Chapitre 6

Les grands théorèmes de l'analyse fonctionnelle

6.1 Le théorème de Baire

On peut donner deux énoncés équivalents du théorème de Baire.

Théorème (de Baire)**.** *Soit* (X, d) *un espace métrique complet et* $(F_n)_{n\in\mathbb{N}}$ *une suite de fermés de* X*, d'intérieur vide. Alors* $\bigcup_{n\in\mathbb{N}} F_n$ *est d'intérieur vide*[a].

[a] mais pas forcément fermé !

Théorème (de Baire)**.** *Soit* (X, d) *un espace métrique complet et* $(\Omega_n)_{n\in\mathbb{N}}$ *une suite d'ouverts denses de* X*. Alors* $\bigcap_{n\in\mathbb{N}} \Omega_n$ *est dense*[a] *dans* X*.*

[a] mais pas forcément ouvert !

Pour voir que les deux énoncés sont équivalents, il suffit de passer au complémentaire et d'utiliser le fait que pour toute partie A d'un espace topologique $(W, \mathcal{O})$, l'adhérence de $W \setminus A$ est égale au complémentaire de $\mathring{A}$. Donc un ensemble est dense si et seulement si son complémentaire est d'intérieur vide.

Donnons maintenant une démonstration du deuxième énoncé du théorème de Baire. Soit $(\Omega_n)_{n\in\mathbb{N}}$ une suite d'ouverts denses de X et V un ouvert non vide de X. Il s'agit de montrer que $V \cap (\bigcap_{n\in\mathbb{N}} \Omega_n)$ n'est pas vide.

Étape 0. Par densité de Ω_0, l'ouvert $\Omega_0 \cap V$ est non vide. Donc il existe $x_0 \in X$ et $r_0 > 0$ tels que $\overline{B}(x_0, r_0) \subset \Omega_0 \cap V$.

Étape 1. De même, par densité de Ω_1, il existe $x_1 \in X$ et $r_1 > 0$ tels que

$$\overline{B}(x_1, r_1) \subset \Omega_1 \cap B(x_0, r_0) \subset \Omega_0 \cap \Omega_1 \cap V.$$

Quitte à diminuer r_1, on peut toujours supposer que $r_1 \leq r_0/2$.

Étape n. Supposons construits une famille $(x_0, x_1, \cdots, x_n)$ de points de X, et une famille $(r_0, r_1, \cdots, r_n)$ de réels strictement positifs tels que pour $j = 1, \cdots, n$ on ait

$$\overline{B}(x_j, r_j) \subset B(x_{j-1}, r_{j-1}) \cap \Omega_j \quad \text{et} \quad r_j \leq r_{j-1}/2.$$

Comme $\Omega_{n+1} \cap B(x_n, r_n)$ n'est pas vide, on peut choisir un $x_{n+1} \in X$ et un réel r_{n+1} de $]0, r_n/2]$ tels que

$$\overline{B}(x_{n+1}, r_{n+1}) \subset \Omega_{n+1} \cap B(x_n, r_n).$$

On a donc construit une suite $(x_n)_{n\in\mathbb{N}}$ de points de X, et une suite $(r_n)_{n\in\mathbb{N}}$ de réels strictement positifs vérfiant

$$\forall n \in \mathbb{N}, \overline{B}(x_n, r_n) \subset \Omega_0 \cap \cdots \cap \Omega_n \cap V, \quad \overline{B}(x_{n+1}, r_{n+1}) \subset \overline{B}(x_n, r_n) \quad \text{et} \quad r_{n+1} \leq r_n/2.$$

Il est clair que la suite $(x_n)_{n\in\mathbb{N}}$ construite est de Cauchy. Comme X est complet, elle converge vers un élément x de X qui, par construction, appartient à toutes les boules fermées $\overline{B}(x_n, r_n)$ donc à $V \cap (\bigcap_{n\in\mathbb{N}} \Omega_n)$. ■

En relation avec le théorème de Baire, nous pouvons introduire la définition suivante :

Définition. On dit qu'un espace topologique séparé $(X, \mathcal{O})$ a la **propriété de Baire** si pour toute suite $(\Omega_n)_{n\in\mathbb{N}}$ d'ouverts denses de X, l'ensemble $\bigcap_{n\in\mathbb{N}} \Omega_n$ est dense.

Remarque. Le théorème de Baire assure que tout espace de Banach a la propriété de Baire. Mais on peut construire des e.v.n. (forcément non complets !) qui n'ont pas la propriété de Baire. C'est le cas par exemple de $E = \mathcal{C}([0,1]; \mathbb{R})$ muni de la norme $\|f\|_{L^1} = \int_0^1 |f(t)|\, dt$. En effet, si l'on définit

$$F_n = \left\{ f \in E \mid \sup_{t\in[0,1]} |f(t)| \leq n \right\},$$

on a clairement $\bigcup_{n\in\mathbb{N}} F_n = E$ bien que chaque F_n soit fermé d'intérieur vide. En effet, si $f \in F_n$ alors on constate que pour tout $j \in \mathbb{N}$, la fonction $f + f_j$ avec f_j continue affine par morceaux sur $[0,1]$, nulle en 0 et sur $[2^{-j}, 1]$ et égale à $2(n+1)$ en 2^{-j-1} appartient à $E \setminus F_n$ (car $f + f_j$ vaut au moins $n+2$ en 2^{-j-1}). Cependant $\|f_j\|_{L^1} = 2^{-j}(n+1)$ donc $f + f_j$ tend vers f dans E quand j tend vers l'infini. En conséquence, il n'existe pas de voisinage de f qui soit inclus dans F_n.

Remarque. Sachant que $\mathbb{R}$ est un espace métrique complet, le théorème de Baire permet de retrouver le fait que $\mathbb{R}$ n'est pas dénombrable (indication : considérer l'intersection de tous les ouverts $\mathbb{R} \setminus \{a\}$ avec $a \in \mathbb{R}$).

Proposition. *Soit X un espace topologique séparé ayant la propriété de Baire et Ω un ouvert de X. Alors Ω muni de la topologie induite par X a aussi la propriété de Baire.*

Preuve : Soit $(\Omega_n)_{n\in\mathbb{N}}$ une suite d'ouverts denses de Ω. Comme Ω est ouvert, on vérifie facilement que $\big(\Omega_n \cup (X \setminus \overline{\Omega})\big)_{n\in\mathbb{N}}$ est une suite d'ouverts denses de X. On est donc assuré que $\bigcap_{n\in\mathbb{N}} \big(\Omega_n \cup (X \setminus \overline{\Omega})\big)$ est dense dans X.
Maintenant si V est un ouvert arbitraire de Ω, c'est un ouvert de X tel que $V \cap (X \setminus \overline{\Omega}) = \emptyset$. Donc

$$V \cap \Big(\bigcap_{n\in\mathbb{N}} \big(\Omega_n \cup (X \setminus \overline{\Omega})\big)\Big) = V \cap \Big(\big(\bigcap_{n\in\mathbb{N}} \Omega_n\big) \cup \big(X \cap \overline{\Omega}\big)\Big) = V \cap \big(\bigcap_{n\in\mathbb{N}} \Omega_n\big)$$

est non vide, d'où le résultat. ■

Théorème. *Soit X un espace topologique de Baire et $(F_n)_{n\in\mathbb{N}}$ une suite de fermés de X telle que $\bigcup_{n\in\mathbb{N}} F_n = X$. Alors $\bigcup_{n\in\mathbb{N}} \mathring{F}_n$ est dense dans X.*

Preuve : Soit Ω un ouvert non vide de X. Comme X est un espace de Baire, Ω aussi. Chaque ensemble $F_n \cap \Omega$ est un fermé de Ω pour la topologie induite. En raisonnant par l'absurde, on en déduit facilement que $\bigcup_{n\in\mathbb{N}} \big(\mathring{F}_n \cap \Omega\big)$ n'est pas vide. ■

Corollaire. *Soit Ω un ouvert non vide et non borné de $\mathbb{R}^+$. Alors il existe $T > 0$ tel que*

$$\{n \in \mathbb{N} \mid nT \in \Omega\}$$

ait une infinité d'éléments.

Preuve : On veut trouver un $T>0$ tel que

$$T \in \bigcap_{n\geq 1}\Big(\bigcup_{p\geq n}\frac{1}{p}\,\Omega\Big).$$

On vérifie que pour tout $n\in\mathbb{N}^*$, l'ensemble $\Omega_n=\bigcup_{p\geq n}\frac{1}{p}\,\Omega$ est un ouvert dense de $\mathbb{R}^+$. Comme $\mathbb{R}^+$ est un espace métrique complet, l'ensemble $\cap_{n\geq1}\Omega_n$ est donc dense dans $\mathbb{R}^+$. En particulier, il contient un élément $T>0$. ∎

6.2 Le théorème de Banach-Steinhaus

Théorème (de Banach-Steinhaus)**.** *Soit E un Banach, F un e.v.n. et $(T_i)_{i\in I}$ une famille d'applications de $\mathcal{L}(E;F)$. On suppose que pour tout $x\in E$, l'ensemble $\{T_i(x)\mid i\in I\}$ est borné dans F.*
Alors il existe $M\geq 0$ tel que $\forall i\in I,\ \|T_i\|_{\mathcal{L}(E;F)}\leq M$.

Preuve : Pour tout $n\in\mathbb{N}$, on définit l'ensemble $A_n \overset{\text{déf}}{=} \{x\in E\,|\,\forall i\in I, \|T_i(x)\|_F\leq n\}$. L'ensemble A_n est est un fermé de E (car intersection de tous les fermés $T_i^{-1}(\overline{B}_F(0,n))$) et l'on a $\bigcup_{n\in\mathbb{N}} A_n=E$. Comme E est complet et, manifestement, n'est pas d'intérieur vide, le théorème de Baire garantit l'existence d'un $n_0\in\mathbb{N}$ tel que A_{n_0} ne soit pas d'intérieur vide. Il existe donc $x_0\in E$ et $r_0>0$ tels que $\overline{B}(x_0,r_0)\subset A_{n_0}$. Autrement dit,

$$\forall i\in I, \forall y\in \overline{B}_E(0,1),\ \|T_i(x_0+r_0y)\|_F\leq n_0.$$

En conséquence, en vertu de la linéarité des T_i,

$$\forall i\in I,\ \forall y\in \overline{B}_E(0,1),\ \|T_i(x_0)+r_0T_i(y)\|_F\leq n_0$$

d'où, grâce à la deuxième inégalité triangulaire,

$$\sup_{y\in\overline{B}_E(0,1)}\|T_i(y)\|_F\leq r_0^{-1}\big(n_0+\|T_i(x_0)\|_F\big).$$

On a donc bien le résultat voulu : il suffit de choisir $M=r_0^{-1}\big(n_0+\|T_i(x_0)\|_F\big)$. ∎

Corollaire. *Soit E un Banach, F un e.v.n, et $(T_n)_{n\in\mathbb{N}}$ une suite de $\mathcal{L}(E;F)$. On suppose que pour tout $x\in E$, la suite $(T_n(x))_{n\in\mathbb{N}}$ converge dans F vers une limite notée $T(x)$. Alors $T\in\mathcal{L}(E;F)$, la suite $(T_n)_{n\in\mathbb{N}}$ est bornée et*

$$\|T\|_{\mathcal{L}(E;F)}\leq \liminf \|T_n\|_{\mathcal{L}(E;F)}. \tag{6.1}$$

Preuve : Tout d'abord, les opérations linéaires étant des opérations continues, l'application T est linéaire. Ensuite, sachant que toute suite convergente est bornée, le théorème de Banach-Steinhaus assure que la suite $(T_n)_{n\in\mathbb{N}}$ est bornée : il existe un $M\geq 0$ tel que

$$\forall n\in\mathbb{N}, \forall x\in E,\ \|T_n(x)\|_F\leq M\|x\|_E.$$

Du fait que pour tout $x\in E$ on a

$$\forall n\in\mathbb{N},\ \|T_n(x)\|_F\leq \|T_n\|_{\mathcal{L}(E;F)}\|x\|_E,$$

on en déduit que

$$\forall x\in E,\ \|T(x)\|_F\leq \|x\|_E \liminf \|T_n\|_{\mathcal{L}(E;F)}.$$

Le théorème de continuité des applications linéaires permet maintenant de conclure que T est continue et vérifie (6.1). ∎

Attention : Pour appliquer le théorème de Banach-Steinhaus, il est essentiel que **l'espace de départ soit complet**.

En effet, considérons $E = \mathcal{C}^1([0,1];\mathbb{R})$ et $F = \mathcal{C}([\frac{1}{4},\frac{3}{4}];\mathbb{R})$. On munit E et F de la norme de la convergence uniforme. Pour tout $h \in]0,\frac{1}{4}[$ et $f \in E$ on note $T_h(f)$ la fonction définie sur $[\frac{1}{4},\frac{3}{4}]$ par $[T_h(f)](x) = h^{-1}(f(x+h)-f(x))$.

On constate alors que chaque T_h est linéaire continue de E sur F, et que

$$\|T_h(f)\|_F \leq \|f'\|_{L^\infty} \quad \text{et} \quad \lim_{h\to 0} T_h(f) = f'.$$

Autrement dit, en notant $T : f \mapsto f'$, on remarque que $(T_{2^{-n}}(f))_{n\geq 2}$ converge vers $T(f)$ pour tout $f \in F$. Il est par ailleurs facile de montrer que $T \notin \mathcal{L}(E;F)$, ce qui semble contredire le corollaire ci-dessus.

En fait, toutes les hypothèses du corollaire sont vérifiées sauf une : la complétude de E.

6.3 Le théorème de l'application ouverte

Définition. Soit $(X,\mathcal{O})$, $(X',\mathcal{O}')$ deux espaces topologiques. On dit que $f : X \to X'$ est une **application ouverte** si l'image de tout ouvert de X par f est un ouvert de X'.

Attention : Si f est une fonction continue de X sur X' alors l'image réciproque de tout ouvert de X' est un ouvert de X. On ne peut rien dire a priori sur la nature topologique de l'image (directe) d'un ouvert de X par f.

Pour s'en persuader, on peut considérer le cas d'une application constante de $\mathbb{R}$ dans $\mathbb{R}$. Pour tout ouvert non vide $\Omega \subset \mathbb{R}$, $f(\Omega)$ est un singleton, et n'est donc pas ouvert (pour la topologie usuelle de $\mathbb{R}$), bien que f soit continue.

Théorème (de l'application ouverte). *Soit E et F deux Banach, et $T \in \mathcal{L}(E;F)$ surjective. Alors T est une application ouverte.*

Preuve : **Première étape :** on cherche un $c > 0$ tel que $B_F(0,4c) \subset \overline{T(B_E(0,1))}$.

Soit $X_n \stackrel{\text{déf}}{=} T(B_E(0,n))$. Comme T est surjectif, on a $F = \bigcup_{n\in\mathbb{N}^*} X_n$ et donc a fortiori $F = \bigcup_{n\in\mathbb{N}^*} \overline{X}_n$. L'espace F étant complet, le théorème de Baire assure l'existence d'un entier $n_0 \in \mathbb{N}^*$ tel que l'intérieur de $\overline{X}_{n_0}$ ne soit pas vide.

Soit donc $y_0 \in F$ et $r_0 > 0$ tels que

$$B_F(y_0,r_0) \subset \overline{T(B_E(0,n_0))}.$$

Par linéarité de T, on vérifie que

$$B_F(n_0^{-1}y_0, n_0^{-1}r_0) \cup B_F(-n_0^{-1}y_0, n_0^{-1}r_0) \subset \overline{T(B_E(0,1))}.$$

La linéarité de T assure aussi que $\overline{T(B_E(0,1))}$ est convexe.

On en déduit que

$$B_F(0,n_0^{-1}r_0) = \frac{1}{2}B_F(n_0^{-1}y_0, n_0^{-1}r_0) + \frac{1}{2}B_F(-n_0^{-1}y_0, n_0^{-1}r_0) \subset \overline{T(B_E(0,1))}.$$

On peut donc prendre $c = r_0/(4n_0)$.

Deuxième étape : on montre que $B_F(0,c) \subset T(B_E(0,1))$.

Soit $y \in B_F(0,c)$. On construit par récurrence une suite $(x_n)_{n\in\mathbb{N}^*} \in E^{\mathbb{N}}$ telle que

$$\forall n \in \mathbb{N}^*,\ \|x_n\|_E < 2^{-(n+1)} \quad \text{et} \quad \|y - T(x_1+\cdots+x_n)\|_F < 2^{-n}c.$$

Cela est possible car l'étape précédente implique que $B_F(0,c) \subset \overline{T(B_E(0,1/4))}$. Il existe donc $x_1 \in E$ tel que $\|x_1\|_E < 1/4$ et $\|y - T(x_1)\|_F < c/2$.

Ensuite, si $x_1, \cdots, x_n$ ont été construits, on remarque que

$$\|2^n y - T(2^n x_1 + \cdots + 2^n x_n)\|_F < c$$

et l'on obtient comme précédemment un $x_{n+1} \in E$ tel que $\|2^n x_{n+1}\|_E < 1/4$ et $\|2^n y - T(2^n x_1 + \cdots + 2^n x_n) - T(2^n x_{n+1})\|_F < c/2$.

Par complétude de E, comme la série $\sum x_n$ est absolument convergente, la série $\sum_{n=1}^{\infty} x_n$ converge vers un élément x de E vérifiant $\|x\|_E \leq 1/2 < 1$. De plus, par continuité de T, on a $y = T(x)$.

Donc $B_F(0,c) \subset T(\overline{B}_E(0,\frac{1}{2})) \subset T(B_E(0,1))$.

Troisième étape : Conclusion.

On veut montrer que $T(\Omega)$ est ouvert dans F, pour tout Ω ouvert de E. On peut toujours supposer que Ω n'est pas vide. Soit donc $y \in T(\Omega)$ et $x \in \Omega$ tel que $y = T(x)$. Comme x est intérieur à Ω, il existe $r > 0$ tel que $B_E(x,r) \subset \Omega$. En utilisant la deuxième étape et la linéarité de T, on en déduit que

$$B_F(y, cr) = T(x) + rB_F(0,c) \subset T(x) + rT(B_E(0,1)) = T(B_E(x,r)) \subset T(\Omega).$$

En conséquence y est bien un point intérieur à $T(\Omega)$, et $T(\Omega)$ est ouvert. ■

Remarque. La complétude de F est utilisée dans la première étape alors que celle de E intervient dans la deuxième étape.

Corollaire 1. *Soit E et F deux Banach, et $T \in \mathcal{L}(E;F)$ bijective. Alors T^{-1} est continue.*

Preuve : Soit Ω un ouvert de E. Comme T est bijective, on a $(T^{-1})^{-1}(\Omega) = T(\Omega)$. Mais $T(\Omega)$ est ouvert en vertu du théorème précédent. On peut donc maintenant conclure que l'image réciproque de tout ouvert de E par T^{-1} est ouvert. Donc T^{-1} est continue. ■

Corollaire 2. *Soit E un e.v.n. que l'on munit de deux normes $\|\cdot\|_1$ et $\|\cdot\|_2$. On suppose que $(E, \|\cdot\|_1)$ et $(E, \|\cdot\|_2)$ sont complets et qu'il existe $c > 0$ tel que*

$$\forall x \in E, \ \|x\|_1 \leq c\|x\|_2.$$

Alors les normes $\|\cdot\|_1$ et $\|\cdot\|_2$ sont équivalentes.

Preuve : Considérons l'application identité Id de $(E, \|\cdot\|_2)$ dans $(E, \|\cdot\|_1)$. Vues les hypothèses, le corollaire 1 assure que Id^{-1} est continue. Il existe donc $c' > 0$ telle que

$$\forall x \in E, \ \|x\|_2 \leq c'\|x\|_1,$$

d'où l'équivalence des normes. ■

6.4 Le théorème du graphe fermé

Définition. Soit E, F deux e.v.n. et T une application linéaire de E vers F. On appelle **graphe** de T (noté $G(T)$) le sous-ensemble de $E \times F$ suivant :

$$G(T) = \{(x,y) \in E \times F \mid y = T(x)\}.$$

Il est facile de vérifier que T continue entraîne que $G(T)$ est fermé (au sens de la topologie produit de $E \times F$). Sous de bonnes hypothèses, il y a en fait équivalence entre la continuité et la fermeture du graphe :

Théorème (du graphe fermé)**.** *Soit E et F, deux Banach, et T une application linéaire de E vers F. Alors T est continue si et seulement si $G(T)$ est un fermé de $E \times F$.*

Preuve : Seule l'implication réciproque n'est pas triviale. Supposons donc que $G(T)$ est fermé dans $E \times F$. Comme E et F sont complets, on vérifie aisément que $E \times F$ muni de la norme $\|(x,y)\|_{E\times F} = \max(\|x\|_E, \|y\|_F)$ est aussi complet puis que l'e.v.n. $G(T)$ muni de la norme induite est complet (c'est ici qu'intervient l'hypothèse que $G(T)$ est fermé).

Définissons

$$P \ : \ \begin{cases} G(T) & \longrightarrow & E \\ (x,y) & \longmapsto & x. \end{cases}$$

L'application P est bijective par définition de $G(T)$. De plus, pour tout $(x,y) \in G(T)$, on a

$$\|P(x,y)\|_E = \|x\|_E \leq \|(x,y)\|_{E\times F}.$$

Donc P est continue. Le corollaire du théorème de l'application ouverte assure donc que P^{-1} aussi est continue. Autrement dit, il existe $C > 0$ tel que

$$\forall x \in E,\ \|T(x)\|_F \leq \max(\|x\|_E, \|T(x)\|_F) \leq C\|x\|_E,$$

d'où le résultat. ∎

Remarque. Soit E et F des Banach et T linéaire de E vers F. Pour montrer la continuité de T, il suffit donc d'établir que pour toute suite $(x_n)_{n\in\mathbb{N}}$ de E telle que $(x_n, T(x_n))_{n\in\mathbb{N}}$ converge vers (x,y), on a $y = T(x)$.

Exercice : Soit T un endomorphisme sur $L^2([0,1])$. On suppose T continu de $(L^2([0,1]); \|\cdot\|_{L^2})$ dans $(L^2([0,1]); \|\cdot\|_{L^1})$. À l'aide du théorème du graphe fermé, montrer que T est aussi continu de $(L^2([0,1]); \|\cdot\|_{L^2})$ dans $(L^2([0,1]); \|\cdot\|_{L^2})$.

6.5 Autour du théorème de Hahn-Banach (hors-programme)

6.5.1 Théorème de Hahn-Banach, forme analytique

Au chapitre 3, nous avons vu que toute application linéaire continue définie sur un sous-espace dense F d'un e.v.n. E et à valeurs dans un Banach admettait un unique prolongement continu à E tout entier, et que ce prolongement avait de plus la même norme.

Dans cette section, nous cherchons à établir que, sans hypothèse particulière, toute *forme linéaire continue* sur F admet un prolongement continu[1] sur E.

Nous verrons que ce résultat est l'une des multiples conséquences du **théorème de Hahn-Banach** que nous énonçons ci-dessous.

1. pas nécessairement unique

Théorème (Hahn-Banach, version analytique). *Soit E un espace vectoriel sur $\mathbb{R}$ et p une fonction de E dans $\mathbb{R}^+$ vérifiant :*

(i) $\forall \lambda \geq 0,\ \forall x \in E,\ p(\lambda x) = \lambda p(x),$

(ii) $\forall (x,y) \in \mathbb{R}^2,\ p(x+y) \leq p(x) + p(y).$

Soit F un s.e.v. de E et L une forme linéaire sur F telle que

$$\forall x \in F,\ L(x) \leq p(x).$$

Alors il existe une forme linéaire $\widetilde{L}$ définie sur E, qui prolonge L (i.e. $\widetilde{L}(x) = L(x)$ pour tout $x \in F$) et telle que

$$\forall x \in E,\ \widetilde{L}(x) \leq p(x).$$

La démonstration de ce théorème repose sur un résultat célèbre de la théorie des ensembles, le lemme de Zorn qui est lui-même équivalent à l'axiome du choix.

Avant d'énoncer le lemme de Zorn, nous avons besoin de rappeler quelques notions provenant de la théorie des ensembles.

Définition. Soit $\mathcal{E}$ un ensemble. On dit que $\leq$ est une **relation d'ordre** sur $\mathcal{E}$ si

(i) $\forall (x,y,z) \in \mathcal{E}^3,\ \big(x \leq y \text{ et } y \leq z\big) \Rightarrow x \leq z,$

(ii) $\forall (x,y) \in \mathcal{E}^2,\ \big(x \leq y \text{ et } y \leq x\big) \Rightarrow x = y,$

(iii) $\forall x \in \mathcal{E},\ x \leq x.$

On dit que l'ordre est **total** si de plus pour tout $(x,y) \in \mathcal{E}^2$, on a $x \leq y$ ou $y \leq x$.

Un couple $(\mathcal{E}, \leq)$ avec $\leq$ relation d'ordre sur E est appelé **ensemble ordonné**.

Exemples.

1. L'ensemble $\mathbb{N}$ muni de la relation d'ordre usuel est totalement ordonné.
2. Si A est un ensemble ayant au moins deux éléments, l'ensemble $\mathcal{F}(A;\mathbb{R})$ muni de la relation d'ordre
$$f \leq g \quad \text{si} \quad f(x) \leq g(x) \quad \text{pour tout} \quad x \in A$$
n'est pas totalement ordonné.

Définition. Soit $(\mathcal{E}, \leq)$ un ensemble ordonné et X une partie de $\mathcal{E}$. On dit que x est un **majorant** de X si

$$\forall y \in X,\ y \leq x.$$

On dit que x est un élément **maximal** de X si

$$\forall y \in \mathcal{E},\ x \leq y \Longrightarrow y = x.$$

Définition. On dit qu'un ensemble ordonné $(\mathcal{E}, \leq)$ est **inductif** si toute partie totalement ordonnée de $\mathcal{E}$ admet un majorant.

Lemme (de Zorn). *Tout ensemble non vide inductif admet un élément maximal.*

Preuve du théorème de Hahn-Banach :

On note $\mathcal{E}$ l'ensemble des couples (V, u) tels que V soit un sous-espace vectoriel de E contenant F, et u une forme linéaire sur V qui coïncide avec L sur F et vérifie $u(x) \leq p(x)$ pour tout $x \in V$. On munit $\mathcal{E}$ de la relation d'ordre $\leq$ définie par

$$(V_1, u_1) \leq (V_2, u_2) \quad \text{si} \quad V_1 \subset V_2 \quad \text{et} \quad u_2 = u_1 \quad \text{sur} \quad V_1.$$

L'ensemble $(\mathcal{E}, \leq)$ est ordonné par construction. De plus, il est non vide car contient L et inductif car si $\{(V_i, u_i), i \in I\}$ est une partie totalement ordonnée de $\mathcal{E}$ alors $\cup_{i\in I} V_i$ est un sous-espace vectoriel de E contenant F et l'endomorphisme u définit sur $\cup_{i\in I} V_i$ par $u(x) = u_i(x)$ si $x \in V_i$ est bien un majorant de la partie $\{(V_i, u_i), i \in I\}$

D'après le lemme de Zorn, $\mathcal{E}$ admet donc un élément maximal $(V, \widetilde{L})$. Supposons par l'absurde que $V \neq E$. Alors $E \setminus V$ contient au moins un élément x non nul. Pour a réel donné, on définit alors une forme linéaire $\widetilde{L}_a$ sur $V \oplus \mathbb{R}x$ par

$$\widetilde{L}_a(y) = \widetilde{L}(y) \quad \text{si} \quad y \in V, \quad \text{et} \quad \widetilde{L}_a(x) = a.$$

Grâce aux propriétés de $\widetilde{L}$ et de p, on peut ajuster a de telle sorte que

$$\forall y \in V,\, \forall \lambda \in \mathbb{R},\ \widetilde{L}_a(y + \lambda x) = \widetilde{L}(y) + \lambda a \leq p(y + \lambda x). \tag{6.2}$$

En effet, (6.2) est clairement vérifiée pour tout $y \in V$ si $\lambda = 0$. Par ailleurs, l'inégalité (6.2) restreinte aux $\lambda > 0$ est équivalente à

$$\forall y \in V,\ \widetilde{L}(y) + a \leq p(y + x)$$

alors que pour $\lambda < 0$, elle est équivalente à

$$\forall z \in V,\ \widetilde{L}(z) - a \leq p(z - x).$$

Finalement (6.2) est donc vérifiée si et seulement si a est choisi de telle sorte que

$$\sup_{z\in V}\Big(\widetilde{L}(z) - p(z - x)\Big) \leq a \leq \inf_{y\in V}\Big(p(y + x) - \widetilde{L}(y)\Big).$$

Comme pour tout $(y, z) \in V^2$, on a $\widetilde{L}(y) + \widetilde{L}(z) \leq p(y + z) \leq p(y + x) + p(z - x)$, un tel choix de a est possible.

On a donc $(V, \widetilde{L}) \leq (V \oplus \mathbb{R}x, \widetilde{L}_a)$. Comme $x \notin V$, cela contredit la maximalité de $(V, \widetilde{L})$. ■

Avant de d'énoncer un corollaire fondamental du théorème de Hahn-Banach, définissons le concept de semi-norme.

Définition. Soit E un e.v. sur $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$. On dit qu'une application $p : E \to \mathbb{R}^+$ est une **semi-norme** si

(*i*) $\forall \lambda \in \mathbb{K},\ \forall x \in E,\ p(\lambda x) = |\lambda| p(x)$,

(*ii*) $\forall (x, y) \in E^2,\ p(x + y) \leq p(x) + p(y)$.

Exemples. (*i*) Une norme est toujours une semi-norme.

(*ii*) Soit $E = \mathcal{C}^1([0, 1]; \mathbb{R})$. Pour $f \in E$, posons

$$p(f) = \sup_{t\in[0,1]} |f'(t)|.$$

Alors p est une semi-norme mais pas une norme car $p(f) = 0$ n'entraîne pas $f = 0$ (en fait dans cet exemple, $p(f) = 0$ ssi f est constante).

Corollaire 1. *Sous les hypothèses du théorème de Hahn-Banach, si de plus p est une semi-norme alors on peut trouver une forme linéaire $\widetilde{L}$ qui prolonge L sur E et telle que*

$$\forall x \in E,\ |\widetilde{L}(x)| \leq p(x).$$

Preuve : Notons $\widetilde{L}$ le prolongement fourni par le théorème de Hahn-Banach. Comme $\widetilde{L}(-x) = -\widetilde{L}(x)$ pour tout $x \in E$ et comme $p(x) = p(-x)$, on obtient clairement $|\widetilde{L}(x)| \leq p(x)$. ■

Corollaire 2. *Le corollaire précédent reste valable pour les espaces vectoriels sur $\mathbb{C}$ à condition de supposer de plus que $|L(x)| \leq p(x)$ pour tout $x \in F$ et d'interpréter $|\cdot|$ comme le module.*

Preuve : On considère E comme un $\mathbb{R}$-espace vectoriel. Il est alors clair que l'application linéaire $\mathcal{R}e\, L$ vérifie les hypothèses du corollaire 1. Soit ϕ une forme linéaire de E sur $\mathbb{R}$ qui prolonge $\mathcal{R}e\, L$ sur E entier et vérifie

$$\forall x \in E,\ |\phi(x)| \leq p(x).$$

Pour tout $x \in E$, on pose alors $\widetilde{L}(x) = \phi(x) - i\phi(ix)$. Il est immédiat que L est une forme linéaire de E sur $\mathbb{C}$ et que $\widetilde{L}$ coïncide avec L sur F. Enfin, si l'on note $\widetilde{L}(x) = |\widetilde{L}(x)|e^{i\theta}$, on a $e^{-i\theta}\widetilde{L}(x) \in \mathbb{R}$ donc

$$|\widetilde{L}(x)| = \widetilde{L}(e^{-i\theta}x) = \phi(e^{-i\theta}x) \leq p(e^{-i\theta}x) = p(x).$$

■

Corollaire 3. *Soit F un s.e.v. de E et L une forme linéaire continue sur F. Alors on peut prolonger L en une forme linéaire $\widetilde{L}$ continue sur E et de même norme que L.*

Preuve : On applique l'un des deux corollaires précédents avec $p(x) = \|L\|_{\mathcal{L}(F;\mathbb{K})}\|x\|_E$. ■

Définition. On appelle **dual topologique** de E l'ensemble des formes linéaires continues sur E. On note E' le dual topologique.

Corollaire 4. *Soit E un e.v.n. et $x_0 \in E$. Alors il existe $f \in E'$ telle que*

$$\|f\|_{E'} = \|x_0\|_E \quad \textit{et} \quad f(x_0) = \|x_0\|_E^2.$$

Preuve : Après avoir écarté le cas $x_0 = 0$ qui est trivial, on applique le corollaire précédent à $F = \mathbb{K}x_0$ et L définie sur F par $L(\lambda x_0) = \lambda\|x_0\|_E^2$ pour tout $\lambda \in \mathbb{K}$. ■

On en déduit que pour tout $x \in E$ non nul, on peut trouver une forme linéaire f continue sur E telle que $f(x) \neq 0$. On a en fait bien mieux :

Corollaire 5. *Soit E un e.v.n. Pour tout $x \in E$, on a*

$$\|x\|_E = \sup_{L \in E',\, \|L\|_{E'}=1} |L(x)|,$$

et le sup *est atteint.*

Preuve : Le résultat est trivial si $x = 0$. Supposons donc que $x \neq 0$. Soit $L \in E'$ de norme 1. Alors, par définition de la norme sur E', on a $|L(x)| \leq \|x\|_E$ pour tout $x \in E$. Donc $\sup_{L \in E',\, \|L\|_{E'}=1} |L(x)| \leq \|x\|_E$.

Réciproquement, le corollaire 4 assure l'existence de $L \in E'$ telle que $\|L\|_{E'} = \|x\|_E$ et $|L(x)| = \|x\|_E^2$. La forme linéaire $\widetilde{L} \overset{\text{déf}}{=} \|x\|^{-1}L$ est continue, de norme 1 et telle que $\widetilde{L}(x) = \|x\|_E$. D'où le résultat. ■

6.5.2 Le théorème de Hahn-Banach, forme géométrique

Dans toute cette section, les espaces vectoriels considérés sont réels.

Définition. On dit que $M \subset E$ est un **sous-espace affine** de E s'il existe $x_0 \in E$ et un s.e.v. M_0 de E tels que $M = x_0 + M_0$. On dit alors que M est le sous-espace affine passant par x_0 et dirigé par M_0.

Définition. On dit que H est un **hyperplan** de E s'il existe une forme linéaire L non nulle et un réel α tels que

$$H = \{x \in E \mid L(x) = \alpha\}.$$

Remarque. Un hyperplan est donc un espace affine dirigé par le noyau d'une forme linéaire. Si de plus cette forme linéaire est continue alors l'hyperplan est nécessairement fermé.

En fait, on a un résultat plus précis :

Lemme 1. *Tout hyperplan de E est ou bien dense dans E ou bien fermé dans E.*

Preuve : Soit $H = \{x \in E \,|\, \phi(x) = \alpha\}$ avec $\alpha \in \mathbb{R}$ et $\phi \in L(E; \mathbb{R})$ non nulle. Supposons que H ne soit pas dense dans E. Il existe alors $x_0 \in E$ et $r_0 > 0$ tels que $\overline{B}(x_0, r_0) \subset E \setminus H$. On a donc

$$\forall x \in \overline{B}(0,1),\ \phi(x) \neq \frac{\alpha - \phi(x_0)}{r_0}.$$

Comme $\overline{B}(0,1)$ est convexe et ϕ, linéaire, l'ensemble $\phi(\overline{B}(0,1))$ est un convexe de $\mathbb{R}$. On en déduit que $\phi - \left(\frac{\alpha - \phi(x_0)}{r_0}\right)$ garde un signe constant sur $\overline{B}(0,1)$ puis que ϕ est bornée sur la boule unité, donc continue. En conséquence, H est fermé. ∎

Définition. Soit A et B deux sous-ensembles de E et H un hyperplan de E. On dit que H **sépare** A et B s'il existe une forme linéaire L sur E et un réel α tels que

$$\forall x \in A,\ L(x) \leq \alpha \quad \text{et} \quad \forall x \in B,\ L(x) \geq \alpha.$$

On dit que H sépare A et B strictement si les inégalités ci-dessus sont strictes.

Théorème (Hahn-Banach, version géométrique)**.** *Soit A et B deux parties convexes disjointes de E avec A fermé et B compact. Alors il existe un hyperplan fermé séparant strictement A et B. Plus précisément, il existe $\alpha \in \mathbb{R}$ et $L \in E'$ tels que*

$$\forall x \in A,\ L(x) < \alpha \quad \textit{et} \quad \forall x \in B,\ L(x) > \alpha.$$

La preuve de ce théorème repose sur les deux résultats intermédiaires suivants :

1 Si A est une partie convexe ouverte non vide de E et M est un sous-espace affine de E, disjoint de A alors il existe un hyperplan fermé contenant M et disjoint de A.

2 Si A et B sont convexes disjoints avec A ouvert alors il existe un hyperplan fermé séparant A et B.

Preuve du premier résultat intermédiaire

Supposons pour simplifier que $0 \in A$. On introduit alors la **fonction de Minkowski** p définie pour tout $x \in E$ par

$$p(x) = \inf\{\lambda > 0 \,|\, \lambda^{-1} x \in A\}.$$

En utilisant le fait que A est ouvert et convexe, on montre que p vérifie les hypothèses du théorème de Hahn-Banach et que $A = p^{-1}([0,1[)$.

Soit $x_0 \in E$ et M_0 un s.e.v. de E tels que $M = x_0 + M_0$. Comme $0 \in A$ et $M \cap A = \emptyset$, on sait que $x_0 \notin M_0$. Le s.e.v. V engendré par x_0 et M_0 est donc $V = M_0 \oplus \mathbb{R}x_0$. On définit alors la forme linéaire L sur V par $L(x) = \lambda$ pour $x = \lambda x_0 + y$ avec $\lambda \in \mathbb{R}$ et $y \in M_0$.

On vérifie facilement que $L \leq p$ sur V. En effet, si $\lambda \leq 0$ et $y \in M_0$, il est évident que $L(\lambda x_0 + y) \leq p(\lambda x_0 + y)$ puisque le membre de gauche est négatif alors que le membre de droite est positif. Si $\lambda > 0$, on a $p(\lambda x_0 + y) = \lambda p(x_0 + \lambda^{-1}y) \geq \lambda = L(\lambda x_0 + y)$ puisque $x_0 + \lambda^{-1}y \notin A$.

Le théorème de Hahn-Banach nous permet donc de prolonger L en $\widetilde{L} \in E'$ telle que $\widetilde{L} \leq p$ sur E. Il est clair que l'hyperplan $H = \widetilde{L}^{-1}(\{1\})$ contient M et est disjoint de A. Enfin, cet hyperplan est fermé. En effet, s'il ne l'était pas, il serait dense d'après le lemme 1 et donc intersecterait l'ouvert A.

Le cas général où A ne contient pas 0 peut se ramener au cas précédent par translation. Les détails sont laissés au lecteur.

Remarque. De ce premier résultat, on déduit facilement que si $0 \in M$ alors il existe $L \in E'$ telle que $L = 0$ sur M et $L > 0$ sur A.

Preuve du deuxième résultat intermédiaire

Sous les hypothèses A ouvert convexe, B convexe et $A \cap B = \emptyset$, l'ensemble $C \stackrel{\text{déf}}{=} A - B$ est ouvert, convexe, non vide et ne contient pas 0.

Grâce à la remarque précédente (appliquée avec $M = \{0\}$ et $A = C$), il existe $L \in E'$ telle que $L > 0$ sur C. Autrement dit,

$$\forall (a,b) \in A \times B,\ L(a) > L(b).$$

On pose $\alpha = \inf_{a \in A} L(a)$. L'hyperplan $L^{-1}(\{\alpha\})$ répond alors à la question.

Conclusion

Soit A et B vérifiant les hypothèses du théorème de Hahn-Banach (forme géométrique). Pour tout $\varepsilon > 0$, on pose

$$A_\varepsilon \stackrel{\text{déf}}{=} A + B(0,\varepsilon), \qquad B_\varepsilon \stackrel{\text{déf}}{=} B + B(0,\varepsilon).$$

Les ensembles A_ε et B_ε sont ouverts, convexes, et non vides. De plus, les hypothèses A compact, B fermé et $A \cap B = \emptyset$ assurent que $d(A,B) > 0$. Donc A_ε et B_ε sont disjoints pour ε suffisamment petit. Fixons un tel ε.

Soit $\alpha \in \mathbb{R}$ et $L \in E'$ tels que l'hyperplan $H \stackrel{\text{déf}}{=} L^{-1}(\{\alpha\})$ sépare A_ε et B_ε. On a donc

$$\forall (x,y) \in A \times B, \forall (z,z') \in B(0,1)^2,\ L(x + \varepsilon z) \leq \alpha \leq L(y + \varepsilon z').$$

On en déduit alors facilement que

$$\forall (x,y) \in A \times B,\ L(x) + \varepsilon \|L\|_{E'} \leq \alpha \leq L(y) - \varepsilon \|L\|_{E'},$$

d'où le résultat.

Corollaire 1. *Soit F un sous-espace vectoriel de E non dense. Alors F est inclus dans un hyperplan fermé de E.*

Preuve : Soit $x_0 \in E \setminus \overline{F}$. On applique le théorème de Hahn-Banach avec $A = \overline{F}$ et $B = \{x_0\}$. Il existe donc $\alpha \in \mathbb{R}$ et $L \in E'$ tels que

$$\forall x \in \overline{F},\ L(x) < \alpha < L(x_0).$$

On a donc $L(\lambda x) < \alpha$ pour tout $x \in F$ et $\lambda \in \mathbb{R}$. Cela entraîne la nullité de L sur F. Par ailleurs $L \neq 0$ puisque $L(x_0) > L(x)$ pour tout $x \in F$. Donc F est inclus dans l'hyperplan fermé $\operatorname{Ker} L$. ■

Donnons un dernier corollaire fort utile.

Corollaire 2. *Soit F un s.e.v. de E. Alors F est dense si et seulement si toute forme linéaire $L \in E'$ s'annulant sur F s'annule aussi sur E tout entier.*

Preuve : L'implication directe est évidente car une application continue s'annulant sur une partie dense de E est forcément nulle.

Montrons la réciproque par contraposition en supposant que F est un s.e.v. non dense de E. Alors le corollaire précédent assure l'existence d'une forme linéaire continue L non nulle telle que $F \subset \operatorname{Ker} L$, d'où le résultat. ■

6.5.3 Supplémentaires topologiques

Dans toute cette section, E est un Banach.

Proposition. *Soit F et G deux sous-espaces vectoriels fermés de E tels que $F+G$ soit aussi fermé. Alors il existe $C \in \mathbb{R}^+$ tel que tout élément z de $F+G$ puisse se décomposer en $z = x+y$ avec*

$$\|x\| \leq C\|z\| \quad \textit{et} \quad \|y\| \leq C\|z\|.$$

Preuve : On munit $F \times G$ de la norme $\|(x,y)\|_{F\times G} = \|x\|_E + \|y\|_E$ et $F+G$, de la norme induite par E. L'application

$$T \ : \left\{ \begin{array}{lcl} F \times G & \longrightarrow & F+G \\ (x,y) & \longrightarrow & x+y \end{array} \right.$$

est linéaire, continue et surjective.

Comme $F \times G$ et $F+G$ sont fermés, le théorème de l'application ouverte s'applique. Il existe donc $c > 0$ tel que

$$B_{F+G}(0,c) \subset T\big(B_{F\times G}(0,1)\big).$$

Autrement dit, pour tout $z \in F+G$ tel que $\|z\|_E < c$, il existe $(x,y) \in F \times G$ tel que $\|x\|_E + \|y\|_E < 1$ et $z = x+y$.

Pour $z \in F+G$ non nul arbitraire, on pose $z' = \frac{c}{2\|z\|_E} z$. On a $z' \in B_{F+G}(0,c)$. Donc il existe $(x,y) \in F\times G$ tel que $\|x\|_E + \|y\|_E < 1$ et $\widetilde{z} = x+y$. En multipliant par $2c^{-1}\|z\|_E$, on obtient le résultat voulu. ■

Définition. Soit F un s.e.v. fermé de E. On dit que F admet un **supplémentaire topologique** G s'il existe un s.e.v. fermé G de E tel que $E = F \oplus G$.

Définition. Soit F et G deux s.e.v. de E tels que $E = F \oplus G$. On appelle **projecteur** sur F parallèlement à G l'endomorphisme p sur E défini par

$$\forall x \in E,\ \big(x = y+z,\ y \in F,\ z \in G\big) \Longrightarrow p(x) = y.$$

Grâce au théorème ci-dessus, on en déduit le

Corollaire. *Soit F un s.e.v. fermé de E admettant un supplémentaire topologique G. Alors le projecteur p sur F parallèlement à G est continu, ainsi que le projecteur q sur G parallèlement à F.*

Exemples. 1. Tout s.e.v. de dimension finie admet un supplémentaire topologique.

En effet, si F est de dimension finie, on sait déjà que F est fermé. On fixe alors $(e_1, \cdots, e_n)$ une base de F et la base duale $(e_1^*, \cdots, e_n^*)$ de F'. Le théorème de Hahn-Banach permet de prolonger cette base en une famille $(\phi_1, \cdots, \phi_n)$ de formes linéaires continues sur E telles que si $x = \sum_{j=1}^n x_j e_j$ alors $\phi_i(x) = x_i$.

On vérifie aisément que $\bigcap_{i=1}^n \operatorname{Ker} \phi_i$ est un supplémentaire topologique.

2. Tout s.e.v. fermé de codimension finie admet un supplémentaire topologique (prendre un supplémentaire algébrique qui, étant de dimension finie, est forcément fermé). En particulier tout s.e.v. de E s'écrivant comme intersection finie de noyaux d'éléments de E' rentre dans cette catégorie. En effet, soit $F = \cap_{i=1}^{p} \mathrm{Ker}\, L_i$ avec $L_i \in E'$. On peut supposer sans perte de généralité que $(L_1, \cdots, L_p)$ est une famille libre de E'. On définit alors

$$\Phi \ : \ \left\{ \begin{array}{ccc} E & \longrightarrow & \mathbb{R}^p \\ x & \longmapsto & (L_1(x), \cdots, L_p(x)). \end{array} \right.$$

On montre facilement que Φ est linéaire, continue et surjective. Soit $(\varepsilon_1, \cdots, \varepsilon_p)$ la base canonique de $\mathbb{R}^p$. Alors le sous-espace vectoriel G de E engendré par $(\Phi^{-1}(\varepsilon_1), \cdots, \Phi^{-1}(\varepsilon_p))$ est un supplémentaire topologique de F.

Théorème. *Soit $T \in \mathcal{L}(E; F)$ avec E et F espaces de Banach. On suppose de plus que T est surjective. Alors les deux propositions suivantes sont équivalentes :*

(i) *T admet un inverse à droite continu (i.e. il existe $S \in \mathcal{L}(F; E)$ tel que $T \circ S = \mathrm{Id}_F$).*

(ii) $\mathrm{Ker}\, T$ *admet un supplémentaire topologique.*

Preuve : Si T admet un inverse à droite S qui est continu, on constate que $S(F)$ est un supplémentaire algébrique de $\mathrm{Ker}\, T$. En effet, il est clair que $S(F)$ et $\mathrm{Ker}\, T$ sont en somme directe, et comme tout $x \in E$ peut se décomposer en

$$x = \big(x - S(T(x))\big) + S(T(x)),$$

on a bien $\mathrm{Ker}\, T \oplus S(F) = E$.

Le fait que $S(F)$ soit fermé se vérifie facilement à l'aide de la propriété $T \circ S = \mathrm{Id}_F$.

Réciproquement, supposons que $\mathrm{Ker}\, T$ admette un supplémentaire topologique G.

Soit p la projection sur G parallèlement à $\mathrm{Ker}\, T$. Soit $y \in F$ et $x \in E$ tel que $T(x) = y$ (existe car T est surjective). On pose alors $S(y) = p(x)$.

On vérifie facilement que S est bien définie (i.e. ne dépend pas du choix de x) et linéaire, et que $T \circ S = \mathrm{Id}_F$.

Enfin, d'après le théorème de l'application ouverte, il existe $c > 0$ tel que

$$\forall x \in E, \ \|T(x)\|_F < c \Longrightarrow \|x\|_E < 1.$$

En prenant x de la forme $x = S(y)$ avec $y \in F$, on en déduit que

$$\forall y \in B_F(0, c), \ \|S(y)\|_E < 1.$$

Cela assure la continuité de S. ∎

Chapitre 7

Espaces de Hilbert

7.1 Produit scalaire

Dans toute cette section, E désigne un espace vectoriel sur $\mathbb{R}$ ou sur $\mathbb{C}$. Nous souhaitons munir cet espace d'une forme bilinéaire symétrique particulière appelée produit scalaire. Pour faciliter la lecture, nous distinguerons le cas réel du cas complexe.

7.1.1 Le cas réel

Dans toute cette partie, E désigne un espace vectoriel *réel*.

Définition. On dit qu'une **forme bilinéaire**[1] b sur E est un **produit scalaire** si elle est
- **symétrique** : $\forall (x,y) \in E^2,\ b(y,x) = b(x,y)$,
- **définie positive** : $\forall x \in E \setminus \{0\},\ b(x,x) > 0$.

Exemples.

1. Sur $\mathbb{R}^n$, l'application
$$(x,y) \longmapsto \sum_{j=1}^{n} x_j y_j$$
est un produit scalaire, appelé **produit scalaire canonique** de $\mathbb{R}^n$.
2. Sur l'ensemble $\ell^2(\mathbb{R})$ des suites réelles de carrés sommables, l'application
$$(x,y) \longmapsto \sum_{j=0}^{\infty} x_j y_j$$
est un produit scalaire.
3. Soit Ω un ouvert de $\mathbb{R}^n$. Sur l'ensemble $\mathcal{C}_b(\Omega;\mathbb{R})$ des fonctions continues bornées sur Ω et à valeurs réelles, l'application
$$(f,g) \longmapsto \int_\Omega f(x)g(x)\,dx$$
est un produit scalaire.

Définition. Un espace vectoriel réel muni d'un produit scalaire est appelé **espace préhilbertien réel**. Le produit scalaire $b(x,y)$ est alors noté $(x \mid y)$ et l'on pose $\|x\| = \sqrt{(x \mid x)}$.

1. Rappelons que, par définition, une forme bilinéaire sur un espace vectoriel réel est une application bilinéaire de $E \times E$ dans $\mathbb{R}$.

Remarque : Un espace préhilbertien réel de dimension finie est aussi appelé **espace euclidien**.

Proposition. *Soit E un espace préhilbertien réel. Alors* **l'inégalité de Cauchy-Schwarz** *est vérifiée :*

$$\forall (x, y) \in E^2,\ |(x \mid y)| \le \|x\|\,\|y\|$$

avec égalité si et seulement si x et y sont colinéaires.

Preuve : Fixons x et y dans E. Éliminons d'emblée le cas $y = 0$ qui est trivial.

Pour $\lambda \in \mathbb{R}$, on pose alors $f(\lambda) = (x + \lambda y | x + \lambda y)$. En développant l'expression de f, on trouve

$$f(\lambda) = \lambda^2 (y|y) + 2\lambda (x|y) + (x|x). \tag{7.1}$$

Comme $(y|y) > 0$ (puisque l'on a supposé que $y \neq 0$), la fonction f est un polynôme de degré 2 en λ qui reste positif pour tout $\lambda \in \mathbb{R}$. Son discriminant réduit

$$\Delta' = |(x|y)|^2 - (x|x)\,(y|y)$$

est donc négatif, ce qui montre l'inégalité voulue.

Les cas d'égalité correspondent à $\Delta' = 0$ ce qui revient à dire que f peut s'annuler sur $\mathbb{R}$, i.e. il existe $\lambda_0 \in \mathbb{R}$ tel que $x + \lambda_0 y = 0$. ∎

Corollaire. *L'application de E dans $\mathbb{R}$ qui à $x \in E$ associe $\|x\|$ est une norme sur E. On dit que c'est la norme associée au produit scalaire.*

Preuve : Il est clair que la fonction $x \mapsto \|x\|$ est bien définie sur E entier, positive, et qu'elle ne s'annule que si $x = 0$. Il est aussi immédiat que $\|\lambda x\| = |\lambda|\|x\|$. Reste à prouver l'inégalité triangulaire. Pour ce, on calcule

$$\|x + y\|^2 = \|x\|^2 + 2(x|y) + \|y\|^2.$$

Par l'inégalité de Cauchy-Schwarz, on a $(x|y) \le \|x\|\|y\|$. Donc

$$\|x + y\|^2 \le \|x\|^2 + 2\|x\|\|y\| + \|y\|^2 = (\|x\| + \|y\|)^2.$$

D'où le résultat en prenant la racine carrée. ∎

Remarque : Dans le cas où E est un espace euclidien, la norme associée au produit scalaire est appelée **norme euclidienne** .

Proposition. *Dans tout espace préhilbertien,* **l'identité du parallélogramme** *suivante est vérifiée :*

$$\boxed{\forall (x, y) \in E^2,\ \|x + y\|^2 + \|x - y\|^2 = 2\|x\|^2 + 2\|y\|^2}$$

ainsi que **l'identité de polarisation** *:*

$$\boxed{\forall (x, y) \in E^2,\ (x \mid y) = \frac{1}{4}\Big(\|x + y\|^2 - \|x - y\|^2\Big).}$$

Preuve : Il suffit de remarquer que, par définition de la norme, on a

$$\begin{aligned} \|x + y\|^2 &= \|x\|^2 + 2(x \mid y) + \|y\|^2, \\ \|x - y\|^2 &= \|x\|^2 - 2(x \mid y) + \|y\|^2. \end{aligned}$$

Il est clair qu'additionner les deux égalités donne l'identité du parallélogramme, et que retrancher la deuxième à la première permet d'obtenir l'identité de polarisation. ∎

Exercice : Démontrer la réciproque : si E est un espace vectoriel normé réel muni d'une norme vérifiant l'identité du parallélogramme alors c'est un espace préhilbertien.

Hint : utiliser l'identité de polarisation pour définir le produit scalaire.

Il est bien connu que les isométries du plan ou de l'espace conservent les angles (ou, de façon équivalente, le produit scalaire). Cette propriété se généralise aisément à tout espace préhilbertien réel :

Proposition. *Soit E un espace préhilbertien réel et f une isométrie sur E. Alors on a*

$$\forall (x,y) \in E^2,\ (f(x)|f(y)) = (x|y).$$

Preuve : Grâce à l'identité de polarisation, on a pour tout $(x,y) \in E^2$,

$$4(f(x)|f(y)) = \|f(x)+f(y)\|^2 - \|f(x)-f(y)\|^2.$$

Par linéarité de f, on a

$$f(x)+f(y) = f(x+y) \quad \text{et} \quad f(x)-f(y) = f(x-y),$$

puis, comme f conserve la norme,

$$\|f(x+y)\| = \|x+y\| \quad \text{et} \quad \|f(x-y)\| = \|x-y\|.$$

En revenant à l'égalité initiale, on obtient donc

$$4(f(x)|f(y)) = \|x+y\|^2 - \|x-y\|^2$$

et l'identité de polarisation appliquée aux vecteurs x et y permet de conclure. ■

7.1.2 Le cas complexe

Dans toute cette partie, E désigne un espace vectoriel *complexe*.

Définition. On dit qu'une application $h : E \times E \to \mathbb{C}$ est un **produit scalaire** hermitien si elle est [2]

- **linéaire par rapport à la première variable** : $h(x+\lambda x', y) = h(x,y) + \lambda h(x',y)$ pour tout $(x,x',y) \in E^3$ et $\lambda \in \mathbb{C}$,
- **antilinéaire par rapport à la deuxième variable** : $h(x, y+\lambda y') = h(x,y) + \overline{\lambda} h(x,y')$ pour tout $(x,y,y') \in E^3$ et $\lambda \in \mathbb{C}$,
- **hermitienne** : $\forall (x,y) \in E^2,\ h(y,x) = \overline{h(x,y)}$,
- **définie positive** :
$$\forall x \in E \setminus \{0\},\ h(x,x) > 0.$$

Attention : Certains auteurs préfèrent imposer l'antilinéarité *par rapport à la première variable.* En mathématiques, on convient en général de l'antilinéarité par rapport à la deuxième variable, alors qu'en physique c'est le contraire. Tout est affaire de convention...

Exemples. 1. Sur $\mathbb{C}^n$, l'application

$$(x,y) \longmapsto \sum_{j=1}^{n} x_j \overline{y_j}$$

est un produit scalaire, appelé **produit scalaire canonique** de $\mathbb{C}^n$.

2. Sur l'ensemble $\ell^2(\mathbb{C})$ des suites complexes de carrés sommables, l'application

$$(x,y) \longmapsto \sum_{j=0}^{\infty} x_j \overline{y_j}$$

est un produit scalaire.

2. Une application vérifiant les deux premières propriétés est dite *sesquilinéaire*.

3. Soit Ω un ouvert de $\mathbb{R}^n$. Sur l'ensemble $\mathcal{C}_b(\Omega;\mathbb{C})$ des fonctions continues bornées sur Ω et à valeurs complexes, l'application
$$(f,g) \longmapsto \int_\Omega f(x)\overline{g(x)}\,dx$$
est un produit scalaire.

Définition. Un espace vectoriel complexe muni d'un produit scalaire est appelé **espace préhilbertien complexe**. Le produit scalaire hermitien $h(x,y)$ est alors noté $(x \mid y)$ et l'on pose $\|x\| = \sqrt{(x \mid x)}$.

Remarque : Un espace préhilbertien complexe de dimension finie est aussi appelé **espace hermitien**.

Proposition. *Soit E un espace préhilbertien complexe. Alors* **l'inégalité de Cauchy-Schwarz** *est vérifiée :*
$$\forall (x,y) \in E^2,\ |(x \mid y)| \leq \|x\|\,\|y\|$$
avec égalité si et seulement si x et y sont colinéaires.

Preuve : La preuve diffère un peu de celle du cas réel. Le cas $y = 0$ étant trivial, on suppose désormais que $y \neq 0$. On pose alors $f(\lambda) = \|x+\lambda y\|^2$ pour $\lambda \in \mathbb{C}$ et l'on constate que
$$f(\lambda) = \|x\|^2 + |\lambda|^2\|y\|^2 + 2\mathcal{R}e\left(\overline{\lambda}(x \mid y)\right).$$
Écrivons que $(x \mid y) = |(x \mid y)|\,e^{i\theta_0}$ et posons $g(\rho) = f(\rho e^{i\theta_0})$ pour $\rho \in \mathbb{R}$. On a donc
$$\forall \rho \in \mathbb{R},\ g(\rho) = \rho^2\|y\|^2 + 2\rho|(x \mid y)| + \|x\|^2.$$
Comme g est une fonction positive, on peut conclure comme dans le cas réel à l'inégalité de Cauchy-Schwarz.

Les cas d'égalité sont laissés au lecteur. ∎

Corollaire. *L'application de E dans $\mathbb{R}$ qui à $x \in E$ associe $\|x\| \stackrel{\text{déf}}{=} \sqrt{(x|x)}$ est une norme sur E. On dit que c'est la norme associée au produit scalaire.*

Preuve : Il est clair que la fonction $x \mapsto \|x\|$ est bien définie sur E entier, positive, et qu'elle ne s'annule que si $x = 0$. Il est aussi immédiat que $\|\lambda x\| = |\lambda|\|x\|$. Reste à prouver l'inégalité triangulaire. Pour ce, on calcule
$$\|x+y\|^2 = \|x\|^2 + 2\mathcal{R}e\,(x \mid y) + \|y\|^2.$$
Par l'inégalité de Cauchy-Schwarz, on a $\mathcal{R}e\,(x \mid y) \leq |(x \mid y)| \leq \|x\|\|y\|$ et l'on peut donc conclure comme dans le cas réel. ∎

Remarque : Dans le cas où E est un espace hermitien, la norme associée au produit scalaire est appelée **norme hermitienne**.

Proposition. *Dans tout espace préhilbertien,* **l'identité du parallélogramme** *suivante est vérifiée :*
$$\boxed{\forall (x,y) \in E^2,\ \|x+y\|^2 + \|x-y\|^2 = 2\|x\|^2 + 2\|y\|^2}$$
ainsi que **l'identité de polarisation** *:*
$$\boxed{\forall (x,y) \in E^2,\ (x \mid y) = \frac{1}{4}\Big(\|x+y\|^2 - \|x-y\|^2 + i\|x+iy\|^2 - i\|x-iy\|^2\Big).}$$

Preuve : Il suffit de remarquer que, par définition de la norme, on a

$$\begin{array}{rcl} \|x+y\|^2 &=& \|x\|^2 + 2\mathcal{R}e\,(x \mid y) + \|y\|^2, \\ \|x-y\|^2 &=& \|x\|^2 - 2\mathcal{R}e\,(x \mid y) + \|y\|^2, \\ \|x+iy\|^2 &=& \|x\|^2 + 2\mathcal{I}m\,(x \mid y) + \|y\|^2, \\ \|x-iy\|^2 &=& \|x\|^2 - 2\mathcal{I}m\,(x \mid y) + \|y\|^2. \end{array}$$

Additionner les deux égalités donne l'identité du parallélogramme. On constate aussi que

$$\|x+y\|^2 - \|x-y\|^2 = 4\mathcal{R}e\,(x \mid y) \quad \text{et} \quad \|x+iy\|^2 - \|x-iy\|^2 = 4\mathcal{I}m\,(x \mid y),$$

ce qui permet d'obtenir l'identité de polarisation. ■

Exercice : Démontrer la réciproque : si E est un espace vectoriel normé complexe muni d'une norme vérifiant l'identité du parallélogramme alors c'est un espace préhilbertien.
Hint : utiliser l'identité de polarisation pour définir le produit scalaire.

En utilisant l'identité de polarisation, le lecteur pourra établir que dans le cas complexe aussi, les isométries conservent les angles :

Proposition. *Soit E un espace préhilbertien complexe et f une isométrie sur E. Alors on a*

$$\forall (x,y) \in E^2,\ (f(x)|f(y)) = (x|y).$$

7.2 Orthogonalité

Dans toute cette partie E désigne un espace préhilbertien réel ou complexe.

Définition. On dit que deux éléments x et y de E sont **orthogonaux** si $(x \mid y) = 0$. On note alors $x \perp y$.

Théorème (de Pythagore)**.** *Supposons que $x \perp y$. Alors on a* $\|x+y\|^2 = \|x\|^2 + \|y\|^2$.

Preuve : On a $\|x+y\|^2 = \|x\|^2 + 2\mathcal{R}e\,(x \mid y) + \|y\|^2$. Or $(x \mid y) = 0$ d'où le résultat. ■

Définition. Pour tout $x \in E$, on définit l'**orthogonal** de x par la formule

$$x^\perp \stackrel{\text{déf}}{=} \{y \in E \,/\, (x \mid y) = 0\}.$$

Plus généralement, pour tout sous-ensemble A non vide de E, on définit l'orthogonal de A par

$$A^\perp \stackrel{\text{déf}}{=} \{y \in E \,/\, \forall x \in A,\ (x \mid y) = 0\}.$$

Proposition. *Pour tout $A \subset E$ non vide, l'ensemble $A^\perp$ est un sous-espace vectoriel fermé de E, et l'on a $A \cap A^\perp \subset \{0\}$.*

Preuve : Montrons d'abord que $A^\perp$ est fermé. Soit $(x_n)_{n\in\mathbb{N}}$ une suite convergente d'éléments de $A^\perp$ et $x \in E$, sa limite. On a pour tout $y \in A$ et $n \in \mathbb{N}$,

$$\big|(x|y)\big| = \big|(x_n|y) + (x - x_n|y)\big| = \big|(x - x_n|y)\big| \leq \|x - x_n\|\|y\|.$$

Par convergence de la suite, le dernier terme tend vers 0 quand n tend vers l'infini. Cela permet de conclure que $x \in A^\perp$.

Montrons maintenant que $A^\perp$ est un sous-espace vectoriel de E. Il est évident que $A^\perp$ contient 0. Il suffit donc de vérifier que $A^\perp$ est stable par combinaisons linéaires. C'est une conséquence immédiate de la linéarité (ou de l'antilinéarité) du produit scalaire par rapport à chaque variable.

Enfin, si $A \cap A^\perp$ contient un élément x alors x est orthogonal à lui-même donc est nul. ■

Proposition. *Pour tout sous-ensemble A de E, on a*

$$\overline{A} \subset (A^\perp)^\perp.$$

Preuve : Il est évident que $A \subset (A^\perp)^\perp$. De plus, d'après la proposition précédente, un orthogonal est toujours fermé, donc $(A^\perp)^\perp$ doit contenir l'adhérence de A. ■

Attention : Si E est de dimension finie et A est un sous-espace vectoriel de E alors $A = (A^\perp)^\perp$ (exercice : le démontrer). Mais en dimension infinie, l'inclusion $\overline{A} \subset (A^\perp)^\perp$ peut être stricte. Nous reviendrons plus loin sur les cas d'égalité.

Définition. On dit qu'une famille $(x_i)_{i\in I}$ d'éléments de E est **orthogonale** si l'on a

$$\forall (i,j) \in I^2, (i \neq j) \Longrightarrow (x_i \mid x_j) = 0.$$

On dit que $(x_i)_{i\in I}$ est **orthonormale** si elle est orthogonale et si de plus $\|x_i\| = 1$ pour tout $i \in I$.

Proposition. *Soit $(x_1, \cdots, x_p)$ une famille orthogonale constituée de vecteurs tous non nuls. Alors cette famille est libre.*

Preuve : Supposons donc que $\sum_{i=1}^p \lambda_i x_i = 0$. En prenant le produit scalaire de cette égalité avec x_j, tous les termes disparaissent, sauf celui correspondant à $i = j$. On obtient donc $\lambda_j \|x_j\|^2 = 0$. Mais comme $x_j \neq 0$, on conclut que $\lambda_j = 0$. Donc finalement tous les λ_i sont nuls, ce qui donne le résultat voulu. ■

Proposition. *Soit $(e_1, \cdots, e_n)$ une famille orthonormale de E et $x \in \mathrm{Vect}\,(e_1, \cdots, e_n)$. Alors*

$$x = \sum_{i=1}^n (x \mid e_i) e_i.$$

Preuve : Par hypothèse, il existe un n-uplet $(x_1, \cdots, x_n)$ tel que $x = \sum_{i=1}^n x_i e_i$. On en déduit que

$$(x \mid e_j) = \sum_{i=1}^n x_i (e_i \mid e_j) = x_j.$$

■

Corollaire. *Soit $(e_1, \cdots, e_n)$ une famille orthonormale de E et (x, y) un couple d'éléments de* $\mathrm{Vect}\,(e_1, \cdots, e_n)$*. Alors on a*

$$\forall (x,y) \in E^2,\ (x \mid y) = \sum_{i=1}^n (x \mid e_i)\overline{(y \mid e_i)}.$$

Preuve : Si l'on note $x = \sum_{i=1}^n x_i e_i$ et $y = \sum_{i=1}^n y_i e_i$, un calcul immédiat donne

$$(x \mid y) = \sum_{i=1}^n x_i \overline{y_i}.$$

D'après la proposition précédente, on a $x_i = (x \mid e_i)$ et $y_i = (y \mid e_i)$ d'où le résultat. ■

Théorème (Orthonormalisation de Gram–Schmidt). *Soit* $(a_1, \cdots, a_p)$ *une famille libre de* E. *Alors il existe une unique famille orthonormale* $(e_1, \cdots, e_p)$ *telle que*

i) $\forall j \in \{1, \cdots, p\}$, $\mathrm{Vect}\,(e_1, \cdots, e_j) = \mathrm{Vect}\,(a_1, \cdots, a_j)$,

ii) $\forall j \in \{1, \cdots, p\}$, $(a_j \mid e_j) \in \mathbb{R}_*^+$.

Preuve : La démonstration est presque plus intéressante que le résultat car elle fournit un algorithme permettant de construire $(e_1, \cdots, e_p)$ à partir de la famille $(a_1, \cdots, a_p)$.

Pour ce faire, on utilise une récurrence limitée. Tout d'abord, on pose $e_1 = a_1/\|a_1\|$. Ce vecteur est clairement de norme 1, son produit scalaire avec a_1 vaut $\|a_1\|$ donc est strictement positif, et bien sûr e_1 et a_1 engendrent la même droite vectorielle.

Supposons que l'on ait construit une famille orthonormale $(e_1, \cdots, e_k)$ vérifiant *i)* et *ii)* pour $j \in \{1, \cdots, k\}$. Si $k = p$, la preuve est terminée. Sinon, on cherche e_{k+1} sous la forme

$$e_{k+1} = \lambda\Big(a_{k+1} + \sum_{i=1}^{k} \alpha_i e_i\Big) \quad \text{avec} \quad \lambda \neq 0.$$

Prenons le produit scalaire de cette égalité avec e_i (pour $i \in \{1, \cdots, k\}$). Pour que e_{k+1} soit orthogonal à e_i, il est nécessaire et suffisant que $(a_{k+1} \mid e_i) + \alpha_i = 0$. Donc

$$e_{k+1} = \lambda\Big(a_{k+1} - \sum_{i=1}^{k} (a_{k+1} \mid e_i)\, e_i\Big).$$

Comme $a_{k+1} \notin \mathrm{Vect}\,(e_1, \cdots, e_k)$ (car $(a_1, \cdots, a_{k+1})$ est libre), le terme entre parenthèses n'est pas nul.

Pour rendre e_{k+1} de norme 1, il suffit donc de choisir $\lambda = \|a_{k+1} - \sum_{i=1}^{k} (a_{k+1} \mid e_i) e_i\|^{-1}$. En conséquence,

$$e_{k+1} = \frac{a_{k+1} - \sum_{i=1}^{k} (a_{k+1} \mid e_i) e_i}{\|a_{k+1} - \sum_{i=1}^{k} (a_{k+1} \mid e_i) e_i\|}.$$

En prenant le produit scalaire de e_{k+1} avec cette égalité, on obtient de plus

$$1 = \frac{(a_{k+1} \mid e_{k+1})}{\|a_{k+1} - \sum_{i=1}^{k} (a_{k+1} \mid e_i) e_i\|},$$

ce qui montre que $(a_{k+1} \mid e_{k+1}) > 0$. Ceci achève la preuve de l'existence.

L'unicité se démontre en reprenant la construction précédente et en vérifiant qu'à chaque étape il n'y a pas d'autre choix possible que celui que l'on a fait ci-dessus. ∎

Corollaire. *En dimension finie, toute famille orthonormale peut être complétée en une base orthonormale. En particulier, tout espace euclidien (ou hermitien) admet une base orthonormale.*

Preuve : Soit $(f_1, \cdots, f_p)$ une famille orthonormale de E (avec éventuellement $p = 0$). On commence par compléter cette famille en une base $(f_1, \cdots, f_p, f_{p+1}, \cdots, f_n)$ de E. Le procédé d'orthonormalisation de Gram–Schmidt permet alors de transformer cette base en une base orthonormale de E. Clairement, les p premiers vecteurs de la base resteront inchangés. ∎

Exercice : En déduire que si E est de *dimension finie* et si A est un s.e.v. de E, alors on a

$$\dim A + \dim A^\perp = \dim E.$$

En effectuant une récurrence complète, on obtient une version *infinie* du procédé d'orthonormalisation de Schmidt :

Théorème. *Soit* $(a_p)_{p\in\mathbb{N}}$ *une famille libre de* E. *Alors il existe une famille orthonormale unique* $(e_p)_{p\in\mathbb{N}}$ *de* E *telle que*

(i) $\forall j \in \mathbb{N},\ \mathrm{Vect}\,(e_1,\cdots,e_j) = \mathrm{Vect}\,(a_1,\cdots,a_j)$,

(ii) $\forall j \in \mathbb{N},\ (a_j \mid e_j) \in \mathbb{R}_*^+$.

7.3 Espaces de Hilbert

Définition. Un espace préhilbertien complet pour la norme associée au produit scalaire est appelé **espace de Hilbert**.

Exemples. 1. Tout espace euclidien ou hermitien est complet car de dimension finie. C'est donc un espace de Hilbert. En particulier, $\mathbb{R}^n$ et $\mathbb{C}^n$ munis du produit scalaire canonique sont des espaces de Hilbert.

2. L'ensemble des suites réelles ou complexes de carrés sommables muni du produit scalaire

$$(u|v) = \sum_{n=0}^{+\infty} u_n \overline{v_n}$$

est un espace de Hilbert.

3. Nous démontrerons plus loin que pour toute partie mesurable A de $\mathbb{R}^n$ l'ensemble des classes d'équivalence [3] de fonctions mesurables sur A à valeurs dans $\mathbb{R}$ ou $\mathbb{C}$, de module au carré intégrable, muni du produit scalaire

$$(f \mid g) = \int_A f(x)\overline{g(x)}\,dx$$

est un espace de Hilbert.

En revanche, même dans le cas A compact, $\mathcal{C}(A;\mathbb{R}^n)$ muni de ce même produit scalaire n'est pas complet (exercice : donner un contre-exemple).

Un espace de Hilbert est en particulier un espace de Banach. De ce fait, toute série absolument convergente y est convergente. Mais on dispose d'une propriété plus forte qui peut être vue comme une généralisation du théorème de Pythagore à *une suite* de vecteurs orthogonaux :

Proposition. *Soit* $(x_n)_{n\in\mathbb{N}}$ *une famille orthogonale d'un espace de Hilbert* H. *Alors* $\sum x_n$ *converge dans* H *si et et seulement si* $\sum_{n\in\mathbb{N}} \|x_n\|^2 < \infty$. *Si cette dernière condition est vérifiée, on a alors l'***égalité de Parseval** *:*

$$\left\|\sum_{n=0}^{+\infty} x_n\right\|^2 = \sum_{n=0}^{+\infty} \|x_n\|^2. \tag{7.2}$$

Preuve : Notons $S_n = \sum_{k=0}^{n} x_k$. Pour $p > n$, on a $S_p - S_n = \sum_{k=n+1}^{p} x_k$. Par orthogonalité de la famille $(x_k)_{k\in\mathbb{N}}$, on a donc

$$\|S_p - S_n\|^2 = \sum_{k=n+1}^{p} \|x_k\|^2.$$

3. pour la relation d'égalité presque partout au sens de la mesure de Lebesgue.

Si l'on suppose que $\sum_{k\in\mathbb{N}} \|x_k\|^2 < \infty$ alors le terme de droite tend vers 0 (unifomément en $p > n$) quand n tend vers $+\infty$. Donc $(S_n)_{n\in\mathbb{N}}$ est une suite de Cauchy de H, donc converge vers une limite S car H est complet. En reprenant le calcul ci-dessus, on obtient

$$\|S_n\|^2 = \sum_{k=0}^{n} \|x_k\|^2. \tag{7.3}$$

En faisant tendre n vers $+\infty$, on conclut que $\|S\|^2 = \sum_{k\in\mathbb{N}} \|x_k\|^2$. Cela achève la preuve de la réciproque de la proposition, et de (7.2).

Pour la partie directe, on utilise le fait que si $\sum x_n$ converge alors on peut passer à la limite dans (7.3). ■

7.4 Projecteurs orthogonaux

Dans cette section et les suivantes, sauf mention contraire, H désignera un espace de Hilbert réel ou complexe.

Théorème. *Soit K un convexe fermé non vide de H. Alors pour tout $x \in H$, il existe un unique point $p_x \in K$ tel que*

$$\|x - p_x\| = d(x, K) \stackrel{\text{déf}}{=} \inf_{y\in K} \|x - y\|.$$

Le point p_x est appelé **projection** *de x sur K. C'est l'unique point de K vérifiant*

$$\forall y \in K,\ \mathcal{R}e\,(x - p_x \mid y - p_x) \le 0. \tag{7.4}$$

Preuve : L'unicité de p_x provient du fait que si $p'_x \in K$ vérifie aussi $\|x - p'_x\| = d(x, K)$ alors on a d'après l'**identité de la médiane** (conséquence facile de l'identité du parallélogramme) :

$$\|x - p_x\|^2 + \|x - p'_x\|^2 = \frac{1}{2}\|p'_x - p_x\|^2 + 2\|x - \tfrac{p'_x + p_x}{2}\|^2,$$

et donc $p'_x \neq p_x$ entraînerait $\|x - \frac{p'_x+p_x}{2}\| < d(x, K)$. Comme, par convexité, le point $\frac{p'_x+p_x}{2}$ est dans K, cela contredirait la définition de $d(x, K)$.

Démontrons maintenant l'existence. Notons $d = d(x, K)$. Par définition de d, il existe une suite $(x_n)_{n\in\mathbb{N}}$ de points de K telle que $\lim_{n\to+\infty} \|x - x_n\| = d$. D'après l'identité de la médiane, on a pour tout $(m, n) \in \mathbb{N}^2$,

$$\|x - x_m\|^2 + \|x - x_n\|^2 = \frac{1}{2}\|x_n - x_m\|^2 + 2\|x - \tfrac{x_n + x_m}{2}\|^2.$$

Par convexité de K, on a $\frac{x_n+x_m}{2} \in K$ donc $\|x - \frac{x_n+x_m}{2}\| \ge d$. En conséquence, on a

$$\frac{1}{2}\|x_n - x_m\|^2 \le \|x - x_m\|^2 + \|x - x_n\|^2 - 2d^2,$$

et la suite $(x_n)_{n\in\mathbb{N}}$ est donc de Cauchy. Sa limite p_x appartient à K car K est fermé.

Il ne reste plus qu'à vérifier (7.4). Soit donc $y \in K$. Pour tout $t \in [0, 1]$, le point $y_t \stackrel{\text{déf}}{=} p_x + t(y - p_x)$ appartient à K. Donc on a

$$\|x - p_x\|^2 \le \|x - y_t\|^2 = \|x - p_x\|^2 + t^2\|y - p_x\|^2 + 2t\,\mathcal{R}e\,(p_x - x|y - p_x).$$

En faisant tendre t vers 0, on en déduit que $\mathcal{R}e\,(x-p_x|y-p_x)\leq 0$. Réciproquement, si un point $x_0\in K$ vérifie $\mathcal{R}e\,(x-x_0|y-x_0)\leq 0$ pour tout $y\in K$ alors on a, compte-tenu de $\mathcal{R}e\,(x_0-p_x|x-p_x)\leq 0$,

$$\|x_0-p_x\|^2=\mathcal{R}e\,(x_0-p_x|x_0-x)+\mathcal{R}e\,(x_0-p_x|x-p_x)\leq 0.$$

Donc $x_0=p_x$. ■

Corollaire. *Soit F un sous-espace vectoriel fermé de H. Pour tout $x\in H$, la projection de x sur F est l'unique point p_x de F tel que $x-p_x\in F^\perp$.*

Preuve : On sait déjà que p_x est l'unique point de F tel que

$$\forall z\in F,\ \mathcal{R}e\,(x-p_x\mid z-p_x)\leq 0.$$

Pour $y\in F$ fixé, en appliquant cette inégalité à $z=p_x-y$ puis à $z=p_x+y$, (et aussi à $z=p_x\pm iy$ dans le cas complexe), on obtient $(x-p_x\mid y)=0$.
Soit p'_x un point de F tel que $x-p'_x\in F^\perp$. Comme p'_x-p_x est dans F, on a

$$(x-p_x\mid p'_x-p_x)=0\quad\text{et}\quad(x-p'_x\mid p'_x-p_x)=0.$$

En retranchant la deuxième égalité à la première, on trouve que $\|p'_x-p_x\|^2=0$. Donc $p_x=p'_x$. ■

Si F est un sous-espace vectoriel de l'espace de Hilbert H, on a toujours $F\cap F^\perp=\{0\}$ mais il n'y pas de raison en général pour que l'on ait $F+F^\perp=H$. De même, on n'a pas nécessairement $F=(F^\perp)^\perp$: seule l'inclusion $F\subset(F^\perp)^\perp$ est toujours vérifiée.

Dans la proposition ci-dessous, nous donnons une condition nécessaire suffisante pour que ces deux identités soient vérifiées.

Proposition. *Soit F un sous-espace vectoriel de H. Les trois énoncés suivants sont équivalents :*

(i) *F est fermé,*

(ii) *$H=F\oplus F^\perp$,*

(iii) *$(F^\perp)^\perp=F$.*

Preuve : $(i)\Rightarrow(ii)$: Supposons F fermé et montrons que $H=F+F^\perp$. Soit donc $x\in H$ et p_x la projection de x sur le s.e.v. fermé F. On a bien sûr $x=p_x+(x-p_x)$. Par définiton de p_x et d'après le corollaire précédent, on a $p_x\in F$ et $x-p_x\in F^\perp$. En conséquence $x\in F+F^\perp$, ce qui assure que $H=F\oplus F^\perp$.
$(ii)\Rightarrow(iii)$: Soit x un élément de $(F^\perp)^\perp$. Écrivons $x=y+z$ avec $y\in F$ et $z\in F^\perp$. On a donc

$$(x\mid z)=(y\mid z)+\|z\|^2.$$

Comme par hypothèse x est orthogonal à $F^\perp$, le terme de gauche est nul. Mais comme $y\in F$ alors que $z\in F^\perp$, on a aussi $(y\mid z)=0$. On conclut que $z=0$, autrement dit $x\in F$. Donc $(F^\perp)^\perp\subset F$. Comme l'autre inclusion est toujours vraie, on a $(F^\perp)^\perp=F$.
$(iii)\Rightarrow(i)$: Par hypothèse $F=(F^\perp)^\perp$. Or un orthogonal est toujours fermé. Donc F est fermé. ■

Corollaire. *Pour tout s.e.v. F d'un espace de Hilbert, on a $\overline{F}=(F^\perp)^\perp$.*

Preuve : Seule l'inclusion $(F^\perp)^\perp\subset\overline{F}$ est à établir. Comme $\overline{F}$ est fermé, la proposition précédente assure que $((\overline{F})^\perp)^\perp=\overline{F}$. Mais sachant que $F\subset\overline{F}$, on a $F^\perp\supset(\overline{F})^\perp$ puis $(F^\perp)^\perp\subset((\overline{F})^\perp)^\perp=\overline{F}$. ■

Nous sommes maintenant en mesure de définir les projecteurs orthogonaux.

Définition. Soit F un s.e.v. fermé de H. Le s.e.v. $F^\perp$ est appelé **supplémentaire orthogonal** de F. Tout élément x de H se décompose alors de manière unique en

$$x = y + z \quad \text{avec} \quad y \in F \quad \text{et} \quad z \in F^\perp.$$

Le vecteur y est la projection orthogonale de x sur F et l'application $x \mapsto y$ est appelée **projection orthogonale** ou **projecteur orthogonal** sur F.

Remarque. Comme $F = (F^\perp)^\perp$, le point z de la décomposition ci-dessus est égal à la projection orthogonale de x sur $F^\perp$.

Sachant qu'un s.e.v. de dimension finie est fermé, on peut toujours lui associer un projecteur orthogonal. Cela motive les deux résultats suivants.

Proposition. *Soit F un s.e.v. de dimension finie, et $(e_1, \cdots, e_n)$ une base orthonormale de F. Le projecteur orthogonal p sur F est donné par la formule :*

$$\forall x \in F,\ p(x) = \sum_{i=1}^{n} (x \mid e_i)\, e_i. \tag{7.5}$$

Preuve : Notons p le projecteur défini par la formule (7.5). Il est clair que $\operatorname{Im} p = F$ et que $\operatorname{Ker} p = F^\perp$. La proposition est donc démontrée. ■

En combinant ce résultat avec la définition de la projection d'un point sur F en termes de distance, on obtient le résultat suivant.

Corollaire. *Soit F un s.e.v. de E de dimension finie. Soit $(e_1, \cdots, e_n)$ une base orthonormale de F. Alors on a*

$$\forall x \in E,\ d(x, F) = \left\| x - \sum_{i=1}^{n} (x \mid e_i)\, e_i \right\|.$$

De plus l'unique point de F où cette distance est atteinte est $p(x) = \displaystyle\sum_{i=1}^{n} (x \mid e_i)\, e_i$.

7.5 Deux résultats importants sur les espaces de Hilbert

Dans le théorème ci-dessous, nous allons établir qu'il existe une isométrie bijective entre l'espace de Hilbert H et son dual topologique H'. Cette propriété est bien connue en dimension finie. Elle demeure vraie pour les espaces de Hilbert. C'est une conséquence facile du théorème suivant :

Théorème (de représentation de Riesz-Fréchet)**.** *Soit H un Hilbert. Alors pour tout $f \in H'$, il existe un unique $x \in H$ tel que*

$$\forall y \in H,\ f(y) = (y \mid x). \tag{7.6}$$

De plus, $\|x\| = \|f\|_{H'}$.

Preuve : Tout d'abord, si $f(y) = (y \mid x) = (y \mid x')$ pour tout $y \in H$, alors on a

$$\forall y \in H,\ (y \mid x - x') = 0,$$

et donc $x - x' = 0$. Cela donne l'unicité.

L'existence dans le cas $f = 0$ est évidente (prendre $x = 0$). Supposons maintenant que $f \in H' \setminus \{0\}$. Alors $\operatorname{Ker} f$ est un hyperplan *fermé* de H et admet donc un supplémentaire orthogonal $(\operatorname{Ker} f)^\perp$ qui n'est pas réduit à $\{0\}$. Fixons un vecteur x_0 non nul de $(\operatorname{Ker} f)^\perp$. On a donc $f(x_0) \neq 0$ et on constate que

$$\forall y \in H,\ y - \frac{f(y)}{f(x_0)} x_0 \in \operatorname{Ker} f.$$

En conséquence, on a

$$\forall y \in H,\ (y \mid x_0) = \frac{f(y)}{f(x_0)} \|x_0\|^2.$$

Il ne reste plus qu'à poser $x = \frac{x_0 \overline{f(x_0)}}{\|x_0\|^2}$ pour établir (7.6).

Enfin, puisque $f(y) = (y \mid x)$ pour tout $y \in H$, l'inégalité de Cauchy-Schwarz assure que $|f(y)| \leq \|x\|\|y\|$, et donc $\|f\|_{H'} \leq \|x\|$. Mais comme $f(x) = \|x\|^2$, on a en fait $\|f\|_{H'} = \|x\|$. ■

Théorème (de Lax-Milgram). *Soit H un espace de Hilbert et a une forme bilinéaire continue sur H. On suppose que a est* **coercive**, *c'est-à-dire qu'il existe $c_0 > 0$ tel que*

$$\forall x \in H,\ \mathcal{R}e\, a(x,x) \geq c_0 \|x\|^2.$$

Alors pour tout $f \in H'$ il existe un unique $x \in H$ vérifiant

$$\forall y \in H,\ a(y,x) = f(y).$$

Preuve : L'unicité de x est une conséquence facile de la coercivité. Démontrons l'existence. Comme a est, par hypothèse, continue, pour tout $x \in H$, l'application

$$y \longrightarrow a(y,x)$$

est une forme linéaire continue sur H. En conséquence, le théorème de Riesz-Fréchet assure l'existence d'un unique élément A_x de H tel que

$$\forall y \in H,\ a(y,x) = (y \mid A_x).$$

Il est clair que l'application $A : x \longmapsto A_x$ est linéaire dans le cas réel et antilinéaire dans le cas complexe. De plus, par continuité de a, il existe un réel positif C tel que

$$\forall x \in H,\ \|A_x\|^2 = (A_x \mid A_x) = a(A_x, x) \leq C \|A_x\| \|x\|.$$

Cela assure que l'application A est continue de H dans H.

Par ailleurs, toujours d'après le théorème de Riesz-Fréchet, il existe un $x_0 \in H$ tel que

$$\forall y \in H,\ f(y) = (y \mid x_0).$$

Nous sommes donc ramenés à la résolution de l'équation $A(x) = x_0$.

Pour $\rho > 0$ considérons l'application

$$S_\rho : \left\{ \begin{array}{lcl} H & \longrightarrow & H \\ x & \longmapsto & x + \rho(x_0 - A(x)). \end{array} \right.$$

Clairement à ρ fixé résoudre $A(x) = x_0$ revient à trouver un point fixe pour S_ρ.

On a pour tout $(x, x') \in H^2$,

$$\|S_\rho(x) - S_\rho(x')\|^2 = \|x - x'\|^2 + 2\rho \mathcal{R}e \left(x - x' \mid A(x' - x) \right) + \rho^2 \|A(x' - x)\|^2.$$

Donc, en notant c la norme de l'application linéaire A et en utilisant la coercivité de a,

$$\forall (x, x') \in H^2, \ \|S_\rho(x) - S_\rho(x')\|^2 \leq (1 - 2\rho c_0 + \rho^2 c^2) \|x - x'\|^2.$$

On choisit $\rho > 0$ suffisamment proche de 0 pour que $1 - 2\rho c_0 + \rho^2 c^2 < 1$. Le calcul ci-dessus montre alors que S_ρ est *contractante*. Comme H est un espace métrique complet (car c'est un Hilbert), le théorème du point fixe permet de conclure qu'il existe un unique $x \in H$ tel que $S_\rho(x) = x$. ■

7.6 Bases hilbertiennes et espaces de Hilbert séparables

Définition. On dit qu'un sous-ensemble A d'un espace de Hilbert H est **total** si le sous-espace vectoriel $\operatorname{Vect} A$ engendré par A est dense dans H.

Proposition. *Soit A un sous-ensemble de H. Alors A est total si et seulement si $A^\perp = \{0\}$.*

Preuve : Supposons d'abord que A soit total. Soit $x \in A^\perp$. Alors par linéarité du produit scalaire par rapport à la première variable, on déduit que x est aussi orthogonal à toute combinaison linéaire d'éléments de A, donc à $\operatorname{Vect} A$ qui, par hypothèse est dense dans H. Soit $(x_n)_{n \in \mathbb{N}}$ une suite d'éléments de $\operatorname{Vect} A$ qui converge vers x. On a bien sûr $(x \mid x_n) = 0$ pour tout $n \in \mathbb{N}$. En passant à la limite, on en conclut que $\|x\|^2 = 0$. Donc $x = 0$.

Réciproquement, supposons que $A^\perp = \{0\}$. Alors on a $(A^\perp)^\perp = H$. Mais il clair que $A^\perp = (\operatorname{Vect} A)^\perp$. En conséquence, on a

$$\overline{\operatorname{Vect} A} = \left((\operatorname{Vect} A)^\perp \right)^\perp = (A^\perp)^\perp = H.$$

Donc A est total. ■

> *Remarque.* Comme cas particulier très important, on obtient le fait qu'un s.e.v. F est dense si et seulement si $F^\perp = \{0\}$.

Rappelons qu'un espace topologique est dit **séparable** s'il contient une partie dénombrable dense.

Attention : Ne pas confondre séparable et séparé ! Tout espace de Hilbert est séparé car muni de la topologie associée à une norme, mais on peut construire des espaces de Hilbert qui ne sont pas séparables.

Proposition. *Un espace de Hilbert est séparable si et seulement si il admet un sous-ensemble total dénombrable.*

Preuve : La partie directe est triviale.

Réciproquement, soit A un sous-ensemble total dénombrable de l'espace de Hilbert H considéré. Alors on vérifie aisément que l'ensemble des combinaisons linéaires d'éléments de A à coefficients rationnels (cas réel) ou à parties réelles et imaginaires rationnelles (cas complexe) est dense dans H. Cet ensemble est dénombrable, donc H est bien séparable. ■

Définition. Soit H un Hilbert et $(e_n)_{n \in \mathbb{N}}$ une suite d'éléments de H. On dit que $(e_n)_{n \in \mathbb{N}}$ est une **base hilbertienne** de H si (*i*) $(e_n)_{n\in\mathbb{N}}$ est une famille orthonormale de H,

(*ii*) l'ensemble $\{e_n\,,\, n \in \mathbb{N}\}$ est total.

Proposition. *Un espace de Hilbert de dimension infinie est séparable si et seulement si il admet une base hilbertienne.*

Preuve : Si l'espace de Hilbert H admet une base hilbertienne $(e_k)_{k\in\mathbb{N}}$, alors l'ensemble constitué par ces vecteurs est total et dénombrable. Donc H est séparable.

Réciproquement, supposons H séparable. Soit $(a_n)_{n\in\mathbb{N}}$ une suite totale de H. Quitte à supprimer des éléments de cette suite, on peut se ramener au cas où cette famille est linéairement indépendante : il suffit de raisonner par récurrence en supprimant de la suite $(a_n)_{n\in\mathbb{N}}$ tout vecteur qui est combinaison linéaire des précédents. La famille ainsi obtenue est libre, et dénombrable non finie (sinon H serait de dimension finie).

Notons $(b_k)_{k\in\mathbb{N}}$ cette famille. Le **procédé d'orthonormalisation de Schmidt** permet alors de construire à partir de $(b_k)_{k\in\mathbb{N}}$ une suite orthonormale $(e_k)_{k\in\mathbb{N}}$ telle que pour tout $n \in \mathbb{N}$, on ait

$$\mathrm{Vect}\,(e_0, \cdots, e_n) = \mathrm{Vect}\,(b_0, \cdots, b_n).$$

Vérifions que l'ensemble $\{e_k\,,\, k \in \mathbb{N}\}$ est total. Soit $x \in H$ et $\varepsilon > 0$. Comme $\{b_k\,,\, k \in \mathbb{N}\}$ est total, il existe $n \in \mathbb{N}$ et $y \in \mathrm{Vect}\,(b_0, \cdots, b_n) = \mathrm{Vect}\,(e_0, \cdots, e_n)$ tel que $\|x - y\| \le \varepsilon$. Cela achève la démonstration. ■

Proposition. *Soit H un espace de Hilbert séparable et $(e_n)_{n\in\mathbb{N}}$ une base hilbertienne de H. Alors pour tout $x \in H$, on a*

$$x = \sum_{n\in\mathbb{N}} (x \mid e_n) e_n \quad \text{et} \quad \|x\|^2 = \sum_{n\in\mathbb{N}} \big|(x \mid e_n)\big|^2 \quad (\textbf{égalité de Parseval}).$$

Réciproquement, si $(\alpha_n)_{n\in\mathbb{N}}$ est une suite de ℓ^2 alors $\sum \alpha_n e_n$ converge dans H et l'on a

$$\forall m \in \mathbb{N},\ \Big(\sum_{n\in\mathbb{N}} \alpha_n e_n \,\Big|\, e_m\Big) = \alpha_m.$$

Preuve : Soit $x \in H$. Posons $x_n = \sum_{k=1}^{n} (x \mid e_k) e_k$. On a, d'après le théorème de Pythagore,

$$(x \mid x_n) = \sum_{k=0}^{n} \big|(x \mid e_k)\big|^2 = \|x_n\|^2.$$

À l'aide de l'inégalité de Cauchy-Schwarz, on en déduit que $\|x_n\|_H \le \|x\|_H$ pour tout $n \in \mathbb{N}$. En conséquence, la série $\sum \big|(x \mid e_k)\big|^2$ est convergente. La proposition de la page 74 assure donc la convergence de $\sum (x \mid e_k) e_k$. De plus, en notant y la somme de cette série, on a

$$\|y\|^2 = \sum_{n\in\mathbb{N}} \big|(x \mid e_n)\big|^2.$$

Il est aussi clair que $(y - x \mid e_k) = 0$ pour tout $k \in \mathbb{N}$. Comme $(e_n)_{n\in\mathbb{N}}$ est totale, on en conclut que $y - x = 0$.

Reste à démontrer la réciproque. La convergence de la série $\sum_{n\in\mathbb{N}} \alpha_n e_n$ est un cas particulier de la proposition de la page 74. Maintenant, si la série converge, on a, par continuité du produit scalaire, et grâce à $(e_k \mid e_m) = \delta_{km}$,

$$\Big(\sum_{k\in\mathbb{N}} \alpha_k e_k \,\Big|\, e_m\Big) = \lim_{n\to+\infty} \Big(\sum_{k=0}^{n} \alpha_k e_k \,\Big|\, e_m\Big) = \alpha_m.$$

■

Exercice : Sous les hypothèses de la proposition précédente, vérifier que

$$\forall (x,y) \in H^2,\ (x \mid y) = \sum_{n\in\mathbb{N}} (x|e_n)\,\overline{(y|e_n)}.$$

Remarques : 1) La proposition ci-dessus fournit une isométrie bijective naturelle entre tout espace de Hilbert séparable et ℓ^2, une fois fixée une base hilbertienne $(e_n)_{n\in\mathbb{N}}$. Il suffit de considérer

$$\begin{cases} \ell^2 & \longrightarrow & H \\ (\alpha_n)_{n\in\mathbb{N}} & \longmapsto & \sum_{n\in\mathbb{N}} \alpha_n e_n. \end{cases}$$

Par conséquent, l'étude des espaces de Hilbert séparables (les seuls qui apparaissent dans les applications) peut se ramener à celle de l'espace ℓ^2.

2) On prendra garde au fait qu'une base hilbertienne n'est *jamais* une base algébrique : en reprenant les notations précédentes, il n'est pas vrai que tout vecteur de H peut s'exprimer comme combinaison linéaire des $(e_n)_{n\in\mathbb{N}}$. Pour "la plupart" des vecteurs, il faut utiliser une infinité d'éléments de e_n.

7.7 La convergence faible dans les espaces de Hilbert

L'un des problèmes de base que l'on rencontre dans les espaces de Hilbert de dimension infinie comme l'espace L^2 des fonctions de carré sommable est que les ensembles bornés sont très loin d'être d'adhérence compacte. À titre d'exemple, considérons la boule unité d'un espace de Hilbert séparable de dimension infinie H, et une base hilbertienne $(e_j)_{j\in\mathbb{N}}$ de H. D'après la formule de Parseval, pour tout élément x de H, la suite $(x|e_j)_{j\in\mathbb{N}}$ est de carré sommable. Donc son terme général tend vers 0. Donc si la suite $(e_j)_{j\in\mathbb{N}}$ admet une valeur d'adhérence, ce ne peut être que 0. Or $\|e_j\| = 1$ pour tout j. Donc 0 ne saurait être valeur d'adhérence d'une telle suite. Cette constatation motive la définition suivante.

Définition. Soient $(x_n)_{n\in\mathbb{N}}$ une suite d'éléments d'un espace de Hilbert H et x un élément de H. On dit que la $(x_n)_{n\in\mathbb{N}}$ **converge faiblement** vers x si

$$\forall h \in H\,,\ \lim_{n\to\infty} (h|x_n) = (h|x).$$

On utilise alors la notation $x_n \rightharpoonup x$.

Remarque. Nous laissons au lecteur le soin de vérifier l'unicité de la limite faible.

Exemple. Le raisonnement ci-dessus montre que dans un espace de Hilbert séparable, toute base hilbertienne converge faiblement vers 0.

Dans le théorème suivant, on donne quelques conséquences de la propriété de convergence faible.

Théorème. *Soit $(x_n)_{n\in\mathbb{N}}$ et $(y_n)_{n\in\mathbb{N}}$ deux suites d'éléments d'un espace de Hilbert H et x, y deux éléments de H. On a alors :*

$$x_n \rightharpoonup x \implies (x_n)_{n\in\mathbb{N}} \text{ est bornée et } \|x\| \leq \liminf \|x_n\|\,; \tag{7.7}$$

$$x_n \to x \implies x_n \rightharpoonup x\,; \tag{7.8}$$

$$x_n \rightharpoonup x \text{ et } \lim_{n\to\infty} \|x_n\| = \|x\| \implies \lim_{n\to\infty} \|x_n - x\| = 0. \tag{7.9}$$

$$\big(x_n \to x \quad \text{et} \quad y_n \rightharpoonup y\big) \implies \lim_{n\to\infty} (x_n|y_n) = (x|y). \tag{7.10}$$

Preuve : La preuve du premier point du théorème découle du théorème de Banach-Steinhaus. En effet, notons T_n l'application définie sur H par $T_n(h) = (h \mid x_n)$. Il s'agit clairement d'une forme linéaire continue sur H. Par ailleurs, pour h fixé, la suite de terme général $(T_n)(h)$ est convergente. En conséquence, le théorème de Banach-Steinhaus (ou plutôt le corollaire qui suit) assure que la suite $(T_n)_{n\in\mathbb{N}}$ est bornée et que l'application linéaire limite $T : h \mapsto (h|x)$ est continue et vérifie

$$\|T\|_{H'} \leq \liminf \|T_n\|_{H'}.$$

Mais, d'après le théorème de Riesz-Fréchet, on a $\|T\|_{H'} = \|x\|_H$ et $\|T_n\| = \|x_n\|_H$, ce qui achève la démonstration de la première propriété.

Le deuxième point résulte simplement du fait que

$$|(h|x_n) - (h|x)| \leq \|h\| \, \|x_n - x\|.$$

Pour le troisième point, il suffit d'écrire que

$$\|x_n - x\|^2 = \|x_n\|^2 - 2\mathcal{R}e\,(x|x_n) + \|x\|^2.$$

Comme $(x_n)_{n\in\mathbb{N}}$ tend faiblement vers x, on a

$$-2 \lim_{n\to\infty} \mathcal{R}e\,(x|x_n) = -2\|x\|^2.$$

Cela assure (7.9).

La démonstration de la dernière propriété est très simple. Il suffit d'écrire que

$$\begin{aligned} |(x_n|y_n) - (x|y)| &\leq |(x_n - x|y_n)| + |(x|y_n - y)| \\ &\leq \|x_n - x\| \, \|y_n\| + |(x|y_n - y)|. \end{aligned}$$

Le théorème précédent affirme que la $(y_n)_{n\in\mathbb{N}}$ est bornée. Donc, on a

$$|(x_n|y_n) - (x|y)| \leq C\|x_n - x\| + |(x|y_n - y)|,$$

d'où la proposition. ■

Proposition. *En dimension finie, la convergence faible est équivalente à la convergence forte.*

Preuve : D'après le théorème précédent, la convergence forte entraîne toujours la convergence faible. Réciproquement, supposons que l'espace hilbertien H soit de dimension finie et donnons nous une base orthonormale $(e_1, \cdots, e_p)$ de H, et une suite faiblement convergente $(x_n)_{n\in\mathbb{N}}$. Soit x sa limite faible. On a

$$\|x_n - x\|^2 = \sum_{i=1}^{p} |(e_i \mid x_n - x)|^2.$$

Par convergence faible, on a $\lim_{n\to+\infty}(e_i \mid x_n - x) = 0$ pour tout $i \in \{1, \cdots, p\}$. Donc $\lim_{n\to+\infty} \|x_n - x\| = 0$. ■

Donnons un autre cas où les notions de convergence forte et faible se rejoignent :

Proposition. *Soit C un ensemble convexe de l'espace de Hilbert H. Alors les deux énoncés suivants sont équivalents :*

- *l'ensemble C est fermé,*

– *pour toute suite $(x_n)_{n\in\mathbb{N}}$ faiblement convergente d'éléments de C, la limite faible x est dans C.*

Preuve : Sachant que toute suite fortement convergente est faiblement convergente, l'implication réciproque est évidente.

Supposons donc C fermé et considérons une suite $(x_n)_{n\in\mathbb{N}} \in C^{\mathbb{N}}$ convergeant faiblement vers $x \in H$. Il s'agit de démontrer que x est dans C. Comme C est convexe fermé, le point x admet une projection p_x sur C. Cette projection est l'unique point de C tel que

$$\forall y \in C,\ \mathcal{R}e\,(x - p_x \mid y - p_x) \le 0.$$

En appliquant cette relation à $y = x_n$ puis en faisant tendre n vers l'infini (c'est ici qu'intervient l'hypothèse de convergence faible), on obtient

$$\|x - p_x\|^2 \le 0.$$

En conséquence $x = p_x \in C$. ∎

En dimension finie, le théorème de Riesz assure que toute suite bornée admet une sous-suite convergente. Le théorème suivant (dont les applications notamment en EDP sont multiples) assure que la propriété demeure vraie pour la convergence faible dans n'importe quel espace de Hilbert.

Théorème (de compacité faible). *De toute suite bornée d'un espace de Hilbert, on peut extraire une sous-suite faiblement convergente.*

Preuve : Soit $(x_n)_{n\in\mathbb{N}}$ une suite bornée de l'espace de Hilbert H. Notons M une borne strictement positive de la suite. Pour simplifier, supposons dans un premier temps que H soit *séparable.* Nous écartons d'emblée le cas de la dimension finie qui est couvert par le théorème de Riesz et fixons une base hilbertienne $(e_j)_{j\in\mathbb{N}}$ de H. Pour montrer la convergence faible de la suite, nous allons faire appel une fois de plus au procédé d'extraction diagonal de Cantor.

La suite $(e_0 \mid x_n)_{n\in\mathbb{N}}$ est une suite bornée de $\mathbb{K}$. Il existe donc un élément λ_0 de $\mathbb{K}$ et une extraction φ_0 de $\mathbb{N}$ dans $\mathbb{N}$ tels que l'on ait

$$\lim_{n\to\infty} (e_0 \mid x_{\varphi_0(n)}) = \lambda_0.$$

Supposons construites une suite finie $(\varphi_j)_{0\le j\le m}$ de fonctions strictement croissantes de $\mathbb{N}$ dans $\mathbb{N}$ et une suite finie $(\lambda_j)_{1\le j\le m}$ de scalaires telles que, pour tout $j \le m$, on ait

$$\lim_{n\to+\infty} (e_j \mid x_{\varphi_0\circ\cdots\circ\varphi_j(n)}) = \lambda_j.$$

La suite $(e_{m+1} \mid x_{\varphi_0\circ\cdots\circ\varphi_m(n)})_{n\in\mathbb{N}}$ est une suite bornée de $\mathbb{K}$. Il existe donc une fonction strictement croissante φ_{m+1} de $\mathbb{N}$ dans $\mathbb{N}$ et un élément λ_{m+1} de $\mathbb{K}$ tels que

$$\forall j \le m+1\,,\ \lim_{n\to\infty} (e_j \mid x_{\varphi_0\circ\cdots\circ\varphi_{m+1}(n)}) = \lambda_j.$$

Finalement, si l'on pose $\psi(n) = \varphi_0 \circ \cdots \circ \varphi_n(n)$, toutes les suites $(e_j \mid x_{\psi(n)})_{n\in\mathbb{N}}$ sont convergentes.

Soit V le sous-espace vectoriel engendré par tous les e_j, c'est-à-dire l'ensemble des combinaisons linéaires (finies) de e_j. Puisque $(e_j)_{j\in\mathbb{N}}$ est une base hilbertienne de H, l'ensemble V est dense dans H. Considérons l'application linéaire L définie par

$$L\ :\ \begin{cases} V & \to \quad \mathbb{K} \\ y & \mapsto \quad \lim\limits_{n\to\infty} (y|x_{\psi(n)}). \end{cases}$$

Notons que la défininition de L ne pose pas de problème puisque tout élément de V est combinaison linéaire finie des e_j et que la suite $(e_j \mid x_{\psi(n)})$ converge pour chaque $j \in \mathbb{N}$. De plus, L est continue car, pour tout y de V, nous avons

$$|L(y)| \leq (\sup_{n\in\mathbb{N}} \|x_n\|)\|y\| \leq M\|y\|.$$

En vertu du théorème de la page 31, la densité de V nous permet alors de prolonger la forme linéaire L à l'espace H tout entier. Notons $\widetilde{L}$ la forme linéaire continue sur H ainsi obtenue.

D'après le théorème de Riesz-Fréchet, il existe un élément x de H tel que

$$\forall y \in H,\ \widetilde{L}(y) = (y \mid x).$$

Ainsi nous avons en particulier que

$$\forall y \in V,\ \lim_{n\to\infty}(y|x_{\psi(n)}) = (y|x).$$

Il reste à démontrer la convergence pour tout $z \in H$. Fixons donc $z \in H$ et $\varepsilon > 0$. Par densité de V dans H, il existe $y \in V$ tel que $2(M + \|x\|)\|y - z\| \leq \varepsilon$. En écrivant que $(z|x_{\psi(n)} - x) = (y|x_{\psi(n)} - x) + (z - y|x_{\psi(n)} - x)$, et en se souvenant que la suite est bornée par $M > 0$, on obtient

$$|(z|x_{\psi(n)} - x)| \leq |(y|x_{\psi(n)} - x)| + \varepsilon/2.$$

Pour n assez grand, on aura donc $|(z|x_{\psi(n)} - x)| \leq \varepsilon$. Cela achève la démonstration de la convergence faible de $(x_{\psi(n)})_{n\in\mathbb{N}}$.

Pour terminer, considérons le cas d'un espace de Hilbert H *non séparable*. On définit alors F comme étant l'adhérence du sous-espace vectoriel engendré par $(x_n)_{n\in\mathbb{N}}$. L'espace préhilbertien F est un espace de Hilbert car est un fermé d'un espace de Hilbert. Il est bien sûr séparable, par construction. La démonstration précédente assure l'existence d'une sous-suite $(x_{\psi(n)})_{n\in\mathbb{N}}$ et d'un élément x de F tels que $(x_{\psi(n)})_{n\in\mathbb{N}}$ tende vers x *pour la convergence faible sur* F.

Soit maintenant $y \in H$ quelconque. Sachant que $H = F \oplus F^{\perp}$ (car F est fermé), on peut écrire $y = y_1 + y_2$ avec $y_1 \in F$ et $y_2 \in F^{\perp}$. Il est alors clair que pour tout $n \in \mathbb{N}$, on a $(y|x_{\psi(n)}) = (y_1|x_{\psi(n)})$ et que $(y|x) = (y_1|x)$. On a donc

$$\lim_{n\to+\infty}(y|x_{\psi(n)}) = (y|x).$$

Cela achève la démonstration du cas général. ∎

Comme application, donnons le résultat suivant qui généralise un théorème bien connu en dimension finie (voir le cours de calcul différentiel de licence) :

Théorème. *Soit H un espace de Hilbert et et A un convexe fermé non vide de H. Soit $\varphi \in \mathcal{C}(A;\mathbb{R})$. On suppose que φ est convexe et vérifie*

$$\lim_{\substack{x\in A\\ \|x\|\to+\infty}} \varphi(x) = +\infty.$$

Alors φ est minorée et atteint son minimum absolu : il existe $a \in A$ tel que

$$\forall x \in A,\ \varphi(x) \geq \varphi(a).$$

Preuve : Soit $m = \inf_{x\in A}\varphi(x)$. Supposons par l'absurde que

$$\forall x \in A,\ \varphi(x) > m.$$

Soit $(x_n)_{n\in\mathbb{N}}$ une suite d'éléments de A telle que

$$\lim_{n\to+\infty}\varphi(x_n) = m.$$

Comme φ tend vers l'infini à l'infini, la suite $(x_n)_{n\in\mathbb{N}}$ est bornée donc admet une sous-suite $(x_{\psi(n)})_{n\in\mathbb{N}}$ faiblement convergente en vertu de théorème de compacité faible. Notons a la limite de cette sous-suite.

Pour $x \in A$ notons $A_x = \{y \in A\,|\,\varphi(y) \le \varphi(x)\}$. Cet ensemble est convexe fermé car φ est convexe continue, et il contient les termes de la suite $(x_{\psi(n)})_{n\in\mathbb{N}}$ à partir d'un certain rang, donc a en vertu de la proposition de la page 82. On a donc en particulier $\varphi(a) \le \varphi(x)$. En d'autres termes, φ atteint sa borne inférieure. Cela achève la preuve du théorème. ■

7.8 Application à l'espace L^2 et aux séries de Fourier

Dans toute cette section, nous ferons appel à des notions classiques de théorie de l'intégration vues en cours de licence[4].

7.8.1 Les espaces L^p

Soit A un borélien de $\mathbb{R}$ de mesure non nulle. Pour $p \in [1,+\infty[$, on note

$$\mathcal{L}^p(A;\mathbb{K}) \overset{\text{déf}}{=} \Big\{f : A \to \mathbb{K}\ \mid\ f \text{ Lebesgue mesurable et } \int_A |f(x)|^p\,dx < \infty\Big\}.$$

L'application

$$f \longmapsto \|f\|_{L^p} \overset{\text{déf}}{=} \left(\int_A |f(x)|^p\,dx\right)^{\frac{1}{p}}$$

a toutes les propriétés d'une norme sauf celle qui stipule que $\|f\|_{L^p} = 0$ si et seulement si $f = 0$. En effet $\|f\|_{L^p} = 0$ entraîne seulement que $f = 0$ presque partout. De ce fait $\|\cdot\|_{L^p}$ n'est qu'une semi-norme sur $\mathcal{L}^p(A;\mathbb{K})$.

Pour pallier ce grave défaut, on considère les *classes d'équivalence* des fonctions de $\mathcal{L}^p(A;\mathbb{K})$ pour la relation d'équivalence d'égalité presque partout au sens de la mesure de Lebesgue. On note $L^p(A;\mathbb{K})$ (ou plus simplement L^p en l'absence d'ambiguïté) l'ensemble de ces classes d'équivalence et en pratique, on confond les éléments de L^p (qui sont des classes d'équivalence de fonctions) avec leurs représentants (qui sont des fonctions de puissance p-ième sommable).

Théorème. *Pour tout $p \in [1,+\infty[$, l'espace $L^p(A;\mathbb{K})$ muni de la norme $\|\cdot\|_{L^p}$ est complet.*

Preuve : Soit $(f_n)_{n\in\mathbb{N}}$ une suite de Cauchy de $L^p(A;\mathbb{K})$. On veut montrer que cette suite converge dans $L^p(A;\mathbb{K})$.

Tout d'abord remarquons que l'on peut construire par récurrence une sous-suite $(f_{\varphi(n)})_{n\in\mathbb{N}}$ telle que

$$\forall n \in \mathbb{N},\ \|f_{\varphi(n+1)} - f_{\varphi(n)}\|_{L^p} \le 2^{-n}.$$

Si l'on pose $u_n \overset{\text{déf}}{=} f_{\varphi(n+1)} - f_{\varphi(n)}$, on constate donc que la série $\sum u_n$ vérifie

$$\sum_{n\in\mathbb{N}} \|u_n\|_{L^p} < \infty\,^5.$$

4. pour plus de détails on pourra lire avec profit [1], [8] ou les deux premiers chapitres de [9].
5. On dit que la série $\sum u_n$ est **absolument convergente**.

Si l'on parvient à démontrer que $\sum u_n$ converge dans $L^p(A;\mathbb{K})$ alors on en déduira que $(f_{\varphi(n)})_{n\in\mathbb{N}}$ est une suite convergente. En d'autres termes, $(f_n)_{n\in\mathbb{N}}$ aura une valeur d'adhérence. Mais comme c'est aussi une suite de Cauchy, on pourra finalement conclure que $(f_n)_{n\in\mathbb{N}}$ converge.

Reste donc à établir que $\sum u_n$ converge. Par l'inégalité triangulaire, on a pour tout $n\in\mathbb{N}$,

$$\left\|\sum_{k=0}^{n}|u_k|\right\|_{L^p}\leq\sum_{k=0}^{n}\|u_k\|_{L^p}\leq\sum_{k=0}^{\infty}\|u_k\|_{L^p}.$$

Donc, d'après le théorème de convergence monotone,

$$\int_A\left(\sum_{k=0}^{+\infty}|u_k(x)|\right)^p dx<\infty.$$

Cela assure que $\sum_{k=0}^{+\infty}|u_k|$ est fini presque partout sur A et appartient à $L^p(A;\mathbb{K})$. On en déduit que $\sum_{k=0}^{+\infty}u_k$ converge aussi pour presque tout $x\in A$ et appartient à $L^p(A;\mathbb{K})$.

Reste à montrer la convergence de $\sum_{k=0}^{n}u_k$ vers $\sum_{k=0}^{+\infty}u_k$ dans $L^p(A;\mathbb{K})$. Pour cela, il suffit d'écrire que

$$\int_A\left|\sum_{k=0}^{+\infty}u_k(x)-\sum_{k=0}^{n}u_k(x)\right|^p dx=\int_A\left|\sum_{k=n+1}^{+\infty}u_k(x)\right|^p dx.$$

L'intégrand du membre de droite tend vers 0 presque partout et est majoré par la fonction $\left(\sum_{k=0}^{+\infty}|u_k|\right)^p$ qui est intégrable. En conséquence, le théorème de convergence dominée permet de conclure que le membre de gauche tend vers 0 quand n tend vers l'infini. Autrement dit, $\sum u_k$ est convergente dans $L^p(A;\mathbb{K})$. ■

Afin d'aborder l'étude des séries de Fourier, nous aurons aussi besoin du résultat de densité suivant.

Théorème. *Pour tout $p\in[1,+\infty[$, l'ensemble $\mathcal{C}_c(\mathbb{R};\mathbb{K})$ des fonctions continues de $\mathbb{R}$ dans $\mathbb{K}$ à* **support compact**[6] *est dense dans $L^p(\mathbb{R};\mathbb{K})$. De plus, si $f\in L^p(\mathbb{R};\mathbb{K})$ est nulle en dehors de l'intervalle $[a,b]$ alors on peut trouver une suite de fonctions continues f_n sur $\mathbb{R}$ nulles en dehors de $[a,b]$ telle que*

$$\lim_{n\to+\infty}\|f-f_n\|_{L^p}=0.$$

Preuve : Soit $f\in L^p(\mathbb{R};\mathbb{K})$. On veut construire une suite de fonctions de $\mathcal{C}_c(\mathbb{R};\mathbb{K})$ qui converge vers f dans L^p. À l'aide du théorème de convergence dominée, il est facile de voir que pour toute fonction f dans L^p, on a

$$\lim_{n\to+\infty}\int_{|x|\geq n}|f(x)|^p\,dx=0.$$

En conséquence, il suffit de traiter le cas où f est à support compact.

Traitons d'abord le cas $p=1$. Supposons donc que $f\in L^1(\mathbb{R};\mathbb{K})$ est nulle en dehors d'un intervalle $[a,b]$. Rappelons qu'on appelle **fonction simple** toute combinaison linéaire de fonctions de type 1_A avec A borélien de mesure finie et que, par construction de l'intégrale de Lebesgue, on peut trouver une suite de fonctions simples nulles en dehors de $[a,b]$ qui converge vers f dans $L^1(\mathbb{R};\mathbb{K})$. Il suffit donc de montrer le résultat de densité dans le cas

6. C'est-à-dire nulles en dehors d'un compact de $\mathbb{R}$.

$f = 1_B$ avec B borélien borné. Mais par construction de la mesure de Lebesgue, pour tout $\varepsilon > 0$ on peut trouver un ouvert Ω de $\mathbb{R}$ contenant B et tel que

$$0 \leq \mu(\Omega) - \mu(B) = \int_{\mathbb{R}} \big(1_\Omega(x) - 1_B(x)\big)\, dx \leq \varepsilon.$$

Donc il suffit de traiter le cas où B est un ouvert borné de $\mathbb{R}$. On utilise ensuite le fait que tout ouvert borné de $\mathbb{R}$ s'écrit comme réunion *finie ou dénombrable* d'intervalles ouverts bornés et deux à deux disjoints.

En effet, pour $x \in B$, on note I_x la réunion de tous les intervalles contenant x et inclus dans B. Comme B est ouvert et tous les I_x sont connexes et contiennent x, l'ensemble I_x est un connexe ouvert de $\mathbb{R}$ inclus dans B. C'est donc un intervalle ouvert. Il est aussi aisé de vérifier que si x et y sont deux points de B alors ou bien $I_x = I_y$, ou bien $I_x \cap I_y = \emptyset$. Cela permet de définir la relation d'équivalence suivante :

$$x \sim y \quad \text{si} \quad I_x = I_y.$$

Comme $\mathbb{Q}$ est dense dans $\mathbb{R}$, on constate que chaque classe d'équivalence pour la relation $\sim$ contient un nombre rationnel. En conséquence, si l'on choisit un nombre rationnel x_n dans chaque classe d'équivalence, on peut écrire

$$B = \bigcup_n I_{x_n}.$$

La réunion est finie et dénombrable puisque $\mathbb{Q}$ est dénombrable. Enfin, par construction, les intervalles sélectionnés sont deux à deux disjoints. Cela permet d'écrire que

$$1_B = \sum_n 1_{I_{x_n}}.$$

Comme 1_B est dans $L^1(\mathbb{R};\mathbb{K})$, le théorème de convergence dominée assure que la série ci-dessus converge au sens de la norme $\|\cdot\|_{L^1}$. Cela permet de se ramener au cas où B est réunion *finie* d'intervalles. Enfin, comme le résultat que nous souhaitons démontrer est stable par combinaison linéaire, il suffit de considérer le cas où B est un intervalle ouvert borné.

En résumé, il s'agit finalement de montrer que la fonction $1_{]a,b[}$ peut être approchée au sens de la norme $\|\cdot\|_{L^1}$ par une suite de fonctions continues nulles en dehors de $[a,b]$. Pour cela, on peut considérer (pour n suffisamment grand) les fonctions f_n continues affines par morceaux, nulles en dehors de $[a,b]$ et valant 1 sur $[a+2^{-n}, b-2^{-n}]$. Un calcul évident (ou même un dessin) permet de vérifier que

$$\|1_{]a,b[} - f_n\| = 2^{1-n},$$

d'où le résultat dans ce cas particulier, et donc le théorème dans le cas $p = 1$.

Traitons maintenant que le cas $p \in]1,\infty[$. Soit donc une fonction f de $L^p(\mathbb{R};\mathbb{K})$ à support compact. Quitte à séparer f en partie réelle et partie imaginaire, on peut se restreindre au cas $\mathbb{K} = \mathbb{R}$. Maintenant, si f est à valeurs réelles alors on peut écrire

$$f = f^+ - f_- \quad \text{avec} \quad f^+ = \sup(0,f) \;\text{ et }\; f^- = \sup(0,-f),$$

et l'on a f dans $L^p(\mathbb{R};\mathbb{R})$ si et seulement si f_- et f_+ le sont. Donc on peut se limiter au cas où f est dans L^p et est positive et nulle en dehors d'un intervalle $[a,b]$.

Pour $N \in \mathbb{N}$, notons $f_N = \inf(f, N)$. On a

$$\mu\big(f^{-1}(]N, +\infty[)\big) = \int_{\{x\in\mathbb{R}\,/\,1<|f(x)|^p/N^p\}} 1\,dx \le \frac{1}{N^p}\|f\|_{L^p}^p.$$

Donc $\lim_{N\to+\infty} \mu\big(f^{-1}(]N, +\infty[)\big) = 0$. Une fois de plus par convergence dominée, on en déduit que $(f_N)_{N\in\mathbb{N}}$ converge vers f dans L^p.

On peut donc se limiter au cas où f est positive, bornée et nulle en dehors d'un intervalle $[a, b]$. Il est alors évident que f est intégrable sur $\mathbb{R}$. Donc il existe une suite de fonctions $(f_n)_{n\in\mathbb{N}}$ continues à support compact (inclus dans $[a, b]$ si f est nulle en dehors de $[a, b]$) qui converge vers f dans L^1. Soit

$$g_n = \sup\big(0, \inf(f_n, \|f\|_{L^\infty})\big).$$

On constate que $|g_n - f| \le |f_n - f|$. Donc $(g_n)_{n\in\mathbb{N}}$ est une suite de fonctions continues à support compact tendant vers f dans L^1 et vérifiant de plus $0 \le g_n \le \|f\|_{L^\infty}$. Sachant que

$$\|f - g_n\|_{L^p}^p = \int_a^b |f(x) - g_n(x)|^p\,dx = \int_a^b |f(x) - g_n(x)|\,|f(x) - g_n(x)|^{p-1}\,dx,$$

on en conclut que

$$\|f - g_n\|_{L^p}^p \le \|f\|_{L^\infty}^{p-1}\|f - g_n\|_{L^1}.$$

Donc $(g_n)_{n\in\mathbb{N}}$ converge vers f dans L^p. ∎

7.8.2 Séries de Fourier

Nous supposons désormais que $p = 2$. Visiblement, $\|\cdot\|_{L^2}$ est la norme associée au produit scalaire

$$(f \mid g) = \int_A f(x)\overline{g(x)}\,dx. \tag{7.11}$$

On en déduit le résultat fondamental suivant :

Théorème. *Pour tout borélien A de $\mathbb{R}$, l'ensemble $L^2(A; \mathbb{K})$ muni du produit scalaire défini en (7.11) est un espace de Hilbert.*

Nous supposons désormais que $A = [a, b]$. Pour des raisons qui apparaîtront un peu plus loin, nous munissons $L^2([a, b]; \mathbb{K})$ du produit scalaire

$$(f \mid g) \overset{\text{déf}}{=} \frac{1}{T}\int_a^b f(x)\overline{g(x)}\,dx \quad \text{avec} \quad T = b - a,$$

ce qui évidemment ne change rien aux propriétés topologiques énoncées jusqu'à présent.

Théorème. *L'ensemble $L^2([a, b]; \mathbb{C})$ est un espace de Hilbert séparable et les fonctions e_p définies par*

$$e_p \;:\; \left\{\begin{array}{lcl} [a, b] & \longrightarrow & \mathbb{C} \\ t & \longmapsto & e^{ip\omega t} \end{array}\right. \qquad \text{avec} \;\; \omega = \frac{2\pi}{T} \quad \text{et} \quad p \in \mathbb{Z}$$

forment une base hilbertienne de $L^2([a, b]; \mathbb{C})$.

Preuve : Par définition du produit scalaire et de la suite $(e_p)_{p\in\mathbb{Z}}$, on pour tout $(p,q)\in\mathbb{Z}^2$,

$$(e_p \mid e_q) = \frac{1}{T}\int_a^b e^{i(p-q)\omega t}\,dt.$$

En calculant l'intégrale, on vérifie que $(e_p)_{p\in\mathbb{Z}}$ est une famille orthonormale.
Pour montrer que cette famille constitue une base hilbertienne, il suffit d'établir que l'orthogonal de $\{e_n \,/\, n\in\mathbb{Z}\}$ est réduit à $\{0\}$. Soit donc $f\in L^2([a,b];\mathbb{C})$ telle que

$$\forall n\in\mathbb{Z},\ (f\mid e_n)=0. \tag{7.12}$$

Fixons $\varepsilon>0$. Si l'on prolonge f par 0 en dehors de $[a,b]$, on obtient une fonction de $L^2(\mathbb{R};\mathbb{C})$. D'après le théorème de densité de la section 7.8.1 il existe donc $g\in\mathcal{C}(\mathbb{R};\mathbb{C})$ nulle en dehors de $]a,b[$ et telle que $\|f-g\|_{L^2}\leq\frac{\varepsilon}{2}$. Comme $g(a)=g(b)=0$, on peut prolonger g en une fonction continue sur $\mathbb{R}$ par périodicité puis la considérer comme une fonction $\widetilde{g}$ définie sur l'ensemble[7] $\mathbb{R}/T\mathbb{Z}$.
L'ensemble $\mathbb{R}/T\mathbb{Z}$ muni de la distance induite par celle de $\mathbb{R}$ est un ensemble métrique compact. À toute fonction e_n, on peut associer son représentant $\widetilde{e}_n$ sur $\mathbb{R}/T\mathbb{Z}$. L'ensemble $\mathcal{T}$ des polynômes trigonométriques de période T (c'est-à-dire le s.e.v. engendré par les $\widetilde{e}_n$) est clairement stable par conjugaison, combinaison linéaire et multiplication. De plus, il contient les fonctions constantes et sépare les points (exercice : vérifier cette dernière propriété). Le théorème de Stone-Weierstrass (cas complexe) assure donc la densité de $\mathcal{T}$ dans $\mathcal{C}(\mathbb{R}/T\mathbb{Z};\mathbb{C})$. En conséquence, il existe un polynôme trigonométrique P tel que $\|g-P\|_{L^2}\leq\frac{\varepsilon}{2}$.
On a donc finalement $\|f-P\|_{L^2}\leq\varepsilon$ avec P polynôme trigonométrique. Écrivons que

$$\|f\|_{L^2}^2=(f\mid f-P)+(f\mid P).$$

D'après (7.12), on a $(f\mid P)=0$. Une application immédiate de l'inégalité de Cauchy-Schwarz assure donc que

$$\|f\|_{L^2}\leq\varepsilon.$$

Le raisonnement étant valable pour tout $\varepsilon>0$, on peut conclure que $f=0$. Donc $(e_n)_{n\in\mathbb{Z}}$ est une base hilbertienne de $L^2([a,b];\mathbb{C})$. Comme conséquence immédiate, on obtient le fait que $L^2([a,b];\mathbb{C})$ est séparable. ■

Maintenant que nous savons que $L^2([a,b];\mathbb{C})$ est un espace de Hilbert séparable et que $(e_n)_{n\in\mathbb{Z}}$ en est une base hilbertienne, la proposition de la page 80 nous donne le résultat suivant :

Corollaire. *Soit $a<b$ et $\omega=2\pi/(b-a)$. Toute fonction f de $L^2([a,b];\mathbb{C})$ peut se décomposer de façon unique en*

$$f=\sum_{p\in\mathbb{Z}}c_p(f)e_p \quad \textit{avec} \quad e_p(t)=e^{ip\omega t},\quad c_p(f)=(f\mid e_p)=\frac{1}{b-a}\int_a^b e^{-ip\omega t}f(t)\,dt, \tag{7.13}$$

et l'on a l'identité de Bessel-Parseval suivante :

$$\frac{1}{b-a}\int_a^b |f(t)|^2\,dt=\sum_{p\in\mathbb{Z}}|c_p(f)|^2.$$

Réciproquement, pour toute suite $(\gamma_p)_{p\in\mathbb{Z}}\in\ell^2(\mathbb{Z})$, la série $\sum\gamma_p e_p$ converge dans $L^2([a,b];\mathbb{C})$ vers une fonction f telle que $c_p(f)=\gamma_p$.

7. Rappelons que cet ensemble est celui des classes d'équivalence de réels pour la relation de congruance modulo T : dire que $x=y$ dans $\mathbb{R}/T\mathbb{Z}$ signifie que $x-y$ est un multiple de T.

Remarque : L'égalité (7.13) doit être comprise dans le sens suivant :

$$\lim_{n\to+\infty} \sum_{p=-n}^{n} c_p(f)e^{ip\omega t} = f$$

au sens de la norme de $L^2([a,b];\mathbb{C})$ (appelée parfois **norme de la convergence en moyenne quadratique**). Cela entraîne bien sûr la convergence en presque tout point. Pour avoir vraiment une convergence en tout point, il faut faire des hypothèses nettement plus fortes sur f (vues en L3 dans le cours *suites et séries de fonctions*).

En identifiant les fonctions T périodiques à leur restriction sur un intervalle d'amplitude T (par exemple $[0,T]$), le corollaire ci-dessus donne :

Corollaire. *Toute fonction f périodique de période T et de carré sommable sur $[0,T]$ peut se décomposer de façon unique en*

$$f = \sum_{p\in\mathbb{Z}} c_p(f)e^{ip\omega t} \quad \textit{avec} \quad c_p(f) = \frac{1}{T}\int_0^T e^{-ip\omega t} f(t)\,dt, \quad \textit{et} \ \ \omega = \frac{2\pi}{T},$$

l'égalité ayant lieu au sens de la convergence des sommes partielles de $-n$ à n dans $L^2([0,T];\mathbb{C})$. De plus on a l'identité de Bessel-Parseval

$$\frac{1}{T}\int_0^T |f(t)|^2\,dt = \sum_{p\in\mathbb{Z}} |c_p(f)|^2.$$

Chapitre 8

Opérateurs compacts sur les espaces de Hilbert

8.1 Opérateurs compacts

Dans toute cette section, H_1 et H_2 sont des espaces de Hilbert, et l'on note $\overline{B}_{H_1}$ et $\overline{B}_{H_2}$ leur boule unité respective.

Définition. On dit qu'une application linéaire T de H_1 dans H_2 est **compacte** si l'image de la boule unité de H_1 par T est relativement compacte dans H_2.

Notation. On note $\mathcal{K}(H_1, H_2)$ l'ensemble des opérateurs compacts de H_1 dans H_2. Dans le cas $H_1 = H_2 = H$, on adopte la notation condensée $\mathcal{K}(H)$.

Retenons la caractérisation suivante, fort utile, des applications linéaires compactes (appelées aussi **opérateurs compacts**).

Proposition. *Une application linéaire T de H_1 dans H_2 est compacte si et seulement si pour toute suite bornée $(x_n)_{n\in\mathbb{N}}$ de H_1, la suite $(T(x_n))_{n\in\mathbb{N}}$ a une valeur d'adhérence dans H_2.*

Preuve : Supposons que T soit compact et considérons une quelconque suite bornée $(x_n)_{n\in\mathbb{N}}$ de H_1. Soit $M > 0$ une borne de cette suite. La suite $(T(M^{-1}x_n))_{n\in\mathbb{N}}$ est incluse dans $T(\overline{B}_{H_1})$ dont l'adhérence est compacte. On peut donc extraire de $(T(M^{-1}x_n))_{n\in\mathbb{N}}$ une sous-suite convergente. Il en va de même pour $(T(x_n))_{n\in\mathbb{N}}$.

Réciproquement, soit $(y_n)_{n\in\mathbb{N}}$ une suite d'éléments de l'adhérence de $T(\overline{B}_{H_1})$. Il existe une suite $(z_n)_{n\in\mathbb{N}}$ d'éléments de $T(\overline{B}_{H_1})$ telle que

$$\forall n \in \mathbb{N},\ \|y_n - z_n\| \leq 2^{-n}.$$

Par hypothèse, la suite $(z_n)_{n\in\mathbb{N}}$ admet une valeur d'adhérence, donc la suite $(y_n)_{n\in\mathbb{N}}$ aussi. ■

Proposition. *L'ensemble $\mathcal{K}(H_1; H_2)$ est un s.e.v. fermé de $\mathcal{L}(H_1; H_2)$.*

Preuve : Tout d'abord, si $T \in \mathcal{K}(H_1, H_2)$ alors $T(\overline{B}_{H_1})$ est un borné de H_2 car est relativement compact dans H_2. Cela assure que $T \in \mathcal{L}(H_1, H_2)$. Le fait que $\mathcal{K}(H_1, H_2)$ soit stable par combinaison linéaire est facile à établir (raisonner avec des suites extraites). Bien sûr $\mathcal{K}(H_1, H_2)$ n'est pas vide (car contient l'application linéaire nulle par exemple) donc $\mathcal{K}(H_1, H_2)$ est bien un s.e.v. de $\mathcal{L}(H_1, H_2)$.

Soit maintenant $(T_n)_{n\in\mathbb{N}}$ une suite d'opérateurs compacts qui tend vers T dans $\mathcal{L}(H_1, H_2)$. Soit $\varepsilon > 0$. Il s'agit de montrer que $T(\overline{B}_{H_1})$ peut être recouvert par un nombre *fini* de boules de H_2 de rayon ε. Fixons un $n_0 \in \mathbb{N}$ tel que $\|T - T_{n_0}\|_{\mathcal{L}(E,F)} < \frac{\varepsilon}{3}$. Comme T_{n_0} est un opérateur compact, il existe une famille finie $(x_1, \cdots, x_N)$ d'éléments de $\overline{B}_{H_1}$ telle que

$$T_{n_0}(\overline{B}_{H_1}) \subset \bigcup_{i=1}^{N} B_{H_2}\big(T_{n_0}(x_i), \tfrac{\varepsilon}{3}\big).$$

Fixons $x \in \overline{B}_{H_1}$ et $i \in \{1, \cdots, N\}$ tel que $\|T_{n_0}(x) - T_{n_0}(x_i)\|_{H_2} \leq \frac{\varepsilon}{3}$. En écrivant que

$$\|T(x) - T(x_i)\|_{H_2} \leq \|T(x) - T_{n_0}(x)\|_{H_2} + \|T_{n_0}(x) - T_{n_0}(x_i)\|_{H_2} + \|T_{n_0}(x_i) - T(x_i)\|_{H_2},$$

on conclut que

$$T(\overline{B}_{H_1}) \subset \bigcup_{i=1}^{N} B_{H_2}(T(x_i), \varepsilon),$$

ce qui achève la preuve de la compacité de T. ■

Proposition. *La composée d'un opérateur compact et d'une application linéaire continue (dans un sens ou dans l'autre) est encore un opérateur compact.*

Preuve : Soit H_1, H_2 et H_3 trois espaces de Hilbert, $T_1 \in \mathcal{L}(H_1, H_2)$ et $T_2 \in \mathcal{K}(H_2, H_3)$. Comme T_1 est continu, il existe $c > 0$ tel que $T_1(c\overline{B}_{H_1}) \subset \overline{B}_{H_2}$. Par compacité de T_2, on en déduit que $T_2 \circ T_1(c\overline{B}_{H_1})$ est relativement compact dans H_3. Il en est bien sûr de même pour $T_2 \circ T_1(\overline{B}_{H_1})$. Donc $T_2 \circ T_1$ est un opérateur compact.

Le cas $T_1 \in \mathcal{K}(H_1, H_2)$ et $T_2 \in \mathcal{L}(H_2, H_3)$ est laissé en exercice. ■

Définition. On dit que $T \in \mathcal{L}(H_1, H_2)$ est un opérateur **de rang fini** si $\dim \operatorname{Im} T < \infty$.

Proposition. *Un opérateur de rang fini est compact.*

Preuve : Il suffit de remarquer que si T est de rang fini alors $T(\overline{B}_{H_1})$ est un borné de l'espace vectoriel de dimension finie $\operatorname{Im} T$, donc est relativement compact. ■

Théorème. *Une application linéaire entre espaces de Hilbert est compacte si et seulement si elle est la limite d'une suite d'opérateurs de rang fini.*

Preuve : La partie directe est facile : un opérateur de rang fini étant compact et l'ensemble des opérateurs compacts étant un fermé de $\mathcal{L}(H_1; H_2)$, une limite d'une suite d'opérateurs de rang fini est bien un opérateur compact.

Montrons la réciproque. Soit donc $T \in \mathcal{K}(H_1; H_2)$. À $n \in \mathbb{N}$ fixé, il existe un nombre fini N_n de points x_i^n de H_1 tels que

$$T(\overline{B}_{H_1}) \subset \bigcup_{i=1}^{N_n} B_{H_2}\big(T(x_i^n), 2^{-n}\big).$$

Notons G_n le s.e.v. (fermé car de dimension finie) engendré par $\big(T(x_1^n), \cdots, T(x_{N_n}^n)\big)$, p_n le projecteur orthogonal sur G_n et $T_n = p_n \circ T$. Par construction, T_n est continu et de rang fini. De plus, compte tenu de $p_n(T(x_i^n)) = T(x_i^n)$,

$$\forall x \in H_1, \forall i \in \{1, \cdots, N_n\},\ \|T_n(x) - T(x)\|_{H_2} \leq \|p_n(T(x) - T(x_i^n))\|_{H_2} + \|T(x_i^n) - T(x)\|_{H_2}.$$

Si x est dans $\overline{B}_{H_1}$, on peut choisir $i \in \{1, \cdots, N_n\}$ de telle sorte que $T(x)$ soit dans $B_{H_2}(T(x_i^n), 2^{-n})$. Alors, comme un projecteur orthogonal est toujours de norme inférieure ou égale à 1 [1], on obtient finalement,

$$\forall x \in \overline{B}_{H_1},\ \|T_n(x) - T(x)\|_{H_2} \leq 2^{1-n}.$$

En conséquence, la suite d'opérateurs de rang fini $(T_n)_{n\in\mathbb{N}}$ tend vers T. ■

Exemple. Soit $a < b$ deux réels. Rappelons que l'ensemble $L^2 \stackrel{\text{déf}}{=} L^2([a,b];\mathbb{C})$ muni du produit scalaire

$$(f \mid g)_{L^2} = \frac{1}{b-a}\int_a^b f(t)\overline{g}(t)\,dt$$

est un espace de Hilbert séparable, et que $(e_p)_{p\in\mathbb{Z}}$ avec $e_p(t) = e^{ip\omega t}$ et $\omega = 2\pi/(b-a)$ en est une base hilbertienne. Autrement dit, toute fonction f de L^2 s'écrit

$$f = \sum_{p\in\mathbb{Z}} c_p(f)e_p \quad \text{avec} \quad c_p(f) = \frac{1}{b-a}\int_a^b f(t)e^{-ip\omega t}\,dt$$

et l'on a d'après l'égalité de Parseval,

$$(f \mid g)_{L^2} = \sum_{p\in\mathbb{Z}} c_p(f)\overline{c_p(g)}.$$

Soit H^1 l'ensemble des fonctions f de L^2 telles que

$$\sum_{p\in\mathbb{Z}}(1+\omega^2p^2)|c_p(f)|^2 < \infty.$$

Il n'est pas difficile de montrer (exercice : le faire) que H^1 muni du produit scalaire

$$(f \mid g)_{H^1} = \sum_{p\in\mathbb{Z}}(1+\omega^2p^2)\,c_p(f)\overline{c_p(g)}$$

est aussi un espace de Hilbert séparable, et que la suite constituée des vecteurs $\frac{e_p}{\sqrt{1+\omega^2p^2}}$ en est une base hilbertienne. Enfin, il est évident que $\|f\|_{L^2} \leq \|f\|_{H^1}$ pour tout $f \in H^1$.

Il n'est pas difficile de démontrer à l'aide du théorème de Dirichlet pour les séries de Fourier que si f et g admettent un prolongement périodique de classe C^1 et C^2 par morceaux alors

$$(f \mid g)_{H^1} = (f \mid g)_{L^2} + (f' \mid g')_{L^2}.$$

On verra dans le cours d'EDP que H^1 est l'ensemble des éléments de L^2 dont la dérivée *au sens des distributions* appartient à L^2. L'espace H^1 est appelé **espace de Sobolev**.

Proposition. *L'application identité de H^1 dans L^2 est compacte.*

Preuve : Notons $T(f) = f$ pour $f \in H^1$. Comme $\|f\|_{L^2} \leq \|f\|_{H^1}$, l'application T est linéaire continue. Nous allons montrer que T est limite d'opérateurs de rang fini. Cela donnera le résultat voulu.

Pour $f \in H^1$, nous posons donc

$$T_n(f) = \sum_{p=-n}^{n} c_p(f)e_p.$$

1. En effet, si p est un projecteur orthogonal de l'espace de Hilbert H alors le théorème de Pythagore assure que pour tout $x \in H$, on a $\|x\|^2 = \|p(x)\|^2 + \|x - p(x)\|^2$, et donc $\|p(x)\| \leq \|x\|$.

Il est évident que T_n est de rang fini (son image est le s.e.v. de dimension $2n+1$ engendré par $(e_{-n}, \cdots, e_n)$). De plus, pour tout $f \in H^1$, on a en vertu de l'égalité de Parseval,

$$\begin{array}{rcl} \|f - T_n(f)\|_{L^2}^2 & = & \sum_{|p|>n} |c_p(f)|^2, \\ & \leq & (1+\omega^2 n^2)^{-1} \sum_{|p|>n} (1+\omega^2 p^2)|c_p(f)|^2, \\ & \leq & (1+\omega^2 n^2)^{-1} \|f\|_{H^1}^2. \end{array}$$

Cela assure la convergence de la suite $(T_n)_{n\in\mathbb{N}}$ vers T. ■

8.2 Opérateurs adjoints

Dans toute cette section, H_1 et H_2 désignent deux espaces de Hilbert. On dira que $T : H_1 \to H_2$ est un **opérateur borné** de H_1 dans H_2 si T est une application linéaire continue de H_1 dans H_2. La définition ci-dessous généralise aux espaces de Hilbert la notion d'endomorphisme adjoint vue en L2 dans le cadre des espaces euclidiens.

Théorème. *Soit T un opérateur borné de H_1 dans H_2. Il existe un et un seul opérateur borné T^* de H_2 dans H_1 vérifiant*

$$\forall (f,g) \in H_1 \times H_2,\ (T(f) \mid g)_{H_2} = (f \mid T^*(g))_{H_1}.$$

L'opérateur T^* est appelé **opérateur adjoint** *de T, et l'on a $\|T\|_{\mathcal{L}(H_1;H_2)} = \|T^*\|_{\mathcal{L}(H_2;H_1)}$.*

Preuve : À $g \in H_2$ fixé, considérons la forme linéaire L sur H_1 définie par $L_g(f) = (T(f) \mid g)$. D'après l'inégalité de Cauchy-Schwarz dans H_2, $|L_g(f)| \leq \|T\|_{\mathcal{L}(H_1;H_2)} \|g\|_{H_2} \|f\|_{H_1}$. Donc L_g est continue. Le théorème de représentation de Riesz-Fréchet assure donc l'existence d'un unique élément $T^*(g)$ de H_1 tel que

$$\forall f \in H_1,\ (T(f) \mid g)_{H_2} = (f \mid T^*(g))_{H_1}.$$

De plus, ce même théorème donne $\|T^*(g)\|_{H_1} = \|L_g\|_{H_1'} \leq \|T\|_{\mathcal{L}(H_1;H_2)} \|g\|_{H_2}$. Vérifions la linéarité de l'application $g \mapsto T^*(g)$. Pour $(\lambda_1, \lambda_2) \in \mathbb{K}^2$ et $(g_1, g_2) \in H_2^2$, on a par définition de T^*,

$$\begin{array}{rcl} \forall f \in H_1,\ \big(f \mid T^*(\lambda_1 g_1 + \lambda_2 g_2)\big)_{H_1} & = & (T(f) \mid \lambda_1 g_1 + \lambda_2 g_2)_{H_2}, \\ & = & \overline{\lambda_1}(T(f) \mid g_1)_{H_2} + \overline{\lambda_2}(T(f) \mid g_2)_{H_2}, \\ & = & \overline{\lambda_1}(f \mid T^*(g_1))_{H_1} + \overline{\lambda_2}(f \mid T^*(g_2))_{H_1}, \\ & = & \big(f \mid (\lambda_1 T^*(g_1) + \lambda_2 T^*(g_2))\big)_{H_1}. \end{array}$$

Donc $T^*(\lambda_1 g_1 + \lambda_2 g_2) = \lambda_1 T^*(g_1) + \lambda_2 T^*(g_2)$.
Reste à montrer que T et T^* ont la même norme. On a déjà vu que

$$\forall g \in H_2,\ \|T^*(g)\|_{H_1} \leq \|T\|_{\mathcal{L}(H_1;H_2)} \|g\|_{H_2}.$$

Donc $\|T^*\|_{\mathcal{L}(H_2;H_1)} \leq \|T\|_{\mathcal{L}(H_1;H_2)}$.
Mais en utilisant l'inégalité de Cauchy-Schwarz et la définition de l'adjoint, on obtient

$$\|T(f)\|_{H_2}^2 = \big|(T(f)|T(f))_{H_2}\big| = \big|(f|T^*(T(f)))\big| \leq \|T^*\|_{\mathcal{L}(H_2;H_1)} \|f\|_{H_1} \|T(f)\|_{H_2},$$

ce qui assure visiblement que $\|T\|_{\mathcal{L}(H_1;H_2)} \leq \|T^*\|_{\mathcal{L}(H_2;H_1)}$. ■

Remarque : Retenons également que $(T^*)^* = T$. En effet, par définition de $(T^*)^*$, on a pour tout $(f,g) \in H_1 \times H_2$,

$$\big(g|(T^*)^*(f)\big)_{H_2} = (T^*(g)|f)_{H_1}.$$

Mais en prenant le conjugué de cette égalité et en utilisant la sesquilinéarité, on obtient

$$\big((T^*)^*(f)|g\big)_{H_2} = (f|T^*(g))_{H_1}.$$

Par définition de T^*, le membre de gauche vaut $(T(f)|g)$. Donc on a $(T^*)^*(f) = T(f)$.

Exemple. Soit I un intervalle de $\mathbb{R}$ et a une fonction bornée sur I et à valeurs dans $\mathbb{C}$. Prenons $H = L^2(I;\mathbb{C})$. Il est clair que l'application linéaire $T : f \mapsto af$ est linéaire continue de H dans H. En effet, d'après l'inégalité de Hölder,

$$\forall f \in H,\ \|T(f)\|_{L^2}^2 = \int_I |a(x)f(x)|^2\,dx \le \|a\|_{L^\infty}^2 \|f\|_{L^2}^2.$$

De plus, on a

$$\forall (f,g) \in H^2,\ (T(f) \mid g) = \int_I a(x)f(x)\,\overline{g(x)}\,dx = \int_I f(x)\,\overline{\overline{a}(x)g(x)}\,dx.$$

Donc T^* est l'opérateur de multiplication par la fonction $\overline{a}$.

Définition. On dit qu'une application linéaire $T : H_1 \to H_2$ est **faiblement continue** si pour toute suite $(x_n)_{n\in\mathbb{N}} \in H_1^{\mathbb{N}}$, on a $x_n \rightharpoonup x$ implique $T(x_n) \rightharpoonup T(x)$.

La définition de l'adjoint va nous permettre de montrer que pour les applications linéaires, la continuité faible est équivalente à la continuité forte.

Proposition. *Une application linéaire est continue si et seulement si elle est* **faiblement continue**.

Preuve : Soit $T : H_1 \to H_2$ continue et $(x_n)_{n\in\mathbb{N}} \in H_1^{\mathbb{N}}$ telle que $x_n \rightharpoonup x$. On a donc pour tout $y \in H_2$,

$$(T(x_n) \mid y)_{H_2} = (x_n \mid T^*(y))_{H_1} \longrightarrow_{n\to+\infty} (x \mid T^*(y))_{H_1} = (T(x) \mid y)_{H_2}.$$

Donc T est faiblement continue.

Réciproquement, supposons T faiblement continue. Soit $(x_n)_{n\in\mathbb{N}}$ une suite d'éléments de H_1 telle que $(x_n, T(x_n))_{n\in\mathbb{N}}$ converge fortement vers (x,y) dans $H_1 \times H_2$. Alors on a aussi $x_n \rightharpoonup x$ et donc $T(x_n) \rightharpoonup T(x)$. Mais par hypothèse $T(x_n) \to y$, et donc a fortiori $T(x_n) \rightharpoonup y$. Par unicité de la limite faible, on en déduit que $y = T(x)$. En conséquence, le graphe de T est fermé. Comme H_1 et H_2 sont complets, le théorème du graphe fermé permet de conclure que l'opérateur T est continu. ■

Examinons maintenant l'effet d'un opérateur *compact* sur la convergence faible.

Proposition. *Soit T un opérateur continu de l'espace de Hilbert H_1 dans l'espace de Hilbert H_2. Les deux énoncés suivants sont équivalents :*

(i) L'opérateur T est compact.

(ii) Pour toute suite $(x_n)_{n\in\mathbb{N}} \in H_1^{\mathbb{N}}$, on a $\big(x_n \rightharpoonup x\big) \Longrightarrow \big(T(x_n) \to T(x)\big)$.

Preuve : Supposons d'abord que T soit compact. Soit $(x_n) \in H_1^{\mathbb{N}}$ une suite faiblement convergente vers $x \in H_1$. Un opérateur compact étant continu, on a $T(x_n) \rightharpoonup T(x)$ d'après la proposition précédente. Mais toute suite faiblement convergente est bornée. Donc $(T(x_n))_{n\in\mathbb{N}}$ admet une sous-suite fortement convergente. La limite de cette sous-suite ne peut être que $T(x)$. Finalement $(T(x_n))_{n\in\mathbb{N}}$ est à valeurs dans le compact $\overline{T(\overline{B}_{H_1})}$ et a pour unique valeur d'adhérence $T(x)$. Donc la suite toute entière converge fortement vers $T(x)$.

Réciproquement supposons que la propriété (ii) soit vérifiée. Soit $(x_n)_{n\in\mathbb{N}}$ une quelconque suite bornée de H_1. Alors il existe $x \in H_1$ et une sous-suite $(x_{\varphi(n)})_{n\in\mathbb{N}}$ telles que $x_{\varphi(n)} \rightharpoonup x$. D'après l'hypothèse, on a donc $T(x_{\varphi(n)}) \to T(x)$. En conséquence, l'opérateur T est compact. ■

Proposition. *Soit T une application linéaire continue de H_1 dans H_2. Alors T est compact si et seulement si T^* est compact.*

Preuve : Supposons que $T : H_1 \to H_2$ soit compact. Soit $(y_n)_{n\in\mathbb{N}}$ une suite faiblement convergente d'éléments de H_2. Notons y sa limite faible. Comme T^* est faiblement continu car linéaire continu, on a $T^*(y_n) \rightharpoonup T^*(y)$. Donc, par compacité de T, on a aussi $T(T^*(y_n)) \to T(T^*(y))$ Par ailleurs,

$$\|T^*(y_n)\|_{H_1}^2 = (T^*(y_n) \mid T^*(y_n))_{H_1} = (y_n \mid T(T^*(y_n)))_{H_2} \to (y \mid T(T^*(y)))_{H_2} = \|T^*(y)\|_{H_1}^2$$

car $y_n \rightharpoonup y$ et $T(T^*(y_n)) \to T(T^*(y))$.
On a donc à la fois $\|T^*(y_n)\|_{H_1} \to \|T^*(y)\|_{H_1}$ et $T^*(y_n) \rightharpoonup T^*(y)$, ce qui entraîne $T^*(y_n) \to T^*(y)$. Donc T^* est compact.

La réciproque se démontre en appliquant le raisonnement précédent à T^* et en utilisant le fait que $(T^*)^* = T$. ■

8.3 Alternative de Fredholm

Pour simplifier, on suppose dans toute cette section que $H_1 = H_2 = H$ espace de Hilbert [2].

Proposition. *Soit T un opérateur borné sur H. On a les propriétés suivantes.*

$$\begin{array}{llll} (i) & \operatorname{Ker} T = (\operatorname{Im} T^*)^{\perp}, & (ii) & \operatorname{Ker} T^* = (\operatorname{Im} T)^{\perp}, \\ (iii) & \overline{\operatorname{Im} T} = (\operatorname{Ker} T^*)^{\perp}, & (iv) & \overline{\operatorname{Im} T^*} = (\operatorname{Ker} T)^{\perp}. \end{array}$$

Preuve : Le fait que $(T^*)^* = T$ assure que les propriétés (i) et (ii) (resp. (iii) et (iv)) sont équivalentes. Nous nous bornerons donc à la démonstration de (i) et de (iii).

– Démonstration de (i) : Soit $x \in \operatorname{Ker} T$ et $z \in \operatorname{Im} T^*$. Fixons $y \in H$ tel que $z = T^*(y)$. Puisque $T(x) = 0$, on a

$$(x \mid z) = (x \mid T^*(y)) = (T(x) \mid y) = 0.$$

Donc $x \in (\operatorname{Im} T^*)^{\perp}$.
Réciproquement, soit $x \in (\operatorname{Im} T^*)^{\perp}$. Alors pout tout $y \in H$, on a

$$0 = (x \mid T^*(y)) = (T(x) \mid y).$$

Donc $T(x) = 0$. Donc $(\operatorname{Im} T^*)^{\perp} \subset \operatorname{Ker} T$.

2. Mais la proposition suivante demeure valable si $H_1 \neq H_2$.

– Démonstration de (iii) : En prenant l'orthogonal de l'égalité (ii), on obtient

$$\left(\operatorname{Ker} T^*\right)^{\perp} = \left((\operatorname{Im} T)^{\perp}\right)^{\perp}.$$

Or dans un espace de Hilbert, tout sous-espace vectoriel F vérifie $(F^{\perp})^{\perp} = \overline{F}$, d'où le résultat.

■

Notation. dans tout ce qui suit, I désigne l'opérateur identité de H défini par $I(x) = x$.

Théorème (Alternative de Fredholm)**.** *Soit T un opérateur compact sur l'espace de Hilbert H. Alors les propriétés suivantes sont vérifiées.*

(i) $\operatorname{Ker}(I - T)$ *est de dimension finie.*

(ii) $\operatorname{Im}(I - T)$ *est fermé.*

(iii) Alternative de Fredholm : $\operatorname{Ker}(I - T) = \{0\}$ *si et seulement si* $\operatorname{Im}(I - T) = H$.

Preuve : Soit $(x_n)_{n\in\mathbb{N}}$ une suite bornée de $\operatorname{Ker}(I - T)$. On a donc $x_n = T(x_n)$ pour tout $n \in \mathbb{N}$. Mais T est un opérateur compact donc la suite $(T(x_n))_{n\in\mathbb{N}}$ admet une sous-suite $(T(x_{\varphi(n)}))_{n\in\mathbb{N}}$ convergente. Cela assure que $(x_{\varphi(n)})_{n\in\mathbb{N}}$ est convergente. On en déduit que tout ensemble borné de $\operatorname{Ker}(I - T)$ est relativement compact. Donc $\operatorname{Ker}(I - T)$ est de dimension finie (cf th. de Riesz).

Pour montrer (ii), nous considérons une suite $(x_n)_{n\in\mathbb{N}}$ telle que $((I-T)(x_n))_{n\in\mathbb{N}}$ converge vers $y \in H$. Il s'agit de démontrer que $y \in \operatorname{Im}(I - T)$.

Soit z_n la projection orthogonale de x_n sur le sous-espace vectoriel fermé $\operatorname{Ker}(I - T)$. On a donc $d(x_n, \operatorname{Ker}(I - T)) = \|x_n - z_n\|$ et $(I - T)(x_n - z_n) = (I - T)(x_n)$. Notons $x'_n = x_n - z_n$. Si l'on parvient à démontrer que la suite $(x'_n)_{n\in\mathbb{N}}$ est bornée alors on obtient $y \in \operatorname{Im}(I - T)$. En effet, par compacité de T, il existera une extraction φ telle que $T(x'_{\varphi(n)})_{n\in\mathbb{N}}$ converge, ce qui entraînera aussi la convergence de $(x'_{\varphi(n)})_{n\in\mathbb{N}}$.

Supposons par l'absurde que $(x'_n)_{n\in\mathbb{N}}$ n'est pas borné. Alors il existe une extraction φ telle que $(\|x'_{\varphi(n)}\|)_{n\in\mathbb{N}}$ soit une suite strictement positive tendant vers l'infini. Soit $\omega_n = x'_{\varphi(n)}/\|x'_{\varphi(n)}\|$. Il est clair que

$$\lim_{n\to+\infty}(I - T)(\omega_n) = 0. \tag{8.1}$$

La suite $(\omega_n)_{n\in\mathbb{N}}$ étant bornée, on peut supposer, quitte à extraire à nouveau une sous-suite que $(T(\omega_n))_{n\in\mathbb{N}}$ est convergente. L'égalité (8.1) implique alors que $(\omega_n)_{n\in\mathbb{N}}$ converge aussi vers un élément ω appartenant à $\operatorname{Ker}(I - T)$. Mais, par construction $\omega_n \in \operatorname{Ker}(I - T)^{\perp}$ donc $\omega \in \operatorname{Ker}(I - T)^{\perp}$, d'où $\omega = 0$. Comme tous les ω_n sont de norme 1, on doit avoir $\|\omega\| = 1$, ce qui est impossible.

Démontrons (iii). Supposons d'abord que $\operatorname{Ker}(I-T) = \{0\}$. Supposons par l'absurde que $I - T$ ne soit pas surjectif. Notons $E_1 \stackrel{\text{déf}}{=} \operatorname{Im}(I - T)$, puis $E_2 \stackrel{\text{déf}}{=} (I - T)(E_1)$, etc. D'après *(ii)*, $(E_n)_{n\in\mathbb{N}}$ est une suite de s.e.v. fermés de H. Comme $I-T$ est injectif et non surjectif, la suite $(E_n)_{n\in\mathbb{N}}$ est strictement décroissante (au sens de l'inclusion). En effet, si tel n'était pas le cas, il existerait un indice $n_0 \geq 1$ tel que $(I - T)(E_{n_0}) = (I - T)(E_{n_0-1}) = E_{n_0}$ mais E_{n_0} strictement inclus dans E_{n_0-1}. Cela contredirait l'injectivité.

Finalement donc la suite $(E_n)_{n\in\mathbb{N}}$ est strictement décroissante, et l'on peut appliquer le théorème de Riesz du chapitre 1. On obtient une suite $(x_n)_{n\in\mathbb{N}}$ de H telle que[3]

$$\forall n \in \mathbb{N},\ \|x_n\| = 1, \quad x_n \in E_n \quad \text{et} \quad d(x_n, E_{n+1}) \geq \frac{1}{2}.$$

3. En fait dans un Hilbert, on peut se passer du théorème de Riesz et remplacer le facteur $\frac{1}{2}$ ci-dessous par 1 (exo : pourquoi ?)

Pour $n > m$, on a

$$T(x_n) - T(x_m) = \underbrace{(I-T)(x_m) - (I-T)(x_n) + x_n}_{\in E_{m+1}} - x_m.$$

Donc $\|T(x_n) - T(x_m)\| \geq \frac{1}{2}$ et $(T(x_n))_{n\in\mathbb{N}}$ n'a donc pas de valeur d'adhérence. Cela contredit la compacité de T.

Réciproquement, supposons que $\mathrm{Im}\,(I-T) = H$. Alors la proposition précédente montre que $\mathrm{Ker}\,(I-T^*) = \{0\}$. Mais l'opérateur T^* est aussi compact, donc en appliquant le raisonnement ci-dessus à T^* au lieu de T, on obtient $\mathrm{Im}\,(I-T^*) = H$ puis, en passant à l'orthogonal, $\mathrm{Ker}\,(I-T) = \{0\}$. ∎

Proposition. *Soit T un opérateur compact sur l'espace de Hilbert H. Alors on a*

$$\dim(\mathrm{Ker}\,(I-T)) = \dim(\mathrm{Ker}\,(I-T^*)).$$

Preuve : Comme T et T^* sont compacts, on sait déjà que les deux s.e.v. considérés sont de dimension finie. Notons d et d^* leurs dimensions respectives.

Comme $\mathrm{Im}\,(I-T)$ est fermé, on peut décomposer H en

$$H = (\mathrm{Im}\,(I-T))^\perp \oplus \mathrm{Im}\,(I-T) = \mathrm{Ker}\,(I-T^*) \oplus \mathrm{Im}\,(I-T).$$

Supposons par l'absurde que $d < d^*$. Il existe alors une application linéaire Λ continue injective (mais non surjective) de $\mathrm{Ker}\,(I-T)$ dans $\mathrm{Ker}\,(I-T^*)$. Soit $S = T + \Lambda \circ P$ où P désigne le projecteur orthogonal sur $\mathrm{Ker}\,(I-T)$. Clairement S est compact car somme d'un opérateur compact et d'un opérateur de rang fini. Par ailleurs, si $(I-S)(x) = 0$ alors

$$(I-T)(x) = \Lambda \circ P(x).$$

Mais le membre de gauche appartient à $\mathrm{Im}\,(I-T)$ alors que le membre de droite appartient au sous-espace $\mathrm{Ker}\,(I-T^*)$ qui lui est orthogonal. Donc $\Lambda(P(x)) = 0$ puis $P(x) = x = 0$ par injectivité de Λ et parce que $x \in \mathrm{Ker}\,(I-T)$. Donc, en vertu de l'alternative de Fredholm, $(I-S)$ est bijective. Soit donc $y \in \mathrm{Ker}\,(I-T^*) \setminus \mathrm{Im}\,\Lambda$ et $x \in H$ tel que $(I-S)(x) = y$. On a

$$(I-S)(x) = \underbrace{(I-T)(x)}_{\in \mathrm{Im}\,(I-T)} - \underbrace{\Lambda(P(x))}_{\in \mathrm{Ker}\,(I-T^*)} = y.$$

Comme $\mathrm{Im}\,(I-T) \cap \mathrm{Ker}\,(I-T^*) = \{0\}$ et $y \in \mathrm{Ker}\,(I-T^*)$, cela entraîne que $(I-T)(x) = 0$, puis $y \in \mathrm{Im}\,\Lambda$, ce qui est contraire à l'hypothèse sur y. Conclusion : $d \geq d^*$.

L'autre inégalité se démontre en appliquant ce qui précède à T^*. ∎

8.4 Spectre des opérateurs

Définition. Soit E un espace vectoriel sur $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$, et T un endomorphisme de E. On appelle :

- **spectre** de T l'ensemble $\sigma(T)$ des $\lambda \in \mathbb{K}$ tels que $T - \lambda I$ ne soit pas bijectif de E dans E,
- **spectre ponctuel** de T l'ensemble $\sigma_p(T)$ des valeurs propres de T,
- **ensemble résolvant** de T le complémentaire $\rho(T)$ du spectre $\sigma(T)$.

Remarque. En dimension finie, on a bien sûr $\sigma(T) = \sigma_p(T)$. C'est une conséquence triviale du théorème du rang. En dimension infinie en revanche, l'ensemble $\sigma_p(T)$ peut être strictement inclus dans $\sigma(T)$ (par exemple pour $E = \mathbb{R}[X]$ et $T : P \mapsto XP$, on a $0 \in \sigma(T) \setminus \sigma_P(T)$).

Proposition 1. *Soit E un Banach sur $\mathbb{K}$ et $T \in \mathcal{L}(E)$. Alors $\sigma(T)$ est un compact de $\mathbb{K}$ borné par $\|T\|_{\mathcal{L}(E)}$.*

Preuve : Soit $\lambda \in \mathbb{R}$ tel que $|\lambda| > \|T\|_{\mathcal{L}(E)}$. On a

$$T - \lambda I = -\lambda(I - \lambda^{-1}T).$$

Du fait que $\|\lambda^{-1}T\|_{\mathcal{L}(E)} < 1$, la série $\sum(\lambda^{-1}T)^k$ est absolument convergente donc convergente car E est un Banach (voir la section 3.4). On vérifie aisément que la somme de cette série est un isormorphisme continu de E qui est l'inverse de $I - \lambda^{-1}T$. En conséquence $I - \lambda^{-1}T$ est inversible et donc $\lambda \notin \sigma(T)$.

Montrons maintenant que $\mathbb{R} \setminus \sigma(T)$ est un ouvert. Soit donc $\lambda_0 \in \mathbb{R} \setminus \sigma(T)$. On a

$$T - \lambda I = T - \lambda_0 I + (\lambda_0 - \lambda)I.$$

Par hypothèse, l'endomorphisme $S \stackrel{\text{déf}}{=} T - \lambda_0 I$ est inversible. Son inverse est continu (grâce au théorème de Banach). En raisonnant comme précédemment, on en déduit que $T - \lambda I$ est également inversible dès que $|\lambda_0 - \lambda|\|S^{-1}\|_{\mathcal{L}(E)} < 1$.

On peut maintenant conclure que $\sigma(T)$ est un fermé borné de $\mathbb{R}$, donc un compact. ■

Théorème. *Soit H un espace de Hilbert et $T \in \mathcal{K}(H)$.*

(i) Si H est de dimension infinie alors $0 \in \sigma(T)$.

(ii) Si $\lambda \in \sigma(T) \setminus \{0\}$ alors λ est valeur propre de T de multiplicité finie.

(iii) Si $\sigma(T)$ contient une suite d'éléments deux à deux distincts alors cette suite tend vers 0.

Preuve :

(i) Supposons que $0 \notin \sigma(T)$. Alors T est inversible d'inverse continu (grâce au théorème de Banach). En conséquence, $I = T^{-1} \circ T$ est compact car composée d'un opérateur compact et d'un opérateur continu. Donc $\overline{B}_H = I(\overline{B}_H)$ est compacte. Le théorème de Riesz assure alors que H est de dimension finie.

(ii) Soit $\lambda \in \mathbb{R} \setminus \{0\}$. Si λ n'est pas valeur propre de T alors $I - \lambda^{-1}T$ est injectif, donc aussi surjectif d'après l'alternative de Fredhlom. Donc $\lambda \notin \sigma(T)$. Dans le cas contraire, on a vu que λ est de multiplicité finie.

(iii) Soit $(\lambda_n)_{n\in\mathbb{N}}$ une suite d'éléments deux à deux distincts de $\sigma(T)$. On sait que cette suite est bornée donc admet une valeur d'adhérence. Quitte à extraire, on peut supposer que $\lambda_n \neq 0$ pour tout $n \in \mathbb{N}$ et que $(\lambda_n)_{n\in\mathbb{N}}$ tend vers une limite λ.

Fixons pour tout $n \in \mathbb{N}$, un vecteur e_n de norme 1 tel que $T(e_n) = \lambda_n e_n$. On montre aisément (comme en dimension finie) que $(e_n)_{n\in\mathbb{N}}$ est une famille libre. Posons $E_n = \operatorname{Vect}(e_0, \cdots, e_n)$. La suite $(E_n)_{n\in\mathbb{N}}$ est une suite strictement croissante de s.e.v. fermés de H. Appliquons le théorème de Riesz : il existe $(x_n)_{n\in\mathbb{N}^*}$ telle que

$$\forall n \in \mathbb{N},\ \|x_n\| = 1, \quad x_n \in E_n \quad \text{et} \quad d(x_n, E_{n-1}) \geq \frac{1}{2}.$$

En écrivant que

$$\frac{T(x_n)}{\lambda_n} - \frac{T(x_m)}{\lambda_m} = \underbrace{\frac{T(x_n) - \lambda_n x_n}{\lambda_n} - \frac{T(x_m) - \lambda_m x_m}{\lambda_m} - x_m}_{\in E_{n-1}} + x_n,$$

on en déduit que $\|\lambda_n^{-1}T(x_n) - \lambda_n^{-1}T(x_m)\| \geq \frac{1}{2}$ pour $n > m$.

Donc la suite $(\lambda_n^{-1}T(x_n))_{n\in\mathbb{N}}$ n'a pas de valeur d'adhérence. Comme T est compact, cela est incompatible avec le fait que $\lambda \neq 0$. Donc $\lambda = 0$.

■

8.5 Les opérateurs auto-adjoints

Définition. On dit que $T \in \mathcal{L}(H)$ est un **opérateur auto-adjoint** si $T = T^*$, c'est-à-dire :

$$\forall (x,y) \in H^2,\ (T(x) \mid y) = (x \mid T(y)).$$

Nous laissons le soin au lecteur de vérifier que si T est auto-adjoint alors pour tout $x \in H$, $(T(x) \mid x)$ est réel (même si H est un espace de Hilbert *complexe*).

Comme en dimension finie avec les endomorphismes symétriques, on peut définir les notions de positivité et de négativité pour les opérateurs auto-adjoints.

Définition. On dit qu'un opérateur auto-adjoint T est **positif** si

$$\forall x \in H,\ (T(x) \mid x) \geq 0.$$

On dit qu'il est **défini positif** si l'inégalité ci-dessus est stricte pour $x \neq 0$.

On dit que T est **négatif** (resp. **défini négatif**) si $-T$ est positif (resp. défini positif).

Proposition. *Si T est opérateur auto-adjoint alors toutes ses valeurs propres sont réelles. Si de plus il est positif (resp. défini positif) alors toutes ses valeurs propres sont positives (resp. strictement positives).*

Preuve : Soit $\lambda \in \sigma_p(T)$ et $x \neq 0$ tel que $T(x) = \lambda x$. Montrons d'abord que si T est auto-adjoint alors λ est réelle. On a

$$\lambda \|x\|^2 = (\lambda x \mid x) = (T(x) \mid x) = (x \mid T(x)) = (x \mid \lambda x) = \overline{\lambda} \|x\|^2,$$

donc $\lambda = \overline{\lambda}$.

Si de plus T est défini positif alors on a

$$\lambda \|x\|^2 = (T(x) \mid x) > 0,$$

d'où le résultat. Les autres cas se traitent de façon similaire. ■

Proposition 2. *Soit T un opérateur auto-adjoint sur un espace de Hilbert réel. Notons*

$$m = \inf_{\|x\|=1} (T(x) \mid x) \quad \textit{et} \quad M = \sup_{\|x\|=1} (T(x) \mid x).$$

Alors $\sigma(T)$ est inclus dans l'intervalle $[m, M]$, et contient m et M.

Preuve : Concentrons-nous sur la borne supérieure M, la démonstration pour m étant similaire. De la définition de M, on déduit que $(T(x)|x) \leq M\|x\|_H^2$ pour tout $x \in H$. Pour $\lambda > M$, on a donc

$$\forall x \in H,\ (\lambda x - T(x)|x) \geq (\lambda - M)\|x\|^2. \tag{8.2}$$

Introduisons la forme bilinéaire a défine par

$$\forall (x,y) \in H^2,\ a(x,y) = (\lambda x - T(x) \mid y).$$

Elle est clairement continue et symétrique (car T est borné auto-adjoint), et coercive grâce à (8.2). En conséquence, le théorème de Lax-Milgram assure que pour tout $z \in H$, il existe un unique $x \in H$ tel que

$$\forall y \in H,\ (\lambda x - T(x) \mid y) = a(x,y) = (z \mid y).$$

Cela montre la bijectivité de $\lambda I - T$. Donc $\lambda \notin \sigma(T)$.

Montrons maintenant que $M \in \sigma(T)$. Soit $b(x,y) = (Mx - T(x) \mid y)$. La forme b est bilinéaire symétrique positive et l'on dispose donc de l'inégalité de Cauchy-Schwarz suivante [4] :

$$\forall (x,y) \in H^2,\ |b(x,y)| \leq \sqrt{b(x,x)b(y,y)}.$$

En prenant $y = Mx - T(x)$ et en exploitant la continuité de b, on en déduit que

$$\forall x \in H, \|Mx - T(x)\|_H \leq C\sqrt{(Mx - T(x)|x)}.$$

Soit $(x_n)_{n\in\mathbb{N}}$ une suite d'éléments de H de norme 1 telle que $(T(x_n) \mid x_n)$ converge vers M. L'inégalité ci-dessus assure que $T(x_n) \to Mx_n$. Si l'on suppose (par l'absurde) que $M \notin \sigma(T)$ alors on a $(MI - T)^{-1}$ continu et

$$x_n = (MI - T)^{-1}(MI - T)(x_n),$$

et donc $x_n \to 0$, ce qui contredit le fait que $\|x_n\| = 1$ pour tout $n \in \mathbb{N}$. Donc $M \in \sigma(T)$. ■

Corollaire. *Sur un espace de Hilbert réel, le seul opérateur auto-adjoint T tel que $\sigma(T) = 0$ est l'opérateur nul.*

Preuve : D'après la proposition précédente, on a $m = M = 0$, et donc $(T(x) \mid x) = 0$ pour tout $x \in H$. Cela entraîne la nullité de T. ■

En dimension finie, il est bien connu que tout endomorphisme symétrique admet une base orthonormale de vecteurs propres. En dimension infinie, ce théorème fondamental se généralise aux opérateurs auto-adjoints compacts comme suit :

Théorème (spectral)**.** *Soit H un espace de Hilbert réel séparable de dimension infinie, et T un opérateur auto-adjoint compact sur H. Il existe alors une base hilbertienne de H constituée de vecteurs propres de T.*

Preuve : D'après le théorème de la page 99, le spectre $\sigma(T)$ de T est borné et a au plus un point d'accumulation (à savoir 0). Il est donc fini ou dénombrable [5]. Soit $E_\lambda \overset{\text{déf}}{=} \ker(T - \lambda I)$ le sous-espace propre de T pour la valeur propre λ. On vérifie (comme dans le cas de la dimension finie) que les sous-espaces propres sont en somme directe orthogonale. Notons F le s.e.v. engendré par tous les vecteurs propres de T, et montrons que F est dense dans H. Il est clair que F est stable par T. Le s.e.v. $F^\perp$ est aussi stable par T. En effet, comme T est autoadjoint et $T(F) \subset F$,

$$\forall (x,y) \in F^\perp \times F,\ (T(x)|y) = (x|T(y)) = 0.$$

On peut donc définir l'application linéaire T_0 induite par T sur $F^\perp$. Notons que $F^\perp$ est un sous-espace fermé de l'espace de Hilbert H. C'est donc un Hilbert. Par ailleurs, l'opérateur T_0 est clairement compact auto-adjoint sur $F^\perp$ car T est compact auto-adjoint. Remarquons enfin que $\sigma(T_0)$ ne peut pas contenir d'élément non nul. En effet, dans le cas contraire, cet élément serait une valeur propre de T_0 et donc de T aussi. Tout vecteur propre associé devrait donc se trouver dans F, ce qui est impossible puisque $F \cap F^\perp = \{0\}$. Donc $\sigma(T_0) = \{0\}$. D'après le corollaire précédent, on a donc $T_0 = 0$. Donc $F^\perp \subset \operatorname{Ker} T \subset F$ puis $F^\perp = \{0\}$. Autrement dit, F est dense dans H.

4. qui se montre comme dans le cas classique en considérant le discriminant de la fonction $\lambda \mapsto b(x+\lambda y, x+\lambda y)$.

5. en effet, comme $\sigma(T)$ est compact, pour tout $n \geq 1$, l'ensemble $\sigma(T) \cap \big(]-\infty, -1/n] \cup [1/n, +\infty[\big)$ a un nombre fini d'éléments.

On sait par ailleurs que chaque sous-espace propre correspondant à une valeur propre *non nulle* est de dimension finie, donc possède une base orthonormale. Si 0 est valeur propre, l'espace E_0 est ou bien de dimension finie et donc admet une base orthonormale, ou bien Hilbert séparable (car s.e.v. fermé de H Hilbert séparable) donc admet une base hilbertienne. Pour construire une base hilbertienne de H constituée de vecteurs propres de T, il suffit donc de considérer la réunion (au plus dénombrable) des bases orthonormales (ou hilbertiennes) de tous les sous-espaces propres de T. ■

Remarque : Si T est compact, auto-adjoint et défini positif, on peut de plus affirmer que chaque vecteur de la base hilbertienne construite ci-dessus correspond à une valeur propre strictement positive.

8.6 Quelques applications

8.6.1 Résolution d'une équation différentielle

Considérons l'équation différentielle suivante :

$$(E)\qquad \begin{cases} u'' = f, \\ u(a) = u(b) = 0. \end{cases}$$

Lorsque f est continue sur $[a, b]$, cette équation peut se résoudre explicitement. En effet, en intégrant une première fois, il vient, pour une constante C à déterminer

$$u'(t) = \int_a^t f(s)\,ds - C, \tag{8.3}$$

puis, en intégrant une deuxième fois, sachant que l'on veut que $u(a) = 0$,

$$u(x) = \int_a^x \int_a^t f(s)\,ds\,dt - C(x - a).$$

La deuxième condition $u(b) = 0$ permet de déterminer C de manière unique. On obtient finalement :

$$u(x) = \int_a^x \int_a^t f(s)\,ds\,dt - \frac{x-a}{b-a} \int_a^b \int_a^b f(s)\,ds\,dt. \tag{8.4}$$

On s'intéresse maintenant à l'opérateur T qui à f associe u. On munit $\mathcal{C}([a, b]; \mathbb{R})$ de la norme associée au produit scalaire

$$(g \mid h) = \frac{1}{b-a} \int_a^b g(x)h(x)\,dx.$$

Avec ce choix de norme, on a clairement, en vertu de l'inégalité de Cauchy-Schwarz,

$$\forall g \in \mathcal{C}([a, b]; \mathbb{R}),\ \|g\|_{L^2} \leq \|g\|_{L^\infty}.$$

En appliquant l'inégalité de Cauchy-Schwarz dans la formule (8.4), il est alors facile d'établir que T est borné de $\mathcal{C}([a, b]; \mathbb{R})$ dans $L^2 \overset{\text{déf}}{=} L^2([a, b]; \mathbb{R})$ les deux espaces étant munis de la norme définie ci-dessus. Par densité de $\mathcal{C}([a, b]; \mathbb{R})$ dans L^2, on peut donc prolonger T en un opérateur borné de L^2 dans L^2. On note encore T le prolongement obtenu.

Montrons que T est compact. Soit $\overline{B}$ la boule unité de L^2 et $f \in \overline{B}$. En utilisant l'égalité (8.3) et l'inégalité de Cauchy-Schwarz, on constate que u' est continue. Donc u est C^1 sur $[a, b]$. De plus, toujours à l'aide de l'inégalité de Cauchy-Schwarz, on montre l'existence d'un réel positif M indépendant de $f \in \overline{B}$ tel que $\|T(f)\|_{L^2} + \|(T(f))'\|_{L^2} \leq M$. En conséquence,

$$T(\overline{B}) \subset \left\{ g \in \mathcal{C}^1([a, b]; \mathbb{R}) \,/\, \|g\|_{L^2}^2 + \|g'\|_{L^2}^2 \leq M \right\}.$$

Dans le chapitre 5, on a vu que l'ensemble de droite est relativement compact dans $\mathcal{C}([a,b];\mathbb{R})$ et donc a fortiori dans L^2. Donc T est bien un opérateur compact.

Vérifions ensuite que T est auto-adjoint.

En intégrant par parties deux fois et en utilisant (E), on a pour tout couple de fonctions f et g continues sur $[a,b]$,

$$
\begin{aligned}
(T(f)\mid g) &= \int_a^b T(f)(x)\,g(x)\,dx,\\
&= \int_a^b T(f)(x)\,(T(g))''(x)\,dx,\\
&= -\int_a^b (T(f))'(x)\,(T(g))'(x)\,dx,\\
&= \int_a^b (T(f))''(x)\,T(g)(x)\,dx,\\
&= (f\mid T(g)).
\end{aligned}
$$

Par densité de $\mathcal{C}([a,b];\mathbb{R})$ dans L^2, l'égalité demeure vraie pour tout $(f,g)\in L^2\times L^2$. Donc T est auto-adjoint. Remarquons aussi au passage que la troisième égalité ci-dessus appliquée avec $g=f$ montre que T est négatif. De plus, à l'aide de la formule (8.4), il n'est pas difficile de montrer que $\operatorname{Ker} T=\{0\}$. Donc toutes les valeurs propres de T sont strictement négatives.

Le théorème spectral assure donc qu'il existe une base hilbertienne de L^2 constituée de vecteurs propres pour T. Comme les valeurs propres sont strictement négatives, on peut les chercher sous la forme $\lambda=-\alpha^{-2}$ avec $\alpha>0$. Soit donc f un vecteur propre associé à une telle valeur propre. Comme $T(f)=\lambda f$ et $T(f)$ est de classe C^1, il en est de même pour f. En conséquence f vérifie l'équation différentielle

$$
(F)\qquad \left\{\begin{array}{l} f''+\alpha^2 f=0,\\ f(a)=f(b)=0.\end{array}\right.
$$

Il s'agit maintenant de déterminer pour quelles valeurs de α le système (F) admet une solution non triviale.

La solution générale de $f''+\alpha^2 f=0$ s'écrit

$$f(t)=C\sin(\alpha t+\varphi)\quad\text{avec}\quad (C,\varphi)\in\mathbb{R}^2.$$

Pour $C\neq 0$, les conditions $f(a)=f(b)=0$ sont assurées si et seulement si

$$\alpha a+\varphi\equiv 0\,[\pi]\quad\text{et}\quad \alpha b+\varphi\equiv 0\,[\pi].$$

On en déduit que $\alpha=k\omega/2$ avec $k\in\mathbb{N}^*$ et $\varphi=-ak\omega/2$. Comme d'habitude, on a posé $\omega=2\pi/(b-a)$. Enfin, on choisit $C=\sqrt{2}$ afin d'avoir $\|f\|_{L^2}=1$.

Nous pouvons résumer les résultats obtenus ainsi :

Proposition. *L'opérateur T défini ci-dessus est compact, auto-adjoint, et défini négatif. La suite des valeurs propres $(\lambda_k)_{k\in\mathbb{N}^*}$ et la base hilbertienne correspondante $(f_k)_{k\in\mathbb{N}^*}$ sont*

$$\lambda_k=-\frac{4}{k^2\omega^2}\quad\textit{et}\quad f_k(t)=\sqrt{2}\sin\left(\frac{k\omega(t-a)}{2}\right).$$

8.6.2 Opérateurs de Hilbert-Schmidt

Notons I l'intervalle $[a,b]$ et Q le carré $[a,b]\times[a,b]$. Notons $L^2(I)=L^2(I;\mathbb{R})$ et $L^2(Q)=L^2(Q;\mathbb{R})$. On munit $L^2(I)$ du produit scalaire

$$(f\mid g)_{L^2(I)}=\int_a^b f(x)g(x)\,dx$$

et $L^2(Q)$ du produit scalaire

$$(F\mid G)_{L^2(Q)}=\int_a^b\int_a^b F(x,y)G(x,y)\,dx\,dy.$$

Fixons une fonction $K\in L^2(Q)$. Pour $u\in\mathcal{C}([a,b];\mathbb{R})$, on définit (formellement)

$$Tu(x)=\int_a^b K(x,y)u(y)\,dy.$$

Lorsque K et u sont continues, cette expression définit une fonction continue. Vérifions que sous nos hypothèses, l'opérateur T est continu de $L^2(I)$ dans $L^2(I)$. Tout d'abord, on peut vérifier à l'aide du théorème de Fubini et de l'inégalité de Cauchy-Schwarz que la fonction $(x,y)\mapsto K(x,y)u(y)$ est intégrable sur Q. Donc Tu est une fonction mesurable.

Ensuite, grâce à l'inégalité de Cauchy-Schwarz, on a

$$\int_a^b\left|\int_a^b K(x,y)u(y)\,dy\right|^2\leq\int_a^b\left(\int_a^b K^2(x,y)\,dy\right)\left(\int_a^b u^2(y)\,dy\right)dx=\|K\|^2_{L^2(Q)}\|u\|^2_{L^2(I)}.$$

Donc T est un opérateur borné de $L^2(I)$ dans $L^2(I)$, et $\|T\|_{\mathcal{L}(L^2(I))}\leq\|K\|_{L^2(Q)}$.

Nous allons maintenant montrer que T est compact. Pour cela, nous allons écrire T comme limite d'une suite d'opérateurs de rang fini à l'aide d'une base hilbertienne $(e_k)_{k\in\mathbb{N}}$ de $L^2(I)$. Nous aurons besoin du résultat suivant :

Lemme. *La famille tensorisée $(e_j\otimes e_k)_{k\in\mathbb{N}}$ définie sur Q par*

$$(e_j\otimes e_k)(x,y)=e_j(x)e_k(y)$$

est une base hilbertienne de $L^2(Q)$.

Preuve : Un calcul immédiat montre que $(e_j\otimes e_k)_{k\in\mathbb{N}}$ est orthonormale. Il s'agit maintenant de démontrer que cette famille est totale, et pour cela il suffit d'établir que le seul élément F de $L^2(Q)$ orthogonal à tous les $(e_j\otimes e_k)_{k\in\mathbb{N}}$ est 0. Supposons donc que

$$\forall(j,k)\in\mathbb{N}^2,\ \int_a^b\int_a^b F(x,y)e_j(x)e_k(y)\,dx\,dy=0. \tag{8.5}$$

Considérons deux fonctions f et g continues sur I. Soit $f\otimes g:(x,y)\longmapsto f(x)g(y)$. Comme $(e_j)_{j\in\mathbb{N}}$ est une base hilbertienne de $L^2(I)$, on peut écrire

$$f=\sum_{j\in\mathbb{N}}(f\mid e_j)_{L^2(I)}e_j\quad\text{et}\quad g=\sum_{k\in\mathbb{N}}(g\mid e_k)_{L^2(I)}e_k$$

où les égalités doivent être comprises au sens de la norme de $L^2(I)$.

Il est alors facile de vérifier que, au sens de la norme de $L^2(Q)$,

$$f\otimes g=\sum_{(j,k)\in\mathbb{N}^2}(f\mid e_j)_{L^2(I)}(g\mid e_k)_{L^2(I)}e_j\otimes e_k.$$

Comme le produit scalaire dans $L^2(Q)$ est continu par rapport à ses deux arguments, l'identité (8.5) combinée au résultat de convergence ci-dessus permet de conclure que

$$\int_a^b \int_a^b F(x,y)f(x)g(y)\,dx\,dy = 0. \tag{8.6}$$

D'autre part, le s.e.v. $\mathcal{F}$ engendré par l'ensemble des fonctions de type $(x,y) \mapsto f(x)g(y)$ avec f et g continues sur I est stable par combinaison linéaire, multiplication, contient les fonctions constantes et sépare les points. Comme Q est un espace métrique compact, le théorème de Stone-Weirstrass assure la densité de $\mathcal{F}$ dans $\mathcal{C}(Q;\mathbb{R})$ au sens de la norme $L^\infty(Q)$ et donc, a fortiori, au sens de la norme $L^2(Q)$. Mais comme $\mathcal{C}(Q;\mathbb{R})$ est dense dans $L^2(Q;\mathbb{R})$ (démonstration similaire à celle de la densité de $\mathcal{C}(I;\mathbb{R})$ dans $L^2(I;\mathbb{R})$), on en déduit que $\mathcal{F}$ est également dense dans $L^2(Q)$.

Soit $\varepsilon > 0$ et $G \in \mathcal{F}$ tel que $\|F - G\|_{L^2(Q)} \leq \varepsilon$. Écrivons que

$$\|F\|^2_{L^2(Q)} = (F \mid F - G)_{L^2(Q)} + (F \mid G)_{L^2(Q)}.$$

Comme G est combinaison linéaire de fonctions du type $f \otimes g$ avec f et g continues sur I, l'égalité (8.6) assure que le dernier terme ci-dessus est nul. En utilisant l'inégalité de Cauchy-Schwarz dans $L^2(Q)$ afin de majorer le premier terme du membre de droite, on conclut que $\|F\|_{L^2(Q)} \leq \varepsilon$. Comme notre raisonnement est valable pour tout $\varepsilon > 0$, on conclut que $F = 0$. Donc $(e_j \otimes e_k)_{(j,k)\in\mathbb{N}^2}$ est totale. ■

Grâce à ce lemme, en vertu de l'égalité de Parseval dans $L^2(I)$ puis dans $L^2(Q)$, on peut donc écrire que

$$\begin{aligned}
\sum_{j\in\mathbb{N}} \|T(e_j)\|^2_{L^2(I)} &= \sum_{j\in\mathbb{N}}\sum_{k\in\mathbb{N}} |(T(e_j) \mid e_k)_{L^2(I)}|^2, \\
&= \sum_{(j,k)\in\mathbb{N}^2} \left| \int_a^b \int_a^b K(x,y)e_j(y)e_k(x)\,dx\,dy \right|^2, \\
&= \|K\|^2_{L^2(Q)}. \qquad (8.7)
\end{aligned}$$

En observant que tout $f \in L^2(I)$ se décompose en

$$f = \sum_{j\in\mathbb{N}} (f \mid e_j)_{L^2(I)}\, e_j,$$

il vient, par continuité de T,

$$T(f) = \sum_{j\in\mathbb{N}} (f \mid e_j)_{L^2(I)}\, T(e_j).$$

Il est donc naturel de poser

$$\forall f \in L^2(I),\ T_n(f) = \sum_{j<n} (f \mid e_j)_{L^2(I)}\, T(e_j).$$

L'opérateur T_n obtenu est clairement borné sur $L^2(I)$ et de rang fini. Par ailleurs

$$\forall f \in L^2(I),\ (T - T_n)(f) = \sum_{j\geq n} (f \mid e_j)_{L^2(I)} T(e_j).$$

Donc, par l'inégalité de Cauchy-Schwarz dans ℓ^2 et l'égalité de Parseval,

$$\begin{aligned}\|(T-T_n)(f)\|_{L^2(I)}^2 &\leq \Big(\sum_{j\geq n}|(f\mid e_j)_{L^2(I)}|\,\|T(e_j)\|_{L^2(I)}\Big)^2,\\ &\leq \sum_{j\geq n}|(f\mid e_j)_{L^2(I)}|^2\sum_{j\geq n}|\|T(e_j)\|_{L^2(I)}^2,\\ &\leq \|f\|_{L^2(I)}^2\sum_{j\geq n}\|T(e_j)\|_{L^2(I)}^2.\end{aligned}$$

La somme de la dernière ligne est le reste d'une série convergente donc tend vers 0 quand n tend vers l'infini. En conséquence, $T_n \to T$ dans $\mathcal{L}(L^2(I))$ ce qui assure que T est compact.

Supposons de plus que la fonction K soit symétrique :

$$K(x,y)=K(y,x) \quad \text{pour presque tout} \quad (x,y)\in\mathbb{R}^2,$$

et vérifions que T est alors auto-adjoint.

Pour $(f,g)\in L^2(I)\times L^2(I)$, le théorème de Fubini, le changement de variables $(x,y)\mapsto(y,x)$ et le fait que $K(x,y)=K(y,x)$ permettent d'écrire que

$$\begin{aligned}(Tf\mid g)_{L^2(I)} &= \int_a^b\int_a^b K(x,y)f(y)\,g(x)\,dy\,dx,\\ &= \int_a^b\int_a^b K(y,x)f(x)g(y)\,dx\,dy,\\ &= \int_a^b\int_a^b K(x,y)g(y)\,f(x)\,dx\,dy,\\ &= (f\mid Tg)_{L^2(I)}.\end{aligned}$$

Donc T est bien auto-adjoint.

Résumons le résultat obtenu.

Proposition. *Soit $K\in L^2(Q)$ tel que $K(x,y)=K(y,x)$. L'opérateur de Hilbert-Schmidt T défini de $L^2(I)$ dans $L^2(I)$ par*

$$Tf(x)=\int_a^b K(x,y)f(y)\,dy$$

est compact auto-adjoint et admet donc une base hilbertienne de vecteurs propres.

Remarque. Si l'on choisit pour $(f_j)_{j\in\mathbb{N}}$ la base hilbertienne associée à la suite de valeurs propres $(\lambda_j)_{j\in\mathbb{N}}$ de T, on obtient, grâce à (8.7),

$$\begin{aligned}\|K\|_{L^2(I)}^2 &= \sum_j\|T(f_j)\|_{L^2(I)}^2,\\ &= \sum_{j\in\mathbb{N}}\lambda_j^2.\end{aligned}$$

Chapitre 9

La topologie faible

9.1 Le cadre abstrait

Soit X un espace topologique, et $(Y_i, \mathcal{O}_i)_{i\in I}$ une famille d'espaces topologiques. On se donne une famille d'applications $\varphi_i : X \to Y_i$.

On cherche à déterminer une topologie sur X qui rende toutes les applications φ_i continues. Il est clair que si l'on munit X de la topologie discrète (celle qui rend toutes les parties de X ouvertes) alors toutes les applications φ_i sont continues. Mais cette topologie n'est pas très intéressante : elle rend continue n'importe quelle application sur X.

On cherche à déterminer la topologie sur X *la plus économique* (c'est-à-dire la moins fine) rendant toutes les applications φ_i continues. Cette topologie est décrite dans la proposition suivante.

Proposition. *L'ensemble $\mathcal{O}$ constitué par les réunions arbitraires d'intersections finies de $\varphi_i^{-1}(\omega_i)$ avec ω_i ouvert de Y_i est une topologie sur X. C'est la topologie la moins fine rendant toutes les applications φ_i continues.*

Preuve : Tout d'abord, remarquons qu'une condition nécessaire pour que toutes les applications φ_i soient continues est que pour tout $i \in I$ et ouvert ω_i de Y_i, l'ensemble $\varphi_i^{-1}(\omega_i)$ soit un ouvert de X. Donc une topologie rendant toutes les applications φ_i continues doit nécessairement contenir tous les $\varphi_i^{-1}(\omega_i)$ avec ω_i ouvert de X_i puis les réunions quelconques d'intersections finies de tels ensembles, donc $\mathcal{O}$.

Par ailleurs, il est clair que $\mathcal{O}$ contient X et $\emptyset$, et est stable par intersection finie et réunion quelconque. Enfin, par construction, pour tout $i \in I$ et ω_i ouvert de Y_i, l'ensemble $\varphi_i^{-1}(\omega_i)$ est dans $\mathcal{O}$. Donc chaque application φ_i est continue. ∎

Attention : Faire l'opération contraire (intersection finie de réunions quelconques) ne donne pas forcément une topologie (exercice : pourquoi ?)

Dans $(X, \mathcal{O})$, on dispose d'un critère simple pour déterminer si une suite converge :

Proposition 6. *Une suite $(x_n)_{n\in\mathbb{N}}$ d'éléments de X converge vers $x \in X$ au sens de la topologie $\mathcal{O}$ si et seulement si chaque suite $\big(\varphi_i(x_n)\big)_{n\in\mathbb{N}}$ converge vers $\varphi_i(x)$.*

Preuve : Supposons que $(x_n)_{n\in\mathbb{N}} \in X^{\mathbb{N}}$ converge vers $x \in X$. Comme chaque φ_i est continue, on a bien $\varphi_i(x_n) \to \varphi_i(x)$.

Réciproquement, supposons que $\varphi_i(x_n) \to \varphi_i(x)$ pour tout $i \in I$, $x \in X$ et suite $(x_n)_{n\in\mathbb{N}} \in X^{\mathbb{N}}$ convergente vers x. Soit Ω un voisinage de x. Il existe alors un nombre fini d'éléments $i_1, \cdots, i_k$ de I et d'ouverts $\omega_{i_1}, \cdots, \omega_{i_k}$ appartenant respectivement à

$Y_{i_1}, \cdots, Y_{i_k}$ tels que

$$x \in \bigcap_{j=1}^{k} \varphi_{i_j}^{-1}(\omega_{i_j}) \subset \Omega.$$

Mais il existe des entiers $N_1, \cdots, N_{i_k}$ tels que

$$\forall j \in \{1, \cdots, k\},\ n \geq N_j \Longrightarrow \varphi_{i_j}(x_n) \in \omega_{i_j}.$$

Pour $n \geq \max(N_1, \cdots, N_k)$, on a donc $x_n \in \Omega$. D'où la convergence de la suite $(x_n)_{n\in\mathbb{N}}$ vers x. ∎

Sur le même modèle, on dispose du critère suivant pour la continuité.

Proposition 7. *Soit Z un espace topologique et $\psi : Z \to X$. L'application ψ est continue si et seulement si chaque application $\varphi_i \circ \psi : Z \to X_i$ est continue.*

La preuve est laissée au lecteur à titre d'exercice.

9.2 La topologie faible

Dans toute cette section, E est un espace de Banach. On rappelle que E' désigne le dual topologique de E, c'est-à-dire l'ensemble des formes linéaires continues sur E.

On veut munir E de la topologie la moins fine possible rendant continues toutes les applications de E'.

Définition. On appelle **topologie faible** sur E (notée $\sigma(E, E')$) la topologie la moins fine rendant continues toutes les formes linéaires de E'.

En appliquant la construction de la partie précédente à la famille comprenant tous les éléments de E', on voit que les ouverts pour la topologie faible sont les réunions quelconques d'intersections finies d'ensembles du type $L^{-1}(\omega)$ avec ω ouvert de $\mathbb{R}$ (ou de $\mathbb{C}$ si E est un e.v. sur $\mathbb{C}$)

Étant donné que les intervalles ouverts [1] constituent une base de voisinages pour $\mathbb{R}$, on en déduit le résultat suivant.

Bases de voisinages pour la topologie faible : Tout voisinage de $x_0 \in E$ pour la topologie $\sigma(E, E')$ contient un ouvert du type

$$\bigcap_{i=1}^{N} \big\{ x \in E \,\big|\, |f_i(x) - f_i(x_0)| < \alpha_i \big\}$$

avec $\alpha_i > 0$ et $f_i \in E'$.

Proposition. *La topologie faible est séparée.*

Preuve : Soit x_1 et x_2 deux éléments distincts de E. D'après la forme géométrique du théorème de Hahn-Banach, il existe $L \in E'$ telle que $L(x_1) \neq L(x_2)$. Posons

$$\varepsilon = \frac{|L(x_2) - L(x_1)|}{4}, \quad V_1 = L^{-1}\big(]L(x_1)-\varepsilon, L(x_1)+\varepsilon[\big) \text{ et } V_2 = L^{-1}\big(]L(x_2)-\varepsilon, L(x_2)+\varepsilon[\big).$$

Les ouverts V_1 et V_2 contiennent respectivement x_1 et x_2, et sont disjoints. ∎

On peut donc munir E de deux topologies séparées différentes :

(i) la "topologie forte" associée à la norme de E,

1. ou les disques ouverts si l'on travaille dans un e.v.n. sur $\mathbb{C}$

(*ii*) la "topologie faible" notée $\sigma(E,E')$.

Par construction, la topologie $\sigma(E,E')$ est moins fine que la topologie forte : tout ouvert pour la topologie faible est un ouvert pour la topologie forte.

Proposition. *En dimension finie, les deux topologies sont les mêmes.*

Preuve : Soit E un e.v.n. de dimension finie. Fixons une base $(e_1,\cdots,e_n)$ de E et notons $(e_1^*,\cdots,e_n^*)$ la base duale. Puisqu'en dimension finie toutes les normes sont équivalentes, on peut toujours supposer que la norme sur E est donnée par

$$\forall x\in E,\ \|x\|_E=\max_{1\leq i\leq n}|e_i^*(x)|.$$

On sait déjà que tout ouvert pour la topologie faible est un ouvert pour la topologie forte. Il suffit donc d'établir que tout ouvert pour la topologie forte est aussi un ouvert pour la topologie faible.

Soit Ω un ouvert de E pour la topologie forte, et $x_0\in\Omega$. Fixons un $r_0>0$ tel que la boule ouverte $B(x_0,r_0)$ au sens de la norme introduite ci-dessus soit incluse dans Ω. On constate que

$$B(x_0,r_0)=\big\{x\in E\mid\forall i\in\{1,\cdots,n\},\,|e_i^*(x-x_0)|<r_0\big\}.$$

Donc $B(x_0,r_0)$ est aussi un voisinage de x_0 pour la topologie faible. Cela permet de conclure que Ω est un ouvert pour la topologie faible. ■

En dimension infinie, il existe des ouverts pour la topologie forte qui ne sont pas des ouverts pour la topologie faible. C'est une conséquence immédiate du résultat suivant :

Proposition. *En dimension infinie, la boule unité ouverte $B(0,1)$ est d'intérieur vide pour la topologie faible.*

Preuve : Soit E un Banach de dimension infinie. Supposons par l'absurde que $B(0,1)$ admette un point intérieur x_0. Alors il existe un nombre fini k de formes linéaires continues $\varphi_1,\cdots,\varphi_k$ et de réels strictement positifs $\alpha_1,\cdots,\alpha_k$ tels que

$$V\stackrel{\text{déf}}{=}\big\{x\in E\mid\forall i\in\{1,\cdots,k\},\ |\varphi_i(x-x_0)|<\alpha_i\big\}\subset B(0,1).$$

L'application

$$\Psi\ :\ \left\{\begin{array}{lcl} E & \longrightarrow & \mathbb{R}^k\\ x & \longmapsto & (\varphi_1(x),\cdots,\varphi_k(x))\end{array}\right.$$

est linéaire et non injective (sinon E serait de dimension finie). Il existe donc $y_0\in E\setminus\{0\}$ tel que $\Psi(y_0)=0$. Par linéarité, on en déduit que $x_0+\lambda y_0$ appartient à V pour tout $\lambda\in\mathbb{R}$. Cela contredit le caractère borné de V donc le fait que $V\subset B(0,1)$. ■

Par passage au complémentaire, on conclut qu'en dimension infinie la topologie forte contient aussi plus de fermés que la topologie faible.

Dans le cas convexe cependant, les deux notions se rejoignent :

Proposition. *Soit C un convexe de E. Alors C est fermé pour la topologie forte ssi C est fermé pour la topologie faible.*

Preuve : Seule la partie directe est non triviale. On considère donc un convexe C fermé pour la topologie forte. On peut supposer de plus que $C\neq E$ (sinon le résultat est trivial) et on va montrer que $E\setminus C$ est ouvert. Soit donc $x_0\notin C$. D'après la forme géométrique du théorème de Hahn-Banach, il existe $f\in E'$ et $\alpha\in\mathbb{R}$ tels que

$$\forall x\in C,\ f(x)<\alpha<f(x_0).$$

L'ensemble $f^{-1}(]\alpha,+\infty[)$ est un ouvert pour la topologie faible qui contient x_0 et est inclus dans $E\setminus C$. Donc $E\setminus C$ est ouvert pour la topologie faible. ■

Avons de nous intéresser plus spécifiquement aux propriétés des suites convergentes au sens de la topologie faible, introduisons les notations suivantes :

Soit $(x_n)_{n\in\mathbb{N}}$ une suite d'éléments de E, et $x \in E$. On note

- Convergence forte : $x_n \to x$.
- Convergence faible : $x_n \rightharpoonup x$.

Proposition. *Soit $(x_n)_{n\in\mathbb{N}}$ une suite de E. On a les propriétés suivantes :*

(i) $x_n \rightharpoonup x$ *ssi* $f(x_n) \to f(x)$ *pour tout* $f \in E'$.

(ii) $x_n \to x \implies x_n \rightharpoonup x$.

(iii) Si $x_n \rightharpoonup x$ *alors* $(x_n)_{n\in\mathbb{N}}$ *est bornée dans* E *et* $\|x\|_E \leq \liminf \|x_n\|_E$.

(iv) $\varphi_n \to \varphi$ *dans* E' *et* $x_n \rightharpoonup x \implies \varphi_n(x_n) \to \varphi(x)$.

Preuve : La première propriété résulte de la proposition 6.

Le deuxième point est également immédiat : si $x_n \to x$ alors pour toute fonction f continue, on a $f(x_n) \to f(x)$. On a donc a fortiori $f(x_n) \to f(x)$ pour tout $f \in E'$.

Démontrons *(iii)*. Pour tout $n \in \mathbb{N}$, on pose

$$T_n \ : \left\{ \begin{array}{ccc} E' & \longrightarrow & \mathbb{R} \\ \varphi & \longmapsto & \varphi(x_n). \end{array} \right.$$

Toutes les applications T_n sont linéaires et continues sur E'. De plus, par hypothèse, pour tout $\varphi \in E'$, la suite $(T_n(\varphi))_{n\in\mathbb{N}}$ est convergente donc bornée.

D'après le théorème de Banach-Steinhaus, la suite $(T_n)_{n\in\mathbb{N}}$ est donc bornée : il existe $M \in \mathbb{R}^+$ tel que

$$\forall n \in \mathbb{N}, \forall \varphi \in E',\ |T_n(\varphi)| = |\varphi(x_n)| \leq M\|\varphi\|_{E'}.$$

Mais d'après le corollaire 5 du théorème de Hahn-Banach, on a

$$\|x_n\|_E = \sup_{\varphi \in E', \|\varphi\|_{E'}=1} |\varphi(x_n)|.$$

Donc on peut conclure que

$$\forall n \in \mathbb{N},\ \|x_n\|_E \leq M.$$

En passant à la limite inférieure dans l'inégalité

$$|\varphi(x_n)| \leq \|\varphi\|_{E'}\|x_n\|_E$$

puis en appliquant à nouveau le corollaire 5 du théorème de Hahn-Banach, on obtient l'inégalité souhaitée.

Démontrons le dernier point. Pour cela, écrivons

$$\varphi_n(x_n) - \varphi(x) = \big(\varphi_n(x_n) - \varphi(x_n)\big) + \big(\varphi(x_n) - \varphi(x)\big).$$

Par hypothèse, le dernier terme tend vers 0. De plus,

$$|\varphi_n(x_n) - \varphi(x_n)| \leq \|x_n\|_E\|\varphi_n - \varphi\|_{E'}.$$

D'après *(iii)*, la suite $(x_n)_{n\in\mathbb{N}}$ est bornée. Donc on a aussi

$$\varphi_n(x_n) - \varphi(x_n) \to 0,$$

ce qui achève la preuve de la proposition. ■

Terminons par un résultat spécifique aux applications linéaires.

Proposition. *Soit E et F deux Banach et $T : E \to F$ linéaire. Alors T est continue de E vers F pour la topologie forte (des deux espaces) ssi T est continue de E vers F au sens de la topologie faible.*

Preuve : $(i) \Rightarrow (ii)$: Soit $T : E \to F$ linéaire continue au sens de la topologie forte. Soit V un ouvert de $(F, \sigma(F, F'))$ et $x_0 \in T^{-1}(V)$. Comme $T(x_0)$ appartient à l'ouvert V, il existe un nombre fini k de formes linéaires $\varphi_1, \cdots, \varphi_k$ de E' et de réels strictement positifs $\alpha_1, \cdots, \alpha_k$ tels que

$$\bigcap_{j=1}^{k} \big\{y \in F \mid \forall i \in \{1, \cdots, k\},\, |\varphi_i(y - T(x_0))| < \alpha_i\big\}.$$

On constate que l'ensemble

$$U \stackrel{\text{déf}}{=} \bigcap_{j=1}^{k} \big\{x \in E \mid \forall i \in \{1, \cdots, k\},\, |\varphi_i \circ T(x - x_0)| < \alpha_i\big\}$$

vérifie $U \subset T^{-1}(V)$.

Comme T est (fortement) continue, chaque $\varphi_i \circ T$ est une forme linéaire continue sur E. Donc U est un ouvert de $\sigma(E, E')$.

$(ii) \Rightarrow (i)$: En vertu du théorème du graphe fermé, il suffit d'établir que le graphe de T est fermé pour la topologie forte.

Soit donc $(x_n, T(x_n))_{n\in\mathbb{N}}$ qui converge vers $(x, y) \in E \times F$. En particulier $x_n \to x$, donc a fortiori $x_n \rightharpoonup x$. Donc, puisque T est continue pour la topologie faible, $T(x_n) \rightharpoonup T(x)$. Mais par hypothèse $T(x_n) \to y$. Donc $T(x_n) \rightharpoonup y$. La topologie faible étant séparée, il y a unicité de la limite et, par conséquent, $y = T(x)$. ∎

9.3 Un exemple

Soit ℓ^∞ l'ensemble des suites réelles bornées et ℓ^1 l'ensemble des suites réelles sommables :

$$\ell^1 \stackrel{\text{déf}}{=} \Big\{(x_n)_{n\in\mathbb{N}} \in \mathbb{R}^{\mathbb{N}} \mid \sum_{n=0}^{\infty} |x_n| < \infty\Big\}.$$

Rappelons que les espaces vectoriels ℓ^1 et ℓ^∞ munis respectivement de la norme $\|x\|_1 = \sum_{n=0}^{\infty} |x_n|$ et $\|x\|_\infty = \sup_{n\in\mathbb{N}} |x_n|$ sont des Banach. Nous allons montrer que le dual topologique de ℓ^1 peut être identifié à ℓ^∞.

Tout d'abord, remarquons qu'à toute suite $y = (y_n)_{n\in\mathbb{N}}$ de ℓ^∞, on peut associer une forme linéaire continue T_y sur ℓ^1 en posant

$$\forall x \in \ell^1,\ T_y(x) = \sum_{n=0}^{\infty} x_n y_n.$$

En effet, comme

$$\forall (x, y) \in \ell^1 \times \ell^\infty,\ |T_y(x)| \leq \|x\|_1 \|y\|_\infty, \tag{9.1}$$

l'application T_y est bien une forme linéaire continue sur ℓ^1.

Proposition. *L'application $T : y \mapsto f_y$ est linéaire continue et bijective de ℓ^∞ dans $(\ell^1)'$. De plus elle conserve la norme.*

Preuve : La linéarité de T est évidente et l'inégalité (9.1) assure de plus que T_y est de norme inférieure ou égale à $\|y\|_\infty$. On a en fait égalité : en effet pour $\varepsilon \in]0, \|y\|_\infty[$, on considère un indice $n_0 \in \mathbb{N}$ tel que $|y_{n_0}| \geq \|y\|_\infty - \varepsilon$ et l'on constate que

$$|T_y(e^{n_0})| = |y_{n_0}| \geq \|e^{n_0}\|_1\big(\|y\|_\infty - \varepsilon\big),$$

où e^{n_0} désigne la suite dont le $n_0$ème terme vaut 1 et dont les autres termes sont nuls.

Donc l'application T conserve la norme comme annoncé. Reste à établir la surjectivité. Soit donc $\varphi \in (\ell^1)'$. On note y la suite définie par $y_n = \varphi(e_n)$ et l'on constate que y est bornée et vérifie $T(y) = \varphi$. ■

Remarque. Il existe donc une **isométrie** bijective entre ℓ^∞ et $(\ell^1)'$. De ce fait, on a l'habitude d'identifier $(\ell^1)'$ et ℓ^∞, et l'on dit couramment que ℓ^∞ *est le dual de* ℓ^1. On note d'ailleurs $\sigma(\ell^1, \ell^\infty)$ (au lieu de $\sigma(\ell^1, (\ell^1)'))$ la topologie faible sur ℓ^1.

Bien que la topologie forte sur ℓ^1 ait strictement plus d'ouverts que la topologie faible (on est en dimension infinie), on a le résultat suivant [2] :

Théorème (de Schur). *Soit $(x^n)_{n\in\mathbb{N}}$ une suite d'éléments de ℓ_1, et $x \in \ell_1$. Alors*

$$x^n \to x \iff x_n \rightharpoonup x.$$

Preuve : Seule l'implication réciproque est à prouver. et l'on peut visiblement se limiter à l'étude des suites tendant vers 0.

Soit donc $(x^n)_{n\in\mathbb{N}}$ une suite d'éléments de ℓ^1 telle que $x^n \rightharpoonup 0$. Supposons par l'absurde que $(x^n)_{n\in\mathbb{N}}$ ne tende pas fortement vers 0. Quitte à extraire une sous-suite, on peut se ramener au cas où il existe $\alpha > 0$ tel que

$$\forall n \in \mathbb{N}, \ \|x^n\|_1 \geq \alpha.$$

En posant $y^n = x^n/\|x^n\|_1$ on obtient alors $y^n \rightharpoonup 0$ et $\|y^n\|_1 = 1$ pour tout $n \in \mathbb{N}$. On procède alors comme suit :

– On choisit $j_1 \in \mathbb{N}^*$ tel que $\sum_{j=0}^{j_1-1} |y_j^0| \geq \frac{3}{4}$.

  Étant donné que $y^n \rightharpoonup 0$, la suite de réels $(y_j^n)_{n\in\mathbb{N}}$ tend vers 0 à j fixé. On peut donc trouver $n_1 \in \mathbb{N}$ tel que

$$\forall n \geq n_1, \ \sum_{j=0}^{j_1-1} |y_j^n| \leq \frac{1}{8}.$$

  Autrement dit, pour $n \geq n_1$, toutes les suites y^n sont telles que $\sum_{j\geq j_1} |y_j^n| \geq \frac{7}{8}$.

– On considère maintenant $j_2 > j_1$ tel que $\sum_{j=j_1}^{j_2-1} |y_j^{n_1}| \geq \frac{3}{4}$ puis $n_2 > n_1$ tel que

$$\forall n \geq n_2, \ \sum_{j=0}^{j_2-1} |y_j^n| \leq \frac{1}{8}.$$

Par récurrence, on obtient ainsi deux suites d'entiers naturels $(j_k)_{k\in\mathbb{N}}$ et $(n_k)_{k\in\mathbb{N}}$ strictement croissantes et telles que

$$\forall k \in \mathbb{N}, \sum_{j=j_k}^{j_{k+1}-1} |y_j^{n_k}| \geq \frac{3}{4} \quad \text{et} \quad \sum_{j=0}^{j_k-1} |y_j^n| \leq \frac{1}{8} \quad \text{pour} \quad n \geq n_k.$$

2. Le caractère surprenant de ce résultat provient du fait que l'ensemble ℓ^1 muni de la topologie faible *n'est pas métrisable* c'est-à-dire que l'on ne peut pas trouver de distance sur ℓ^1 dont la topologie associée coïncide avec $\sigma(\ell^1, \ell^\infty)$. De de fait, les deux topologies ont les mêmes suites convergentes bien qu'elles ne soient pas égales !

Soit alors $(\gamma_j)_{j\in\mathbb{N}}$ la suite définie par $\gamma_j = \operatorname{sgn}(y_j^{n_k})$ si $j_k \leq j \leq j_{k+1} - 1$. C'est clairement une suite bornée telle que $\|\gamma\|_\infty = 1$. En écrivant que

$$\sum_{j=0}^{\infty} \gamma_j y_j^{n_k} = \sum_{j=0}^{j_k - 1} \gamma_j y_j^{n_k} + \sum_{j=j_k}^{j_{k+1}-1} \gamma_j y_j^{n_k} + \sum_{j=j_{k+1}}^{\infty} \gamma_j y_j^{n_k},$$

on constate que $\sum_{j=0}^{\infty} \gamma_j y_j^{n_k} \geq \frac{3}{8}$ pour tout $k \in \mathbb{N}$.
Cela contredit la convergence faible de $(y^{n_k})_{k\in\mathbb{N}}$ vers 0. ■

Chapitre 10

La topologie faible étoile

10.1 Définition

Soit E un Banach. On sait que E' est aussi un espace de Banach. On note E'' l'ensemble des formes linéaires continues sur E'. C'est encore un espace de Banach appelé **bidual** de E.

La construction du chapitre précédent permet de munir E' de deux topologies (séparées) :

- la "topologie forte" associée à la norme de E',
- la "topologie faible" sur E', notée $\sigma(E', E'')$.

La deuxième topologie est en général strictement plus faible que la première : elle a moins d'ouverts et de fermés. En contrepartie, elle possède plus de compacts et de suites convergentes. On souhaiterait munir E' d'une troisième topologie séparée suffisamment faible pour que $\overline{B}_{E'}(0,1)$ devienne compacte même si E est de dimension infinie.

Avant d'introduire cette nouvelle topologie, nous allons définir un sous-espace vectoriel de E'' par le procédé suivant. À tout $x \in E$, on associe une forme linéaire ξ_x sur E' en posant

$$\xi_x \;:\; \left\{ \begin{array}{lcl} E' & \longrightarrow & \mathbb{R} \\ f & \longmapsto & f(x). \end{array} \right.$$

On a pour tout $(x, f) \in E \times E'$,

$$|\xi_x(f)| = |f(x)| \leq \|f\|_{E'} \|x\|_E.$$

Donc ξ_x est une forme linéaire continue sur E' de norme inférieure ou égale à $\|x\|_E$. En fait, le corollaire 5 du théorème de Hahn-Banach assure que $\|\xi_x\|_{E''} = \|x\|_E$.

On note J l'application $x \mapsto \xi_x$. Il est évident que J est linéaire. De plus, on vient de voir que J conserve la norme donc J est continue. Autrement dit, J appartient à $\mathcal{L}(E; E'')$.

L'application J est bien sûr injective car conserve la norme. En revanche, si E est de dimension infinie, il n'y a pas de raison pour que J soit surjective. En général $J(E)$ est donc un sous-espace strict de E''.

Définition. On appelle **topologie faible *** la topologie la moins fine rendant toutes les applications ξ_x continues. La topologie faible * se note $\sigma(E', E)$

On peut donc maintenant munir E' de trois topologies différentes (que l'on classe ci-dessous de la plus forte à la moins forte) :

(*i*) la "topologie forte" associée à la norme de E',

(*ii*) la "topologie faible" sur E', notée $\sigma(E', E'')$,

(*iii*) la "topologie faible *" sur E', notée $\sigma(E', E)$.

10.2 Propriétés

Par définition, tout ouvert de $\sigma(E', E)$ est du type

$$\bigcup_{\text{quelconque}} \bigcap_{\text{finie}} \xi_{x_i}(\omega_i)$$

où les ω_i sont des ouverts de $\mathbb{R}$ et les x_i des vecteurs de E.

On en déduit le résultat suivant sur les **bases de voisinages pour la topologie faible *.**

Proposition. *Tout voisinage de $f_0 \in E'$ pour la topologie $\sigma(E', E)$ contient un ouvert du type*

$$\bigcap_{i=1}^{N} \big\{ f \in E' \,\big|\, |f(x_i) - f_0(x_i)| < \alpha_i \big\}$$

avec $\alpha_i > 0$ et $x_i \in E$.

On peut maintenant facilement prouver le résultat important suivant :

Proposition. *La topologie faible * est séparée.*

Preuve : Soit (f_1, f_2) un couple d'éléments distincts de E'. Alors il existe $x \in E$ tel que $f_1(x) \neq f_2(x)$. Posons

$$\varepsilon = \frac{|f_2(x) - f_1(x)|}{4} \quad \text{et} \quad \omega_i =]f_i(x) - \varepsilon, f_i(x) + \varepsilon[\quad \text{pour} \quad i = 1, 2.$$

Il est clair que ω_1 et ω_2 sont deux ensembles disjoints contenant respectivement f_1 et f_2. De plus, ce sont des ouverts pour la topologie faible *. ∎

Notations pour la convergence des suites : Soit $(f_n)_{n\in\mathbb{N}}$ une suite d'éléments de E', et $f \in E'$. On note

- Convergence en norme : $f_n \to f$.
- Convergence faible : $f_n \rightharpoonup f$.
- Convergence faible * : $f_n \rightharpoonup f$ faible *.

Proposition. *Soit $(f_n)_{n\in\mathbb{N}}$ une suite de E'. On a les propriétés suivantes :*

(i) $f_n \rightharpoonup f$ *faible * ssi* $f_n(x) \to f(x)$ *pour tout* $x \in E$.

(ii) $f_n \to f \implies f_n \rightharpoonup f \implies f_n \rightharpoonup f$ *faible *.*

(iii) Si $f_n \rightharpoonup f$ *faible * alors* $(f_n)_{n\in\mathbb{N}}$ *est bornée dans* E' *et* $\|f\|_{E'} \leq \liminf \|f_n\|_{E'}$.

(iv) $x_n \to x$ *et* $f_n \rightharpoonup f$ *faible ** $\implies f_n(x_n) \to f(x)$.

Preuve : La propriété *(i)* découle de la définition de $\sigma(E', E)$ et de la proposition 6 page 107.
La première implication de *(ii)* est déjà connue. Maintenant, si $f_n \rightharpoonup f$ alors pour tout $\xi \in E''$, on a $\xi(f_n) \to \xi(f)$. En choisissant $\xi = \xi_x$ avec $x \in E$, on obtient $f_n(x) \to f(x)$ et l'on peut conclure à l'aide de *(i)*.

Supposons que $f_n \rightharpoonup f$ faible *. Alors pour tout $x \in E$, la suite $(f_n(x))_{n\in\mathbb{N}}$ est convergente donc bornée. Le théorème de Banach-Steinhaus assure donc que $(f_n)_{n\in\mathbb{N}}$ est bornée.
Par ailleurs, pour tout $x \in E$, on a

$$|f_n(x)| \leq \|f_n\|_{E'} \|x\|_E.$$

En passant à la limite inf, on obtient donc

$$|f(x)| \leq \big(\liminf \|f_n\|_{E'}\big) \|x\|_E.$$

De par la définition de la norme $\|\cdot\|_{E'}$ on a donc $\|f\|_{E'} \leq \liminf \|f_n\|_{E'}$.

Pour prouver *(iv)*, on écrit

$$f_n(x_n) - f(x) = (f_n - f)(x) + f_n(x_n - x).$$

Le premier terme tend vers 0 car $f_n \rightharpoonup f$ faible *. Le deuxième aussi car d'après *(iii)* la suite $(f_n)_{n\in\mathbb{N}}$ est bornée dans E', et $\|x_n - x\|_E \to 0$. ■

Proposition 8. *Soit $\varphi : E' \to \mathbb{R}$ une application linéaire qui est continue pour la topologie faible *. Alors $\varphi \in J(E)$ (i.e. il existe $x \in E$ tel que pour tout $f \in E'$, on ait $\varphi(f) = f(x)$).*

Preuve : Comme φ est continue en 0, il existe un voisinage ouvert Ω de 0 pour la topologie faible * sur E' tel que

$$\forall f \in \Omega,\ |\varphi(f)| < 1.$$

Cet ouvert Ω contient un élément de la base de voisinage de 0 décrite ci-dessus. Autrement dit, il existe des réels strictement positifs $\alpha_1, \cdots, \alpha_N$, et $(x_1, \cdots, x_N) \in E^N$ tels que

$$\forall g \in E',\ \Big(\forall i \in \{1, \cdots, N\},\ |g(x_i)| < \alpha_i\Big) \Longrightarrow |\varphi(g)| < 1. \tag{10.1}$$

Quitte à supprimer des x_i et à diminuer les α_i, on peut de plus supposer que $(x_1, \cdots, x_N)$ est libre. Introduisons alors la fonction

$$\Lambda\ :\ \left\{\begin{array}{lcl} E' & \longrightarrow & \mathbb{R}^{N+1} \\ f & \longmapsto & \big(\varphi(f), f(x_1), \cdots, f(x_N)\big). \end{array}\right.$$

L'application Λ est linéaire. De plus, elle n'est pas surjective. En effet, si f est une fonction de E' qui s'annule sur tous les x_i alors, d'après (10.1), on a $|\varphi(f)| < 1$. En conséquence, $(1, 0, \cdots, 0) \notin \operatorname{Im}\Lambda$.
Donc $\operatorname{Im}\Lambda$ est un s.e.v. strict de $\mathbb{R}^{N+1}$, et l'on peut conclure que $\operatorname{Im}\Lambda$ est inclus dans un hyperplan de $\mathbb{R}^{N+1}$. Autrement dit, il existe $(\lambda_0, \cdots, \lambda_N) \in \mathbb{R}^{N+1} \setminus \{0\}$ tel que

$$\forall f \in E',\ \lambda_0\varphi(f) + \sum_{i=1}^{N} \lambda_i f(x_i) = 0. \tag{10.2}$$

Si $\lambda_0 \neq 0$, il n'y a plus qu'à poser $x \overset{\text{déf}}{=} -\sum_{i=1}^N \frac{\lambda_i}{\lambda_0} x_i$ et l'on obtient $\varphi = \xi_x$.
Reste à justifier que λ_0 ne peut pas être nul. Supposons par l'absurde que $\lambda_0 = 0$. D'après (10.2), on a donc

$$\forall f \in E',\ f\left(\sum_{i=1}^{N} \lambda_i x_i\right) = 0.$$

D'après le corollaire 5 du théorème de Hahn-Banach (analytique), on a donc $\sum_{i=1}^N \lambda_i x_i = 0$, d'où, puisque $(x_1, \cdots, x_N)$ est libre, $\lambda_1 = \cdots = \lambda_N = 0$. Ceci est contraire à la définition de $(\lambda_0, \cdots, \lambda_N)$. ■

Théorème. *Soit H un hyperplan de E', fermé pour la topologie faible *. Alors il existe $x \in E$ et $\alpha \in \mathbb{R}$ tels que*

$$H = \big\{f \in E' \mid f(x) = \alpha\big\}.$$

Preuve : Par définition d'un hyperplan de E', il existe $\varphi \in E''$ non nulle et $\alpha \in \mathbb{R}$ tels que $H = \varphi^{-1}(\{\alpha\})$.
Fixons un f_0 dans $E' \backslash H$. Comme E' est fermé pour la topologie faible *, l'ensemble $E' \backslash H$ est ouvert. Il existe donc des réels strictement positifs $\alpha_1, \cdots, \alpha_N$ et $(x_1, \cdots, x_N) \in E^N$ tels que

$$V \overset{\text{déf}}{=} \big\{f \in E' \mid \forall i \in \{1, \cdots, n\},\ |f(x_i) - f_0(x_i)| < \alpha_i\big\} \subset E' \setminus H.$$

Notons $W \stackrel{\text{déf}}{=} V - f_0$. Il est clair que

$$W = \{g \in E' \mid \forall i \in \{1, \cdots, n\},\ |g(x_i)| < \alpha_i\}.$$

L'ensemble W est convexe et symétrique autour de 0. Comme φ est une fonction linéaire, on en déduit que $\varphi(V)$ est un intervalle de $\mathbb{R}$ symétrique autour de 0. Cet intervalle est en fait borné car, dans le cas contraire, il contiendrait $\alpha - \varphi(f_0)$ et il existerait donc $h \in V$ tel que $\varphi(h) = \alpha$.

Finalement, il existe donc $C > 0$ tel que

$$\forall g \in W,\ |\varphi(g)| \leq C.$$

Fixons $\varepsilon > 0$ et $g_0 \in E'$. L'inégalité ci-dessus assure que

$$\forall g \in \Big(g_0 + \frac{\varepsilon}{C} W\Big),\ |\varphi(g) - \varphi(g_0)| \leq \varepsilon.$$

Donc φ est continue pour la topologie $\sigma(E', E)$. En vertu de la proposition 8, il existe donc $x \in E$ tel que $\varphi = \xi_x$. ■

Remarque. Si $\xi \in E'' \setminus J(E)$ alors $\{f \in E' \mid \xi(f) = 0\}$ est un hyperplan qui est fermé pour la topologie $\sigma(E', E'')$ mais qui n'est pas pas fermé pour la topologie $\sigma(E', E)$.

Dans le cas où $J(E) \neq E''$ la topologie faible contient donc strictement plus de fermés que la topologie faible *.

10.3 Un résultat de compacité

Dans cette partie, nous allons voir que la topologie faible étoile rend compacte la boule unité fermée.

Théorème (Banach-Alaoglu-Bourbaki)**.** *La boule unité fermée de E' (i.e. $\{f \in E' \mid \|f\|_{E'} \leq 1\}$) est compacte pour la topologie faible *.*

Ce résultat fondamental repose sur le **théorème de Tychonov** (admis) qui établit que le produit cartésien d'une famille arbitraire d'ensembles compacts est compact pour la topologie produit (c'est-à-dire pour la topologie qui rend continues toutes les projections sur l'une des composantes du produit).

Plus concrètement, donnons-nous une famille quelconque $(X_\alpha)_{\alpha \in I}$ d'espaces topologiques et notons

$$X = \prod_{\alpha \in I} X_\alpha = \Big\{(x_\alpha)_{\alpha \in I} \mid \forall \alpha \in I,\ x_\alpha \in X_\alpha\Big\}.$$

On munit X de la topologie la moins fine rendant continues toutes les **projections canoniques** définies par :

$$P_\beta : \left\{ \begin{array}{ccc} X & \longrightarrow & X_\beta \\ (x_\alpha)_{\alpha \in I} & \longmapsto & x_\beta. \end{array} \right.$$

Théorème (de Tychonov)**.** *Si pour tout $\alpha \in I$ l'espace topologique X_α est compact alors X (muni de la topologie la moins fine rendant continues toutes les projections canoniques) aussi.*

La preuve du théorème de Tychonov repose sur le lemme de Zorn, donc sur l'axiome du choix.

Nous pouvons maintenant démontrer le théorème de Banach-Alaoglu-Bourbaki.

Pour ce faire, nous allons appliquer le théorème de Tychonov avec $I = E$ et $X_i = \mathbb{R}$ pour tout $i \in I$. En effet, on a alors $X = \mathbb{R}^E = \prod_{x\in E} \mathbb{R}$. Rappelons que, modulo l'usage de l'application

$$\Phi \ : \ \left\{ \begin{array}{ccc} \mathcal{F}(E;\mathbb{R}) & \longrightarrow & \mathbb{R}^E \\ f & \longmapsto & (f(x))_{x\in E}. \end{array} \right.$$

(qui est bien sûr bijective), l'ensemble $\mathbb{R}^E$ peut être identifié à l'ensemble des fonctions de E dans $\mathbb{R}$.

Notons Ψ la bijection réciproque de Φ restreinte à l'ensemble $\Phi(E')$. C'est une application qui est continue de $\Phi(E')$ (muni de la topologie produit de $\mathbb{R}^E$) dans E' muni de la topologie $\sigma(E', E)$.

En effet, en vertu de la proposition 7 page 108, il suffit de vérifier que pour tout $y \in E$, l'application

$$\Lambda_y \ : \ \left\{ \begin{array}{ccc} \Phi(E') & \longrightarrow & \mathbb{R} \\ \omega & \longmapsto & \Psi(\omega)(y) \end{array} \right.$$

est continue.

On constate que pour tout $f \in E'$, on a

$$\Lambda_y(\Phi(f)) = f(y) = P_y(\Phi(f))$$

où $P_y : \mathbb{R}^E \to \mathbb{R}$ désigne la projection canonique suivant la $y-$ième composante.

Donc Λ_y n'est autre que la restriction de P_y à $\Phi(E')$ donc est continue puisque P_y l'est.

Pour conclure, nous allons montrer que $\overline{B}_{E'}(0, 1)$ est l'image par Ψ (qui est continue) d'un compact de $\mathbb{R}^E$.

Introduisons $K = \big\{(\omega_x)_{x\in E} \mid \forall x \in E,\ |\omega_x| \leq \|x\|\big\}$. Comme $K = \prod_{x\in E}[-\|x\|_E, \|x\|_E]$, le théorème de Tychonof assure que K est compact.

Soit $F = \big\{\omega \in X \mid \forall (x, y) \in E^2, \forall (\lambda, \mu) \in \mathbb{R}^2,\ \omega_{\lambda x+\mu y} = \lambda\omega_x + \mu\omega_y\big\}$.

Pour $(x, y) \in E^2$ et $(\lambda, \mu) \in \mathbb{R}^2$ fixés, l'application

$$L^{\lambda,\mu}_{x,y} \ : \ \left\{ \begin{array}{ccc} X & \longrightarrow & \mathbb{R} \\ \omega & \longrightarrow & \omega_{\lambda x+\mu y} - \lambda\omega_x - \mu\omega_y \end{array} \right.$$

est continue car les projections canoniques le sont ainsi que l'addition et la multiplication dans l'ensemble des réels.

Donc F est fermé comme intersection d'images réciproques de $\{0\}$ par des fonctions continues. Cela entraîne la compacité de $K \cap F$.

Il est clair que $\overline{B}_{E'}(0, 1) = \Psi(K \cap F)$, ce qui achève la preuve du théorème. ■

Chapitre 11

Espaces réflexifs et espaces séparables

11.1 Définition et exemples

En dimension finie, on a clairement $\dim E = \dim E' = \dim E''$. Du fait de l'injectivité de J, on en déduit que $J(E) = E''$. L'application J est alors une isométrie bijective de E sur E'' ce qui fait qu'en pratique on identifie toujours E avec son bidual E''.

En dimension infinie, l'inclusion $J(E) \subset E''$ est, en général, stricte, comme on va le voir dans l'exemple suivant.

Proposition. *L'ensemble c_0 des suites réelles tendant vers 0 à l'infini muni de la norme $\|x\|_\infty = \sup_{n\in\mathbb{N}} |x_n|$ est un espace de Banach non réflexif.*

Preuve : Soit ℓ^∞ l'ensemble des suites réelles bornées muni de la norme $\|\cdot\|_\infty$. Nous laissons au lecteur le soin de montrer que c_0 est un sous-espace fermé de ℓ^∞. Comme ℓ^∞ est complet, c_0 est donc un espace de Banach. Dans le chapitre 9, nous avons vu que si l'on munit ℓ^1 de la norme $\|x\|_1 = \sum_{n\in\mathbb{N}} |x_n|$, il existe une isométrie bijective S entre ℓ^∞ et $(\ell^1)'$. Nous allons montrer qu'il existe aussi une isométrie bijective T entre ℓ^1 et c_0'. Pour cela, considérons l'application T qui à tout $y \in \ell^1$ associe la fonction

$$T_y \ : \ \left\{ \begin{array}{lcl} c_0 & \longrightarrow & \mathbb{R} \\ x & \longmapsto & \sum_{n\in\mathbb{N}} x_n y_n. \end{array} \right.$$

Il est clair que T_y est bien définie et linéaire, et vérifie

$$|T_y x| \leq \sum_{n\in\mathbb{N}} |x_n||y_n| \leq \|x\|_\infty \|y\|_1.$$

Donc $T_y \in c_0'$. En conséquence T est une application de ℓ^1 dans c_0'. La linéarité de T est claire. De plus, on a

$$\forall y \in \ell^1, \ \|T(y)\|_{c_0'} \leq \|y\|_1.$$

L'injectivité de T est évidente. Montrons la surjectivité. Soit donc $\varphi \in c_0'$. Soit y la suite de terme général $y_n = \varphi(e^n)$ (avec e^n définie comme au chapitre 9). Pour tout $N \in \mathbb{N}$, on a

$$\sum_{n=0}^{N} |y_n| = \sum_{n=0}^{N} \operatorname{sgn}(\varphi(e^n))\varphi(e^n) = \varphi\Big(\sum_{n=0}^{N} \operatorname{sgn}(\varphi(e^n))e^n\Big).$$

Donc on a

$$\sum_{n=0}^{N} |y_n| \leq \|\varphi\|_{c_0'}.$$

En conséquence $y \in \ell^1$ et $\|y\|_1 \leq \|\varphi\|_{c_0'}$. Enfin, il est clair que $T(y) = \varphi$ donc l'application T est bien une isométrie bijective entre ℓ^1 et c_0'.
Considérons maintenant la transposée T' de T. Rappelons que T' est l'application linéaire de c_0'' dans $(\ell^1)'$ définie par

$$\forall \xi \in c_0'', \forall \varphi \in \ell^1,\ [T'(\xi)](\varphi) = \xi(T(\varphi)).$$

Sachant que T est un isomorphisme continu, il en est de même pour T' (exercice : le montrer). En notant $\widetilde{S} : c_0 \to (\ell^1)'$ la restriction de S au s.e.v. c_0, et i l'**injection canonique** de $J(c_0)$ dans c_0'' définie par $i(\xi) = \xi$ pour tout $\xi \in J(c_0)$, on obtient le diagramme suivant :

$$\begin{array}{ccc} c_0 & \xrightarrow{J} & J(c_0) \\ \widetilde{S}\big\downarrow & & \big\downarrow i \\ (\ell^1)' & \xleftarrow[T']{} & c_0'' \end{array}$$

Sachant que J et T' sont des bijections, mais que $\widetilde{S}$ n'est pas surjective (car S est bijective de ℓ^∞ dans $(\ell^1)'$ et c_0 est un sous-ensemble strict de ℓ^∞), on peut maintenant conclure que i n'est pas surjective. Autrement dit, $J(c_0)$ est un sous-espace strict de c_0''.

■

Définition. On dit qu'un espace de Banach E est **réflexif** si $J(E) = E''$.

Remarque. Si E est réflexif, on identifie E'' à E en confondant chaque $\varphi \in E''$ avec l'unique $x \in E$ tel que $\varphi = J(x)$.

11.2 Résultats

Pour alléger la présentation, on convient dorénavant de noter $\overline{B}_F$ la boule unité fermée de l'e.v.n. F.

Théorème. *Soit E un Banach. Alors E est réflexif si et seulement si $\overline{B}_E$ est compacte pour la topologie faible $\sigma(E, E')$.*

La preuve de ce théorème repose sur les trois lemmes suivants :

Lemme 1. *Un Banach E est réflexif si et seulement si $J(\overline{B}_E) = \overline{B}_{E''}$.*

Preuve : Supposons $J(\overline{B}_E) = \overline{B}_{E''}$. Soit $\xi \in E''$. Il est clair que l'antécédent de 0 par J est 0. On peut donc supposer que $\xi \neq 0$. On a alors $\xi/\|\xi\|_{E''} \in \overline{B}_{E''}$ donc il existe $x \in \overline{B}_E$ tel que $\xi/\|\xi\|_{E''} = J(x)$. Il est clair que $\xi = J(\|\xi\|_{E''} x)$ et l'on peut donc conclure que E est réflexif.

Réciproquement, supposons E réflexif. Comme J conserve la norme, il est clair que $J(\overline{B}_E) \subset \overline{B}_{E''}$. Inversement, si $\xi \in \overline{B}_{E''}$, il existe par réflexivité de E, un $x \in E$ tel que $\xi = J(x)$. Sachant J conserve la norme, on doit avoir $\|x\|_E = \|\xi\|_{E''}$, donc $x \in \overline{B}_E$. Donc on a $\overline{B}_{E''} \subset J(\overline{B}_E)$, d'où $J(\overline{B}_E) = \overline{B}_{E''}$. ■

Lemme 2. *Soit E un e.v.n, $L_1, \cdots, L_n$ des formes linéaires continues sur E, et $\alpha_1, \cdots, \alpha_n$ des réels. Les deux assertions suivantes sont équivalentes :*

(i) $\forall \varepsilon > 0,\ \exists x_\varepsilon \in \overline{B}_E,$ *tel que* $\sup_{1\leq i\leq n} |L_i(x_\varepsilon) - \alpha_i| < \varepsilon\big)$,

(ii) $\forall (\beta_1, \cdots, \beta_n) \in \mathbb{R}^n,\ |\sum_{i=1}^n \beta_i \alpha_i| \leq \| \sum_{i=1}^n \beta_i L_i\|_{E'}$.

Preuve : $(i) \Longrightarrow (ii)$. Soit $(\beta_1, \cdots, \beta_n) \in \mathbb{R}^n$, $\varepsilon > 0$ et $x_\varepsilon \in \overline{B}_E$ tel que $|L_i(x_\varepsilon) - \alpha_i| < \varepsilon$ pour tout $i \in \{1, \cdots, n\}$. On a

$$\begin{aligned} |\textstyle\sum_{i=1}^n \beta_i \alpha_i| &\leq \left|\textstyle\sum_{i=1}^n \beta_i(\alpha_i - L_i(x_\varepsilon))\right| + \left|\Big(\textstyle\sum_{i=1}^n \beta_i L_i\Big)(x_\varepsilon)\right|, \\ &\leq \varepsilon \textstyle\sum_{i=1}^n |\beta_i| + \left\|\textstyle\sum_{i=1}^n \beta_i L_i\right\|_{E'} \quad \text{car} \quad \|x_\varepsilon\|_E \leq 1. \end{aligned}$$

Il ne reste plus qu'à faire tendre ε vers 0.

$(ii) \Longrightarrow (i)$. Notons α le vecteur de $\mathbb{R}^n$ de composantes $(\alpha_1, \cdots, \alpha_n)$ et

$$\varphi \ : \left\{ \begin{array}{lcl} E & \longrightarrow & \mathbb{R}^n \\ x & \longmapsto & (L_1(x), \cdots, L_n(x)). \end{array} \right.$$

On voit que montrer (i) revient à établir que α est dans l'adhérence dans $\mathbb{R}^n$ de $\varphi(\overline{B}_E)$ pour la norme $\|\cdot\|_\infty$.

Raisonnons par l'absurde : supposons que $\alpha \notin \overline{\varphi(\overline{B}_E)}$. L'ensemble $\overline{\varphi(\overline{B}_E)}$ étant convexe fermé et le singleton $\{\alpha\}$, convexe compact, la version géométrique du théorème de Hahn-Banach assure qu'il existe un hyperplan H de $\mathbb{R}^n$ qui sépare strictement $\{\alpha\}$ et $\overline{\varphi(\overline{B}_E)}$. Notons $\sum_{i=1}^n \beta_i y_i = \gamma$ l'équation de cet hyperplan. On a donc

$$\forall x \in \overline{B}_E,\ \sum_{i=1}^n \beta_i L_i(x) < \gamma < \sum_{i=1}^n \beta_i \alpha_i.$$

Donc on a

$$\Big\|\sum_{i=1}^n \beta_i L_i\Big\|_{E'} \leq \gamma < \sum_{i=1}^n \beta_i \alpha_i,$$

ce qui contredit (ii). ∎

Lemme 3. *L'ensemble $J(\overline{B}_E)$ est dense un sous-ensemble dense dans $\overline{B}_{E''}$ pour la topologie faible $*$ sur E''.*

Preuve : Comme J conserve la norme, il est clair que $J(\overline{B}_E) \subset \overline{B}_{E''}$.

Soit $\xi \in \overline{B}_{E''}$ et V un voisinage de ξ pour la topologie $\sigma(E'', E')$. Quitte à réduire V, on peut supposer qu'il existe $\varepsilon > 0$ et $L_1, \cdots, L_n$ dans E' tels que

$$V = \big\{\eta \in E'' \mid \sup_{1\leq i\leq n} |\eta(L_i) - \xi(L_i)| < \varepsilon\big\}.$$

Notons $\alpha_i = \xi(L_i)$. Étant donné que $\|\xi\|_{E''} \leq 1$, on voit que pour tout $(\beta_1, \cdots, \beta_n) \in \mathbb{R}^n$, on a

$$\begin{aligned} \left|\textstyle\sum_{i=1}^n \alpha_i \beta_i\right| &= \left|\xi\Big(\textstyle\sum_{i=1}^n \beta_i L_i\Big)\right|, \\ &\leq \|\xi\|_{E''} \left\|\textstyle\sum_{i=1}^n \beta_i L_i\right\|_{E'}, \\ &\leq \left\|\textstyle\sum_{i=1}^n \beta_i L_i\right\|_{E'}. \end{aligned}$$

Le lemme précédent assure donc qu'il existe $x_\varepsilon \in \overline{B}_E$ tel que

$$\forall i \in \{1, \cdots, n\},\ |L_i(x_\varepsilon) - \alpha_i| = |(J(x_\varepsilon))(L_i) - \xi(L_i)| < \varepsilon.$$

Donc $J(x_\varepsilon) \in V$ avec $x_\varepsilon \in \overline{B}_E$. ∎

Remarque. Comme J conserve la norme, il est facile de montrer que $J(\overline{B}_E)$ est toujours fermé pour la topologie forte de E''. En effet, si $J(x_n) \to \xi$ pour la topologie forte de E'', la suite $(J(x_n))_{n\in\mathbb{N}}$ est de Cauchy (car $\|J(x_n) - J(x_m)\|_{E''} = \|x_n - x_m\|_E$) donc converge car E'' est complet.

En conséquence, $J(\overline{B}_E)$ dense pour la topologie forte de E'' entraîne $J(\overline{B}_E) = \overline{B}_{E''}$, et donc E réflexif.

Preuve du théorème :

$\Longleftarrow$ Supposons $\overline{B}_E$ compacte pour la topologie $\sigma(E, E')$. Comme $J : E \to E''$ est continue pour la topologie forte, elle l'est aussi si l'on munit E de la topologie faible $\sigma(E, E')$ et E'' de la topologie faible $\sigma(E'', E''')$ (cf chapitre 9). Comme $\sigma(E'', E')$ est moins fine que $\sigma(E'', E''')$, on en déduit que J est continue de $\big(E; \sigma(E, E')\big)$ dans $\big(E''; \sigma(E'', E')\big)$. En conséquence $J(\overline{B}_E)$ est compact pour $\sigma(E'', E')$ donc fermé. Mais le lemme 3 nous garantit que $J(\overline{B}_E)$ est dense dans $\overline{B}''_E$ pour $\sigma(E'', E')$. Donc $J(\overline{B}_E) = \overline{B}_{E''}$.

Grâce au lemme 1, on peut maintenant conclure que E réflexif.

$\Longrightarrow$ Supposons E réflexif.

D'après le lemme 1, $J(\overline{B}_E) = \overline{B}_{E''}$. De plus, $\overline{B}_{E''}$ est compacte pour $\sigma(E'', E')$. Donc il suffit de prouver que J^{-1} est continue de $\big(E''; \sigma(E'', E')\big)$ dans $\big(E; \sigma(E, E')\big)$. Pour ce faire, considérons un ouvert U de E pour $\sigma(E, E')$. Il s'agit de montrer que $J(U)$ est un ouvert de E'' pour $\sigma(E'', E')$. Soit $x_0 \in U$. Il existe $L_1, \cdots, L_n \in E'$ et $\varepsilon > 0$ tels que

$$V \stackrel{\text{déf}}{=} \{x \in E \mid \forall i \in \{1, \cdots, n\}, |L_i(x - x_0)| < \varepsilon\} \subset U.$$

Par définition de J, on a

$$J(V) = \big\{\xi \in E'' \mid \xi = J(x) \quad \text{et} \quad \forall i \in \{1, \cdots, n\}, |L_i(x) - L_i(x_0)| < \varepsilon\big\}.$$

Mais comme on a $\xi(L_i) = L_i(x)$ pour $\xi = J(x)$, cela peut se récrire

$$J(V) = \big\{\xi \in E'' \mid \forall i \in \{1, \cdots, n\}, |\xi(L_i) - J(x_0)(L_i)| < \varepsilon\big\}.$$

Cet ensemble est un voisinage de $J(x_0)$ pour la topologie $\sigma(E'', E')$. En conséquence, J^{-1} est continue, et le théorème est démontré.

■

Proposition 1. *Soit E un Banach réflexif et M un s.e.v. fermé de E pour la topologie forte. Alors M muni de la norme de E est aussi un Banach réflexif.*

Preuve : Comme l'ensemble M est un Banach (car s.e.v. fermé d'un Banach), on peut le munir de la topologie faible $\sigma(M, M')$.

Montrons que $\sigma(M, M')$ n'est autre que la topologie induite par $\sigma(E, E')$ sur M.

Soit U un ouvert de $\sigma(M, M')$. On peut toujours supposer que U est de la forme

$$U = \big\{x \in M \mid \forall i \in \{1, \cdots, n\}, |L_i(x - x_0)| < \alpha_i\big\} \quad \text{avec} \quad L_i \in M' \quad \text{pour} \quad i = 1, \cdots, n. \tag{11.1}$$

On utilise le théorème de Hahn-Banach pour prolonger les L_i en forme linéaires $\widetilde{L}_i$ continues sur E. Il est alors clair que $U = M \cap V$ avec

$$V = \big\{x \in E \mid \forall i \in \{1, \cdots, n\}, |\widetilde{L}_i(x - x_0)| < \alpha_i\big\}.$$

Comme V est un ouvert de $\sigma(E, E')$, l'ensemble U est un ouvert pour la topologie induite par $\sigma(E, E')$ sur M.

Inversement, si V est un ouvert de $\sigma(E,E')$, on peut toujours supposer qu'il est de la forme

$$V = \{x \in E \mid \forall i \in \{1,\cdots,n\},\ |\widetilde{L}_i(x-x_0)| < \alpha_i\} \text{ avec } \widetilde{L}_i \in E' \text{ pour } i=1,\cdots,n.$$

En notant L_i la restriction de $\widetilde{L}_i$ à M, on constate que $L_i \in M'$ et que $V \cap M$ est donné par (11.1), donc est un ouvert de $\sigma(M,M')$.

Nous pouvons maintenant passer à la preuve de la proposition proprement dite. L'ensemble M est un ensemble convexe (car c'est un s.e.v) fermé pour la topologie forte, donc également fermé pour $\sigma(E,E')$ (voir le chapitre 9).

On a $\overline{B}_M = M \cap \overline{B}_E$ et $\overline{B}_E$ est compacte pour $\sigma(E,E')$ car E est réflexif. Donc $\overline{B}_M$ est compacte pour la topologie induite par $\sigma(E,E')$ sur M, donc pour $\sigma(M,M')$. En appliquant la réciproque du théorème précédent, on peut maintenant conclure que M est réflexif. ∎

La propriété de réflexivité est stable par isomorphisme continu :

Proposition 2. *Soit E_1, E_2 deux Banach et $T \in \mathcal{L}(E_1;E_2)$ bijective. Alors E_1 est réflexif si et seulement si E_2 est réflexif.*

Preuve : Pour $i=1,2$, on note $J_i : E_i \longrightarrow E_i''$ l'isométrie définie par

$$\forall x_i \in E_i, \forall f_i \in E_i',\ (J_i(x_i))(f_i) = f_i(x_i).$$

On considère $T' : E_2' \longrightarrow E_1'$ l'application transposée de T définie par

$$\forall x_1 \in E_1, \forall L_2 \in E_2',\ [T'(L_2)](x_1) = L_2(T(x_1))$$

puis $T'' : E_1'' \longrightarrow E_2''$ l'application transposée de T' définie par

$$\forall \xi_1 \in E_1'', \forall L_2 \in E_2',\ [T''(\xi_1)](L_2) = \xi_1(T'(L_2)).$$

Comme T est bijective continue, il en est de même pour T' et pour T'' (exercice :le vérifier). Nous avons donc le diagramme suivant :

$$\begin{array}{ccc} E_1 & \xrightarrow{J_1} & E_1'' \\ T\big\downarrow & & \big\downarrow T'' \\ E_2 & \xrightarrow[J_2]{} & E_2'' \end{array}$$

Supposons E_1 réflexif de telle sorte que l'application $J_1 : E_1 \to E_1''$ soit une isométrie bijective. Le diagramme ci-dessus montre que J_2 est également bijective, et donc E_2'' est réflexif.

Cela peut se retrouver par le calcul : soit $\xi_2 \in E_2''$. Il s'agit de montrer que ξ_2 a un antécédent par J_2.

Comme J_1 et T'' sont bijectives, il existe $x_1 \in E_1$ tel que $T''^{-1}(\xi) = J_1(x_1)$, puis, comme T est bijective, un élément y_2 de E_2 tel que

$$\xi_2 = T''(J_1(T^{-1}(x_2))).$$

Soit $f_2 \in E_2'$. On a

$$\begin{array}{rcl} \xi_2(f_2) & = & T''(J_1(T^{-1}(x_2)))(f_2), \\ & = & J_1(T^{-1}(x_2))(T'(f_2)), \\ & = & [T'(f_2)](T^{-1}(x_2)), \\ & = & f_2(x_2), \\ & = & [J_2(x_2)](f_2). \end{array}$$

Donc $\xi_2 = J_2(x_2)$.

L'implication réciproque se démontre en échangeant les rôles de E_1 et de E_2 et en remplaçant T par T^{-1}. Cela est licite car le premier corollaire du théorème de l'application ouverte assure la continuité de T^{-1}. ■

De cette proposition et de la précédente, on déduit le résultat important suivant :

Théorème. *Soit E un Banach. Alors E est réflexif si et seulement si E' est réflexif.*

Preuve : Supposons E réflexif. Alors les topologies $\sigma(E', E)$ et $\sigma(E', E'')$ coïncident (pour le voir, le plus simple est de revenir à la définition de ces deux topologies). On sait d'après le théorème de Banach-Alaoglu-Bourbaki que $\overline{B}_{E'}$ est compacte pour $\sigma(E', E)$. Donc elle l'est aussi pour $\sigma(E', E'')$ et le théorème précédent assure donc que E' est réflexif.

Réciproquement, supposons que E' soit réflexif. En reprenant le raisonnement précédent, on voit que E'' est réflexif. Comme $J(E)$ est un sous-espace fermé de E'' pour la topologie forte, la proposition 1 assure que $J(E)$ est réflexif.

Enfin, J est une isométrie bijective de E sur $J(E)$, et la proposition 2 permet de conclure à la réflexivité de E. ■

11.3 Espaces uniformément convexes

Dans cette partie, nous donnons un critère explicite assurant la réflexivité : il s'agit de l'uniforme convexité.

Définition. On dit qu'un Banach E est **uniformément convexe** si

$$\forall \varepsilon > 0,\ \exists \delta > 0,\ \Big(x \in \overline{B}_E,\ y \in \overline{B}_E \text{ et } \|x - y\|_E > \varepsilon\Big) \Longrightarrow \left\| \frac{x+y}{2} \right\| < 1 - \delta.$$

Exemple. Grâce à l'identité, du parallélogramme, on constate que tout espace euclidien est uniformément convexe.

En revanche, $\mathbb{R}^n$ muni de la norme $\|x\|_1 = \sum_{i=1}^n |x_i|$ ou $\|x\|_\infty = \sup_{i=1}^n |x_i|$ ne l'est pas.

Cet exemple simple montre que **l'uniforme convexité dépend de la norme choisie**. En particulier, même en dimension finie, changer la norme en une norme équivalente peut faire perdre la propriété d'uniforme convexité.

Théorème. *Un espace de Banach uniformément convexe est réflexif.*

Preuve : Soit E un espace uniformément convexe. Pour montrer que E est réflexif, il suffit d'établir que tout $\xi \in E''$ de norme 1 est dans $J(\overline{B}_E)$. Sachant que $J(\overline{B}_E)$ est fermé pour la topologie forte de E'', il suffit en fait d'établir que

$$\forall \varepsilon > 0, \exists x_\varepsilon \in \overline{B}_E,\ \|J(x_\varepsilon) - \xi\|_{E''} \leq \varepsilon.$$

Fixons donc $\varepsilon > 0$ et choisissons δ selon la définition de l'uniforme convexité. Comme $\|\xi\|_{E''} = 1$, il existe $L \in E'$ de norme 1 tel que $\xi(L) > 1 - \frac{\delta}{2}$. Soit

$$V \stackrel{\text{déf}}{=} \big\{\eta \in E'' \mid |\eta(L) - \xi(L)| < \frac{\delta}{2}\big\}.$$

Par construction, V est un voisinage de ξ pour la topologie faible * sur E''. On a vu précédemment que $J(\overline{B}_E)$ était dense dans $\overline{B}_{E''}$. Donc il existe $x \in \overline{B}_E$ tel que $J(x) \in V$. Montrons que

$$\|J(x) - \xi\|_{E''} \leq \varepsilon. \tag{11.2}$$

Supposons par l'absurde que cette inégalité n'est pas satisfaite. Alors ξ appartient à l'ensemble $W \stackrel{\text{déf}}{=} E'' \setminus \big(J(x) + \varepsilon \overline{B}_{E''}\big)$. Du fait que $\overline{B}_{E''}$ est compacte donc fermée, cet ensemble est ouvert pour la topologie faible *. Par ailleurs V et W contiennent ξ donc $V \cap W$ rencontre $\overline{B}_{E''}$. Par densité de $J(\overline{B}_E)$, il existe donc $y \in \overline{B}_E$ tel que $J(y) \in V \cap W$. Par construction, on a donc nécessairement

$$\|J(x) - J(y)\|_{E''} > \varepsilon.$$

Par ailleurs, puisque $J(y) \in V$ et $J(x) \in V$, on a

$$|L(x) - \xi(L)| < \frac{\delta}{2} \quad \text{et} \quad |L(y) - \xi(L)| < \frac{\delta}{2},$$

d'où

$$2\xi(L) \le L(x+y) + \delta \le \|x+y\|_E + \delta.$$

Mais comme $\xi(L) > 1 - \frac{\delta}{2}$, on peut maintenant conclure que $\left\|\frac{x+y}{2}\right\|_E \ge 1 - \delta$. Cela contredit l'hypothèse de convexité uniforme.

En conséquence, (11.2) est réalisée et le théorème, démontré. ■

11.4 Les espaces séparables

Définition. Soit (X, d) un espace métrique. On dit que X est **séparable** si X admet une partie dénombrable et dense.

Exemples. 1. Les espaces vectoriels $\mathbb{R}^n$ ou $\mathbb{C}^n$ munis d'une norme quelconque sont séparables : en effet $\mathbb{Q}^n$ est dense dans $\mathbb{R}^n$, et l'ensemble

$$\{(x_1 + iy_1, \cdots, x_n + iy_n) \mid (x_1, \cdots, x_n) \in \mathbb{Q}^n \text{ et } (y_1, \cdots, y_n) \in \mathbb{Q}^n\}$$

est dense dans $\mathbb{C}^n$.

2. Pour $p \in [1, +\infty[$, l'ensemble $\ell^p(\mathbb{R})$ des suites réelles de puissances p-ièmes sommables est séparable. Nous laissons au lecteur le soin d'établir que l'ensemble des suites à termes rationnels et nulles à partir d'un certain rang est dénombrable et dense dans $\ell^p(\mathbb{R})$.

   De même, $\ell^p(\mathbb{C})$ est séparable pour $p \in [1, +\infty[$.

3. L'ensemble $\ell^\infty(\mathbb{R})$ des suites bornées muni de la norme $\|u\|_\infty = \sup_{n\in\mathbb{N}} |u_n|$ n'est pas séparable. Il en est de même pour $\ell^\infty(\mathbb{C})$.

   En effet, supposons que $\ell^\infty(\mathbb{R})$ admette une partie dénombrable dense. Les éléments de cette partie peuvent être rangés en une suite $(u^n)_{n\in\mathbb{N}}$. Définissons $v \in \ell^\infty(\mathbb{R})$ par $v_n = -1$ si $u^n_n > 0$ et $v_n = 1$ sinon. On a

$$\forall n \in \mathbb{N},\ \|v - u^n\|_\infty \ge \|v_n - u^n_n\| \ge 1.$$

   Cela montre que $v \notin \overline{(u^n)_{n\in\mathbb{N}}}$.

Exercice : Montrer que dans les e.v.n, la propriété de séparabilité est stable par isomorphisme continu et, en particulier, par changement de norme en une norme équivalente.

Théorème. *Soit E un Banach. Si E' est séparable alors E l'est aussi.*

Preuve : Soit $(f_n)_{n\in\mathbb{N}}$ une suite dénombrable dense de E'. Pour tout $n \in \mathbb{N}$, on fixe un $x_n \in E$ de norme 1 tel que $|f_n(x_n)| \geq \frac{1}{2}\|f_n\|_{E'}$. On considère alors l'ensemble A des combinaisons linéaires finies à coefficients rationnels de termes de la suite $(x_n)_{n\in\mathbb{N}}$.

On vérifie aisément que A est dénombrable et dense dans l'ensemble $\widetilde{A}$ des combinaisons linéaires finies à coefficients réels de termes de la suite $(x_n)_{n\in\mathbb{N}}$.

L'ensemble $\widetilde{A}$ est un s.e.v. D'après le corollaire 2 du théorème de Hahn-Banach (forme géométrique), pour montrer que $\widetilde{A}$ est dense dans E, il suffit d'établir que le seul élément de E' s'annulant sur $\widetilde{A}$ est la forme linéaire nulle.

Soit donc $f \in E'$ s'annulant sur $\widetilde{A}$. Fixons $\varepsilon > 0$. Soit $n \in \mathbb{N}$ tel que $\|f - f_n\|_{E'} \leq \frac{\varepsilon}{3}$. Comme $f(x_n) = 0$, on a

$$\frac{1}{2}\|f_n\|_{E'} \leq |(f - f_n)(x_n)| \leq \frac{\varepsilon}{3}.$$

Donc

$$\|f\|_{E'} \leq \|f - f_n\|_{E'} + \|f_n\|_{E'} \leq \frac{\varepsilon}{3} + \frac{2\varepsilon}{3} = \varepsilon.$$

En conclusion, $\|f\|_{E'} = 0$ puis $f = 0$. ■

Remarque. La réciproque est en général fausse. En effet, l'ensemble ℓ^1 est séparable mais son dual $(\ell^1)'$ peut s'identifier isométriquement à l'espace ℓ^∞ qui, lui, n'est pas séparable.

On a cependant le résultat suivant :

Corollaire. *Soit E un Banach. Alors E est réflexif et séparable si et seulement si E' est réflexif et séparable.*

Preuve : On sait déjà que E' séparable entraîne E séparable et E' réflexif implique E réflexif, donc l'implication réciproque est vraie.

Soit maintenant E réflexif séparable. Alors $E'' = J(E)$ est réflexif séparable, donc E' aussi. ■

Remarque. Sachant que ℓ^1 est séparable mais que ℓ^∞ ne l'est pas, on en déduit que ℓ^1 n'est pas réflexif.

Théorème. *Soit E un Banach séparable. Alors il existe une distance d définie sur $\overline{B}_{E'}$ et telle que la topologie associée à d soit celle de la topologie induite par $\sigma(E', E)$ sur $\overline{B}_{E'}$. On dit que la boule unité fermée de E' est* **métrisable** *pour la topologie faible *.*

Preuve : Soit $(x_n)_{n\in\mathbb{N}}$ une suite dénombrable dense dans $\overline{B}_E$. On pose pour tout $(f, g) \in E'$,

$$d(f, g) = \sum_{n=0}^{\infty} 2^{-n}|g(x_n) - f(x_n)|.$$

Comme tous les x_n vérifient $\|x_n\|_E \leq 1$, on vérifie facilement que d est à valeurs dans $\mathbb{R}^+$ puis que c'est une distance sur E'. Reste à montrer que la topologie associée à d coïncide avec la topologie induite par $\sigma(E', E)$ sur $\overline{B}_E$.

- Soit $f_0 \in \overline{B}_{E'}$ et V un voisinage de V pour f_0 au sens de la topologie $\sigma(E', E)$. Nous allons montrer que V contient une boule ouverte centrée en f_0 pour la métrique d.
  Comme d'habitude, on peut supposer que

$$V = \{f \in \overline{B}_{E'} \mid \max_{1\leq i\leq n} |f(y_i) - f_0(y_i)| < \varepsilon\}.$$

Pour tout $i \in \{1, \cdots, n\}$, il existe $k_i \in \mathbb{N}$ tel que $\|x_{k_i} - y_i\|_E < \frac{\varepsilon}{4}$. Choisissons $r > 0$ tel que $2^{k_i} r < \frac{\varepsilon}{2}$ pour tout $i \in \{1, \cdots, n\}$. Soit $U \stackrel{\text{déf}}{=} \{f \in \overline{B}_{E'} \mid d(f, f_0) < r\}$. Alors on a pour tout $f \in U$ et $i \in \{1, \cdots, n\}$,

$$\begin{aligned} |f(y_i) - f_0(y_i)| &\leq |f(y_i) - f(x_{k_i})| + |f(x_{k_i}) - f_0(x_{k_i})| + |f_0(x_{k_i}) - f_0(y_i)|, \\ &< \tfrac{\varepsilon}{4} + 2^{k_i} r + \tfrac{\varepsilon}{4}. \end{aligned}$$

Vue l'hypothèse sur r, on a donc $|f(y_i) - f_0(y_i)| < \varepsilon$, d'où $U \subset V$.

- Inversement, fixons $f_0 \in \overline{B}_{E'}$ et $r > 0$, et montrons que

$$U \stackrel{\text{déf}}{=} \{f \in \overline{B}_{E'} \mid d(f, f_0) < r\}$$

est un voisinage de f_0 pour la topologie faible *. On a pour tout $k \in \mathbb{N}$,

$$d(f, f_0) = \sum_{n=0}^{k} 2^{-n} |f(x_n) - f_0(x_n)| + \sum_{n=k+1}^{+\infty} 2^{-n} |f(x_n) - f_0(x_n)|.$$

Choisissons k tel que $2^{k-2} r > 1$ et posons

$$V \stackrel{\text{déf}}{=} \{f \in \overline{B}_{E'} \mid \max_{0 \leq i \leq k} |(f - f_0)(x_i)| < \frac{r}{4}\}.$$

L'ensemble V est un ouvert pour la topologie induite par $\sigma(E', E)$ sur $\overline{B}_{E'}$. de plus, $f \in V$ entraîne

$$d(f, f_0) \leq \frac{r}{4} \sum_{n=0}^{k} 2^{-n} + 2 \sum_{n=k+1}^{\infty} 2^{-n} < \frac{r}{2} + \frac{1}{2^{k-1}} < r.$$

Donc on a $V \subset U$.

■

Remarque. La réciproque du théorème ci-dessus est vraie mais nous ne l'utiliserons pas.

Théorème. *Soit E un Banach réflexif et $(x_n)_{n \in \mathbb{N}}$ une suite bornée d'éléments de E. Alors il existe une sous-suite $(x_{\varphi(n)})_{n \in \mathbb{N}}$ qui converge pour la topologie faible $\sigma(E, E')$.*

Preuve : Quitte à multiplier chaque élément de la suite par une constante positive fixe, on peut supposer que $(x_n)_{n \in \mathbb{N}}$ est une suite d'éléments de $\overline{B}_E$.

Soit M_0 l'espace vectoriel engendré par les x_n, et M l'adhérence de M_0 (pour la topologie forte). Il est clair que M est séparable (l'ensemble des combinaisons linéaires finies à coefficients rationnels de x_n est une partie dense de M), et réflexif (car s.e.v. fermé de E réflexif). Donc $\overline{B}_M$ est compacte pour $\sigma(M, M')$. De plus, comme M est réflexif et séparable, M' aussi, donc, d'après le théorème précédent, $\overline{B}_{M''}$ est métrisable pour $\sigma(M'', M')$. Comme il existe une isométrie bijective entre M et M'', la boule $\overline{B}_M$ est également métrisable pour $\sigma(M, M')$. Cela signifie que $\overline{B}_M$ muni de la topologie faible est un espace métrique compact.

D'après le théorème de Bolzano-Weierstrass, il existe donc une sous-suite $(x_{\varphi(n)})_{n \in \mathbb{N}}$ qui converge pour la topologie faible $\sigma(M, M')$ donc pour $\sigma(E, E')$ (car $\sigma(M, M')$ coïncide avec la topologie induite par $\sigma(E, E')$ sur M). ■

Comma application, donnons le résultat suivant qui généralise un théorème bien connu en dimension finie (voir le cours de calcul différentiel de licence) :

Théorème 1. *Soit E un Banach réflexif et $A \subset E$ un convexe fermé non vide. Soit $\varphi \in \mathcal{C}(A;\mathbb{R})$. On suppose de plus que φ est convexe et que*

$$\lim_{\substack{x\in A\\ \|x\|_E\to+\infty}} \varphi(x) = +\infty.$$

Alors φ est minorée et atteint son minimum absolu : il existe $a \in A$ tel que

$$\forall x \in A,\ \varphi(x) \geq \varphi(a).$$

Preuve : Soit $m = \inf_{x\in A}\varphi(x)$ et $(x_n)_{n\in\mathbb{N}}$ une suite d'éléments de A telle que

$$\lim_{n\to+\infty} \varphi(x_n) = m.$$

Comme φ tend vers l'infini à l'infini, cette suite est bornée donc admet une sous-suite $(x_{\varphi(n)})_{n\in\mathbb{N}}$ faiblement convergente. Notons a la limite de cette sous-suite.
Soit $x \in A$ et $A_x = \{y \in A \mid \varphi(y) \leq \varphi(x)\}$. Cet ensemble est convexe fermé (car φ est convexe continue) donc faiblement fermé. Si $\varphi(x) = m$, le théorème est prouvé. Sinon, l'ensemble A_x contient les termes de la suite $(x_{\varphi(n)})_{n\in\mathbb{N}}$ à partir d'un certain rang donc a car A_x est faiblement fermé. On a donc en particulier $\varphi(a) \leq \varphi(x)$. Cela achève la preuve du théorème. ∎

11.5 Les espaces L^p

Dans toute cette section, A désigne un borélien de mesure non nulle de $\mathbb{R}^n$. Pour la définition des espaces $L^p(A;\mathbb{K})$ dans le cas p fini, on renvoie à la section 7.8.1. Intéressons-nous maintenant au cas $p = \infty$. On définit l'ensemble $\mathcal{L}^\infty(A;\mathbb{K})$ des fonctions mesurables $f : A \to \mathbb{R}$ telles qu'il existe $C \in \mathbb{R}^+$ vérifiant

$$|f(x)| \leq C \ \text{ pour presque tout } \ x \in \Omega,$$

puis l'ensemble $L^\infty(A;\mathbb{K})$ des classes d'équivalence de fonctions de $\mathcal{L}^\infty(A;\mathbb{K})$ pour la relation d'égalité presque partout au sens de la mesure de Lebesgue.

On munit $L^\infty(A;\mathbb{K})$ d'une norme en prenant pour $\|f\|_{L^\infty}$ la borne inférieure de tous les C tels que $|f(x)| \leq C$ presque partout.

Dans la suite de cette section, on utilise la notation raccourcie L^p pour désigner $L^p(A;\mathbb{K})$. En adaptant les arguments utilisés dans le chapitre 7 dans le cas de la dimension 1, on obtient le résultat classique suivant :

Théorème. *Pour tout $p \in [1,\infty]$ l'espace vectoriel normé $(L^p, \|\cdot\|_p)$ est complet.*

Les espaces L^p possèdent la propriété fondamentale suivante :

Théorème. *Pour tout $p \in]1,\infty[$, l'espace $(L^p, \|\cdot\|_{L^p})$ est réflexif.*

Preuve : Dans le cas $p \in [2,\infty[$, la preuve de ce résultat repose sur l'**inégalité de Clarkson**

$$\forall (a,b) \in \mathbb{R}^2,\ \left|\frac{a+b}{2}\right| + \left|\frac{a-b}{2}\right| \leq \frac{1}{2}\Big(|a|^p + |b|^p\Big)$$

que nous admettons provisoirement et qui va nous permettre de montrer que L^p est *uniformément convexe.*

En effet, si $\|f\|_{L^p} \leq 1$, $\|g\|_{L^p} \leq 1$ et $\|f-g\|_{L^p} \geq \varepsilon$ pour un ε donné alors l'inégalité de Clarkson assure que

$$\begin{aligned}
\left\|\tfrac{f+g}{2}\right\|_{L^p}^p &= \int_A \left|\tfrac{f(x)+g(x)}{2}\right|^p dx,\\
&\leq \tfrac{1}{2}\Big(\int_A |f(x)|^p\,dx + \int_A |g(x)|^p\,dx\Big) - \int_A \left|\tfrac{f(x)-g(x)}{2}\right|^p dx,\\
&\leq 1-\left(\tfrac{\varepsilon}{2}\right)^p.
\end{aligned}$$

On peut donc prendre $\delta = 1-\left(1-\left(\frac{\varepsilon}{2}\right)^p\right)^{\frac{1}{p}}$ dans la définition de l'uniforme convexité.

La preuve de l'inégalité de Clarkson repose sur un argument de convexité. Remarquons tout d'abord que pour tout $x\in[0,1]$, et $q\geq 1$, on a

$$x^q+(1-x)^q \leq x+(1-x)=1.$$

En appliquant ceci à $x=\alpha^2/(\alpha^2+\beta^2)$, et $q=p/2$, on obtient

$$\forall(\alpha,\beta)\in\mathbb{R}^2,\ |\alpha|^p+|\beta|^p = (\alpha^2)^{\frac{p}{2}}+(\beta^2)^{\frac{p}{2}} \leq (\alpha^2+\beta^2)^{\frac{p}{2}}.$$

Enfin, en choisissant $\alpha=(a+b)/2$ et $\beta=(a-b)/2$, on en déduit que

$$\left|\frac{a+b}{2}\right| + \left|\frac{a-b}{2}\right| \leq \left(\frac{a^2+b^2}{2}\right)^{\frac{p}{2}}.$$

Par convexité de la fonction $x\mapsto x^{\frac{p}{2}}$, le membre de droite se majore par $\frac{1}{2}\big(|a|^p+|b|^p\big)$.

La preuve de la réflexivité dans le cas $p\in]1,2[$ repose sur un argument de *dualité*. Notons p' l'**exposant conjugué** de p défini par $1/p+1/p'=1$.

Pour $u\in L^p$ et $f\in L^{p'}$, on pose $T_u(f)=\int_A fu\,dx$. L'inégalité de Hölder assure que $T_u(f)$ est bien défini et que

$$|T_u(f)| \leq \|u\|_{L^p}\,\|f\|_{L^{p'}}\,.$$

En conséquence, T_u est une forme linéaire continue sur $L^{p'}$ et l'on a $\|T_u\|_{(L^{p'})'} \leq \|u\|_{L^p}$.

Montrons que l'on a $\|T_u\|_{(L^{p'})'} = \|u\|_{L^p}$.

Fixons $u\in L^p$ et posons $f=|u|^{p-1}\operatorname{sgn} u$. Il est clair que $f\in L^{p'}$ et que $\|f\|_{L^{p'}} = \|u\|_{L^p}^{p-1}$. De plus $T_u(f)=\|u\|_{L^p}^p$. Donc on a bien l'égalité souhaitée.

Cela assure que $T: u\mapsto T_u$ est une isométrie de L^p sur $(L^{p'})'$. On en déduit que T est injective et que $T(L^p)$ est un sous-espace fermé de $(L^{p'})'$.

Comme $p'\geq 2$, nous savons déjà que $L^{p'}$ est réflexif. Mais par ailleurs nous savons également que le dual d'un espace réflexif est réflexif donc $(L^{p'})'$ est réflexif. Ainsi $T(L^p)$ est également réflexif car sous-espace fermé de $(L^{p'})'$.

Finalement les espaces L^p et $T(L^p)$ se déduisent l'un de l'autre par une isométrie bijective, donc L^p est réflexif. ■

Les espaces L^p possèdent la propriété de dualité suivante :

Théorème (de représentation de Riesz)**.** *Soit $q\in]1,\infty[$ et q' l'exposant conjugué de q. Soit φ une forme linéaire continue sur L^q. Alors il existe un unique $u\in L^{q'}$ tel que*

$$\forall f\in L^q,\ \varphi(f)=\int_A uf\,dx.$$

De plus $\|\varphi\|_{(L^q)'} = \|u\|_{L^{q'}}$.

Autrement dit, l'application $T: u\mapsto T_u$ définie pour tout $u\in L^{q'}$ et $f\in L^q$ par $T_u(f)=\int_A fu\,dx$ est une isométrie bijective de $L^{q'}$ sur $\big(L^q\big)'$.

Preuve : Nous avons déjà établi que $T : u \mapsto T_u$ est une isométrie de $L^{q'}$ dans $(L^q)'$ (remplacer p' par q dans la démonstration précédente). Reste à montrer la surjectivité. Soit donc $\xi \in (L^q)''$ tel que $\xi(T_u) = 0$ pour tout $u \in L^{q'}$. La réflexivité de L^q assure l'existence d'un élément h de L^q tel que $\xi = J(h)$. On a donc

$$\forall u \in L^{q'},\ \xi(T_u) = T_u(h) = \int_A hu.$$

En prenant $u = |h|^{q-1} \operatorname{sgn} h$, on trouve que $\int |h|^q = 0$ donc $h = 0$ puis $\xi = 0$. Le corollaire 2 du théorème de Hahn-Banach géométrique permet de conclure que $T(L^{q'})$ est dense dans $(L^q)'$. Comme $T(L^{q'})$ est aussi fermé, on a $T(L^{q'}) = (L^q)'$. ■

Remarque. Voici quelques résultats autres résultats imporants sur les espaces $L^p(A)$:

- Pour $1 < p < \infty$, l'espace L^p est réflexif et séparable.
- L'espace L^1 est séparable mais non réflexif, et son dual peut être identifié à $L^\infty(A)$.
- L'espace L^∞ n'est ni séparable ni réflexif et son dual contient strictement L^1.

Bibliographie

[1] A. Bouziad et J. Calbrix : théorie de la mesure et de l'intégration, Publications de l'Université de Rouen, numéro 185.

[2] H. Brézis : Analyse fonctionnelle. Théorie et applications, **Masson**.

[3] A. Chambert-Loir et S. Fermigier : Agrégation de mathématiques, analyse 1,2 et 3, exercices, **Dunod**.

[4] J. Conway : A course in functional analysis, second edition, Graduate Texts in Mathematics, **Springer-Verlag**.

[5] R. Danchin : Notes de cours d'analyse pour la préparation au CAPES, téléchargeable sur `http ://perso-math.univ-mlv.fr/users/danchin.raphael/`

[6] J. Dieudonné : Éléments d'analyse, tomes 1 et 2, **Gauthier-Villars**.

[7] J. Dixmier : Cours de topologie générale, **Presses Universitaires de France**.

[8] D. Revuz : mesure et intégration, Hermann, 1997.

[9] W. Rudin : Analyse réelle et complexe. Masson.

[10] W. Rudin : Functional analysis. Second edition. International Series in Pure and Applied Mathematics. McGraw-Hill, Inc., New York, 1991.

[11] M. Samuelides et L. Touzillier : Problèmes d'analyse fonctionnelle et d'analyse harmonique, **Cépaduès éditions**.

[12] M. Schechter : Principles of Functional Analysis, second edition, Graduate Studies in Mathematics, **American Mathematical Society**.

Index