Université Paris Dauphine

Notes sur le cours d'Analyse fonctionnelle approfondie

Master mention Mathématiques appliquées 1ère année

2017-2018

Pierre Cardaliaguet

Introduction

Le cours tourne autour de la notion de *compacité* dans des espaces fonctionnels. Comme dans de nombreux autres domaines des mathématiques, la compacité joue ici un rôle central, que ce soit en optimisation (le "calcul des variations"), que pour les théorèmes de point fixe,... La principale difficulté cependant est qu'en dimension infinie (les espaces fonctionnels sont *toujours* de dimension infinie), les boules fermées ne sont jamais compacte.

Deux approches peuvent être envisagée pour surmonter cette difficulté. La plus directe est de chercher des critères de compacité. C'est par exemple le cas du *théorème d'Ascoli*, qui caractérise les sous-ensembles compact de l'ensemble des fonctions continues.

Cette démarche n'est cependant pas suffisante car les critères de compacité sont souvent très restrictifs. Une autre possibilité est de relaxer la topologie et de chercher des ensembles compacts pour une topologie plus faible, associée à une notion de *convergence faible*. Pour cette topologie faible, il y a plus d'ensembles compacts : par exemple les boules fermées deviennent compactes pour cette topologie. Il y a aussi moins de fonctions continues.

Les notions introduites seront illustrées par deux domaines d'applications : d'une part, le calcul des variations, qui n'est rien d'autre que de l'optimisation dans des espaces fonctionnels ; d'autre part l'analyse des opérateurs linéaires compacts (i.e., tels que l'image de tout ensemble borné est relativement compacte), qui partagent de nombreuses similitudes avec les applications linéaires en dimension finie.

Ce texte a largement emprunté aux ouvrages suivants :

- H. Brézis. Analyse fonctionnelle. Collection Mathématiques Appliquées pour la Maîtrise. Masson, 1983.

- L.C. Evans. Partial differential equations, volume 19 of Graduate Studies in Mathematics. American Mathematical Society, Providence, Second edition, 2010.

- L. Schwartz. Analyse. Collection Enseignement des Sciences, Vol. 42. Hermann, Paris, 1991.

Table des matières

1 Compacité, théorème d'Ascoli

1.1 Compacité

On rappelle qu'un espace métrique est une paire (X, d), où X est un ensemble et $d : X \to [0, +\infty[$ est une distance, i.e., une application vérifiant

1. $d(x, y) \geq 0$ pour tout $x, y \in X$ et $\quad [d(x, y) = 0 \ \Leftrightarrow \ x = y]$.
2. $d(x, z) \leq d(x, y) + d(y, z)$ pour tout $x, y, z \in X$.

Définition 1.1 *On dit qu'un espace métrique (X, d) est compact si, de toute suite (x_n) de X on peut extraire une sous-suite qui converge vers un élément de X.*

Notons que, par définition, tout sous-ensemble fermé d'un compact est compact. Il est également facile de voir qu'un ensemble compact est forcément complet.

Soit $(\mathcal{O}_i)_{i\in I}$ une famille (quelconque) d'ouverts de X. On dit que la famille $(\mathcal{O}_i)_{i\in I}$ *recouvre* X (ou est un recouvrement de X) si tout point de X appartient à au moins un des $\mathcal{O}_i$:

$$X = \bigcup_{i\in I} \mathcal{O}_i.$$

Voici une caractérisation des espaces compacts :

Théorème 1.2 *Les assertions suivantes sont équivalentes :*

1. *l'espace métrique (X, d) est compact,*
2. *si $(\mathcal{O}_i)_{i\in I}$ est un recouvrement de X par des ouverts $\mathcal{O}_i$, alors il existe $N \in \mathbb{N}^*$ et $i_1, \dots, i_N \in I$ tels que $(\mathcal{O}_{i_k})_{k\in\{1,\dots,N\}}$ est un recouvrement ouvert fini de X,*
3. *X est complet et, pour tout $\epsilon > 0$, on peut recouvrir X par un nombre fini de boules de rayon ϵ.*

Attention, le résultat est très faux si les ensembles $\mathcal{O}_i$ ne sont pas ouverts. Par exemple, si X est un ensemble infini, il est clair que le résultat ne peut fonctionner avec $I := X$ et $\mathcal{O}_i := \{i\}$ où $i \in X$. Par contre, pour la condition (3), le fait que les boules soient ouvertes ou fermées est indifférent (parce qu'on peut changer le rayon des boules).

Preuve. Commençons par la partie difficile : (1) entraîne (2). Supposons X compact et considérons $(\mathcal{O}_i)_{i\in I}$ un recouvrement ouvert de X. Montrons d'abord qu'il existe $\epsilon > 0$ tel que, pour tout $x \in X$, la boule $B(x, \epsilon)$ est incluse toute entière dans au moins un des $\mathcal{O}_i$. En effet, sinon, il existe pour tout $n \in \mathbb{N}^*$ un point $x_n \in X$ tel que $B(x_n, 1/n)$ n'est contenue dans aucun des $\mathcal{O}_i$. Comme K est compact, il existe une suite extraite (x_{n_k}) qui converge vers un point $\bar{x} \in X$. Fixons maintenant $i \in I$. Comme $B(x_{n_k}, 1/n) \not\subset \mathcal{O}_i$, il existe $y_{n_k} \in B(x_{n_k}, 1/n)\backslash\mathcal{O}_i$. Comme (x_{n_k}) converge vers $\bar{x}$, la suite (y_{n_k}) converge également vers $\bar{x}$. Comme $\mathcal{O}_i$ est ouvert et les (y_{n_k}) ne sont pas dans $\mathcal{O}_i$, $\bar{x}$ n'appartient pas non plus à $\mathcal{O}_i$. Mais alors $\bar{x}$ n'appartient à aucun $\mathcal{O}_i$, ce qui contredit le fait que les $(\mathcal{O}_i)_{i\in I}$ forment un recouvrement de X. On en déduit qu'il existe $\epsilon > 0$ tel que, pour tout $x \in X$, la boule $B(x, \epsilon)$ est incluse toute entière dans au moins un des $\mathcal{O}_i$.

Montrons maintenant qu'on peut recouvrir X par un nombre fini de boules de rayon ϵ. En effet, sinon, on peut construire par récurrence une suite de point (x_n) tels que $x_{n+1} \notin \bigcup_{p=1}^{n} B(x_p, \epsilon)$ pour tout n. Mais alors $d(x_n, x_m) \geq \epsilon$ pour tout $n < m$, ce qui est impossible puisque, par compacité, la suite (x_n) possède une sous-suite qui converge. Donc il existe $N \in \mathbb{N}^*$ et $x_1, \dots, x_N \in X$ tels que la famille $(B(x_n, \epsilon))_{n=1,\dots,N}$ recouvre X. Comme, pour tout n, il existe $i_n \in I$ tel que $B(x_n, \epsilon) \subset \mathcal{O}_{i_n}$, la famille finie $(\mathcal{O}_{i_n})_{n=1,\dots,N}$ recouvre également X. Cela montre que (1) entraîne (2).

Montrons maintenant que (2) entraîne (3). Il suffit de prouver que, sous la condition (2), X est complet, puisque la propriété de recouvrement de (3) est une conséquence directe de l'hypothèse (2). Considérons (x_n) une suite de Cauchy de X. Soit $x \in X$. Si (x_n) ne converge vers x, il existe $\epsilon_x > 0$ et, pour tout $n \in \mathbb{N}$, un entier $p \geq n$ tel que $d(x, x_p) \geq 2\epsilon_x$. Comme la suite (x_n) est de Cauchy, cela entraîne qu'il existe n_x tel que $d(x, x_n) \geq \epsilon_x$ pour tout $n \geq n_x$. Si (x_n) ne convergeait vers aucun $x \in X$, alors on pourrait recouvrir X des boules ouvertes $B(x, \epsilon_x)$ pour $x \in X$. Par hypothèse (2), il existerait alors $y_1, \dots, y_N$ telle que la famille finie $(B(y_i, \epsilon_{y_i})_{i=1,\dots,N}$ recouvre X. Mais alors, pour tout $n \geq \max\{n_{y_1}, \dots, n_{y_N}\}$ et pour tout $i = 1, \dots, N$, on aurait $d(y_i, x_n) \geq \epsilon_{y_i}$, ce qui contredirait le fait que les boules $(B(y_i, \epsilon_{y_i})_{i=1,\dots,N}$ recouvrent X. Donc (2) entraîne (3).

Reste à prouver que (3) entraîne (1). On suppose la propriété (3) et on considère (x_n) une suite de X. Par propriété (3), pour tout $p \in \mathbb{N}^*$ il existe $m_p \in \mathbb{N}^*$ et $(y_{p,k})_{k=1,\dots,m_p}$ tels que la famille finie $(B(y_{p,k}, 1/p))_{k=1,\dots,m_p}$ recouvre X. Donc il existe $k_1 \in \{1, \dots, m_1\}$ tel que $B(y_{1,k_1}, 1)$ contient une infinité de points de la suite (x_n). Par récurrence on construit alors facilement une suite (k_p), avec $k_p \in \{1, \dots, m_p\}$, telle que, pour tout p la boule $B(y_{p+1,k_{p+1}}, 1/(p+1))$ contient une infinité de termes de la famille $\{x_n\} \cap \bigcap_{r=1}^{p} B(y_{r,k_r}, 1/r)$. Toujours par récurrence, on peut alors construire une suite croissante (n_p) et, pour tout $p \in \mathbb{N}^*$, un élément x_{n_p} dans $\bigcap_{r=1}^{p} B(y_{r,k_r}, 1/r)$. Il est clair que la suite (x_{n_p}) est de Cauchy, puisque si $p < q$, alors x_p et x_q sont tout deux dans $B(y_{p,k_p}, 1/p)$ et donc $d(x_p, x_q) \leq 1/p$. Comme X est complet, cela implique que la suite (x_{n_p}) converge dans X : X est donc compact. □

On rappelle qu'on dit qu'un espace métrique (X, d) est séparable s'il existe (au moins) une suite dense dans X.

Proposition 1.3 *Tout espace métrique compact est séparable.*

Preuve. Soit (X, d) un espace métrique compact. Pour tout $n \in \mathbb{N}^*$, on considère le recouvrement de X constitué des boules ouvertes $B(x, 1/n)$ pour $x \in X$. Comme X est compact, on peut en extraire un recouvrement fini : il existe $(x_{n,k})_{k \leq k_n}$ tel que la famille $(B(x_{n,k}, 1/n))_{k=1,\dots,k_n}$ recouvre X. Il est clair que la famille $(x_{n,k})_{n\geq 1,\, k\leq k_n}$ est dense et dénombrable. Si on la ré-indice, la suite ainsi obtenue sera donc également dense. □

On dit qu'un sous-ensemble d'un espace métrique (X, d) est *relativement compact* si sa fermeture dans X est compacte. Par exemple, un sous-ensemble d'un espace euclidien est relativement compact, si et seulement s'il est borné. *Lorsque X est complet,* les assertions suivantes sont équivalentes (la preuve, facile, est laissée en exercice) :

(i) le sous-ensemble E de X est relativement compact,

(ii) de toute suite (x_n) de E on peut extraire une sous-suite (x_{n_k}) qui converge dans X,

(iii) pour tout $\epsilon > 0$ on peut recouvrir E par un nombre fini de boules de rayon ϵ (on dit que E est pré-compact).

Notons que si X est compact, alors tout sous-ensemble de X est pré-compact. En particulier, en dimension finie, les sous-ensembles pré-compacts sont les sous-ensembles bornés (car les compacts sont les fermés bornés). Cette caractérisation est spécifique à la dimension finie :

Théorème 1.4 (Riesz) *Soit E un espace vectoriel normé. Alors E est de dimension finie, si et seulement si, la boule unité fermée de E est compacte.*

Preuve. Il est bien connu que, si X est de dimension finie, tout fermé borné est compact, et donc la boule unité fermée de X est compacte. Montrons la réciproque. Si E est de dimension infinie, alors il existe une famille strictement croissante de sous-espaces vectoriels fermés (E_n) de E. Montrons qu'alors il existe une suite (x_n) de points telle que, pour tout n, $x_n \in E_n$, $\|x_n\| = 1$ et $dist(x_{n+1}, E_n) \geq 1/2$. En effet, soit $v \in E_{n+1} \backslash E_n \neq \emptyset$ et $w \in E_n$ tel que $\|v - w\| \leq 2dist(v, E_n)$ (comme E_n est fermé, on a $d(v, E_n) > 0$). Alors le vecteur $x_{n+1} := (v - w)/\|v - w\|$ vérifie : $x_{n+1} \in E_{n+1}$, $\|x_{n+1}\| = 1$ et, pour tout $y \in E_n$,

$$\|x_{n+1} - y\| = \frac{\|v - w - \|v - w\|y\|}{\|v - w\|} \geq \frac{dist(v, E_n)}{2dist(v, E_n)} = \frac{1}{2}$$

où l'inégalité vient du fait que $w + \|v - w\|y \in E_n$ tandis que $\|v - w\| \leq 2dist(v, E_n)$. Donc $dist(x_{n+1}, E_n) \geq 1/2$, ce qui prouve l'existence des (x_n). Par construction, on a pour tout $n < m$, $x_n \in E_{m-1}$ et donc $\|x_m - x_n\| \geq dist(x_m, E_{m-1}) \geq 1/2$. La suite (x_n) ne possède donc aucune sous-suite convergente, ce qui contredit la compacité de la boule unité. □

1.2 Théorème d'Ascoli

En analyse fonctionnelle, une question essentielle est de trouver des sous-ensembles compacts d'espaces de fonctions. Le théorème d'Ascoli s'intéresse aux sous-ensembles (relativement) compacts d'espaces de fonctions continues. Dans cette partie, on fixe (X_1, d_1) et (X_2, d_2) deux espaces métriques et on note par $C^0(X_1, X_2)$ l'ensemble des applications continues de X_1 dans X_2. Rappelons que $C^0(X_1, X_2)$ est muni de la distance

$$d_\infty(f_1, f_2) := \sup_{x \in X_1} d_2(f_1(x), f_2(x)) \qquad \forall f_1, f_2 \in C^0(X_1, X_2).$$

Définition 1.5 *Soient (X, d_1) et (X_2, d_2) deux espaces métriques et $\mathcal{F}$ un sous-ensemble de $C^0(X_1, X_2)$. On dit que $\mathcal{F}$ est équi-continu si*

$$\forall \epsilon > 0,\ \exists \eta > 0,\ \forall f \in \mathcal{F},\ \forall x, y \in X_1 \text{ avec } d_1(x, y) \leq \eta,\ \text{on a } d_2(f(x), f(y)) \leq \epsilon.$$

Autrement dit, la famille $\mathcal{F}$ possède un module de continuité uniforme : il existe une fonction croissante $\omega : [0, +\infty[\to [0, +\infty[$, qui tend vers 0 en 0, telle que

$$(*) \qquad d_2(f(x), f(y)) \leq \omega(d_1(x, y)) \qquad \forall x, y \in X_1,\ \forall f \in \mathcal{F},$$

Il suffit de prendre

$$(**) \qquad \omega(r) := \sup \{d_2(f(x), f(y)) \mid f \in \mathcal{F},\ x, y \in X_1,\ d_1(x, y) \leq r\}.$$

Remarque : le module de continuité uniforme ω défini par $(**)$ n'est pas continu en général. Pour des raisons techniques, il est utile dele remplacer par un module de continuité soit continu. Il suffit pour cela de remplacer ω défini par $(**)$ par $\tilde{\omega}$ défini par

$$\tilde{\omega}(r) = \sup_{\rho \in]0,2[} \left\{ \omega(\rho) - \frac{|r - \rho|}{r} \right\} \qquad \forall r \in]0, 1[.$$

La fonction $\tilde{\omega}$ ainsi obtenue est croissante, localement lipschitzienne et tend vers 0 en 0 (exercice).

Théorème 1.6 (Théorème d'Ascoli) *On suppose que X_1 est compact et X_2 complet. Soit $\mathcal{F}$ un sous-ensemble de $C^0(X_1, X_2)$. La famille $\mathcal{F}$ est relativement compacte dans $C^0(X_1, X_2)$, si et seulement si, les deux conditions suivantes sont satisfaites :*

(i) pour tout $x \in X_1$, l'ensemble $\{f(x) \mid f \in \mathcal{F}\}$ est relativement compact dans X_2,

(ii) la famille $\mathcal{F}$ est équi-continue.

En particulier, si (f_n) est une suite de fonction continues d'un compact X_1 vers un ensemble complet X_2, qui est équi-continue et qui converge simplement vers une fonction f, alors la convergence de (f_n) vers f est uniforme (et, en particulier, f est continue).

La preuve du théorème d'Ascoli repose sur le procédé diagonal de Cantor :

Proposition 1.7 (Procédé diagonal) *Pour tout $n \in \mathbb{N}$, soit (X^n, d^n) un espace métrique et $(x^n_k)_{k\in\mathbb{N}}$ une suite relativement compacte dans X^n. Alors il existe une suite extraite $(k_l)_{l\in\mathbb{N}}$ et, pour tout $n \in \mathbb{N}$, un élément $x^n \in X^n$ tels que, pour tout $n \in \mathbb{N}$, la suite $(x^n_{k_l})_{l\in\mathbb{N}}$ converge vers x^n lorsque $l \to +\infty$.*

Par relative compacité, on sait que, pour tout n, il existe une suite extraite $(x^n_{k_l})_{l\in\mathbb{N}}$ qui converge, mais cette suite extraite $(k_l)_{l\in\mathbb{N}}$ dépend *a priori* de n. Le point central du procédé diagonal est de construire une suite extraite indépendante de n.

Preuve de la proposition 1.7. Comme, pour $n = 0$, la suite (x^0_k) est relativement compacte dans X^0, il existe une suite extraite $(k^{(0)}_l)_{l\in\mathbb{N}}$ telle que $(x^0_{k^{(0)}_l})_{l\in\mathbb{N}}$ converge vers un certain $x^0 \in X^0$. La suite $(x^1_{k^{(0)}_l})_{l\in\mathbb{N}}$ est relativement compacte dans X^1 et il existe une suite-extraite de $(k^{(0)}_l)$, notée $(k^{(1)}_l)$, telle que $(x^1_{k^{(1)}_l})_{l\in\mathbb{N}}$ converge vers un certain $x^1 \in X^1$. On construit ainsi par récurrence une famille de suites $(k^{(n)}_l)$ telles que $(k^{(n)}_l)$ est une suite extraite de $(k^{(n-1)}_l)$ et la suite $(x^n_{k^{(n)}_l})$ converge vers un certain $x^n \in X^n$ lorsque $l \to +\infty$. Considérons maintenant la suite extraite $(k^{(l)}_l)$. Par définition, pour tout $n \in \mathbb{N}$ fixé et pour tout $l \geq n$, $k^{(l)}_l$ appartient à l'ensemble $\{k^{(n)}_{l'},\ l' \in \mathbb{N}\}$. Donc $(k^{(l)}_l)_{l\geq n}$ est une suite extraite de $(k^{(n)}_l)$, ce qui implique que $(x^n_{k^{(l)}_l})$ converge vers x^n lorsque $l \to +\infty$. On obtient donc le résultat en posant $k_l := k^{(l)}_l$. $\square$

Au cours de la démonstration du théorème d'Ascoli, nous aurons également besoin de la remarque suivante, qui a un intérêt intrinsèque.

Proposition 1.8 *Soient (X_1, d_1) et (X_2, d_2) deux espaces métriques, avec X_2 complet, et S une partie dense de X_1. Soit $\tilde{f} : S \to X_2$ ayant un module de continuité ω :*

$$d_2(\tilde{f}(x), \tilde{f}(y)) \leq \omega(d_1(x, y)) \qquad \forall x, y \in S,$$

où la fonction $\omega :]0, +\infty[\to]0, +\infty[$ est une fonction continue, croissante, qui tend vers 0 en 0. Alors il existe une unique fonction continue $f : X_1 \to X_2$ qui coïncide avec $\tilde{f}$ sur S. De plus f a pour module de continuité ω.

Preuve de la proposition 1.8. Soit x un point de X_1 et (x_n) une suite de S qui converge vers x. La $(\tilde{f}(x_n))$ est de Cauchy puisque

$$d_2(\tilde{f}(x_n), \tilde{f}(x_m)) \leq \omega(d_1(x_n, x_m))$$

avec (x_n) de Cauchy (car convergente) et $\omega(s) \to 0$ lorsque $s \to 0$. Comme X_2 est complet, la suite $(\tilde{f}(x_n))$ possède une limite $f(x) \in X_2$. Montrons que cette limite est indépendante de la suite (x_n) choisie. Soit (x'_n) est une autre suite qui converge vers x, avec $(\tilde{f}(x'_n))$ qui converge vers y. Définissons la suite (x''_n) par $x''_n = x_{n/2}$ si n est pair et $x''_n = x'_{(n-1)/2}$ si n est impair. Alors (x''_n)

converge également vers x. Donc $(\tilde{f}(x''_n))$ possède une limite. Mais $f(x)$ et y sont deux valeurs d'adhérences de la suite $(\tilde{f}(x''_n))$, ce qui prouve que $y = f(x)$. Par conséquent $f(x)$ est défini sans ambiguïté. Notons que $f(x) = \tilde{f}(x)$ si $x \in S$ (prendre la suite constante $(x_n = x)$). Reste à prouver que f est continue : soient $x, x' \in X_1$ et (x_n), (x'_n) deux suites de S convergeant vers x et x' respectivement. Alors

$$d_2(\tilde{f}(x_n), \tilde{f}(x'_n)) \leq \omega(d_1(x_n, x'_n)).$$

Lorsque $n \to +\infty$, on obtient

$$d_2(\tilde{f}(x), \tilde{f}(x')) \leq \omega(d_1(x, x')),$$

ce qui prouve que ω est un module de continuité de f. L'unicité de f découle directement de la construction. $\square$

Preuve du théorème 1.6. Supposons que $\mathcal{F}$ possède les propriétés (i) et (ii) et montrons que $\mathcal{F}$ est relativement compact. La réciproque (qui ne sera pas utilisée par la suite), est laissée à titre d'exercice. On note ω un module de continuité continu de la famille $\mathcal{F}$ (voir la remarque après la définition 1.5).

Comme X_1 est compact, il existe une famille dénombrable dense $S := (x_n)_{n \in \mathbb{N}}$ dans X_1. Par argument diagonal, il existe une suite $(i_k)_{k \in \mathbb{N}}$ telle que, pour tout $n \in \mathbb{N}$, la suite $(f_{i_k}(x_n))$ converge, lorsque $k \to +\infty$, vers un élément de X_2 noté $\tilde{f}(x_n)$. Notons que $\tilde{f}$ possède le même module de continuité ω que la famille $\mathcal{F}$. La proposition 1.8 affirme alors qu'il existe une unique fonction $f \in C^0(X_1, X_2)$ telle que $f(x) = \tilde{f}(x)$ si $x \in S$. De plus, f a pour module de continuité ω.

Reste à montrer que (f_{i_k}) converge uniformément vers f. Soit $\epsilon > 0$ fixé et $\delta > 0$ tel que $\omega(\delta) < \epsilon/3$. Comme X_1 est compact, il existe un nombre fini de point de X_1, $y_1, \ldots, y_L$, tel que $(B(y_l, \delta/2))_{l=1,\ldots,L}$ recouvre X_1. Par densité de S, pour tout $l = 1, \ldots, N$, il existe $x_{n_l} \in S$ tel que $d_1(y_l, x_{n_l}) < \delta/2$. Alors la famille $(B(x_{n_l}, \delta))_{l=1,\ldots,L}$ recouvre X_1. Comme, pour tout $l = 1, \ldots, L$, la suite $(f_{i_k}(x_{n_l}))$ tend vers $f(x_{n_l})$ lorsque $k \to +\infty$, il existe un rang k_0 tel que, pour $k \geq k_0$ et pour tout $l = 1, \ldots, L$, $d_2(f_{i_k}(x_{n_l}), f(x_{n_l})) \leq \epsilon/3$. Alors, pour tout $k \geq k_0$ et pour tout $x \in X_1$, si $x \in B(x_{n_l}, \delta)$, on a

$$\begin{aligned} d_2(f_{i_k}(x), f(x)) &\leq d_2(f_{i_k}(x), f_{i_k}(x_{n_l})) + d_2(f_{i_k}(x_{n_l}), f(x_{n_l})) + d_2(f(x_{n_l}), f(x)) \\ &\leq \omega(d_1(x, x_{n_l})) + \epsilon/3 + \omega(d_1(x, x_{n_l})) \ \leq \ 2\omega(\delta) + \epsilon/3 \ = \ \epsilon, \end{aligned}$$

où on a utilisé le fait que f_{i_k} et f ont pour module de continuité ω dans la seconde inégalité. Cela prouve que

$$\sup_{x \in X_1} d_2(f_{i_k}(x), f(x)) \leq \epsilon \qquad \forall k \geq k_0,$$

i.e., (f_{i_k}) tend uniformément vers f. $\square$

Un cas particulier important est lorsque X_2 est un espace euclidien (disons $X_2 = \mathbb{R}^d$). Par compacité des boules fermées dans un espace euclidien, la première condition demande juste que, pour tout $x \in X_1$,

$$\sup\{\|f(x)\| \mid f \in \mathcal{F}\} \ < \ +\infty.$$

1.3 Une application : le théorème de Cauchy-Péano

Nous allons appliquer le théorème d'Ascoli à l'existence de solutions pour des équations différentielles ordinaires (EDO). Soit $f : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ une application. Etant donné une condition initiale $(t_0, x_0) \in \mathbb{R} \times \mathbb{R}$, on cherche à résoudre l'EDO

$$\begin{cases} x'(t) = f(t, x(t)), & t \in \mathbb{R} \\ x(t_0) = x_0 \end{cases}$$

Rappelons que le théorème de Cauchy-Lipschitz affirme que, si f est C^1 (ou même seulement localement Lipschitzienne en x uniformément en t), alors il existe une solution à cette EDO définie sur un petit intervalle autour de t_0. La solution est construite par le procédé itératif de Banach-Picard. Le résultat n'est plus vrai si f est seulement continue : non seulement le procédé de Banach-Picard ne fonctionne plus, mais il y a des contre-exemples à l'unicité (penser au cas $f(t,x) = (\max\{0,x\})^{\frac{1}{2}}$). Le théorème de Cauchy-Péano affirme qu'il y a cependant existence de solution locale.

Théorème 1.9 *Si f est continue, il existe $\epsilon > 0$ et une solution $x :]t_0 - \epsilon, t_0 + \epsilon[\to \mathbb{R}$ de classe $\mathcal{C}^1$ de l'EDO sur l'intervalle $]t_0 - \epsilon, t_0 + \epsilon[$.*

Comme dans le cadre de Cauchy-Lipschitz, on peut prolonger cette solution locale sur un intervalle maximal, etc...

Preuve. Pour simplifier les notations, on suppose sans perte de généralité que $(t_0, x_0) = (0,0)$ et on construit une solution sur un intervalle de la forme $[0,\epsilon[$, la construction sur $]-\epsilon, 0]$ se faisant de façon symétrique. Soit $M = \max\{|f(t,x)| \mid t \in [0,1],\ |x| \le 1\}$. On pose $\epsilon := 1/M$. Pour tout entier $n \ge 1$, on définit $x_n(\cdot) : [0,\epsilon] \to \mathbb{R}$ de la façon suivante : on définit d'abord par récurrence $x_n(\cdot)$ aux noeuds $t_k := \epsilon k/n$ (pour $k \le n$) :

$$x_n(0) = 0,\ x_n(t_{k+1}) := x_n(t_k) + (\epsilon/n) f(t_k, x_n(t_k)).$$

Puis on interpole les valeurs de $x_n(\cdot)$ linéairement (par exemple) :

$$x_n(t) = x_n(t_k) + \frac{t - t_k}{t_{k+1} - t_k}(x_n(t_{k+1}) - x_n(t_k)) \qquad \text{si } t \in [t_k, t_{k+1}].$$

La fonction $x_n(\cdot)$ ainsi obtenue est continue, C^1 par morceaux. Montrons que $\|x_n(\cdot)\|_\infty \le 1$ et que $(x_n(\cdot))$ est uniformément lipschitzienne. Pour montrer que $\|x_n(\cdot)\|_\infty \le 1$, raisonnons par l'absurde en supposant qu'il existe t tel que $|x_n(t)| > 1$. Soit $\tau := \inf\{t \in [0,\epsilon] \mid |x_n(t)| > 1\}$. Par continuité, on a $|x_n(\tau)| = 1$. Pour tout k tel que $t_k \le \tau$, on a

$$|x_n(t)| \le |x_n(t_k)| + (\epsilon/n)M, \qquad \forall t \in [t_k, t_{k+1}]$$

puisque $|f| \le M$ sur $[0,1] \times [-1,1]$. Donc, par récurrence, on a pour tout $k \le n-1$ tel que $t_k \le \tau$,

$$|x_n(t)| \le (k+1)\epsilon M/n = (k+1)/n, \qquad \forall t \in [t_k, t_{k+1}].$$

En particulier, si $\tau \in [t_k, t_{k+1}]$,

$$\sup_{t \in [0,\tau]} |x_n(t)| \le \tau \epsilon M < 1,$$

ce qui contredit le fait que $|x_n(\tau)| = 1$. D'où $\|x_n(\cdot)\|_\infty \le 1$.

Montrer que la famille $(x_n(\cdot))$ est uniformément lipschitzienne est maintenant facile : en effet, sur chaque intervalle $[t_k, t_{k+1}]$, $x_n(\cdot)$ est lipschitzienne de constante

$$\frac{1}{t_{k+1} - t_k}|x_n(t_{k+1}) - x_n(t_k)| = |f(t_k, x_n(t_k))| \le M$$

puisque $|f| \le M$ sur $[0,1] \times [-1,1]$ et que $\|x_n(\cdot)\|_\infty \le 1$. Donc $x_n(\cdot)$ est globalement lipschitzienne de constante M.

La famille $(x_n(\cdot))$ étant équi-continue et bornée, on en déduit qu'elle admet une sous-suite $(x_{n_p}(\cdot))_{p \in \mathbb{N}}$ qui converge dans $C^0([0,\epsilon], \mathbb{R})$ vers une fonction continue $x(\cdot) : [0,\epsilon] \to \mathbb{R}$. Reste à montrer que $x(\cdot)$ est solution. Pour cela, posons

$$g_n(t) := \frac{1}{t_{k+1} - t_k}(x_n(t_{k+1}) - x_n(t_k)) = f(t_k, x_n(t_k)) \qquad \forall t \in [t_k, t_{k+1}]$$

et remarquons que

$$x_n(t) = \int_0^t g_n(s)ds \qquad \forall t \in [0, \epsilon]. \tag{1}$$

Par (uniforme) continuité de f sur $[0,1] \times [-1,1]$ et convergence de $(x_{n_p}(\cdot))$, (g_{n_p}) converge uniformément vers la fonction $t \to f(t, x(t))$. Donc, en passant à la limite dans (1), on obtient

$$x(t) = \int_0^t f(s, x(s))ds \qquad \forall t \in [0, \epsilon].$$

Cela montre que $x(\cdot)$ est C^1 et, en dérivant, que c'est une solution de l'EDO sur $[0, \epsilon]$. $\square$

1.4 Notes

Il existe de multiples critères de compacité dans d'autres espaces fonctionnels. Par exemple, si $p \in [1, +\infty[$, le théorème de Riesz-Frechet affirme qu'un sous-ensemble borné $\mathcal{F}$ de $L^p(\mathbb{R}^N)$ est relativement compact, si et seulement si,

$$\lim_{h \to 0} \|\tau_h f - f\|_{L^p} = 0$$

uniformément en $f \in \mathcal{F}$, où, pour tout $h \in \mathbb{R}^N$, $\tau_h : L^p(\mathbb{R}^N) \to L^p(\mathbb{R}^N)$ est l'application qui à $f \in L^p(\mathbb{R}^N)$ associe la fonction $\tau_h f(x) = f(x+h)$.

2 Convergence faible

En dimension infinie, il est délicat de trouver des ensembles compacts. Nous avons vu par exemple que, dans l'espace des fonctions continues, la compacité requiert des conditions additionnelles (l'équi-continuité). En pratique, une telle propriété n'est que rarement vérifiée. Pour surmonter cette difficulté, une possibilité est d'affaiblir la topologie, i.e. de passer d'une convergence forte (ou convergence en norme) à une convergence faible que nous définissons plus bas. Nous verrons plus loin que le prix à payer est alors de trouver des fonctions (semi-)continues pour la convergence faible.

2.1 Conséquence de la complétude

Avant d'entamer l'analyse de la convergence faible, il est utile de rappeler certaines propriétés importantes des (familles) d'opérateurs linéaires dans des espaces de Banach (cf. chapitre II du Brézis).

Dans toute cette partie, E et F sont deux espaces de Banach. On note $\mathcal{L}(E,F)$ l'ensemble des opérateurs linéaires continus de E dans F. Lorsque $F=E$, on pose $\mathcal{L}(E):=\mathcal{L}(E,E)$. Les résultats suivants sont des conséquences importantes de la complétude. Pour ne pas alourdir le texte, nous les énonçons sans démonstration.

Théorème 2.1 (Banach-Steinhaus) *Si $(T_i)_{i\in I}$ est une famille quelconque de $\mathcal{L}(E,F)$ telle que*

$$\sup_{i\in I}\|T_i(x)\|_F<+\infty \qquad \forall x\in E,$$

alors la famille (T_i) est équi-continue :

$$\sup_{i\in I}\|T_i\|_{\mathcal{L}(E,F)}<\infty.$$

Autrement dit, dans ce cas, il existe une constante $c>0$ telle que

$$\|T_i(x)\|_F\leq c\|x\|_E \qquad \forall x\in E,\ \forall i\in I.$$

Une conséquence immédiate de ce résultat est qu'une limite simple d'opérateurs linéaires continus est toujours continue.

Théorème 2.2 (De l'application ouverte) *Si $T\in\mathcal{L}(E,F)$ est surjective, alors l'image par T d'un ouvert de E est un ouvert de F. En particulier, si T est bijective, alors T^{-1} est continue.*

Une conséquence de ce résultat est le

Théorème 2.3 (Du graphe fermé) *Une application linéaire $T:E\to F$ est continue, si et seulement si, son graphe est fermé.*

Preuve. Il est évident que, si T est continue, alors son graphe est fermé. Inversement, supposons que le graphe de T soit fermé. Sur E on définit la norme $N(x)=\|x\|_E+\|T(x)\|_F$. Montrons que E est un espace de Banach pour cette norme. Si (x_n) est une suite de Cauchy pour N, alors (x_n) est une suite de Cauchy pour E et $(T(x_n))$ une suite de Cauchy pour F. Par complétude de E et F, (x_n) converge vers $x\in E$ et $(T(x_n))$ converge vers $y\in F$. Le graphe de T étant fermé, on a que $y=T(x)$ et on a donc que

$$N(x-x_n)=\|x-x_n\|_E+\|T(x)-T(x_n)\|_F\to 0.$$

Donc (E,N) est un espace de Banach. L'application identité $I:(E,N)\to(E,\|\cdot\|_E)$ est évidemment continue, bijective. Son inverse $I^{-1}=I$ est donc continue, i.e., il existe $C>0$ tel que $N(x)\leq C\|x\|_E$, d'où $\|T(x)\|\leq C\|x\|_E$. $\square$

2.2 Définition de la convergence faible et propriétés élémentaires

Soit H un espace de Hilbert, muni d'un produit scalaire noté $\langle \cdot, \cdot \rangle_H$ et de norme associée notée $\| \cdot \|_H$.

Définition 2.4 *On dit que la suite (x_n) de H converge faiblement vers $x \in H$ si*

$$\lim_{n \to +\infty} \langle x_n, y \rangle_H = \langle x, y \rangle_H \qquad \forall y \in H.$$

On note $x_n \rightharpoonup x$.

La convergence forte entraîne la convergence faible :

Proposition 2.5 *Si (x_n) converge fortement vers x, alors (x_n) converge faiblement vers x.*

Preuve. En effet, pour tout $y \in H$,

$$|\langle x_n, y \rangle_H - \langle x, y \rangle_H| = |\langle x_n - x, y \rangle_H| \leq \|x_n - x\|_H \|y\|_H \to 0.$$

$\square$

Remarque : La réciproque est fausse en général. Par exemple, il est bien connu que, dans $H := L^2([0, 2\pi])$, la fonction $x_n(t) := \sin(nt)$ vérifie que

$$\lim_{n \to +\infty} \int_0^{2\pi} \sin(nt) y(t) dt = 0 \qquad \forall y \in H.$$

En effet, on vérifie d'abord que c'est vrai pour les fonctions y de classe C^1 (faire une intégration par parties), puis par densité, pour toutes les fonctions $y \in H$. Cela signifie que (x_n) converge faiblement vers la fonction nulle dans H. Mais (x_n) ne tend pas fortement vers la fonction nulle puisque

$$\|x_n\|_H^2 = \int_0^{2\pi} \sin^2(nt) = \frac{1}{n} \int_0^{2n\pi} \sin^2(s) ds = \int_0^{2\pi} \sin^2(s) ds,$$

où le dernier terme est une constante strictement positive indépendante de n.

Proposition 2.6 (Unicité de la limite faible) *Si (x_n) converge faiblement, alors sa limite faible est unique.*

Preuve. Soient x et x' deux limites faibles de (x_n). Alors par définition de la convergence faible, on a

$$\lim_{n \to +\infty} \langle x_n, y \rangle_H = \langle x, y \rangle_H = \langle x', y \rangle_H \qquad \forall y \in H\ ,$$

soit $\langle x - x', y \rangle_H = 0$ pour tout $y \in H$. Donc $x = x'$. $\square$

En dimension finie, convergences faible et forte coïncident :

Proposition 2.7 *Si H est de dimension finie et (x_n) est une suite d'éléments de H, alors (x_n) converge faiblement, si et seulement si, (x_n) converge fortement.*

Preuve. Soit $(e_k)_{k=1,\ldots,N}$ une base orthonormée de H. Si (x_n) converge faiblement vers $x \in H$, alors pour tout $k = 1, \ldots, N$, la suite scalaire $(\langle x_n, e_k \rangle)_{n \in \mathbb{N}}$ converge vers $\langle x, e_k \rangle$. Donc

$$\|x_n - x\|^2 = \sum_{k=1}^{N} (\langle x_n, e_k \rangle - \langle x, e_k \rangle)^2 \to 0.$$

Cela prouve que (x_n) converge fortement vers x.

La réciproque est toujours vraie, même en dimension infinie. $\square$

2.3 Compacité faible de la boule unité

L'intérêt majeur de la convergence faible est de fournir des ensembles compacts :

Théorème 2.8 (Banach-Alaoglu-Bourbaki) *Dans un espace de Hilbert, les boules fermées sont séquentiellement faiblement compactes. Autrement dit, si (x_n) est une suite bornée dans H, alors il existe une sous-suite (x_{n_k}) qui converge faiblement vers un élément $x \in H$.*

Nous ne ferons la preuve que dans un cas particulier (sous l'hypothèse supplémentaire que l'espace est séparable). La preuve dans le cas général figure par exemple dans le livre de Brézis, "Analyse fonctionnelle", Masson (Théorème III.27).

Preuve. Puisque H est un espace de Hilbert séparable, H possède une base Hilbertienne $(e_p)_{p\in\mathbb{N}}$. Soit (x_n) une suite bornée par une constante M. Notons qu'alors, pour tout p, la suite $(\langle x_n, e_p\rangle_H)_{n\in\mathbb{N}}$ est une suite réelle bornée. En utilisant le procédé diagonal de Cantor, on peut donc extraire une sous-suite $(x_{n_k})_{k\in\mathbb{N}}$ telle que, pour tout p, la suite $(\langle x_{n_k}, e_p\rangle_H)_{k\in\mathbb{N}}$ converge vers un réel noté α_p. Notons que, pour tout $P \in \mathbb{N}$,

$$\sum_{p=0}^{P} \alpha_p^2 = \lim_{k\to+\infty} \sum_{p=0}^{P} (\langle x_{n_k}, e_p\rangle_H)^2 \leq \limsup_{k\to+\infty} \|x_{n_k}\|_H^2 \leq M^2$$

(on a utilisé l'inégalité de Parseval). Donc (α_p) est dans ℓ^2 et on peut définir $x := \sum_{p=0}^{\infty} \alpha_p e_p$. Notons que

$$\|x\|_H^2 = \sum_{p=0}^{\infty} \alpha_k^2 \leq M^2.$$

Soit $y \in H$. On veut montrer que $(\langle x_{n_k}, y\rangle)$ tend vers $\langle x, y\rangle$. C'est évident si y est un des e_p puisqu'alors $(\langle x_{n_k}, e_p\rangle)$ tend vers $\alpha_p = \langle x, e_p\rangle_H$. Par linéarité, c'est également vrai si y appartient à l'espace vectoriel engendré par les (e_p). Dans le cas général, on utilise la densité de cet espace vectoriel dans H. Par densité, pour tout $\epsilon > 0$ il existe $y_\epsilon \in Vect(\{e_p\}_{p\in\mathbb{N}})$ tel que $\|y - y_\epsilon\|_H \leq \epsilon$. Alors

$$\limsup_k |\langle x_{n_k}, y\rangle_H - \langle x, y\rangle_H| \leq \limsup_k |\langle x_{n_k}, y_\epsilon\rangle_H - \langle x, y_\epsilon\rangle_H| + 2M\|y - y_\epsilon\|_H \leq 2M\epsilon$$

puisque $(\langle x_{n_k}, y_\epsilon\rangle_H)$ tend vers $\langle x, y_\epsilon\rangle_H$. Donc

$$\limsup_k |\langle x_{n_k}, y\rangle - \langle x, y\rangle| = 0,$$

ce qui est le résultat désiré.

Dans le cas général (i.e., lorsque H n'est pas séparable), la technique consiste à remplacer H par la fermeture de l'espace vectoriel engendré par les $\{x_n, n \in \mathbb{N}\}$, qui est un espace de Hilbert séparable. On peut appliquer à cet espace le résultat précédent. □

Notons qu'une forme de réciproque est vraie :

Proposition 2.9 *Si H est un espace de Hilbert, alors toute suite (x_n) qui converge faiblement est bornée.*

Preuve. C'est une conséquence du théorème de Banach-Steinhaus (voir la partie 2.1). En effet, définissons la forme linéaire $T_n : H \to \mathbb{R}$ par $\phi_n(z) = \langle x_n, z\rangle_H$. Alors par définition de la convergence faible, la suite d'opérateurs T_n converge simplement vers l'opérateur T défini par $T(z) = \langle x, z\rangle_H$.

Donc, d'après Banach-Steinhaus, les normes $\|T_n\|_{\mathcal{L}(X,\mathbb{R})}$ des opérateurs T_n sont bornées par une constante M, avec

$$\|T_n\|_{\mathcal{L}(X,\mathbb{R})} = \sup_{y\in H, \|y\|_H\le 1} |T_n(y)| = \sup_{y\in H, \|y\|_H\le 1} |\langle x_n, y\rangle| = \|x_n\|_H \le M \qquad \forall n\ge 0 \ .$$

□

Une application de la proposition précédente est la remarque :

Corollaire 2.10 *On suppose que (x_n) converge faiblement vers x dans H et que (y_n) converge fortement vers y dans H. Alors la suite réelle $(\langle x_n, y_n\rangle)$ converge vers $\langle x, y\rangle$.*

Preuve. Comme (x_n) converge faiblement, (x_n) est bornée par une constante M : $\|x_n\| \le M$ pour tout $n\in\mathbb{N}$. Alors, par inégalité triangulaire et Cauchy-Schwartz,

$$\begin{aligned} |\langle x_n, y_n\rangle - \langle x, y\rangle| &\le |\langle x_n, y_n\rangle - \langle x_n, y\rangle| + |\langle x_n, y\rangle - \langle x, y\rangle| \\ &\le \|x_n\|\|y_n - y\| + |\langle x_n, y\rangle - \langle x, y\rangle| \\ &\le M\|y_n - y\| + |\langle x_n, y\rangle - \langle x, y\rangle| \end{aligned}$$

Le membre droit de la dernière inégalité tend vers 0 car (y_n) converge fortement vers y et, par convergence faible, $(\langle x_n, y_n\rangle)$ converge vers $\langle x, y\rangle$. On en déduit que $(\langle x_n, y_n\rangle)$ converge vers $\langle x, y\rangle$. □

Si (x_n) converge faiblement vers x dans H, alors la suite des normes $(\|x_n\|_H)$ ne converge pas nécessairement vers $\|x\|_H$. On a cependant le résultat de semi-continuité suivant :

Proposition 2.11 *Soit (x_n) une suite d'un espace de Hilbert H qui converge faiblement vers $x\in H$. Alors*

$$\|x\|_H \le \liminf_{n\to+\infty} \|x_n\|_H \ .$$

Autrement dit, la fonction $x \to \|x\|_H$ est séquentiellement semi-continue inférieurement pour la convergence faible.

Preuve. Soit $y\in H$ avec $\|y\|_H \le 1$. Alors, par définition de la convergence faible, on a

$$\langle x, y\rangle_H = \lim_{n\to+\infty} \langle x_n, y\rangle_H \le \liminf_{n\to+\infty} \|x_n\|_H\|y\|_H \le \liminf_{n\to+\infty} \|x_n\|_H$$

Donc

$$\|x\|_H = \sup_{y\in H, \|y\|_H\le 1} \langle x, y\rangle \le \liminf_{n\to+\infty} \|x_n\|_H.$$

□

L'inégalité dans la proposition précédente est stricte, sauf si la convergence est en fait forte :

Proposition 2.12 *Soit (x_n) une suite d'un espace de Hilbert H qui converge faiblement vers $x\in H$. Si*

$$\|x\|_H = \lim_{n\to+\infty} \|x_n\|_H \ ,$$

alors (x_n) converge fortement vers x.

Preuve. En effet,

$$\|x - x_n\|_H^2 = \langle x - x_n, x - x_n\rangle_H = \langle x - x_n, x\rangle_H - \langle x, x_n\rangle_H + \|x_n\|_H^2.$$

Par convergence faible, les suites $(\langle x - x_n, x\rangle_H)$ et $(\langle x, x_n\rangle_H)$ convergent respectivement vers 0 et $\|x\|_H^2$. Par hypothèse, $(\|x_n\|_H^2)$ tend vers $\|x\|_H^2$. Par conséquent, $(\|x - x_n\|_H^2)$ tend vers 0, ce qui prouve que (x_n) converge fortement vers x. □

2.4 Continuité faible des applications linéaires fortement continues

En général, une application fortement continue n'est pas faiblement continue. Par exemple la norme *n'est pas* faiblement continue, comme l'explique la Proposition 2.12. Le cas des applications linéaires continue est tout à fait particulier puisqu'on a :

Proposition 2.13 *Si $T \in \mathcal{L}(H, H)$ et (x_n) converge faiblement vers x dans H, alors $(T(x_n))$ converge faiblement vers $T(x)$ dans H. Autrement dit, pour les applications linéaires de H dans H, il y a équivalence entre continuité forte et continuité faible.*

Pour montrer la proposition, nous aurons besoin de la notion d'adjoint d'un opérateur linéaire continu :

Proposition 2.14 *Soit H un espace de Hilbert et $T \in \mathcal{L}(H)$. Il existe un unique opérateur $T^* \in \mathcal{L}(H)$ tel que*

$$\langle T(x), y\rangle = \langle x, T^*(y)\rangle \qquad \forall x, y \in H.$$

L'opérateur T^* est l'adjoint de T.

Preuve. Comme, pour tout $y \in H$, l'application $x \to \langle T(x), y\rangle$ est linéaire continue, le théorème de représentation de Riesz affirme qu'il existe un unique vecteur $T^*(y) \in H$ tel que $\langle T(x), y\rangle = \langle x, T^*(y)\rangle$ pour tout $x \in H$. Reste à montrer que l'application $y \to T^*(y)$ est linéaire continue. Par unicité et linéarité de T, elle est clairement linéaire. Montrons la continuité. On a

$$\sup_{\|y\|\leq 1} \|T^*(y)\| = \sup_{\|x\|\leq 1,\ \|y\|\leq 1} \langle x, T^*(y)\rangle = \sup_{\|x\|\leq 1,\ \|y\|\leq 1} \langle T(x), y\rangle = \sup_{\|x\|\leq 1} \|T(x)\| = \|T\|_{\mathcal{L}(H)}.$$

Donc T^* est continu (et de même norme que T ; en fait on montre aisément que l'application $T \to T^*$ est une isométrie de $\mathcal{L}(H)$). □

Preuve de la Proposition 2.13. En effet, pour tout $y \in H$, on a

$$\lim_{n\to+\infty} \langle T(x_n), y\rangle = \lim_{n\to+\infty} \langle x_n, T^*(y)\rangle = \langle x, T^*(y)\rangle = \langle T(x), y\rangle,$$

où on a utilisé la convergence faible dans la seconde égalité. □

2.5 Notes

La notion de topologie faible se généralise aux espaces vectoriels normés (et au-delà). Par exemple, si (x_n) est une suite d'un espace vectoriel normé E, on dit que (x_n) converge faiblement vers $x \in E$ si, pour toute forme linéaire continue f sur E (i.e., $f \in E^*$), la suite $(f(x_n))$ tend vers $f(x)$. Lorsque E est un espace de Banach réflexif, la boule unité de E est alors compacte pour la topologie faible.

Si H est un espace de Hilbert, les applications continues sur H sont rarement continues pour la topologie faible : on a vu que c'était le cas pour les applications linéaires. On peut montrer que, si $f : H \to \mathbb{R}$ est convexe et continue, alors f est semi-continue inférieurement pour la topologie faible :

$$\text{si } x_n \rightharpoonup x, \text{ alors } f(x) \leq \liminf_{n\to+\infty} f(x_n).$$

Ce type de résultat est central en calcul des variations.

3 Introduction au calcul des variations

Le calcul des variations s'intéresse aux problèmes de minimisation dans lesquels la variable à optimiser est une fonction. Pour simplifier, nous ne considérerons ici que des problèmes dans lesquels cette variable est une fonction définie sur un intervalle. Le problème de calcul des variations prend alors la forme

$$(\mathcal{P}) \qquad \inf_{x\in\mathcal{C}^1([0,1];\mathbb{R}),\ x(0)=A,\ x(1)=B} \int_0^1 L(t,x(t),x'(t))dt\ .$$

où $L : [0,1]\times\mathbb{R}\times\mathbb{R}\to\mathbb{R}$ est une fonction continue et $A, B \in \mathbb{R}$ sont des constantes données.

On s'intéresse ici à l'existence d'une solution. C'est une question délicate car, contrairement à ce qui se passe en dimension finie, les sous-ensembles bornés de $\mathcal{C}^1$ ne sont pas compacts. Afin de retrouver cette compacité, il faut changer l'espace sur lequel on travaille et utiliser la topologie faible.

Pour alléger les notations, on a choisi de travailler sur l'intervalle $[0,1]$. Cet intervalle peut être remplacé dans tous les énoncés par un intervalle de la forme $[a,b]$ (quitte parfois à changer les constantes multiplicatives).

3.1 L'espace de Sobolev H^1

Avertissement : cette partie est une introduction éclair à l'espace H^1, qui sera traité plus en détail dans le cours "Analyse fonctionnelle et EDP".

Soit $L^2 = L^2([0,1];\mathbb{R})$ l'espace des fonctions $v : [0,1]\to\mathbb{R}$ de carré intégrable : $\int_0^1 |v(t)|^2dt < +\infty$. On rappelle que L^2 est un espace de Hilbert lorsqu'on le munit du produit scalaire

$$\langle v,w\rangle_{L^2} = \int_0^1 v(t)w(t)\ dt \qquad \forall v,w\in L^2\ ,$$

de norme associée $\|v\|_{L^2} = \langle v,v\rangle_{L^2}^{\frac{1}{2}}$. On note L^1 l'espace des fonctions intégrables sur $[0,1]$ et L^∞ l'espace des fonctions essentiellement bornées sur $[0,1]$. Soit enfin $C_c^1(]0,1[)$ l'ensemble des fonctions de classe C^1 à support compact dans $]0,1[$. Nous utiliserons sans arrêt le fait que, si $\phi\in C_c^1(]0,1[)$, alors $\phi(0)=\phi(1)=0$. Rappelons que $C_c^1(]0,1[)$ est dense dans L^1 et dans L^2.

Définition 3.1 (Dérivée faible) *On dit que $x\in L^1$ possède $y\in L^1$ comme dérivée faible si*

$$\int_0^1 x(t)\phi'(t)dt = -\int_0^1 y(t)\phi(t)dt \qquad \forall \phi\in C_c^1(]0,1[). \tag{2}$$

On remarque facilement que, si x est de classe C^1 sur $[0,1]$, alors x admet pour dérivée faible x' : c'est juste l'intégration par parties. Le résultat suivant affirme que la dérivée faible d'une fonction est définie de façon unique :

Lemme 3.2 *Soit $x\in L^1$ ayant une dérivée faible. Alors il existe une unique fonction y de L^1 pour laquelle (2) a lieu.*

Cette fonction sera désormais notée x'.

Remarque : on montre sans difficulté que, si x_1 et x_2 possèdent une dérivée faible, et si $\lambda\in\mathbb{R}$, alors $\lambda x_1 + x_2$ possède également une dérivée faible et $(\lambda x_1+x_2)' = \lambda x_1' + x_2'$ (linéarité de la dérivation).

Preuve du lemme. Soit y_1 un autre élément de L^1 pour lequel (2) a lieu. On a

$$\int_0^1 (y-y_1)(t)\phi(t)dt = -\int_0^1 x(t)\phi'(t)dt + \int_0^1 x(t)\phi'(t)dt = 0 \qquad \forall\phi\in C_c^1(]0,1[).$$

Le lemme suivant affirme alors que $y - y_1 = 0$ p.p. $\square$

Lemme 3.3 *Soit $w \in L^1$ tel que*

$$\int_0^1 w(t)\phi(t)dt = 0 \qquad \forall \phi \in C_c^1(]0,1[).$$

Alors $w = 0$ p.p.

Preuve. Montrons qu'il existe une suite de fonctions (ϕ_n) dans $C_c^1(]0,1[)$, bornées par 1, telle que (ϕ_n) tend vers $\text{sign}(w)$ pour presque tout $x \in [a,b]$, où

$$\text{sign}(w(t)) = \begin{cases} 1 & \text{si } w(t) > 0 \\ -1 & \text{si } w(t) < 0 \\ 0 & \text{si } w(t) = 0 \end{cases}$$

En effet, comme $C_c^1(]0,1[)$ est dense dans $L^1([0,1])$, il existe une suite de fonction $\psi_n \in C_c^1(]0,1[)$ qui converge vers $\text{sign}(w)$ dans $L^1([0,1])$ et donc, à une sous-suite près encore notée (ψ_n), presque partout sur $[0,1]$. Soit $\theta : \mathbb{R} \to [-1,1]$ une fonction $\mathcal{C}^1$, croissante, telle que $\theta(0) = 0$, $\theta(-1) = -1$ et $\theta(1) = 1$. Alors la suite de fonctions $(\phi_n = \theta \circ \psi_n)$ vérifie les conditions demandées.

Par hypothèse, on a $\displaystyle\int_0^1 w(t)\phi_n(t)dt = 0$ pour tout n. Comme $|w(t)\phi_n(t)| \leq |w(t)|$ et $(w(t)\phi_n(t))$ tend p.p. vers $w(t)\text{sign}(w(t)) = |w(t)|$ dans $[0,1]$, on a, par convergence dominée, $\displaystyle\int_0^1 |w(t)|dt = 0$. Donc $w = 0$ p.p. dans $[0,1]$. $\square$

Définition 3.4 (L'espace de Sobolev H^1) *On dit que x appartient à l'espace de Sobolev $H^1([0,1])$ (le plus souvent abrégé en H^1) si x et sa dérivée faible x' appartiennent à L^2.*

Il est facile de voir que H^1 est un sous-espace vectoriel de L^2. On munit l'espace H^1 du produit scalaire

$$\langle x_1, x_2 \rangle_{H^1} := \int_0^1 x_1(t)x_2(t) + x_1'(t)x_2'(t)\ dt \qquad \forall x_1, x_2 \in H^1$$

et, bien sûr, de la norme associée

$$\|x\|_{H^1} = \left(\int_0^1 x^2(t) + (x'(t))^2\ dt \right)^{1/2} \qquad \forall x_1, x_2 \in H^1.$$

On voit facilement que H^1, muni de ce produit scalaire, est un espace de Hilbert (exercice). Notons que l'application qui à x associe x' est linéaire continue de H^1 dans L^2. Enfin, comme les fonctions continues sont dense dans L^2, les fonctions de classe $\mathcal{C}^1$ sont denses dans H^1 (exercice).

Notons maintenant que les éléments de H^1 sont des fonctions continues.

Lemme 3.5 *Si x appartient à H^1, alors x possède un représentant continu $\tilde{x}$, qui vérifie*

$$\tilde{x}(t) - \tilde{x}(s) = \int_s^t x'(\tau)d\tau \qquad \forall s,t \in [0,1]\ .$$

De plus, $\tilde{x}$ est en fait Hölderienne :

$$|\tilde{x}(t) - \tilde{x}(s)| \leq \|x'\|_2 |t-s|^{1/2} \qquad \forall s,t \in [0,1]\ .$$

Par la suite on *travaillera toujours* avec le représentant continu de x, que l'on notera simplement x. Une conséquence du lemme est que les éléments de H^1 sont exactement les primitives de fonctions de L^2.

Preuve : Soit z définie par $z(t) = \int_0^t x'(s)ds$ pour $t \in [0,1]$. Il est facile de voir que z est bien définie et bornée. On affirme qu'il existe une constante $c \in \mathbb{R}$ telle que $x - z = c$ p.p. Pour cela, notons d'abord que x' est une dérivée faible de z. En effet, on a, par Fubini et pour toute fonction $\phi \in C_c^1(]0,1[)$,

$$\begin{aligned}\int_0^1 z(\tau)\phi'(\tau)d\tau &= \int_0^1 \int_0^\tau x'(s)\phi'(\tau)\ dsd\tau = \int_0^1 x'(s)\int_s^1 \phi'(\tau)d\tau ds \\ &= \int_0^1 x'(s)(\phi(1)-\phi(s))ds = -\int_0^1 x'(s)\phi(s)ds\ .\end{aligned}$$

Donc z a pour dérivée faible x' et on a donc $\displaystyle\int_0^1 (z(s)-x(s))\phi'(s)ds = 0$ pour toute fonction $\phi \in C_c^1(]0,1[)$. Le lemme 3.6 ci-dessous affirme alors qu'il existe une constante c telle que $x - z = c$ p.p. Alors $\tilde{x} := z + c$ est un représentant de x.

Montrons finalement que $\tilde{x}$ est continue dans $[0,1]$. Comme $x' \in L^2$, on a, par Cauchy-Schwartz,

$$|\tilde{x}(t) - \tilde{x}(s)| = \left|\int_s^t x'(s)ds\right| \leq \|x'\|_{L^2}|t-s|^{1/2} \qquad \forall s,t \in [0,1].$$

Ceci prouve que $\tilde{x}$ est Hölderienne, et donc continue, sur $[0,1]$. □

Nous avons utilisé dans la preuve la remarque suivante :

Lemme 3.6 *Soit $w \in L^1$ telle que*

$$\int_0^1 w(t)\phi'(t)dt = 0 \qquad \forall \phi \in C_c^1(]0,1[).$$

Alors w est constante : il existe un réel c tel que $w = c$ p.p. dans I.

Remarque : La réciproque de l'assertion ci-dessus est évidente.

Preuve : Soit $\phi \in \mathcal{C}_c^1(]0,1[)$ et $\eta \in \mathcal{C}_c^1(]0,1[)$ avec $\int_0^1 \eta = 1$. Notons qu'il existe $\psi \in \mathcal{C}_c^1(]0,1[)$ tel que $\psi'(t) = \phi(t) - (\int_0^1 \phi(s)ds)\eta(t)$, puisque la fonction $t \to \phi(t) - (\int_0^1 \phi(s)ds)\eta(t)$ est à support compact et d'intégrale nulle. On applique l'hypothèse à ψ pour obtenir

$$\int_0^1 w(t)\left(\phi(t) - (\int_0^1 \phi(s)ds)\eta(t)\right)dt = 0.$$

Comme, par Fubini,

$$\int_0^1 w(t)(\int_0^1 \phi(s)ds)\eta(t)dt = \int_0^1 \phi(t)(\int_0^1 w(s)\eta(s)ds)dt,$$

on obtient après changement de variable :

$$\int_0^1 \phi(t)\left(w(t) - (\int_0^1 w(s)\eta(s)ds)\right)dt = 0.$$

Comme ceci est vrai pour toute fonction $\phi \in C_c^1(]0,1[)$, on déduit du lemme 3.3 que $w(t) - (\int_0^1 w(s)\eta(s)ds) = 0$ p.p. tout t, et donc que w est égal presque partout à la constante $(\int_0^1 w(s)\eta(s)ds)$. □

Dans H^1, la fonction *évalutation*, qui à la fonction x associe sa valeur $x(t)$ en t, est continue :

Lemme 3.7 *Pour tout $t \in [0,1]$ fixé, l'application Φ_t qui à x associe $x(t)$ est continue de H^1 dans $\mathbb{R}$ et il existe une constante $C > 0$ telle que*

$$\|x\|_\infty \leq C\|x\|_{H^1} \qquad \forall x \in H^1.$$

Preuve. Comme Φ_t est linéaire, il suffit de borner $|\Phi_t(x)|$. Or, d'après le Lemme 3.5 puis Cauchy-Schwarz, on a

$$|\Phi_t(x)| = |x(t)| \leq \int_0^1 (|x(t) - x(s)| + |x(s)|)\, ds \leq \|x'\|_{L^2} \int_0^1 |t-s|^{1/2} ds + \|x\|_{L^2} \leq \sqrt{2}\, \|x\|_{H^1} \; .$$

Par conséquent,

$$\|x\|_\infty \leq \sqrt{2}\|x\|_{H^1}.$$

□

La formule d'intégration par partie reste vraie pour les fonctions de H^1 :

Lemme 3.8 (Formule d'intégration par parties) *Si $x, y \in H^1$, alors, pour tout $0 \leq a \leq b \leq 1$,*

$$\int_a^b x'(t)y(t)\; dt = [x(t)y(t)]_a^b - \int_a^b x(t)y'(t)\; dt \; .$$

Preuve. L'égalité ci-dessus est vraie pour x et y de classe $\mathcal{C}^1$. Si $x, y \in H^1$, on peut approcher x et y par deux suites (x_n) et (y_n) de fonctions de classe $\mathcal{C}^1$, par densité de $\mathcal{C}^1$ dans H^1. Par continuité de l'application $z \to z'$ de H^1 dans L^2 et des applications $z \to z(a)$ et $z \to z(b)$ de H^1 dans $\mathbb{R}$, on déduit facilement le résultat. □

Nous étudions maintenant la convergence faible dans le cas particulier de l'espace de Hilbert H^1.

Proposition 3.9 *Soit (x_n) une suite de H^1 telle que la suite $(x_n(0))$ converge vers $a \in \mathbb{R}$ et la suite (x_n') converge faiblement dans L^2 vers $v \in L^2$. Posons*

$$x(t) = a + \int_0^t v(s)ds \qquad \forall t \in [0,1] \; .$$

Alors la suite de fonctions continues (x_n) converge uniformément vers x.

Preuve. Comme la suite (x_n') converge faiblement, la suite $(\|x_n'\|_{L^2})$ est bornée par une certaine constante M. Montrons dans un premier temps la convergence simple de (x_n) vers x. En effet, pour tout $t \in [0,1]$ fixé, on a, puisque $\mathbf{1}_{[0,t]}$ est dans L^2 :

$$x_n(t) - x_n(0) = \int_0^1 x_n'(s)\mathbf{1}_{[0,t]}\; ds \to \int_0^1 x'(s)\mathbf{1}_{[0,t]}\; ds = x(t) - x(0).$$

D'où

$$\lim_n x_n(t) = \lim_n (x_n(t) - x_n(0)) + \lim_n x_n(0) = x(t) - x(0) + x(0) = x(t) \; .$$

Reste à prouver que la convergence est uniforme. D'après le lemme 3.5, on a

$$|x_n(t) - x_n(s)| \leq \|x_n'\|_{L^2}|t-s|^{\frac{1}{2}} \leq M|t-s|^{\frac{1}{2}} \qquad \forall s,t \in [0,1],\; \forall n \in \mathbb{N} \; . \tag{3}$$

Cela montre que la famille (x_n) est équicontinue. Le théorème d'Ascoli implique alors que la suite (x_n), qui converge simplement vers x, converge en fait uniformément vers x. □

3.2 Existence d'un minimum

Dans cette partie, on étudie l'existence d'un minimum pour une fonctionnelle J de la forme

$$J(x) = \int_0^1 L(t, x(t), x'(t))dt,$$

avec des conditions "au bord" sur x.

Fixons d'abord les hypothèses sur la fonction L. On suppose que $L : [0, 1] \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ est continue, que $L = L(t, x, p)$ est convexe en la variable p, dérivable par rapport à cette variable avec $\frac{\partial L}{\partial p}$ continue dans toutes les variables, et satisfait les conditions de croissance : il existe $C > 0$ avec, pour tout $(t, x, p) \in [0, 1] \times \mathbb{R} \times \mathbb{R}$,

$$C^{-1}p^2 - C \leq L(t, x, p) \leq C(p^2 + 1) \qquad \text{et} \qquad \left|\frac{\partial L}{\partial p}(t, x, p)\right| \leq C(|p| + 1).$$

Par exemple, c'est le cas pour

$$L(t, x, p) = \frac{1}{2}|p|^2 + h(t, x)$$

si la fonction $h : [0, 1] \times \mathbb{R} \to \mathbb{R}$ est continue et bornée.

On définit

$$J(x) := \int_0^1 L(t, x(t), x'(t))dt \qquad \forall x \in H^1.$$

Noter que J est bien définie sur H^1 grâce à la première hypothèse de croissance. Ajoutons maintenant les données au bord. Pour $A, B \in \mathbb{R}$, on pose

$$E = \{x \in H^1 \mid x(0) = A,\ x(1) = B\} .$$

Théorème 3.10 *Sous les hypothèses ci-dessus, la fonctionnelle J admet (au moins) un minimum global sur E.*

Preuve. Soit (x_n) une suite de E telle que

$$\lim_{n\to+\infty} J(x_n) = \inf_{x\in E} J(x) .$$

Par définition de l'inf, une telle suite existe toujours. La preuve se fait en 2 étapes (ce procédé s'appelle *la méthode directe* en calcul des variations) :

1. On montre d'abord que la suite (x'_n) est bornée dans L^2. Par conséquent, (x_n) converge faiblement dans H^1 vers un certain $\bar{x} \in E$.

2. Ensuite on vérifie que la fonctionnelle J est semi-continue inférieurement pour la convergence faible de H^1, et on en déduit que $\bar{x}$ est le minimum du problème.

Première étape. Soit $\tilde{x}$ un élément fixé de E (par exemple $\tilde{x}(t) = (1 - t)A + tB$). Par définition de (x_n), pour tout $n \in \mathbb{N}$ assez grand, on a

$$J(x_n) \leq \inf_{x\in E} J(x) + 1 \leq J(\tilde{x}) + 1 .$$

Comme, par hypothèse de croissance,

$$L(t, x, p) \geq C^{-1}p^2 - C,$$

on a

$$C^{-1}\int_0^1 (x_n'(t))^2 - C \leq J(x_n) \leq J(\tilde{x}) + 1\ .$$

Donc la suite (x_n') est bornée dans L^2. Par compacité faible des boules fermées de L^2, il existe une sous-suite, encore notée (x_n'), qui converge faiblement vers un certain $v \in L^2$. La proposition 3.9 affirme que la suite (x_n) converge uniformément vers $\bar{x} \in H^1$, où

$$\bar{x}(t) = A + \int_0^t v(s)ds \qquad \forall t \in [0,1]\ .$$

En particulier, $B = x_{n_k}(1) \to \bar{x}(1)$, d'où $\bar{x}(1) = B$. Ceci montre que $\bar{x} \in E$.

Deuxième étape. Montrons que

$$\liminf_{n\to+\infty} J(x_n) \geq J(\bar{x}).$$

(la même preuve montre en fait que la fonctionnelle J est semi-continue inférieurement pour la convergence faible de H^1). Par convexité de L par rapport à p, on a, pour tout t, x, p, p_0,

$$L(t,x,p) \geq L(t,x,p_0) + \frac{\partial L}{\partial p}(t,x,p_0)(p-p_0).$$

Fixons $p_0 = p_0(t)$ dans L^2. Alors

$$J(x_n) \geq \int_0^1 L(t,x_n(t),p_0(t)) + \frac{\partial L}{\partial p}(t,x_n(t),p_0(t))(x_n'(t)-p_0(t))\ dt.$$

Par continuité de L et convergence de x_n vers $\bar{x}$, la suite $(L(t,x_n(\cdot),p_0(\cdot)))$ converge p.p. vers $L(t,\bar{x}(\cdot),p_0(\cdot))$. Comme, hypothèse de croissance sur L, on a

$$|L(t,x_n(t),p_0(t))| \leq C(|p_0(t)|^2+1)$$

où le membre droit est dans L^1, la convergence de $(L(t,x_n(\cdot),p_0(\cdot)))$ vers $L(t,\bar{x}(\cdot),p_0(\cdot))$ est dans L^1 par convergence dominée. D'autre part, par hypothèse de croissance sur $\frac{\partial L}{\partial p}$, on a

$$\left|\frac{\partial L}{\partial p}(t,x_n(t),p_0(t))\right| \leq C(|p_0(t)|+1),$$

où le terme de droite est dans L^2. Comme x_n converge vers $\bar{x}$, le théorème de convergence dominée affirme aussi que la suite $(\frac{\partial L}{\partial p}(\cdot,x_n(\cdot),p_0(\cdot)))$ tend vers $\frac{\partial L}{\partial p}(\cdot,\bar{x}(\cdot),p_0(\cdot))$ dans L^2. Donc le produit $\frac{\partial L}{\partial p}(\cdot,x_n(\cdot),p_0(\cdot))p_0(\cdot)$ tend vers $\frac{\partial L}{\partial p}(\cdot,\bar{x}(\cdot),p_0(\cdot))p_0(\cdot)$ dans L^1 tandis que, par convergence faible de (x_n') et corollaire 2.10, le produit scalaire $\langle\frac{\partial L}{\partial p}(\cdot,x_n(\cdot),p_0(\cdot)),x_n'(\cdot)\rangle_{L^2}$ tend vers $\langle\frac{\partial L}{\partial p}(\cdot,\bar{x}(\cdot),p_0(\cdot)),\bar{x}'(\cdot)\rangle_{L^2}$. En mettant ensemble les différentes convergences, on obtient

$$\lim_{n\to+\infty} J(x_n) \geq \int_0^1 L(t,\bar{x}(t),p_0(t)) + \frac{\partial L}{\partial p}(t,\bar{x}(t),p_0(t))(\bar{x}'(t)-p_0(t))\ dt.$$

Cette inégalité est vraie pour tout $p_0 \in L^2$. En particulier, pour $p_0 = \bar{x}'$, on obtient

$$\inf_{x\in E} J(x) = \lim_{n\to+\infty} J(x_n) \geq \int_0^1 L(t,\bar{x}(t),\bar{x}'(t))\ dt = J(\bar{x}).$$

Donc $\bar{x}$ est un minimum de J sur E. □

3.3 Conditions nécessaires d'optimalité pour le minimum

Dans cette partie, on considère la fonctionnelle

$$J(x) = \int_0^1 \frac{1}{2}|x'(t)|^2 + h(t,x(t))dt$$

que l'on minimise sur l'ensemble $E = \{x \in H^1 \mid x(0) = A,\ x(1) = B\}$.

On suppose que la fonction $h : [0,1] \times \mathbb{R} \to \mathbb{R}$ est différentiable par rapport à x et à dérivée $\partial h/\partial x$ continue sur $[0,1] \times \mathbb{R}$. On suppose également que h est bornée, de sorte que, par théorème 3.10, J admet un minimum (pour des conditions au bord données).

Proposition 3.11 *La fonction J est différentiable dans H^1 en tout point $x \in H^1$ et a pour dérivée*

$$dJ(x)(v) = \int_0^1 x'(t)v'(t) + \frac{\partial h}{\partial x}(t,x(t))v(t)\ dt \qquad \forall v \in H^1 .$$

Preuve. Posons $J = J_1 + J_2$ où

$$J_1(x) = \int_0^1 \frac{1}{2}|x'(t)|^2 dt \qquad \text{et} \qquad J_2(x) = \int_0^1 h(t,x(t))dt .$$

Alors J_1 est différentiable car

$$J_1(z) = \int_0^1 \frac{1}{2}|x'(t) + z'(t) - x'(t)|^2 dt = J_1(x) + \int_0^1 x'(t)(z'(t) - x'(t))\ dt + \int_0^1 \frac{1}{2}|z'(t) - x'(t)|^2 dt$$

où $v \to \int_0^1 x'(t)v'(t)\ dt$ est continue dans H^1 et $\int_0^1 \frac{1}{2}|v'(t) - x'(t)|^2 dt = \|v - x\|_{H^1}\epsilon(v)$ où $\epsilon(v)$ tend vers 0 lorsque $\|v\|_{H^1}$ tend vers 0. Donc

$$dJ_1(x)(v) = \int_0^1 x'(t)v'(t)\ dt .$$

Montrons maintenant la différentiabilité de J_2. Soit $x \in H^1$ et $M := \|x\|_\infty + 1$. Comme $\partial h/\partial x$ est continue sur le compact $[0,1] \times [-M,M]$, $\partial h/\partial x$ est uniformément continue avec un module de continuité noté ω. On a alors, pour tout $z \in H^1$ avec $\|z\|_{H^1} \leq 1$,

$$J_2(x+z) = \int_0^1 h(t,x(t)+z(t))dt = J_2(x) + \int_0^1 \frac{\partial h}{\partial x}(t,\xi(t))z(t)dt$$

où $\xi(t)$ est un réel entre $x(t)$ et $x(t)+z(t)$. Comme $z \in H^1([0,1))$ avec $\|z\|_{H^1} \leq 1$, on a $\|z\|_\infty \leq \|z\|_{H^1} \leq 1$ et donc $\|x+z\|_\infty \leq \|x\|_\infty + 1 \leq M$. Donc

$$\begin{aligned}\left|J_2(x+z) - J_2(x) - \int_0^1 \frac{\partial h}{\partial x}(t,x(t))z(t)dt\right| &\leq \int_0^1 \left|\frac{\partial h}{\partial x}(t,\xi(t)) - \frac{\partial h}{\partial x}(t,x(t))\right|\ |z(t)|\ dt \\ &\leq \int_0^1 \omega(|\xi(t) - x(t)|)|z(t)|\ dt \ \leq\ \omega(\|z\|_\infty)\|z\|_\infty\end{aligned}$$

D'où

$$\left|J_2(x+z) - J_2(x) - \int_0^1 \frac{\partial h}{\partial x}(t,x(t))z(t)dt\right| \leq \omega(\|z\|_{H^1})\|z\|_{H^1},$$

ce qui montre la différentiabilité de J_2. □

Théorème 3.12 *Soit x un minimum de J dans $E = \{y \in H^1 \mid y(0) = A,\ y(1) = B\}$. Alors x est de classe $\mathcal{C}^2$ et*

$$x''(t) = \frac{\partial h}{\partial x}(t,x(t)) \qquad \forall t \in [0,1] . \tag{4}$$

La condition (4) s'appelle la condition d'Euler.

Commençons par un lemme :

Lemme 3.13 (Dubois-Raymond) *Soit $\phi \in L^2$. On suppose que, pout tout $v \in H^1$ tel que $v(0) = v(1) = 0$, on a $\int_0^1 \phi(t)v'(t)\ dt = 0$. Alors il existe une constante $c \in \mathbb{R}$ telle que $\phi(t) = c$ pour presque tout $t \in [0,1]$.*

Preuve du Lemme 3.13. Posons $c = \int_0^1 \phi(s)ds$ et $v(t) = \int_0^t \phi(s)ds - ct$. Alors $v \in H^1$ et $v(0) = v(1) = 0$. De plus, $v' = \phi - c$ p.p.. Donc

$$\int_0^1 |\phi(t) - c|^2 dt = \int_0^1 (\phi(t) - c)v'(t)\ dt = \int_0^1 \phi(t)v'(t)\ dt - c\int_0^1 v'(t)\ dt = 0\ .$$

Comme $|\phi(t) - c|^2 \geq 0$, on a $|\phi(t) - c|^2 = 0$ p.p., i.e., $\phi(t) = c$ pour presque tout $t \in [0,1]$. □

Preuve du Théorème 3.12. Soit $y \in H^1$ tel que $y(0) = y(1) = 0$. Alors, pour tout $\lambda \in \mathbb{R}$, $x+\lambda y \in E$. Comme x est un minimum, on a $dJ(x)(y) = 0$, i.e.,

$$\int_0^1 x'(t)y'(t) + \frac{\partial h}{\partial x}(t, x(t))y(t)\ dt = 0\ .$$

En intégrant par parties, cette égalité se réécrit, en utilisant $\psi(t) = \int_0^t \frac{\partial h}{\partial x}(s, x(s))\ ds$,

$$\int_0^1 (x'(t) - \psi(t))y'(t)\ dt = 0\ .$$

Cette égalité étant vraie pour toute fonctions $y \in H^1$ tel que $y(0) = y(1) = 0$, le lemme de Dubois-Raymond dit qu'il existe une constante $c \in \mathbb{R}$ telle que $x'(t) - \psi(t) = c$ p.p. En particulier, comme ψ est de classe $\mathcal{C}^1$, x' possède un représentant de classe $\mathcal{C}^1$. Donc x est de classe $\mathcal{C}^2$. De plus, en dérivant l'égalité $x'(t) - \psi(t) = c$, on obtient l'équation d'Euler. □

4 Opérateurs compacts

4.1 Définitions et exemples

Soit H un espace de Hilbert. On note par $\mathcal{L}(H)$ l'espace vectoriel des applications linéaires continues de H dans H.

Définition 4.1 *On dit que $K \in \mathcal{L}(H)$ est compact si l'ensemble $K(B_1)$ est relativement compact, où B_1 est la boule unité fermée de H.*

Autrement dit K est compact si la fermeture de $K(B_1)$ est compacte. Noter que, par linéarité de K, l'image par K de tout ensemble borné est relativement compacte.

Voici un exemple typique :

Proposition 4.2 *Si $H = L^2([0,1])$ et $k : [0,1] \times [0,1] \to \mathbb{R}$ est continue, alors l'opérateur $K : H \to H$ défini par*

$$K(u)(x) = \int_0^1 k(x,y)u(y)dy \qquad \forall u \in H,\ x \in [0,1],$$

est compact.

Preuve. Comme $[0,1] \times [0,1]$ est compact, k est uniformément continue, de module de continuité ω. Donc, si $u \in B_1$,

$$\begin{aligned} |K(u)(x) - K(u)(x')| &\leq \int_0^1 |k(x,y) - k(x',y)||u(y)|dy \leq \omega(|x-x'|) \int_0^1 |u(y)|dy \\ &\leq \omega(|x-x'|)\|u\|_{L^2} \leq \omega(|x-x'|). \end{aligned}$$

Cela montre non seulement que $K(u)$ est continue, mais aussi que la famille $K(B_1)$ est équi-continue. De plus, pour tout $u \in B_1$ et pour tout $x \in [0,1]$,

$$|K(u)(x)| \leq \int_0^1 |k(x,y)|\ |u(y)|dy \leq \|k\|_\infty \int_0^1 |u(y)|dy \leq \|k\|_\infty \|u\|_2 \leq \|k\|_\infty.$$

La famille $K(B_1)$ est donc bornée. Par Ascoli, $K(B_1)$ est relativement compact dans $C^0([0,1])$. Comme la convergence dans $C^0([0,1])$ entraîne la convergence dans $L^2([0,1]) = H$, $K(B_1)$ est relativement compacte dans H et K est un opérateur compact. □

Bien noter que, si H est de dimension infinie, l'application identité de H n'est pas compacte, puisque la boule unité fermée de H n'est pas compacte.

Proposition 4.3 *Si $K_1, K_2 \in \mathcal{L}(H)$ sont des opérateurs compacts, $\lambda \in \mathbb{R}$, et $T \in \mathcal{L}(H)$, alors les opérateurs $K_1 + \lambda K_2$, $T \circ K_1$ et $K_1 \circ T$ sont compacts.*

Preuve. La preuve est immédiate. □

4.2 Alternative de Fredholm

Dans toute cette partie nous considérons des opérateurs obtenus comme perturbation compacte de l'identité. Plus précisément, nous fixons K un opérateur linéaire, continu et compact sur H et étudions les propriétés de l'opérateur $T = I - K$, où I est l'identité sur H. Il sera aisé de voir que les résultats s'appliquent aussi aux opérateurs de la forme $K - \lambda I$ si $\lambda \neq 0$. Attention, bien garder en tête que l'opérateur $T = I - K$ n'est pas compact dès H est de dimension infinie.

Proposition 4.4 *Le noyau $Ker(T)$ de T est de dimension finie.*

Preuve. Considérons B la boule unité (fermée) de $Ker(T)$. Comme T est continue, $Ker(T)$ est fermé, donc B est également fermée dans H. Par définition de T, on a $x = K(x)$ dans B. Comme K est compact et B borné, $K(B)$ est relativement compact. Mais $K(B) = B$ et donc B est fermé dans H et relativement compact : c'est un compact de H, et donc de $Ker(T)$. Par conséquent, la boule unité fermée de $Ker(T)$ est compacte. Le théorème de Riesz (théorème 1.4) affirme alors que $Ker(T)$ est de dimension finie. □

Proposition 4.5 *L'image $Im(T)$ de l'opérateur T est fermée.*

Preuve. Soit (y_n) une suite dans $Im(T)$ qui converge vers $y \in H$. On veut montrer que $y \in Im(T)$. Comme $y_n \in Im(T)$, il existe $x_n \in H$ tel que $T(x_n) = x_n - K(x_n) = y_n$. Notons par $d : H \to \mathbb{R}$ la distance à $Ker(T)$. Rappelons que d est lipschitzienne, donc continue. On affirme dans un premier temps que la suite $(d(x_n))$ est bornée. Sinon, il existerait une sous-suite telle que $(d(x_{n_k}))$ tend vers $+\infty$. Posons $\tilde{x}_{n_k} := x_{n_k}/d(x_{n_k})$. Comme $Ker(T)$ est un espace vectoriel, l'application d est positivement homogène et donc $d(\tilde{x}_{n_k}) = 1$. Par théorème de projection, il existe $v_{n_k} \in Ker(T)$ tel que $\|\tilde{x}_{n_k} - v_{n_k}\| = 1$. Mais alors, si on pose $z_{n_k} := \tilde{x}_{n_k} - v_{n_k}$, on a

$$z_{n_k} - K(z_{n_k}) = T(z_{n_k}) = T(\tilde{x}_{n_k}) = \frac{y_{n_k}}{d(x_{n_k})} \to 0.$$

Comme la suite (z_{n_k}) est bornée et K compact, la suite $(K(z_{n_k}))$ admet une sous-suite $(K(z_{n_{k'}}))$ qui converge vers une limite z, et donc $(z_{n_{k'}})$ converge vers la même limite z. Comme $Ker(T)$ est de dimension finie et $(v_{n_{k'}})$ bornée, la suite $(v_{n_{k'}})$ admet une sous-suite, notée également $(v_{n_{k'}})$ qui converge vers un élément noté $v \in Ker(T)$. Donc $(\tilde{x}_{n_{k'}}) = (z_{n_{k'}} + v_{n_{k'}})$ converge vers $z + v$. Par continuité, $(d(\tilde{x}_{n_{k'}}))$ tend vers $1 = d(z + v)$. Mais $(T(\tilde{x}_{n_{k'}}))$ tend vers 0, ce qui prouve que $T(z + v) = 0$ et donc $d(z + v) = 0$. Il y a donc une contradiction et la suite $(d(x_n))$ est bornée par une constante C.

Par conséquent il existe $v_n \in Ker(T)$ tel que $\|x_n - v_n\| \leq C$. Posons $z_n := x_n - v_n$. Alors la suite (z_n) est bornée et

$$z_n - K(z_n) = T(z_n) = T(x_n) = y_n.$$

Comme K est compact et (z_n) bornée, il existe une sous-suite $(K(z_{n'}))$ qui converge. Par conséquent, $(z_{n'})$ converge également vers une limite z et on a $T(z) = y$. □

Proposition 4.6 *La suite (croissante) $(Ker(T^n))$ et la suite (décroissante) $Im(T^n)$ sont stationnaires (i.e., constantes à partir d'un certain rang).*

Remarque : l'opérateur T^n est également une perturbation compacte de l'identité ; en effet la formule du binôme implique que $T^n = I + K \circ \sum_{k=1}^{n} C_n^k K^{k-1}(-1)^{n-k}$, qui est de la forme $I - \tilde{K}$ avec $\tilde{K}$ compact. En particulier, $Ker(T^n)$ est de dimension finie et $Im(T^n)$ est fermé.

Preuve. On raisonne par l'absurde en supposant que la suite $Ker(T^n)$ est strictement croissante. Alors pour tout $n \in \mathbb{N}$ il existe $x_{n+1} \in Ker(T^{n+1}) \backslash Ker(T^n)$ tel que $\|x_{n+1}\| = 1$ et $dist(x_{n+1}, Ker(T^n)) = 1$ (il suffit de prendre x_{n+1} dans la sphere unité d'un complémentaire orthogonal de $Ker(T^n)$ dans $Ker(T^{n+1})$). Notons que pour tout $0 \leq n < m$, le vecteur $T(x_m) + K(x_n)$ est dans $Ker(T^{m-1})$ puisque (K et T commutant entre eux)

$$T^{m-1}(T(x_m) + K(x_n)) = T^m(x_m) + K(T^{m-1}(x_n)) = 0 + K(0) = 0.$$

Donc

$$\|K(x_m) - K(x_n)\| = \|x_m - (T(x_m) + K(x_n))\| \geq dist(x_m, Ker(T^{m-1})) \geq 1.$$

On en déduit que la suite $(K(x_n))$ ne possède aucune suite extraite convergente, ce qui contredit la compacité de K puisque la suite (x_n) est bornée.

La preuve de la stationnarité de $Im(T^n)$ se montre de façon similaire. □

Proposition 4.7 *T est injective, si et seulement si, T est surjective.*

Preuve. Supposons que T soit injective, mais pas surjective. Alors $Im(T) \neq H$. Montrons que $Im(T^{n+1}) \neq Im(T^n)$ pour tout n. Pour cela on raisonne par l'absurde en supposant qu'il existe un plus petit indice n tel que $Im(T^{n+1}) = Im(T^n)$. Alors la restriction $\tilde{T}$ de T à $Im(T^n)$ est un opérateur compact surjectif, puisque pour tout $z \in Im(T^n) = Im(T^{n+1})$, il existe $x \in H$ tel que $T^{n+1}(x) = T(T^n(x)) = z$ où $T^n(x) \in Im(T^n)$ et donc $z \in Im(\tilde{T})$. Comme $Im(T^{n-1}) \neq Im(T^n)$, il existe $z \in Im(T^{n-1}) \backslash Im(T^n)$. Mais $T(z)$ est dans $Im(T^n)$ et par surjectivité de $\tilde{T}$ il existe $x \in Im(T^n)$ tel que $T(x) = T(z)$. Comme $x \neq z$ cela contredit l'injectivité de T. La suite $Im(T^n)$ est donc strictement décroissante, ce qui est impossible d'après la proposition 4.6.

Supposons maintenant que T est surjective mais pas injective. Montrons que la suite $Ker(T^n)$ est strictement croissante. Pour cela on raisonne par l'absurde et on suppose qu'il existe un plus petit indice n tel que $Ker(T^n) = Ker(T^{n+1})$. Soit $y \in Ker(T^n) \backslash Ker(T^{n-1})$. Alors par surjectivité de T il existe $x \in H$ tel que $T(x) = y$. Comme $y \in Ker(T^n)$, $T^{n+1}(x) = T^n(y) = 0$, ce qui prouve que $x \in Ker(T^{n+1}) = Ker(T^n)$. Donc $0 = T^n(x) = T^{n-1}(y)$ et on a une contradiction avec l'hypothèse $y \notin Ker(T^{n-1})$. La suite $Ker(T^n)$ est donc strictement croissante, ce qui est impossible d'après la proposition 4.6. □

En résumé nous avons démontré le résultat suivant :

Théorème 4.8 *Soit K un opérateur linéaire, continu et compact sur H. Posons $T = I - K$, où I est l'identité sur H. Alors :*

(i) Le noyau $Ker(T)$ de T est de dimension finie.

(ii) L'image $Im(T)$ de l'opérateur T est fermée.

(iii) Le noyau $Ker(T)$ est réduit au singleton $\{0\}$ si et seulement si l'image $Im(T)$ est égale à l'espace H.

Le théorème ci-dessus s'appelle alternative de Fredholm, au sens où

(i) soit l'équation $y = T(x)$ possède une unique solution x pour tout $y \in H$,

(ii) soit l'équation $y = T(x)$ ne possède pas de solution pour certains $y \in H$, si la solution existe, elle n'est alors pas unique.

Ces propriétés sont familières : ce sont les mêmes qu'en dimension finie! Soulignons qu'elles sont fausses pour des opérateurs quelconques. Par exemple, si $H = \ell^2$ et T est le shift à gauche : $T((x_n)_{n\geq 0}) = (x_{n+1})_{n\geq 0}$, alors $Ker(T)$ est de dimension 1 alors que T est surjective.

4.3 Diagonalisation des opérateurs auto-adjoints compacts

Soit $T \in \mathcal{L}(H)$. On dit que λ appartient à l'ensemble résolvant de T si $(T - \lambda I)^{-1}$ est un opérateur borné. Le spectre de T, noté $\sigma(T)$, est le complémentaire de l'ensemble résolvant. On dit que λ est une valeur propre de T (on note $\lambda \in VP(K)$) s'il existe $v \in H$, $v \neq 0$ tel que $T(v) = \lambda v$. En dimension finie, spectre et valeurs propres coïncident. Il n'en est pas de même en dimension infinie, où un opérateur peut être injectif mais non surjectif et inversement.

Dans toute la suite, H est un espace de Hilbert de dimension infinie.

On dit qu'un opérateur $T \in \mathcal{L}(H)$ est auto-adjoint si $T^* = T$.

Proposition 4.9 *Soit $T \in \mathcal{L}(H)$. Posons*

$$m := \inf_{\|x\|=1} \langle T(x), x\rangle \qquad \text{et} \qquad M := \sup_{\|x\|=1} \langle T(x), x\rangle.$$

Alors $\sigma(T) \subset [m, M]$. Si de plus T est auto-adjoint, alors m et M appartiennent à $\sigma(T)$.

Preuve. Soit $\theta < m$ et $a(x,y) := \langle T(x) - mx, y\rangle$. Alors, par définition de m, a est une forme bilinéaire continue coercive sur $H \times H$ puisque

$$a(x,x) = \langle T(x), x\rangle - \theta\|x\|^2 \geq (m-\theta)\|x\|^2$$

avec $\theta - m > 0$. Par Lax-Milgram, $T - \theta I$ est bijectif et donc $\theta \notin \sigma(T)$. On montre de même que $\theta > M$ implique que $\theta \notin \sigma(T)$, ce qui conclut le fait que $\sigma(T) \subset [m, M]$.

Supposons maintenant que T soit auto-adjoint. Soit $a(x,y) := \langle T(x) - mx, y\rangle$. Par définition de m, a est une forme bilinéaire symétrique continue positive sur $H \times H$. Par Cauchy-Schwartz,

$$|a(x,y)| \leq |a(x,x)|^{\frac{1}{2}}|a(y,y)|^{\frac{1}{2}} \qquad \forall x, y \in H,$$

ce qui implique que, pour tout $x \in H$,

$$\begin{aligned} \|T(x) - mx\| &= \sup_{\|y\|\leq 1} a(x,y) \leq \sup_{\|y\|\leq 1} |a(x,x)|^{\frac{1}{2}}|a(y,y)|^{\frac{1}{2}} \\ &\leq \langle T(x) - mx, x\rangle^{\frac{1}{2}}(\|T\| + |m|). \end{aligned}$$

Par définition de m, il existe $x_n \in H$ avec $\|x_n\| = 1$ et $\langle T(x_n), x_n\rangle \to m$, soit $\langle T(x_n) - mx_n, x_n\rangle \to 0$. Donc, d'après l'inégalité ci-dessus, $T(x_n) - mx_n \to 0$. Si m appartenait à l'ensemble résolvant de T, on aurait $x_n := (T - mI)^{-1}(T(x_n) - mx_n) \to 0$, ce qui est impossible puisque $\|x_n\| = 1$ pour tout $n \in \mathbb{N}$. Donc $m \in \sigma(T)$. La preuve pour M est similaire. □

Théorème 4.10 *Soit $K \in \mathcal{L}(H)$ un opérateur auto-adjoint compact. Alors $0 \in \sigma(K)$ et tout élément non nul de $\sigma(K)$ est une valeur propre de K. Les espaces propres de K associés aux valeurs propres non nulles (i.e., $Ker(K - \lambda I)$ avec $\lambda \in VP(K)\backslash\{0\}$) sont de dimension finie. Les espaces propres de K sont également deux à deux orthogonaux. De plus, si $\sigma(K)\backslash\{0\}$ n'est pas fini, alors son seul point d'accumulation est 0.*

Remarque : sauf pour l'orthogonalité des espaces propres, le résultat reste vrai (avec une preuve légèrement plus délicate) lorsque T n'est pas auto-adjoint.

Preuve. Si K était bijectif, alors $I = K \circ K^{-1}$ serait compact (rappelons que si K est bijective, alors K^{-1} est continue), ce qui contredit le fait que H est de dimension infinie. Donc $0 \in \sigma(K)$.

Soit $\lambda \in \sigma(K)\backslash\{0\}$. Comme, par définition, $I - \frac{1}{\lambda}K$ n'est pas bijective, soit $I - \frac{1}{\lambda}K$ n'est pas injective, soit $I - \frac{1}{\lambda}K$ n'est pas surjective. Par compacité de K, la seconde alternative entraîne la première (Proposition 4.7) et $\lambda \in VP(K)$.

Comme K est compact, la proposition 4.4 affirme que, si $\lambda \in VP(K)\backslash\{0\}$, l'espace propre $Ker(K - \lambda I) = Ker(I - \frac{1}{\lambda}K)$ est de dimension finie. Si λ, μ sont deux valeurs propres distinctes et si v, w sont des valeurs propres associées, alors

$$\lambda\langle v, w\rangle = \langle K(v), w\rangle = \langle v, K(w)\rangle = \mu\langle v, w\rangle.$$

Comme $\lambda \neq \mu$, on a donc $\langle v, w\rangle = 0$: les espaces propres associés à λ et μ sont donc orthogonaux.

Pour prouver la dernière assertion, il suffit de montrer que, pour tout $\epsilon > 0$, K ne possède qu'un nombre fini de valeurs propres (distinctes) de valeur absolue supérieure à ϵ. En effet si (λ_n) est une famille infinie de valeurs propres de K distinctes avec $|\lambda_n| \geq \epsilon$, alors il existe $v_n \in H$ avec $\|v_n\| = 1$ et $K(v_n) = \lambda_n v_n$. Comme $v_n \perp v_m$ si $n \neq m$, on a

$$\|K(v_n) - K(v_m)\|^2 = \|\lambda_n v_n - \lambda_m v_m\|^2 = \lambda_n^2\|v_n\|^2 + \lambda_m^2\|v_m\|^2 \geq 2\epsilon^2.$$

La suite $(K(v_n))$ ne possède donc aucune sous-suite convergente, ce qui contredit la compacité de K puisque la suite (x_n) est bornée. Donc le seul point d'accumulation de $\sigma(K)$ est 0. □

Corollaire 4.11 *Soit $K \in \mathcal{L}(H)$ un opérateur auto-adjoint compact. Si H est séparable, alors il existe une base hilbertienne de H formée de vecteurs propres de K.*

Preuve. Soit (λ_n) la suite des valeurs propres (non nulles) de K. Soit E l'espace vectoriel engendré par $Ker(K)$ et par les espaces propres de K. Notons que, par définition, E est stable par K. Montrons que E est dense dans H. En effet, sinon, considérons la restriction $\tilde{K}$ de K à $E^\perp$. Pour tout $x \in E^\perp$ et pour tout $y \in E$, on a

$$\langle K(x), y\rangle = \langle x, K(y)\rangle = 0$$

puisque $K(y) \in E$ et $x \in E^\perp$. Donc $\tilde{K}$ va de $E^\perp$ dans lui-même et est un opérateur auto-adjoint compact sur $E^\perp$. De plus, par définition de E, $\tilde{K}$ est injectif. Comme $\tilde{K}$ est auto-adjoint et non nulle, la forme quadratique $x \to \langle \tilde{K}(x), x\rangle$ n'est pas identiquement nulle sur $E^\perp$. La proposition 4.9 affirme alors que $\tilde{K}$ admet une valeur singulière non nulle qui est, d'après le théorème 4.10, une valeur propre de $\tilde{K}$. Cette valeur propre est également une valeur propre de K, ce qui contredit la définition de E. En conclusion, E est dense dans H.

Comme H est séparable, on peut trouver une base hilbertienne de $Ker(K)$, que l'on peut compléter par la suite des bases des espaces propres associés à des valeurs propres non nulles (rappelons qu'ils sont de dimension finie). La famille ainsi obtenue est orthonormale car les espaces propres sont orthogonaux deux à deux. De plus, elle est dense puisque E est dense. C'est donc une base hilbertienne de H. □

4.4 Notes

Le prototype d'opérateur compact est $(-\Delta)^{-1}$: plus précisément, soit Ω un ouvert borné de $\mathbb{R}^N$ dont le bord est "régulier". L'opérateur $(-\Delta)^{-1}$ est l'application $L^2(\Omega)$ dans $L^2(\Omega)$ qui à tout $f \in L^2(\Omega)$ associe la solution u de

$$\begin{cases} -\Delta u(x) = f(x) \text{ dans } \Omega \\ u(x) = 0 \text{ sur } \partial\Omega \end{cases}$$

On peut obtenir u en utilisant le théorème de Lax-Milgram (et l'inégalité de Poincaré). Le fait que l'opérateur soit compact repose sur la compacité de la boule unité de $H^1(\Omega)$ dans $L^2(\Omega)$ (Théorème de Rellich-Kondrachov). L'opérateur est également auto-adjoint (faire une intégration par parties). Par conséquent $L^2(\Omega)$ possède une base hilbertienne composée de valeurs propres du laplacien. Cette base hilbertienne est par exemple utilisée dans la méthode de Galerkin pour construire des solutions de l'équation de la chaleur

$$\begin{cases} \partial_t u(t,x) - \Delta u(t,x) = 0 \text{ dans } \Omega \\ u(t,x) = 0 \text{ sur } \partial\Omega \\ u(0,x) = u_0(x) \text{ dans } \Omega \end{cases}$$